Stochastic Processes

Winter Term 2016-2017

Paolo Di Tella
Technische Universität Dresden
Institut für Stochastik

Contents

CHAPTER 1

Preliminaries

In this chapter we collect some of the results of measure theory needed for this lecture notes. In the appendix we summarize well-known result of measure and integration theory, which are assumed to be known to the reader.

1.1 Uniform integrability

Let $(\Omega, \mathscr{F}, \mathbb{P})$ a Probability space and $L^1 = L^1(\mathbb{P}) := L^1(\Omega, \mathscr{F}, \mathbb{P})$.

1.1.1 Exercise. $X \in L^1$ if and only if $\lim_{N\to\infty} \int_{\{|X|\geq N\}} |X| \mathrm{d}\mathbb{P} = 0$.

Exercise 1.1.1 justifies the following definition:

1.1.2 Definition. (i) We say that a family $\mathscr{K} \subseteq L^1$ is uniformly integrable if

$$\sup_{X\in\mathscr{K}} \mathbb{E}[|X| 1_{\{|X|\geq N\}}] \longrightarrow 0 \quad \text{as} \quad N \to +\infty.$$

(ii) A sequence $(X_n)_{n\in\mathbb{N}} \subseteq L^1$ is uniformly integrable if $\mathscr{K} := \{X_1, \ldots\}$ is uniformly integrable.

If $\mathscr{K}$ is dominated in L^1, i.e., there exists $Y \in L^1$ such that $|X| \leq Y$, for every $X \in \mathscr{K}$, then $\mathscr{K}$ is uniformly integrable. Clearly, any finite family of integrable random variables is uniformly integrable. However, if a family $\mathscr{K}$ is uniformly integrable, this does not mean that it is also dominated in L^1. Therefore the following results, for which we refer to Dellacherie & Meyer (1978), Theorem II.21, is a generalization of Lebesgue theorem on dominated convergence (cf. Theorem 1.A.3).

1.1.3 Theorem. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables in L^1 converging almost surely (or only in probability) to a random variable X. Then the following assertions are equivalent:*

(i) *X belongs to L^1 and $(X_n)_{n\in\mathbb{N}}$ converges in L^1 to X.*

(ii) *The family $\mathscr{K} := (X_n)_{n\in\mathbb{N}}$ is uniformly integrable.*

1.1.4 Corollary. *Let $(X_n)_{n\in\mathbb{N}} \subseteq L^1$ be uniformly integrable and converge to X a. s. Then $X \in L^1$ and $\mathbb{E}[X_n]$ converges to $\mathbb{E}[X]$, whenever $n \to +\infty$.*

Proof. Exercise. □

The following result, known as La Vallée–Poussin's Theorem (cf. Dellacherie & Meyer (1978), Theorem II.22), is a characterisation of the uniform integrability

1.1.5 Theorem. *A family $\mathscr{K}$ of integrable random variables is uniformly integrable if and only if there exists a positive, convex and increasing function G on $[0,+\infty]$ into $[0,+\infty)$ such that:* (i) $\frac{G(x)}{x}$ *converges to $+\infty$ as $x \to +\infty$;* (ii) $\sup_{X\in\mathscr{K}} \mathbb{E}[G(|X|)] < +\infty$.

Because of La Vallée–Poussin's Theorem, we can conclude that a family of random variables is uniformly integrable if it is q-integrable and bounded in $L^q(\mathbb{P})$, $q > 1$.

1.2 Monotone class theorems

Monotone class theorems are of different kinds and they are present in the literature in several formulations. We consider only one formulation for sets and one for functions. We refer to Sharpe (1988) and He, Wang & Yan (1992). We start with a monotone class theorem for systems of sets in the same form as He, Wang & Yan (1992), Theorem 1.2. We say that a class $\mathscr{K}$ of subsets of Ω is a *monotone class* if for every monotone sequence $(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{K}$ such that $A_n \uparrow A$ or $A_n \downarrow A$ as $n \to +\infty$, $A \in \mathscr{K}$.

1.2.1 Theorem (Monotone class theorem for sets)**.** *Let $\mathscr{F}$ be an algebra and $\mathscr{K}$ a monotone class of sets of Ω such that $\mathscr{F} \subseteq \mathscr{K}$. Then $\sigma(\mathscr{F}) \subseteq \mathscr{K}$.*

For the formulation of the monotone class theorem for classes of functions we refer to Sharpe (1988), Appendix A0. Let $(\Omega, \mathscr{F})$ be a measurable space. We denote by $\mathbb{B} := \mathbb{B}(\Omega, \mathbb{R})$ the set of bounded measurable functions on $(\Omega, \mathscr{F})$ into $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. If $\mathscr{K}$ is a linear subspace of $\mathbb{B}$ we say that it is a *monotone vector space* if $1 = 1_\Omega \in \mathscr{K}$ and if it is monotonically closed, that is, if $(f_n)_{n\in\mathbb{N}} \subseteq \mathscr{K}$ is such that $0 \leq f_n \leq f_{n+1}$, for all $n \in \mathbb{N}$ and $f \in \mathbb{B}$ is such that $f = \lim_{n\to+\infty} f_n$, then $f \in \mathscr{K}$. We observe that the limit f belongs to $\mathbb{B}$ if and only if $(f_n)_{n\in\mathbb{N}}$ is uniformly bounded. A set $\mathscr{C} \subseteq \mathbb{B}$ is called a *multiplicative class* if it is closed with respect to the multiplication of two elements, meaning that if h and g belong to $\mathscr{C}$, then also their product hg does.

1.2.2 Theorem (Monotone class theorem for functions)**.** *Let $\mathscr{K}$ be a monotone vector space and $\mathscr{C}$ a multiplicative class such that $\mathscr{C} \subseteq \mathscr{K}$ and $\sigma(\mathscr{C}) = \mathscr{F}$. Then $\mathscr{K} = \mathbb{B}$.*

Let $(\Omega, \mathscr{F})$ be a measurable space and μ a measure on it. Given a system $\mathscr{T}$ of numerical functions on $(\Omega, \mathscr{F})$, we denote by $\mathrm{Span}(\mathscr{T})$ its linear hull. Now we fix q in $[1, +\infty)$. If $\mathscr{T} \subseteq L^q(\mu)$ we denote by $\overline{\mathscr{T}}^{(L^q(\mu),\|\cdot\|_q)}$ the closure of $\mathscr{T}$ in $(L^q(\mu), \|\cdot\|_q)$. A system $\mathscr{T} \subseteq L^q(\mu)$ of functions is called *total* in $(L^q(\mu), \|\cdot\|_q)$ if its linear hull is dense, that is, $\overline{\mathrm{Span}(\mathscr{T})}^{(L^q(\mu),\|\cdot\|_q)} = L^q(\mu)$. If μ is a finite measure, then the inclusions $L^q(\mu) \subseteq L^p(\mu)$, $q \geq p$, hold and $L^q(\mu)$ is a total system in $(L^p, \|\cdot\|_p)$. In particular, $L^\infty(\mu)$ is an example of total system in $L^q(\mu)$, for every $q \in [1, +\infty)$. As an application of Theorem 1.2.2, we want to establish a general lemma stating sufficient conditions for a system $\mathscr{T} \subseteq L^q(\mu)$ of bounded functions to be total in $L^q(\mu)$, for $q \in [1, +\infty)$. We recall that we use the notation $\mathbb{B} := \mathbb{B}(\Omega, \mathbb{R})$ to denote the space of bounded measurable functions on $(\Omega, \mathscr{F})$ into $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

1.2.3 Lemma. *Let $\mathscr{T} \subseteq L^q(\mu)$ be a subset of $\mathbb{B}$. Then $\mathscr{T}$ is total in $L^q(\mu)$ if the following conditions are satisfied:*
(i) *$\mathscr{T}$ is stable under multiplication;*
(ii) *$\sigma(\mathscr{T}) = \mathscr{F}$;*
(iii) *There exists a sequence $(h_n)_{n\in\mathbb{N}} \subseteq \mathrm{Span}(\mathscr{T})$ such that $h_n \geq 0$ and $h_n \uparrow 1$ pointwise as $n \to +\infty$.*

Proof. First, we consider the case in which μ is a finite measure. In this case $\mathbb{B} \subseteq L^q(\mu)$ and it is dense in $(L^q(\mu), \|\cdot\|_q)$. We define $\mathscr{H} := \overline{\mathrm{Span}(\mathscr{T})}^{(L^q(\mu),\|\cdot\|_q)}$ and then $\mathscr{K} := \mathscr{H} \cap \mathbb{B}$. Clearly, $\mathscr{K}$ is a closed linear space and $\mathscr{T} \subseteq \mathscr{K}$. By assumption, $h_n \uparrow 1$ pointwise as $n \to +\infty$ and by the finiteness of μ, $1 \in L^q(\mu)$. From the theorem of Lebesgue on dominated convergence (cf. Theorem 1.A.3) h_n converges to 1 in $L^q(\mu)$ and so $1 \in \mathscr{K}$. Moreover, $\mathscr{K}$ is closed under monotone convergence of uniformly bounded nonnegative functions, as a consequence of the finiteness of μ and of Theorem 1.A.3. Consequently, $\mathscr{K}$ is a monotone class and by Theorem 1.2.2, we get $\mathscr{K} = \mathbb{B}$. Hence $\mathbb{B} \subseteq \mathscr{H}$ and since $\mathbb{B}$ is dense and $\mathscr{H}$ is closed, this yields $L^q(\mu) = \mathscr{H}$. Now we consider the case of a *general* measure μ. For $f \in L^q(\mu)$ we put $\mathrm{d}\mu_n := |h_n|^q \,\mathrm{d}\mu$, $n \geq 1$, where $(h_n)_{n\in\mathbb{N}}$ is as in the assumptions of the lemma. Obviously, μ_n is a finite measure for every n. Moreover, with notation in (A.1) below, $\mu_n(|f|^q) = \mu(|fh_n|^q) \leq \mu(|f|^q) < +\infty$ and hence $f \in L^q(\mu_n)$. We choose a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ such that $\varepsilon_n > 0$ and $\varepsilon_n \downarrow 0$ as $n \to +\infty$. By the previous step, there exists a sequence $(g_n)_{n\in\mathbb{N}} \subseteq \mathrm{Span}(\mathscr{T})$ such that

$$\int_\Omega |f\,h_n - g_n\,h_n|^q \,\mathrm{d}\mu = \int_\Omega |f - g_n|^q |h_n|^q \,\mathrm{d}\mu < \varepsilon_n\,, \qquad n \geq 1.$$

The system $\mathscr{T}$ is stable under multiplication so $g_n h_n \in \mathrm{Span}(\mathscr{T})$. On the other side, $\|f - fh_n\|_{L^q(\mu)} = \||f|^q|1-h_n|^q\|_{L^1(\mu)} \longrightarrow 0$, as $n \to +\infty$. Indeed, $|1-h_n|^q \downarrow 0$ as $n \to +\infty$ and $|f|^q|1-h_n|^q \leq |f|^q \in L^1(\mu)$. Theorem 1.A.3 yields the result. Hence $g_n h_n$ converges to f in $L^q(\mu)$ as $n \to +\infty$. □

Notice that the measure μ in the Lemma must be finite.

1.3 Conditional expectation and probability

For this part we refer to Shiryaev (1996), II§7. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, $\mathscr{G}$ a sub-σ-algebra of $\mathscr{F}$ and $X \in L^1 := L^1(\Omega, \mathscr{F}, \mathbb{P})$.

1.3.1 Definition. A random variable $Y \in L^1$ is a version of the conditional expectation of X under the condition (or given) $\mathscr{G}$ if
(i) Y is $\mathscr{G}$-measurable;
(ii) $\int_B Y \mathrm{d}\mathbb{P} = \int_B X \mathrm{d}\mathbb{P}, \quad B \in \mathscr{G}$.

It is well known (cf. Bauer (1996)§15) that a version of the conditional expectation of X given $\mathscr{G}$ always exists and it is a. s. unique (the proof is based on Radon–Nykodim's theorem). If Y is a version of the conditional expectation of X given $\mathscr{G}$ we denote it by $\mathbb{E}[X|\mathscr{G}]$.

1.3.2 Exercise. (i) Let $X \in L^1$. Show that

$$\mathscr{K} := \{\mathbb{E}[X|\mathscr{G}],\ \mathscr{G} \subset \mathscr{F},\ \mathscr{G}\text{sub-}\sigma\text{-algebra}\} \subseteq L^1$$

is uniformly integrable.

(ii) Let Y be $\mathscr{G}$-measurable. Show that $Y = \mathbb{E}[X|\mathscr{G}]$ if and only if $\mathbb{E}[XW] = \mathbb{E}[YW]$, for every $W \geq 0$ bounded and $\mathscr{G}$-measurable.

Conditional probability. For a given measurable sets A, B such that $\mathbb{P}[B] \neq 0$, we have $\mathbb{P}[A|B] := \mathbb{P}[A \cap B]/\mathbb{P}[B]$.

1.3.3 Definition. We call the the function $\mathbb{P}[\cdot|\mathscr{G}] : \mathscr{F} \times \Omega \longrightarrow [0,1]$ defined by $\mathbb{P}[A|\mathscr{G}](\omega) := \mathbb{E}[1_A|\mathscr{G}](\omega)$, $A \in \mathscr{F}$, conditional probability.

Notice that for $A \in \mathscr{F}$, $\mathbb{P}[A|\mathscr{G}]$ is a. s. defined and is a $\mathscr{G}$-measurable random variable. It furthermore fulfils

$$\int_B \mathbb{P}[A|\mathscr{G}]\mathrm{d}\mathbb{P} = \int_B 1_A \mathrm{d}\mathbb{P} = \mathbb{P}[A \cap B], \quad B \in \mathscr{G}.$$

There is an important question connected with the conditional probability. $\mathbb{P}[A|\mathscr{G}]$ is an equivalence class: can we choose a representing element $\widetilde{\mathbb{P}[A|\mathscr{G}]}$ from each equivalence class in such a way that the set function $A \longrightarrow \widetilde{\mathbb{P}[A|\mathscr{G}]}(\omega)$, is a probability measure on $(\Omega, \mathscr{F})$, for every fixed ω? Let $(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{F}$ be a sequence of pairwise disjoint sets. Then because of the property of the conditional expectation we get

$$\mathbb{P}\left[\bigcup_{n=1}^{\infty} A_n |\mathscr{G}\right] = \sum_{n=1}^{\infty} \mathbb{P}[A_n|\mathscr{G}]$$

in the sense of the equivalence classes. However we cannot hope that this will also holds for the representing element $\widetilde{\mathbb{P}[A|\mathscr{G}]}(\omega)$, for each $\omega \in \Omega$ if we are not able to chose this representing element in a very special way.

1.3.4 Definition. A regular version of the conditional probability $\mathbb{P}[\cdot|\mathscr{G}]$ is a mapping $P_{\mathscr{G}}$ on $\mathscr{F} \times \Omega$ such that:

(i) $A \longrightarrow P_{\mathscr{G}}(A, \omega)$ is a probability measure on $(\Omega, \mathscr{F})$ a. s.;

(ii) $\omega \longrightarrow P_{\mathscr{G}}(A, \omega)$ is $\mathscr{G}$-measurable and belongs to the equivalence class of $\mathbb{P}[A|\mathscr{G}]$, for every $A \in \mathscr{F}$.

Of course, if $P_{\mathscr{G}}$ is a regular version of the conditional probability $\mathbb{P}[\cdot|\mathscr{G}]$, then

$$P_{\mathscr{G}}(A, \cdot) = \mathbb{E}[1_A|\mathscr{G}].$$

In particular this yields that

$$\mathbb{E}[X|\mathscr{G}] = \int_\Omega X(\omega')P_{\mathscr{G}}(\mathrm{d}\omega', \cdot),$$

that is we can compute the conditional expectation of X given $\mathscr{G}$ "as the non-conditional" expectation.

For a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ we can always consider the conditional probability $\mathbb{P}[\cdot|\mathscr{G}]$. Therefore it is natural to ask the question if it is also always possible to find a regular version $P_{\mathscr{G}}$ of the conditional probability. As the short discussion before Definition 1.3.4 explains, this is in general not true. However, as we shall see in the next sections, there exist special situations in which this is indeed the case.

Appendix 1.A Measure and integration theory

We consider an arbitrary nonempty set Ω. If $A \subseteq \Omega$ we denote by A^c the complement of A in Ω. Let $(A_n)_{n\geq 1}$ be a sequence of subsets of Ω and $A \subseteq \Omega$. If $A_n \subseteq A_{n+1}$, $n \geq 1$, and $A = \cup_{n=1}^{\infty} A_n$, we write $A_n \uparrow A$. If $A_{n+1} \subseteq A_n$, $n \geq 1$, and $A = \cap_{n=1}^{\infty} A_n$, we write $A_n \downarrow A$.

A system $\mathscr{R}$ of subsets of Ω is called a *semiring* of subsets of Ω if it possesses the following properties: The empty set belongs to $\mathscr{R}$; if A and B belong to $\mathscr{R}$, then their intersection $A \cap B$ does; if A and B belong to $\mathscr{R}$ and $A \subseteq B$, then the set-difference $B \setminus A$ can be written as finite union of pairwise disjoint elements of $\mathscr{R}$. A system $\mathscr{R}$ of subsets of Ω with the following properties is called a *ring*: The empty set belongs to $\mathscr{R}$; if A and B belong to $\mathscr{R}$, then their union $A \cup B$ and their set-difference $A \setminus B$ do. Notice that a ring contains also the intersection of two of its elements because $A \cap B = A \setminus (A \setminus B)$. Obviously a ring is also a semiring.

1.A.1 Definition. A system $\mathscr{F}$ of subsets of Ω is called an *algebra* (in Ω) if it has the following properties:

(i) $\Omega \in \mathscr{F}$;
(ii) if $A \in \mathscr{F}$, then $A^c \in \mathscr{F}$;
(iii) if $A, B \in \mathscr{F}$, then $A \cup B \in \mathscr{F}$.

If (iii) is replaced by

(iii′) if $(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{F}$, then $\cup_{n\in\mathbb{N}} A_n \in \mathscr{F}$,

then $\mathscr{F}$ is denominated a *σ-algebra* (in Ω).

We notice that an algebra is a ring that in addition contains Ω. If $\mathscr{C} \subseteq \Omega$ is a system of sets, the *σ-algebra generated* by $\mathscr{C}$ is denoted by $\sigma(\mathscr{C})$ and is defined as the smallest σ-algebra containing $\mathscr{C}$. If $\mathscr{C} \subseteq \mathscr{F}$ is such that $\sigma(\mathscr{C}) = \mathscr{F}$ we say that $\mathscr{C}$ generates $\mathscr{F}$ and call it a *generator* of $\mathscr{F}$. If $\mathscr{C}$ is a generator of $\mathscr{F}$ which is stable under intersection of two sets, we call it an *∩-stable generator*. If the σ-algebra $\mathscr{F}$ can be generated by a countable system $\mathscr{C}$, we say that it is a *separable* σ-algebra. Let $(\mathscr{C})_{i\in I}$ be a family of systems of subsets in Ω, where I is an arbitrary set of indexes. By $\bigvee_{i\in I} \mathscr{C}_i$ we denote the σ-algebra generated by the union of all the $\mathscr{C}_i$s, that is, $\bigvee_{i\in I} \mathscr{C}_i := \sigma(\bigcup_{i\in I} \mathscr{C}_i)$. Let Ω be a topological space. We denote by $\mathscr{B}(\Omega)$ the *Borel* σ-algebra on Ω, i.e., the σ-algebra generated in Ω by the open sets in the topology of Ω. If, for example, $\Omega = \mathbb{R}$, then $\mathscr{B}(\mathbb{R})$ is separable.

For any σ-algebra $\mathscr{F}$ of Ω, we call the couple $(\Omega, \mathscr{F})$ a *measurable space* and we say that the subsets of Ω which belong to $\mathscr{F}$ are *$\mathscr{F}$-measurable* or simply *measurable*. We consider two measurable spaces $(\Omega, \mathscr{F})$ and $(\Omega', \mathscr{F}')$ and a function f from Ω into Ω'. We say that f is *$(\mathscr{F},\mathscr{F}')$-measurable* or simply *measurable*, if for any $A' \in \mathscr{F}'$ the set $f^{-1}(A') := \{a \in \Omega : f(a) \in A'\}$ is $\mathscr{F}$-measurable. We call the set $f^{-1}(A')$ the inverse image of A' by f. If f is a function on $(\Omega, \mathscr{F})$ into $(\Omega', \mathscr{F}')$ the system of sets $f^{-1}(\mathscr{F}') := \{f^{-1}(A') : A' \in \mathscr{F}'\}$ is a σ-algebra in Ω. Let $\{(\Omega_i, \mathscr{F}_i) : i \in I\}$ be a family of measurable spaces and $\{f_i : i \in I\}$ be a family of functions on Ω such that f_i takes values in Ω_i, for every $i \in I$. The σ-algebra in Ω generated by $\bigcup_{i\in I} f_i^{-1}(\mathscr{F}_i)$ is the smallest σ-algebra $\mathscr{F}'$ with respect to which every f_i is $(\mathscr{F}', \mathscr{F}_i)$-measurable. We designate this σ-algebra by $\sigma(f_i : i \in I)$, that is, $\sigma(f_i : i \in I) := \bigvee_{i\in I} f_i^{-1}(\mathscr{F}_i)$ and we call it the *σ-algebra generated by* $\{f_i : i \in I\}$.

Let $(\Omega, \mathscr{F})$ be a measurable space. A set-function μ on $\mathscr{F}$ into $[0, +\infty]$ such that $\mu(\emptyset) = 0$ and that $\mu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$, for any sequence $(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{F}$ of pairwise-disjoint sets, is called a *measure* on $(\Omega, \mathscr{F})$. If μ takes values in $[-\infty, +\infty]$, then it is called a *signed* measure. If μ is a measure on $(\Omega, \mathscr{F})$, we say that $(\Omega, \mathscr{F}, \mu)$ is a *measure space*. A measure μ such that $\mu(\Omega) < +\infty$ is called a *finite measure*. A *probability measure* is a finite measure μ such that $\mu(\Omega) = 1$. If there exists an increasing sequence $(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{F}$ such that $\mu(A_n) < +\infty$ for every $n \in \mathbb{N}$ and $\bigcup_{n\in\mathbb{N}} A_n = \Omega$, then the measure μ is called *σ-finite*. If Ω is a Hausdorff space with σ-algebra $\mathscr{B}(\Omega)$, we say that μ is *locally finite* if every point of Ω has an open neighbourhood of finite measure μ. The following result holds for σ-finite measures and it is well-known in the literature as *uniqueness theorem* (cf., e.g., Bauer (2001), Theorem I.5.4).

1.A.2 Theorem (Uniqueness theorem)**.** *Let $(\Omega, \mathscr{F})$ be a measurable space, $\mathscr{C} \subseteq \mathscr{F}$ an $\cap$-stable generator of $\mathscr{F}$ and $(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{C}$ satisfying the property $\bigcup_{n\in\mathbb{N}} A_n = \Omega$. We suppose that μ_1 and μ_2 are σ-finite measures on $\mathscr{F}$ such that*
(i) *$\mu_1(A) = \mu_2(A)$, for every $A \in \mathscr{C}$;*
(ii) *$\mu_1(A_n) = \mu_2(A_n) < +\infty$, for every $n \in \mathbb{N}$.*
Then μ_1 and μ_2 are identical on $\mathscr{F}$.

For a measure μ on the measurable space $(\Omega, \mathscr{F})$, it is well understood how to define the integral of a measurable function with values in $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. We introduce the notation

$$f * \mu := \int_\Omega f \,\mathrm{d}\mu := \int_\Omega f(x)\mu(\mathrm{d}x) \tag{A.1}$$

if the integral on the right-hand side exists. In particular, $\mu(f)$ is well defined if f is *nonnegative*. We say that a measurable function f of arbitrary sign is *μ-integrable* or simply *integrable* if $|f| * \mu < +\infty$. We do not go into details and we refer to Bauer (2001), Chapter II. By *functions*, if not otherwise specified, we mean functions with values in $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, that is *numerical functions*. Let f be a measurable function. By $\|f\|_q$ we denote the following norm

$$\|f\|_q := \begin{cases} (|f|^q * \mu)^{\frac{1}{q}}, & q \in [1, +\infty), \\ \\ \operatorname{ess\,sup}_{x\in\Omega} |f(x)|, & q = +\infty, \end{cases}$$

and we put

$$L^q(\mu) := \{f \text{ measurable} : \|f\|_q < +\infty\}, \quad q \in [1, +\infty].$$

We recall that $f \in L^q(\mu)$ is uniquely determined up to equivalence classes μ-a.e. Sometimes, we write $L^q(\Omega, \mathscr{F}, \mu)$ to stress the measure space and $\|\cdot\|_{L^q(\mu)}$ to stress the space $L^q(\mu)$. The space $(L^q(\mu), \|\cdot\|_q)$ is a Banach space. The space $L^2(\mu)$ is especially important because it is a Hilbert space with respect to the scalar product $(f, g)_{L^2(\mu)} := (fg) * \mu$. If $f, g \in L^2(\nu)$ are such that $(f, g)_{L^2(\mu)} = 0$, we say that they are orthogonal (in $L^2(\mu)$) and denote it by $f \perp g$. If $\mathscr{G}$ is a subset of functions in $L^2(\nu)$ and $f \in L^2(\nu)$ is such that $f \perp g$ for every $g \in \mathscr{G}$, we say that f is orthogonal (in $L^2(\mu)$) to $\mathscr{G}$ and we denote it by $f \perp \mathscr{G}$. For a *finite* measure μ, the inclusions $L^q(\mu) \subseteq L^p(\mu)$, $1 \le p \le q$, hold. In particular, $L^\infty(\mu)$ is contained in every $L^q(\mu)$, for $q \in [1, +\infty)$. However, these inclusions are not valid for a general measure μ.

A function f belonging to $L^1(\mu)$ is called *integrable*, while it is called *square integrable* if it belongs to $L^2(\mu)$. In general, we say that f is q-integrable if it belongs to $L^q(\mu)$, $q \in [1, +\infty)$. Let $(f_n)_{n\geq 1}$ be a sequence of measurable functions on the measure space $(\Omega, \mathscr{F}, \mu)$. We say that $(f_n)_{n\geq 1}$ converges (μ-a.e.) *pointwise* to the measurable function f if

$$\lim_{n\to+\infty} |f_n(x) - f(x)| = 0$$

for (μ-almost all) $x \in \Omega$. We write $f_n \longrightarrow f$ *pointwise* to mean that the sequence $(f_n)_{n\geq 1}$ converges pointwise to f. If the sequence $(f_n)_{n\geq 1}$ is monotonically increasing (resp., decreasing), i.e., $f_n \leq f_{n+1}$ (resp., $f_n \geq f_{n+1}$), we write $f_n \uparrow f$ (resp., $f_n \downarrow f$) to mean that it converges pointwise to f. If $(f_n)_{n\geq 1} \subseteq L^q(\mu)$ and $f \in L^q(\mu)$, we say that $(f_n)_{n\geq 1}$ converges to f in $L^q(\mu)$ if

$$\lim_{n\to+\infty} \|f_n - f\|_q = 0.$$

It is important to establish under which conditions a sequence $(f_n)_{n\geq 1} \subseteq L^q(\mu)$ converging a.e. to a measurable function f converges in fact to f in $L^q(\mu)$. Now we state two classical theorems which answer this question: The theorem of Lebesgue on dominated convergence and the theorem of B. Levi on monotone convergence. We refer to Bauer (2001) II§.11 and II§.15. The following is the theorem of Lebesgue on dominated convergence.

1.A.3 Theorem. *We fix $q \in [1, +\infty)$ and consider a sequence $(f_n)_{n\in\mathbb{N}} \subseteq L^q(\mu)$ such that $f_n \longrightarrow f$ μ-a.e. pointwise as $n \to +\infty$. If there exists a function $g \geq 0$ in $L^q(\mu)$ such that $|f_n| \leq g$, for every $n \in \mathbb{N}$, then $f \in L^q(\mu)$ and the convergence takes place also in $L^q(\mu)$.*

Now we state the theorem of B. Levi on monotone convergence.

1.A.4 Theorem. *Let $(f_n)_{n\in\mathbb{N}}$ be a monotone sequence of nonnegative functions such that $f_n \uparrow f$ pointwise as $n \to +\infty$. Then f is measurable and $f_n * \mu \uparrow f * \mu$ as $n \to +\infty$.*

Let $(\Omega, \tilde{\mathscr{F}})$ be a measurable space and let $\mathbb{P}$ be a probability measure on it. We call the measure space $(\Omega, \tilde{\mathscr{F}}, \mathbb{P})$ a *probability space*. By $\mathscr{N}(\mathbb{P})$ we denote the null sets of $\mathbb{P}$, i.e., $\mathscr{N}(\mathbb{P}) := \{A \subseteq \Omega : \exists B \in \tilde{\mathscr{F}},\ A \subseteq B,\ \mathbb{P}(B) = 0\}$. If $\mathscr{N}(\mathbb{P})$ is not contained in $\tilde{\mathscr{F}}$ we enlarge the σ-algebra by setting $\mathscr{F} := \tilde{\mathscr{F}} \vee \mathscr{N}(\mathbb{P})$. We call $\mathscr{F}$ the completion of $\tilde{\mathscr{F}}$ (*in itself*) with respect to $\mathbb{P}$ or simply $\mathbb{P}$-completion of $\tilde{\mathscr{F}}$ and we say that $(\Omega, \mathscr{F}, \mathbb{P})$ is a complete probability space. If not otherwise specified, we assume a probability space to be complete. In the remaining of this chapter we assume that a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is fixed. A measurable mapping X on $(\Omega, \mathscr{F})$ into $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ is called a *random variable*. We denote by $\mathbb{E}$ the expectation with respect to $\mathbb{P}$. If $\mathscr{G}$ is a sub-σ-algebra of $\mathscr{F}$, we denote by $\mathbb{E}[\cdot|\mathscr{G}]$ the conditional expectation with respect to $\mathscr{G}$. Sometimes we write $\mathbb{E}_{\mathbb{P}}$ or $\mathbb{E}_{\mathbb{P}}[\cdot|\mathscr{G}]$ to emphasize the dependence on the probability measure $\mathbb{P}$.

CHAPTER 2

Introduction to the general theory of stochastic processes

We start this chapter with the general definition of stochastic process and study the relation between stochastic processes and finite dimensional distribution of a stochastic process. In particular, we present a very important and deep result of Kolmogorov known as *Kolmogorov Extension Theorem*. Then we discuss some properties of paths of stochastic processes, we introduce the notion of filtration, stopping times and martingales. We then consider Markov processes and show their existence as an application of the Kolmogorov Extension Theorem. We conclude the chapter with a section about processes with independent increments with respect to a filtration and, in particular, we present them as a subclass of Markov processes.

2.1 Kolmogorov extension theorem

For this part we refer to Gihman & Skorohod (1974). Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. A parameter set $T \neq \emptyset$ and a measurable space $(E, \mathscr{E})$, called *state space* or *value space*, are given.

Examples of the parameter set T: $T = [0, +\infty) =: \mathbb{R}_+$; $T = [0, \tau]$, $\tau > 0$; $T = \mathbb{N}$; $T = \mathbb{Z}$; $T \subseteq \mathbb{R}$; $T \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$.

Examples of the state space $(E, \mathscr{E})$: $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$; $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$, $d \in \mathbb{N}$; $E = C([0,1])$, that is the Banach space of continuous functions ($\|z\| = \sup_{t\in[0,1]} |z(t)|$) and $\mathscr{E} := \mathscr{B}(E)$.

2.1.1 Definition. (i) A stochastic process X with state space $(E, \mathscr{E})$ and parameter set T, is an application on $\Omega \times T$ into E such that the mapping $\omega \mapsto X(\omega, t)$ is $(\mathscr{F}, \mathscr{E})$-measurable. We use the notation $X_t(\omega) := X(\omega, t)$ and, we often suppress the dependence on ω, that is write $X_t := X_t(\omega)$. We also sometimes write $X = (X_t)_{t\in T}$.

(ii) Let $(T, \mathscr{T})$ be a measurable space. We say that the process X is measurable, if $X : \Omega \times T \longrightarrow E$ is $(\mathscr{F} \otimes \mathscr{T}, \mathscr{E})$-measurable.

We observe that, if X is a stochastic process with state space E, then X_t is a random variable with values in E, for every $t \in T$, and the notation $X = (X_t)_{t\in T}$ exhibits the stochastic process X as a collection of random variables indexed on T.

If E^n denotes the n-fold Cartesian product of E with itself and $\mathscr{E}^n := \bigotimes_{i=1}^n \mathscr{E}$, then the mapping $\omega \mapsto (X_{t_1}(\omega), \ldots, X_{t_n}(\omega))$, $t_1, \ldots, t_n \in T$, is $(\mathscr{F}, \mathscr{E}^n)$-measurable, $\omega \mapsto X(\omega, t)$ being $(\mathscr{F}, \mathscr{E})$-measurable.

2.1.2 Definition. Let X be a stochastic process.

(i) For $n \in \mathbb{N}$ and $t_1, \ldots, t_n \in T$, by $\mu_{t_1,\ldots,t_n}$ we denote the distribution of the random vector $(X_{t_1}, \ldots, X_{t_n})$.

(ii) The family $\mu := \{\mu_{t_1,\ldots,t_n},\ t_1, \ldots, t_n \in T;\ n \in \mathbb{N}\}$ is called *system of the finite-dimensional distributions* of the stochastic process X.

For a given stochastic process X, the corresponding family μ of the finite-dimensional distributions satisfies some "consistency" conditions. More precisely, we notice that, for a choice of $n \in \mathbb{N}$ and $t_1, \ldots, t_n \in T$, because of the definition of $\mu_{t_1,\ldots,t_n}$, we have

$$\mu_{t_1,\ldots,t_n}(B) = \mathbb{P}\big[(X_{t_1}, \ldots, X_{t_n}) \in B\big], \quad B \in \mathscr{E}^n.$$

Therefore, for $B \in \mathscr{E}^n$ and $t_1, \ldots, t_n, t_{n+1}, \ldots, t_{n+m} \in T$, $n, m \in \mathbb{N}$, we have

$$\begin{aligned}\mu_{t_1,\ldots,t_n,t_{n+1},\ldots,t_{n+m}}(B \times \overbrace{E \times \ldots \times E}^{m\text{-times}}) \\ &= \mathbb{P}\big[\{(X_{t_1}, \ldots, X_{t_n}) \in B\} \cap \{(X_{t_{n+1}}, \ldots, X_{t_{n+m}}) \in E^m\}\big] \\ &= \mathbb{P}\big[\{(X_{t_1}, \ldots, X_{t_n}) \in B\}\big],\end{aligned}$$

where we used $\{(X_{t_{n+1}}, \ldots, X_{t_{n+m}}) \in E^m\} = \Omega$, E being the state space of X. The previous computation shows that, for every $B \in \mathscr{E}^n$ and $t_1, \ldots, t_n, t_{n+1}, \ldots, t_{n+m} \in T$, $n, m \in \mathbb{N}$, the family of the finite dimensional distributions of a stochastic process X satisfies the relation

$$\mu_{t_1,\ldots,t_n,t_{n+1},\ldots,t_{n+m}}(B \times \overbrace{E \times \ldots \times E}^{m\text{-times}}) = \mu_{t_1,\ldots,t_n}(B)\,. \tag{2.1}$$

Moreover, if $B \in \mathscr{E}^n$ has the form $B = B_1 \times \ldots \times B_n$ and $t_1, \ldots, t_n \in T$, for any permutation π of $\{1, \cdots, n\}$, we have

$$\begin{aligned}\mu_{t_{\pi(1)},\ldots,t_{\pi(n)}}(B_{\pi(1)} \times \ldots \times B_{\pi(n)}) \\ &= \mathbb{P}\big[X_{\pi(1)} \in B_{\pi(1)}, \ldots, X_{\pi(n)} \in B_{\pi(n)}\big] \\ &= \mathbb{P}\big[X_1 \in B_1, \ldots, X_n \in B_n\big]\end{aligned}$$

that is, for every rectangle $B \in \mathscr{E}^n$, every $t_1, \ldots, t_n \in T$ and every permutation π of $\{1, \cdots, n\}$,

$$\mu_{t_{\pi(1)},\ldots,t_{\pi(n)}}(B_{\pi(1)} \times \ldots \times B_{\pi(n)}) = \mu_{t_1,\ldots,t_n}(B_1 \times \ldots \times B_n)\,. \tag{2.2}$$

Because of the previous discussion, we can state the following definition and theorem.

2.1.3 Definition. Let $n \in \mathbb{N}$ and assume that with each $t_1, \ldots, t_n \in T$ we can associate a probability measure $\nu_{t_1,\ldots,t_n}$ on $(E^n, \mathscr{E}^n)$.

(i) The family of probabilities $\nu := \{\nu_{t_1,\ldots,t_n},\ t_1, \ldots, t_n \in T;\ n \in \mathbb{N}\}$ is called *a system of finite dimensional distributions* (not necessarily of a stochastic process)

(ii) A system ν of finite dimensional distributions is *consistent* if it satisfies (2.1) and (2.2), called *consistency conditions.*

Notice that in Definition 2.1.2 we started with a stochastic process X and we defined the system of finite dimensional distributions of X. In Definition 2.1.3 we have no stochastic process coming into play but only a by T-parametrized family of probability measures, called system of finite dimensional distributions. At this point it is important to note the difference between "system of finite dimensional distributions *of a process*" and "system of finite dimensional distributions".

2.1.4 Theorem. *Let X be a stochastic process with state space $(E, \mathscr{E})$ and let μ be the family of the finite dimensional distributions of X. Then the family μ is consistent.*

Kolmogorov Extension Theorem is in a certain sense the converse of Theorem 2.1.4. The question is the following. Let T be an index set, $(E, \mathscr{E})$ a measurable space and $\nu := \{\nu_{t_1,\dots,t_n},\ t_1, \dots, t_n \in T;\ n \in \mathbb{N}\}$ a given family of probability measures, such that $\nu_{t_1,\dots,t_n}$ is a probability measure on $(E^n, \mathscr{E}^n)$, for every $t_1, \dots, t_n \in T$, $n \in \mathbb{N}$. There exists a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a stochastic process X on it having $(E, \mathscr{E})$ as state space and such that the family μ of the finite dimensional distributions (with respect to $\mathbb{P}$!!!) of X is equal to ν, that is, such that

$$\mu_{t_1,\dots,t_n}(B) = \mathbb{P}\big[(X_{t_1}, \dots, X_{t_n}) \in B\big] \overset{!!!}{=} \nu_{t_1,\dots,t_n}(B)$$

for every $B \in \mathscr{E}^n$ and every $t_1, \dots, t_n \in T$, $n \in \mathbb{N}$? Clearly, this question is to general to be well formulated: If ν has to be the family of the finite dimensional distributions of a stochastic process, according to Theorem 2.1.4, it has to be, at least, *consistent*. Kolmogorov Extension Theorem claims that, if the state space $(E, \mathscr{E})$ is "regular" enough, then the consistency conditions (2.1) and (2.2) are necessary and sufficient for ν to be the system of finite dimensional distributions of X.

The proof of Kologorov Extension Theorem is *constructive*: the existence of the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and of the process X is proven constructing a measurable space $(\Omega, \mathscr{F})$ and then defining on it a probability measure. On this space a stochastic process X is explicitly constructed and then it is proven that the family μ of the finite dimensional distributions of X are equal to the given consistent system of probabilities ν.

Before we formulate and sketch the proof of Kolmogormov Extension Theorem, we need some preparation.

By E^T we denote the set of the mappings f on T into E, that is $f \in E^T$ if and only if $f : T \longrightarrow E$.

2.1.5 Definition. (i) Let $\tau := (t_1, \dots, t_n)$ be an element of T^n and B of $\mathscr{E}^n$. The set of functions $f \in E^T$ such that $f(\tau) := (f(t_1), \dots, f(t_n))$ belongs to B is called *cylindrical set* (in E^T with basis B over the coordinates τ) and it is denoted by $C_n(\tau, B)$, that is,

$$C_n(\tau, B) := \left\{ f \in E^T :\ (f(t_1), \dots, f(t_n)) \in B \right\}.$$

(ii) By $\mathscr{A}$ we denote the family of cylindrical sets, that is

$$\mathscr{A} := \{C_n(\tau, B),\quad B \in \mathscr{E}^n;\ \tau \in T^n;\ n \in \mathbb{N}\}.$$

(iii) By $\mathscr{E}^T$ we denote the σ-algebra generated by the cylindrical sets $\mathscr{E}^T := \sigma(\mathscr{A})$.

According to Gihman & Skorohod (1974), I.I.§4, Theorem 1 we have

2.1.6 Theorem. *The family $\mathscr{A}$ of cylindrical sets forms an algebra of subsets of E^T.*

Proof. For $B \in \mathscr{E}^n$ and $\tau \in T^n$ the identity $C_n(\tau, B)^c = C_n(\tau, B^c)$ and clearly $E^T = C_n(\tau, E^n)$ is a cylindrical set. The last property to prove is that $\mathscr{A}$ is $\cap$-stable. We do not show this part in details but only give an example: Let us consider the cylindrical sets $\{f : f(t_1) \in B_1\}$ and $\{f : f(t_2) \in B_2\}$. Then setting $\tau := (t_1, t_2)$, we get $\{f : f(t_1) \in B_1\} \cap \{f : f(t_2) \in B_2\} = C_2(\tau, B_1 \times B_2)$. □

We furthermore observe that if $B \in \mathscr{E}^n$, $\tau := (t_1, \dots, t_n) \in T^n$ and $\sigma := (s_1, \dots, s_m) \in T^m$, setting $\tau|\sigma := (t_1, \dots, t_n, s_1, \dots, s_m)$, then

$$C_{n+m}(\tau|\sigma, B \times E^m) = C_n(\tau, B)\,. \tag{2.3}$$

Formula (2.3) shows that different coordinates and basis can generate the same cylindrical set.

Now we are ready to formulate an give a sketch of the proof of the following theorem (cf. Gihman & Skorohod (1974), I.I.§4, Theorem 2).

2.1.7 Theorem (Kolmogorov Extension Theorem)**.** *Let T be a parameter set and $(E, \mathscr{E})$ a measurable space. Let $\mu := \{\mu_{t_1,\dots,t_n},\ t_1, \dots, t_n \in T;\ n \in \mathbb{N}\}$ be a system of finite dimensional distributions. Assume that*

(i) *E is a separable complete metric space and $\mathscr{E} = \mathscr{B}(E)$ is the Borel σ-algebra of E (with respect to the topology of the metric);*

(ii) *the system μ of finite dimensional distributions satisfies the consistency conditions* (2.1) *and* (2.2).

Then there exists a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a stochastic process $X = (X_t)_{t \in T}$ on this probability space and with state space $(E, \mathscr{E})$ admitting μ as associated system of finite dimensional distribution with respect to $\mathbb{P}$, that is, such that for every $B \in \mathscr{E}^n$, $t_1, \dots, t_n \in T$, and $n \in \mathbb{N}$,

$$\mathbb{P}[(X_{t_1}, \dots, X_{t_n}) \in B] = \mu_{t_1,\dots,t_n}(B) \tag{2.4}$$

holds. Furthermore the probability measure $\mathbb{P}$ on $(\Omega, \mathscr{F})$ is uniquely determined by (2.4).

Sketch of the proof. In the proof of the theorem the probability space and the process are concretely constructed.

Step 1: *Definition of Ω and of $\mathbb{P}$ on $\mathscr{A}$.* We define set $\Omega := E^T$. Let $\mathscr{A}$ be the algebra of the cylindrical sets of Ω (cf. Theorem 2.1.6). For $\tau = (t_1, \dots, t_n) \in T^n$ and $B \in \mathscr{E}^n$ we consider $C = C_n(\tau, B)$ and set

$$P'(C) := \mu_{t_1,\dots,t_n}(B). \tag{2.5}$$

Because of the consistency conditions (2.1) and (2.2), (2.5) uniquely defines P' over $\mathscr{A}$ (cf. (2.3)), that is (2.5) does not depend on the special representation of the cylindrical set C.

Step 2: *Construction of $\mathscr{F}$ and construction of $\mathbb{P}$.* Because of the σ-additivity of $\mu_{t_1,\dots,t_n}$ on $\mathscr{E}^n$, P' is finitely additive on the algebra $\mathscr{A}$ of cylindrical sets. In other words, according to Bauer (2001), Definition 3.1, P' is a *content* on the algebra $\mathscr{A}$ (a content ν is a nonnegative finitely additive set function on a ring such that $\nu(\emptyset) = 0$). The main point of this step is to show that P' is a *premeasure* (that is a σ-additive content) on $\mathscr{A}$. To prove this, one has to use the topological properties of $(E, \mathscr{E})$. Notice that (2.5) ensures the finiteness of P'. We can therefore apply the following result: Every σ-finite premeasures $\tilde{\mu}$ on a ring $\mathscr{R}$ can be extended in a unique way to a measure μ on $\sigma(\mathscr{R})$ (cf. Bauer (2001), Theorem 5.6). Let $\mathbb{P}$ be the extension of P' to $\mathscr{E}^T := \sigma(\mathscr{A})$. We set $\mathscr{F} := \mathscr{E}^T$ and denote by $\mathbb{P}$ the unique measure obtained as extension of P' to $\mathscr{F}$. From (2.5) we immediately deduce that $\mathbb{P}$ is a probability measure on $(\Omega, \mathscr{F})$ $(P'(\Omega) := \mu_{t_1,\dots,t_n}(E^n) = 1)$.

Step 3: *Construction of the process* X. Let ω be an element of Ω, that is a function on T into E (recall that by construction $\Omega = E^T$) and define

$$X_t(\omega) := \omega(t)\,. \tag{2.6}$$

The mapping $X : \Omega\times T \longrightarrow E$ defined by (2.6) is clearly a stochastic process on $(\Omega, \mathscr{F}, \mathbb{P})$ with state space $(E, \mathscr{E})$ and for each $B \in \mathscr{E}^n$, $\tau = (t_1, \dots, t_n) \in T^n$, $n \in \mathbb{N}$,

$$\begin{aligned}\mathbb{P}[(X_{t_1}, \dots, X_{t_n}) \in B]& \\ &= \mathbb{P}[(\omega(t_1), \dots, \omega(t_n)) \in B] = \mathbb{P}[C_n(\tau, B)] \\ &= P'[C_n(\tau, B)] = \mu_{t_1,\dots,t_n}(B)\end{aligned}$$

meaning that μ is the system of finite dimensional distributions of the process X and the proof is complete. $\square$

We observe that if we take $T = [0, +\infty)$ and $E = \mathbb{R}^d$, $d \geq 1$, then the consistency conditions (2.1) and (2.2) for the system of finite dimensional probabilities can be simplified to

$$\begin{aligned}&\mu_{t_1,\dots,t_n}(B_1 \times \ldots \times B_{k-1} \times E \times B_{k+1} \ldots \times B_n) = \\ &\qquad \mu_{t_1,\dots,t_{k-1},t_{k+1},\dots,t_n}(B_1 \times \ldots \times B_{k-1} \times B_{k+1} \ldots \times B_n)\end{aligned} \tag{2.7}$$

and the following theorem (cf. Shiryaev (1996), II.§9, Theorem 1) holds

2.1.8 Theorem. *Let* $\mu := \{\mu_{t_1,\dots,t_n},\ t_1 < \ldots < t_n\}$ *be a system of finite dimensional distributions such that* $\mu_{t_1,\dots,t_n}$, $t_1 < \ldots < t_n$, *is a probability measure on* $\big((\mathbb{R}^d)^n, \mathscr{B}(\mathbb{R}^d)^n\big)$, *and satisfying* (2.7). *Then there exists a probability space* $(\Omega, \mathscr{F}, \mathbb{P})$ *and a stochastic process with state space* $\big(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d)\big)$ *such that its associated system of finite dimensional distributions is given by* μ.

We conclude this section fixing some notations: In the remaining chapter of these notes, we shall consider $(E, \mathscr{E})$ equal to $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ or $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ and T equal to $\mathbb{R}_+$ or to $[0, T]$, $T > 0$ (little abuse of notation!).

Existence of a regular version of the conditional probability. At this point we can give an answer to the question about the existence of a regular version of the conditional probability, at least in the case of the conditional distribution of a stochastic process X.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, $\mathscr{G} \subseteq \mathscr{F}$ a sub-σ-algebra of $\mathscr{F}$ and X a stochastic process with state space $(E, \mathscr{E}) := (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$. For the following theorem see Shiryaev (1996), corollary to Theorem 5 in Chapter II, §7.

2.1.9 Theorem. *Let $\mathbb{P}[X_t \in B|\mathscr{G}]$ be the conditional distribution of X given $\mathscr{G}$. Then there exists a regular version of the conditional probability, which we denote by $\mathbb{P}[X_t \in B|\mathscr{G}]$. Furthermore, for every bounded real-valued positive function f, the following formula holds*

$$\mathbb{E}[f(X_t)|\mathscr{G}] = \int_E f(y)\mathbb{P}[X_t \in \mathrm{d}\, y|\mathscr{G}]\,.$$

2.2 Paths of stochastic processes

In this section we denote $(E, \mathscr{E}) := \big(\mathbb{R}, \mathscr{B}(\mathbb{R})\big)$. This will be the state space of all stochastic processes we are going to consider. A probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is given.

Let X be a stochastic process.

2.2.1 Definition. Let X be a stochastic process with values in $(E, \mathscr{E})$.

(i) The mapping $t \mapsto X_t(\omega)$ is called *path* or *trajectory* of the process X.

(ii) The process X is said to be (a.s.) continuous if (almost) all its trajectories are continuous.

(iii) The process X is said to be (a.s.) càdlàg if (almost) all its trajectories are càdlàg.

(iv) The process X is said to be stochastically continuous (or *continuous in probability*) if

$$\lim_{s \to t} \mathbb{P}[|X_t - X_s| > \varepsilon] = 0, \qquad \text{for every } t \geq 0 \text{ and } \varepsilon > 0.$$

2.2.2 Remark. We remark that the word càdlàg is the abbreviation of French *"continue à droite, limité à gauche"*. This means that the paths of the process X are right-continuous and admit finite left-limit in everye point, that is

$$\lim_{s \downarrow t} X_s = X_t\,, \quad \lim_{s \uparrow t} X_s \text{ exists and is finite for every } t > 0.$$

2.2.3 Definition. Let X be a càdlàg process.

(i) We define the random variable X_{t-} for every $t > 0$ as $X_{t-} := \lim_{s \uparrow t} X_s$, which is finite, and $X_{0-} := X_0$. By $X_- = (X_{t-})_{t \geq 0}$ we denote the *left-hand-limit process* associated with X.

(ii) The jump process $\Delta X := (\Delta X_t)_{t \geq 0}$ of X is defined by $\Delta X_t := X_t - X_{t-}, \quad t \geq 0$.

(iii) If, for a *fixed* $t > 0$, $\Delta X_t \neq 0$ we say that the process X has a *fixed discontinuity* at fixed time $t > 0$.

Notice that for any càdlàg process X, $\Delta X_0 = 0$. Furthermore, if X is a continuous stochastic process, then $X_- = X$ and $\Delta X = 0$. For two stochastic processes X and Y, there exist different concepts of equality:

2.2.4 Definition. Let X and Y be stochastic processes. Then X and Y are

(i) *equivalent* if they have the same finite dimensional distributions

(ii) *modifications* if $\mathbb{P}[X_t = Y_t] = 1$, for every $t \geq 0$.

(iii) *indistinguishable* if $\mathbb{P}[X_t = Y_t, \text{ for every } t \geq 0] = 1$.

(iv) (a.s.) *equal* if $X_t(\omega) = Y_t(\omega)$ for every $t \geq 0$ and for (almost) every $\omega \in \Omega$.

If X is a càdlàg and stochastically continuous process, then X has no fixed discontinuities a.s. This means that, the random variable ΔX_t is equal to zero a.s., for every fixed $t > 0$. This is a delicate point: If X is stochastically continuous then the process ΔX is a modification of the process identically equal to zero. However, these two processes are not indistinguishable.

2.2.5 Exercise. ($\star$) (i) Prove that the implications $(iv) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i)$ holds in Definition 2.2.4. Do the implications in the converse direction hold? If no, give an example of two processes which are modifications of each other but not indistinguishable. Prove that if $t \in \mathbb{N}$, then $(ii) \Leftrightarrow (iii)$. Show that the same holds for right continuous processes which are modifications of each other.

(ii) Let X be stochastic continuous and $\varphi_t(u) := \mathbb{E}[\exp(\mathrm{i}\, uX_t)]$, $u \in \mathbb{R}$, $t \geq 0$. Then the mapping $t \mapsto \varphi_t(u)$ is continuous.

The importance of the càdlàg property for stochastic processes will be made clear in the following chapters. We notice here that this property was used to define the process X_- and hence the jump process ΔX. Furthermore the following result, for which we refer to Appelbaum (2009), Lemma 2.9.1, holds.

2.2.6 Theorem. *Let X be a càdlàg process. Then*
(i) *For any $\varepsilon > 0$, the set $S_\varepsilon := \{t > 0 : |\Delta X_t| > \varepsilon\}$ is finite.*
(ii) *The set $S := \{t > 0 : \Delta X_t \neq 0\}$ is at most countable.*

Theorem 2.2.6, (i) asserts that a càdlàg process has only finitely many fixed discontinuities bigger in absolute value of a given positive number, while Theorem 2.2.6, (ii) claims that fixed discontinuities of a càdlàg process are at most countable.

Observe that one important consequence of Theorem 2.2.6 is that for any càdlàg process, the infinite sum

$$\sum_{0 \leq s \leq t} |\Delta X_s| = \sum_{0 \leq s \leq t} |\Delta X_s| 1_{\{\Delta X_s \neq 0\}}, \quad t \geq 0$$

consists at most of countably many terms. Therefore it is a well-defined random variable taking values in $[0, +\infty]$.

We conclude this section with the well-known Continuity Theorem of Kolmogorov, ensuring sufficient conditions for the existence of a (Hölder-)continuous modification of a given stochastic process. We refer to Revuz & Yor (1999), Chapter I, Theorem 1.8.

2.2.7 Theorem. *Let X be a stochastic process with state space $(E, \mathscr{E})$ such that there exist three real numbers $\alpha, \beta, c > 0$ such that for every $s, t \in \mathbb{R}_+$ the estimate*

$$\mathbb{E}\big[|X_t - X_s|^\beta\big] \leq c|t - s|^{1+\alpha}$$

holds. Then there exists a continuous modification Y of X. Furthermore the paths of Y are Hölder-continuous with parameter $\gamma < \alpha/\beta$.

2.3 Filtrations, stopping times and martingales

In this section we denote $(E,\mathscr{E}) := \big(\mathbb{R},\mathscr{B}(\mathbb{R})\big)$. Let $(\Omega,\tilde{\mathscr{F}},\mathbb{P})$ be a probability space. By $\mathscr{N}(\mathbb{P})$ we denote the null sets of $\mathbb{P}$, i.e., $\mathscr{N}(\mathbb{P}) := \{A \subseteq \Omega : \exists\, B \in \tilde{\mathscr{F}},\ A \subseteq B,\ \mathbb{P}(B) = 0\}$. If $\mathscr{N}(\mathbb{P})$ is not contained in $\tilde{\mathscr{F}}$ we enlarge the σ-algebra by setting $\mathscr{F} := \tilde{\mathscr{F}} \vee \mathscr{N}(\mathbb{P})$. We call $\mathscr{F}$ the completion of $\tilde{\mathscr{F}}$ (*in itself*) with respect to $\mathbb{P}$ or simply $\mathbb{P}$-completion of $\tilde{\mathscr{F}}$ and we say that $(\Omega,\mathscr{F},\mathbb{P})$ is a complete probability space. If not otherwise specified, we assume a probability space to be complete. In the remaining of this chapter we assume that a complete probability space $(\Omega,\mathscr{F},\mathbb{P})$ is fixed.

2.3.1 Definition. (i) A family $\mathbb{F} = (\mathscr{F}_t)_{t\geq 0}$ of sub-σ-algebras of $\mathscr{F}$ which is increasing is called a *filtration*, that is, $\mathscr{F}_t \subseteq \mathscr{F}$ is a σ-algebra, $t \geq 0$, and $\mathscr{F}_s \subseteq \mathscr{F}_t$, $0 \leq s \leq t$.

(ii) A filtration $\mathbb{F}$ is called *complete* if $\mathscr{N}(\mathbb{P}) \subseteq \mathscr{F}_t$ for every $t \geq 0$.

(iii) With a filtration $\mathbb{F} = (\mathscr{F}_t)_{t\geq 0}$ we associate the filtration $\mathbb{F}_+ = (\mathscr{F}_{t+})_{t\geq 0}$ by $\mathscr{F}_{t+} := \cap_{\varepsilon>0}\mathscr{F}_{t+\varepsilon}$. The filtration $\mathbb{F}$ is called *right-continuous* if $\mathbb{F} = \mathbb{F}_+$, that is, if $\mathscr{F}_{t+} = \mathscr{F}_t$, $t \geq 0$. Note that $\mathscr{F}_{0+} = \mathscr{F}_0$.

(iv) A filtration $\mathbb{F}$ which is complete and right-continuous is said to satisfy the *usual conditions*.

Notice that if a filtration $\tilde{\mathbb{F}}$ is not complete we can introduce the complete filtration $\mathbb{F} = (\mathscr{F}_t)_{t\geq 0}$ by setting $\mathscr{F}_t := \tilde{\mathscr{F}}_t \vee \mathscr{N}(\mathbb{P})$, $t \geq 0$, and we call the filtration $\mathbb{F}$ the $\mathbb{P}$-*completion* of $\tilde{\mathbb{F}}$ (in $\mathscr{F}$).

Given a filtration $\tilde{\mathbb{F}}$ we can always associate to it a filtration $\mathbb{F}$ satisfying the usual conditions by setting $\mathscr{F}_t := \tilde{\mathscr{F}}_{t+} \vee \mathscr{N}(\mathbb{P})$, $t \geq 0$. In the following of this work, if not otherwise specified, we shall always consider filtrations satisfying the usual conditions.

The mathematical concept of a filtration $\mathbb{F}$ has the following interpretation: The events $B \in \mathscr{F}_t$ are known at time $t \geq 0$. That is $\mathbb{F}$ models the flow of information which increases in time.

We set

$$\mathscr{F} := \mathscr{F}_\infty := \sigma\Big(\textstyle\bigcup_{t\geq 0}\mathscr{F}_t\Big)\,. \tag{2.8}$$

2.3.2 Example. With a stochastic process X, we associate the σ-algebra $\tilde{\mathscr{F}}_t^X := \sigma(X_s,\quad 0 \leq s \leq t)$. We call the filtration $\tilde{\mathbb{F}}^X = (\tilde{\mathscr{F}}_t^X)_{t\geq 0}$ the *filtration generated by* X. Let $\mathscr{F}_t^X := \tilde{\mathscr{F}}_t^X \vee \mathscr{N}(\mathbb{P})$ denote the $\mathbb{P}$-completion of $\tilde{\mathscr{F}}_t^X$. The filtration $\mathbb{F}^X := (\mathscr{F}_t^X)_{t\geq 0}$ is the $\mathbb{P}$-completion of $\tilde{\mathbb{F}}^X$ (in $\mathscr{F}$). By $\tilde{\mathbb{F}}_+^X = (\tilde{\mathscr{F}}_{t+}^X)_{t\geq 0}$ we denote the smallest right-continuous filtration containing the filtration generated by X, i.e., $\tilde{\mathscr{F}}_{t+}^X \supseteq \tilde{\mathscr{F}}_t^X$, $t \geq 0$. In the sequel, the most relevant filtration associated with the stochastic process X will be $\mathbb{F}_+^X$, i.e., the $\mathbb{P}$-completion in $\mathscr{F}$ of $\tilde{\mathbb{F}}_+^X$. The filtration $\mathbb{F}_+^X$ satisfies the usual conditions and we call it the *natural filtration* of X.

A stochastic process X is always adapted to its generated filtration $\tilde{\mathbb{F}}^X$. Moreover, it is adapted to a filtration $\mathbb{F}$ if and only if $\tilde{\mathbb{F}}^X \subseteq \mathbb{F}$.

2.3.3 Definition. (i) A process X is *adapted* to a filtration $\mathbb{F}$ if X_t is $\mathscr{F}_t$-measurable, for every $t \geq 0$.

(ii) A process X is called *progressively measurable* if, for every $u \geq 0$ the mapping $(\omega, t) \mapsto X(\omega, t)$, $t \leq u$, from $\Omega \times [0, u]$ to E is $(\mathscr{F}_u \otimes \mathscr{B}([0, u]), \mathscr{E})$-measurable.

2.3.4 Exercise. ($\star$) Show the following claim: Let X be an $\mathbb{F}$-adapted process and assume that the filtration is complete. Then every modification Y of X is adapted.

Observe that any progressively measurable process is measurable and adapted. The converse claim is, in general, not true. However, the following result holds (cf. Dellacherie (1972), Theorem 11 in Chapter III).

2.3.5 Theorem. *Let X be a stochastic process.*
(i) *If X is right-continuous or left-continuous, then it measurable.*
(ii) *If furthermore X is adapted, then it is progressively measurable.*

Now we come to the important notion of stopping time.

2.3.6 Definition. Let $\mathbb{F}$ be a given filtration.
(i) A mapping τ on Ω into $[0, +\infty]$ is called a *stopping time* (with respect to $\mathbb{F}$) if the set $\{\tau \leq t\} := \{\omega \in \Omega : \tau(\omega) \leq t\}$ is $\mathscr{F}_t$-measurable, for every $t \geq 0$.
(ii) Let τ be a stopping time. We define $\mathscr{F}_\tau := \{A \in \mathscr{F} : A \cap \{\tau \leq t\} \in \mathscr{F}_t, \ \forall t \geq 0\}$.

We now enumerate some (assumed to be known to the reader) properties of stopping times:
(i) $\tau \equiv t$ is a stopping time, $t \geq 0$.
(ii) For a stopping time τ, $\mathscr{F}_\tau$ is a σ-algebra.
(iii) Maximum, minimum and sum of stopping times are stopping times.
(iv) $\mathscr{F}_\sigma \subseteq \mathscr{F}_\tau$ for any stopping times σ and τ with $\sigma \leq \tau$.
(v) τ is a $\mathbb{F}_+$-stopping time if and only if $\{\tau < t\}$ is $\mathscr{F}_t$-measurable, for every $t > 0$.

Let X be a process and τ a finite-valued nonnegative random variable. We denote by X_τ the mapping $\omega \mapsto X(\omega, \tau(\omega))$ on Ω into E. It is clear that, if X is a measurable process, then X_τ is a random variable. Indeed, X_τ is the composition of the measurable mappings $\omega \mapsto (\omega, \tau(\omega))$ of Ω into $\Omega \times \mathbb{R}_+$ and of $(\omega, t) \mapsto X(\omega, t)$ of $\Omega \times \mathbb{R}_+$ into E. In particular, if τ is a finite-valued stopping time, X_τ is a random variable. If furthermore, X is progressively measurable, X_τ is $\mathscr{F}_\tau$-measurable (cf. Dellacherie (1972), Chapter III, Theorem 20 or Revuz & Yor (1999), Chapter I, Proposition 4.9).

2.3.7 Definition. Let X be a measurable $\mathbb{F}$-adapted process and τ a stopping time. The process $X^\tau = (X^\tau_t)_{t \geq 0}$ defined by $X^\tau_t := X_{\tau \wedge t}$ is called *stopped process* (at time τ).

If X is an $\mathbb{F}$-adapted stochastic process and $B \in \mathscr{E}$ is a measurable sets, we define the random variable
$$\tau_B := \inf\{t \geq 0 : X_t \in B\}, \tag{2.9}$$
with the convention $\inf \emptyset := +\infty$. The relevant question is if this random variable is an $\mathbb{F}$-stopping time. We summarise in the next theorem some results concerning this problem

2.3.8 Theorem. (i) *Let X and $\mathbb{F}$ be right-continuous and B be an open set. Then τ_B is a stopping time.*
(ii) *Let X be continuous, $\mathbb{F}$ right-continuous and B closed. Then τ_B is a stopping time.*
(iii) *Let X be progressively measurable and assume that $\mathbb{F}$ satisfies the usual conditions. Then τ_B is a stopping time for every $B \in \mathscr{E}$.*

Proof. We only verify (i) and (ii). To prove claim (iii) a very deep result, called "debut theorem", whose proof is demanding, is needed. We refer Dellacherie (1972), Chapter III, Theorem 23.

We start proving (i). Clearly $\{\tau_B < t\} = \bigcup_{s<t}\{X_s \in B\}$ holds. Because B is open and X is right-continuous, $\bigcup_{s<t}\{X_s \in B\} = \bigcup_{s<t, s\in\mathbb{Q}}\{X_s \in B\}$. Therefore $\bigcup_{s<t}\{X_s \in B\}$ is $\mathscr{F}_t$-measurable, as a countable union of $\mathscr{F}_t$-measurable sets. Hence, $\{\tau_B < t\}$ is $\mathscr{F}_t$-measurable, which is equivalent to $\{\tau_B \leq t\}$ is $\mathscr{F}_t$-measurable, $\mathbb{F}$ being right continuous and the proof of (i) is complete. We now show (ii). Let $d(\cdot,\cdot)$ denote the Euclidean distance in $(E, \mathscr{E})$ and let $U_n := \{x \in E : d(x, B) < 1/n\}$. Then U_n is an open set, its closure is $\overline{U}_n = \{x \in E : d(x, B) \leq 1/n\}$. Because of (i), τ_{U_n} is a stopping time and $\tau_{U_n} \leq \tau_{U_{n+1}}$, $n \geq 1$. Let $\sigma := \lim_{n\to\infty} \tau_{U_n}$. Then σ is a stopping time and because $\tau_{U_n} \leq \tau_B$, we also have $\sigma \leq \tau_B$. We now show $\sigma = \tau_B$ and for this $\sigma \geq \tau_B$ is sufficient. Notice that if $\sigma(\omega) = +\infty$, then $\tau_B(\omega) = +\infty$ and therefore $\sigma(\omega) = \tau_B(\omega)$ on $\{\sigma = \infty\}$. We now consider the case $\sigma < +\infty$. Because of the continuity of X, $X_{\sigma(\omega)}(\omega) = \lim_{n\to+\infty} X_{\tau_n(\omega)}(\omega)$ and $X_{\tau_n(\omega)}(\omega) \in \overline{U}_n \subseteq \overline{U}_m$, for every $n \geq m$. Hence $X_{\sigma(\omega)}(\omega) \in \overline{U}_m$, $m \geq 1$ and therefore $X_{\sigma(\omega)}(\omega) \in \bigcap_{m=1}^{\infty} \overline{U}_m = B$, which implies $\tau_B \leq \sigma$. The proof of the theorem is now complete. □

Now we introduce martingales.

2.3.9 Definition. Let $\mathbb{F}$ be a filtration and X an $\mathbb{F}$-adapted stochastic process. Then X is a non-càdlàg *martingale* (resp. *submartingale*, resp. *supermartingale*) (with respect to $\mathbb{F}$) if

(i) $X_t \in L^1(\mathbb{P})$, for every $t \geq 0$;

(ii) $\mathbb{E}[X_t|\mathscr{F}_s] = X_s$ (resp. $\mathbb{E}[X_t|\mathscr{F}_s] \geq X_s$, resp. $\mathbb{E}[X_t|\mathscr{F}_s] \leq X_s$), for every $0 \leq s \leq t$.

If X is furthermore càdlàg , we call it simply a martingale.

The next theorem ensure the existence of a càdlàg modification of a martingale. For the proof see He, Wang & Yan (1992), Theorem 2.44.

2.3.10 Theorem. *Let X be a non-càdlàg martingale with respect to a filtration $\mathbb{F}$ satisfying the usual conditions. Then there exists a modification of X which is an $\mathbb{F}$-martingale (that is càdlàg).*

We will always assume that the filtration $\mathbb{F}$ satisfies the usual conditions. In this case we always consider martingales (if we have a non-càdlàg martingale we pass to the càdlàg modification, according to Theorem 2.3.10).

2.3.11 Lemma. *Let $(X^n)_{n\in\mathbb{N}}$ be a sequence of $\mathbb{F}$-martingale such that X^n_t converges to X_t in $L^1(\mathbb{P})$ as $n \longrightarrow +\infty$, for every $t \geq 0$. Then X is adapted and satisfies (i) and (ii) of Definition 2.3.9.*

2.3.12 Exercise. ($\star$) (i) Show Lemma 2.3.11.

(ii) Is X càdlàg ?

2.3.13 Theorem. *Let τ be a bounded stopping time ($|\tau| \leq M$, $M \geq 0$, a.s.) and let X be a martingale. Then the stopped process X^τ is again a martingale. Furthermore, for every bounded stopping time σ the relation $\mathbb{E}[X_\tau|\mathscr{F}_\sigma] = X_\sigma$ holds on $\{\sigma \leq \tau\}$.*

2.3.14 Remark. The claim of Theorem 2.3.13 still holds if τ and σ are general stopping times (also not finite) and X is a uniformly integrable martingale. This is in particular true if X_t is dominated in L^1 by a random variable Y (cf. Jacod & Shiryaev (2000), Chapter 1, Theorem 1.39).

Let X be a martingale and $T > 0$. The following estimate is the well-known *Doob's inequality*:

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t|^p\Big] \leq \tfrac{p}{p-1}\,\mathbb{E}\big[|X_T|^p\big]\,, \quad p > 1\,. \tag{2.10}$$

2.3.15 Definition. Let X be a martingale over $[0,T]$, $T > 0$. We say that it is square integrable if X_T (and hence each X_t, $t \leq T$) is square integrable.

2.3.16 Exercise. If X is a square integrable martingale with time set $[0,T]$, is it uniformly integrable? What is the definition of a square integrable martingale if the time set is $\mathbb{R}_+$?

We conclude this section with the important procedure of *localization*.

2.3.17 Definition. Let $\mathscr{C}$ be a class of stochastic process. We denote by $\mathscr{C}_{\mathrm{loc}}$ the localized class of $\mathscr{C}$ defined as such: A process X belongs to $\mathscr{C}_{\mathrm{loc}}$, if there exists a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ such that $\tau_n \uparrow +\infty$, $n \to +\infty$, such that $X^{\tau_n} \in \mathscr{C}$. The sequence $(\tau_n)_{n\in\mathbb{N}}$ is called *a localizing sequence for* X (relative to $\mathscr{C}$).

2.3.18 Exercise. ($\star$) (i) Give the definition of a local martingale.

(ii) Show that any local martingale which is bounded in L^1 is a true martingale. Is a local martingale integrable, in general?

(iii) Define the class of locally bounded processes.

(iv) Let $\mathbb{F}$ satisfies the usual conditions. Show that any continuous adapted process is locally bounded.

Solution. (i) An $\mathbb{F}$-adapted stochastic Process X is called a local martingale if there exists a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ such that $\tau_n \uparrow +\infty$, $n \to +\infty$, such that X^{τ_n} is a martingale.

(ii) We assume that X is a local martingale such that $|X_t| \leq Y \in L^1(\mathbb{P})$. Then, if $(\tau_n)_{n\geq 1}$ is a sequence of stopping times localizing X, applying Lebesgue's theorem on dominated convergence we get, for every $0 \leq s \leq t$,

$$X_s = \lim_{n\to+\infty} X_s^{\tau_n} = \lim_{n\to+\infty} \mathbb{E}[X_t^{\tau_n}|\mathscr{F}_s] = \mathbb{E}[\lim_{n\to+\infty} X_t^{\tau_n}|\mathscr{F}_s] = \mathbb{E}[X_t|\mathscr{F}_s],$$

meaning that X is a true martingale.

(iii) A locally bounded adapted process X is a stochastic process such that there exists a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ such that $\tau_n \uparrow +\infty$, $n \to +\infty$, such that X^{τ_n} is bounded. If X has continuous paths and starts at zero, then $\tau_n := \inf\{t > 0 : |X_t| > n\}$ is a stopping time by Theorem 2.3.8. By continuity of X, $(\tau_n)_{n\geq 1}$ is an increasing sequence of stopping times and, again by continuity, X^{τ_n} is a bounded process. □

We notice that if X is a càdlàg process starting at zero, then the left-limit process is locally bounded. Indeed, $\tau_n := \inf\{t \geq 0 : |X_t| > n\}$ is a localizing sequence: If $t < \tau_n$, then $|X_t| \leq n$. It could happen, that X has a jump in τ_n, therefore this is not a localizing sequence for X but, because X_{t-} is the left limit of X_t, this never happens. Notice that if X is a continuous process, then the sequence $(\tau_n)_{n\in\mathbb{N}}$ is always a localizing sequence. If $X_0 \neq 0$, then $(\tau_n)_{n\in\mathbb{N}}$ is a localizing sequence for $1_{(0,+\infty)}X_-$, in case X is càdlàg, or for $1_{(0,+\infty)}X$ if X is moreover continuous.

2.4 Markov processes

Let as usual $(E,\mathscr{E})$ denote $\big(\mathbb{R},\mathscr{B}(\mathbb{R})\big)$. We consider a stochastic process X on a (complete) probability space $(\Omega,\mathscr{F},\mathbb{P})$ with values in $(E,\mathscr{E})$ and denote by $\tilde{\mathbb{F}}^X$ the filtration generated by X (cf. Example 2.3.2). The filtration $\tilde{\mathbb{F}}^X$ models the flow of information generated by X. Therefore, $\tilde{\mathscr{F}}_t^X := \sigma(X_s,\ s \leq t)$ is the history of the process X up to time $t \geq 0$. Let $s \leq t$ and assume that at this point ("the present") we want to make a prediction about the future of X (that is about X_t), knowing all the history of X up to time $s \geq 0$ (that is $\tilde{\mathscr{F}}_s^X$). If we compare $\sigma(X_s)$ with $\tilde{\mathscr{F}}_s^X$, then clearly the second σ-algebra is, in general, bigger than the first one. Therefore, one could think that making a prediction on X_t at time $s \leq t$ using $\tilde{\mathscr{F}}_s^X$, would issue a "better" result than using only $\sigma(X_s)$. A Markov process is a process for which in the procedure described above, the use of $\tilde{\mathscr{F}}_s^X$ is not more effective than the one of $\sigma(X_s)$. That is, for Markov processes the conditional distribution of X_t given $\tilde{\mathscr{F}}_s^X$, is a measurable function of X_s:

$$\mathbb{P}[X_t \in A|\tilde{\mathscr{F}}_s^X] = g^A(X_s), \quad A \in \mathscr{E}$$

where g^A is a measurable function. In the remaining of this section we try to formalize the intuitive definition of a Markov process given above. Then we shall show the existence of Markov processes as a nice application of Kolmogorov Extension Theorem (cf. Theorem 2.1.7). For this section we refer to Revuz & Yor (1999), Chapter III.

2.4.1 Definition. A *kernel* N on $(E,\mathscr{E})$ is a mapping on $E\times\mathscr{E}$ into $\mathbb{R}_+\cup\{+\infty\}$ such that

(i) $A \mapsto N(x,A)$ is a positive measure on $\mathscr{E}$, for every $x \in E$.

(ii) $x \mapsto N(x,A)$ is $\mathscr{E}$-measurable, for every $A \in \mathscr{E}$.

If furthermore the kernel N is such that $N(x,E) = 1$, for every $x \in E$, then N is called a *transition probability.*

Let N and M be kernels on $(E,\mathscr{E})$, $f \geq 0$ a $\mathscr{E}$-measurable function and $A \in \mathscr{E}$. We define the mappings $x \mapsto Nf(x)$ and $(x,A) \mapsto MN(x,A)$ by

$$Nf(x) := \int_E N(x,\mathrm{d}\,y)f(y), \qquad MN(x,A) := \int_E M(x,\mathrm{d}\,y)N(y,A). \tag{2.11}$$

2.4.2 Exercise. (i) Show that the mapping $x \mapsto Nf(x)$ in (2.11) is measurable and nonnegative. (Hint: Use Theorem 1.2.2 for f bounded...)

(ii) Show that MN in (2.11) is a kernel on $(E,\mathscr{E})$.

2.4.3 Definition. (i) A *transition function* on $(E,\mathscr{E})$ is a family $P = (P_{s,t})_{0\leq s<t\in\mathbb{R}_+}$ of transition probabilities on $(E,\mathscr{E})$ such that for every $A \in \mathscr{E}$, $s < t < v$ and $x \in E$ the relation

$$P_{s,v}(x,A) = \int_E P_{s,t}(x,\mathrm{d}\,y)P_{t,v}(y,A) \tag{CK}$$

holds.

(ii) A transition function P is called homogeneous if $P_{s,t}$ depends on s an t only through their difference $t-s$. In this case we write $P_t := P_{0,t}$ and (CK) reads

$$P_{s+t}(x, A) = \int_E P_s(x, \mathrm{d}\, y) P_t(y, A), \tag{CKH}$$

for every s and t. In other words, $P = (P_t)_{t\geq 0}$ forms a semi-group.

The equations (CK) and (CKH) are called respectively, Chapman–Kolmogormov and homogeneous Chapman–Kolmogormov equation.

2.4.4 Definition. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let $\mathbb{F}$ be a filtration of events of $\mathscr{F}$. Let a probability measure μ be given and X be an adapted stochastic process with value space $(E, \mathscr{E})$ and initial distribution μ (that is, $X_0 \sim \mu$). Let P be a transition function on $(E, \mathscr{E})$.

(i) We say that X is a *Markov process* (relative to $\mathbb{F}$) with transition function P if, for every $f \geq 0$, $0 \leq s \leq t$, we have

$$\mathbb{E}[f(X_t)|\mathscr{F}_s] = P_{s,t} f(X_s), \quad \text{a.s.} \tag{2.12}$$

(ii) If furthermore the transition function is homogeneous, then X is a *homogeneous* Markov process (relative to $\mathbb{F}$) and (2.12) reads

$$\mathbb{E}[f(X_t)|\mathscr{F}_s] = P_{t-s} f(X_s), \quad \text{a.s.} \tag{2.13}$$

Notice that, if (2.12) and (2.13) hold for measurable nonnegative bounded functions, then they also hold for every measurable bounded function.

As we mentioned in the introduction of this section, for a Markov process with respect to a filtration $\mathbb{F}$, we expect that the conditional distribution of X_t given $\mathscr{F}_s$ is a measurable function of X_s. We say that an $\mathbb{F}$-adapted process X possesses the Markov property if

$$\mathbb{P}[X_t \in B|\mathscr{F}_s] = \mathbb{P}[X_t \in B|X_s], \quad 0 \leq s \leq t\,, \quad B \in \mathscr{E}. \tag{MP}$$

2.4.5 Theorem. *Let X be a Markov process relative to $\mathbb{F}$ with transition probability P. Then X possesses the Markov property* (MP).

Proof. Let $f \geq 0$ be a measurable function, then $P_{s,t} f(X_s)$ is $\sigma(X_s)$-measurable. Taking now $B \in \mathscr{E}$ and $f = 1_B$, yields $P_{s,t} f(x) = P_{s,t}(x, B)$. Therefore, using the property of the conditional expectation, X being a Markov process,

$$\mathbb{P}[X_t \in B|X_s] = \mathbb{E}[\mathbb{E}[1_B(X_t)|\mathscr{F}_s]|X_s] = \mathbb{E}[P_{s,t}(X_s, B)|X_s] = P_{s,t}(X_s, B) = \mathbb{P}[X_t \in B|\mathscr{F}_s]$$

and the proof is complete. □

For Markov processes it is easy to determine the family of the finite dimensional distributions, as the following theorem and the subsequent corollary show.

2.4.6 Theorem. *Let X be a Markov process relative to $\mathbb{F}$ with initial distribution and associated transition function μ and P respectively. Then, for every $f_1, \dots, f_n$ measurable, bounded and nonnegative functions, and every $0 = t_0 < t_1 < \ldots < t_n$, $n \in \mathbb{N}$, we have*

$$\mathbb{E}\left[\prod_{i=0}^{n} f_i(X_{t_i})\right] = \int_E \mu(dx_0) f_0(x_0) \int_E P_{0,t_1}(x_0, \mathrm{d}x_1) f_1(x_1) \cdots \int_E P_{t_{n-1},t_n}(x_{n-1}, \mathrm{d}x_n) f_n(x_n). \tag{2.14}$$

Proof. The proof is given by induction. For $n = 1$, using the properties of the conditional expectation and then the definition of a Markov process, we get

$$\begin{aligned}
\mathbb{E}[f_0(X_0) f_1(X_{t_1})] &= \mathbb{E}[f_0(X_0)\mathbb{E}[f_1(X_{t_1})|\mathscr{F}_0]] = \mathbb{E}[f_0(X_0) P_{0,t_1} f_1(X_0)] \\
&= \int_E \mu(\mathrm{d}\, x_0) f_0(x_0) P_{0,t_1} f_1(x_0) \\
&= \int_E \mu(\mathrm{d}\, x_0) f_0(x_0) \int_E P_{0,t_1}(x_0, \mathrm{d}\, x_1) f_1(x_1).
\end{aligned}$$

We now assume that the claim of the theorem holds for $m = n - 1$ and show it for $m = n$. Using the properties of the conditional expectation and then the definition of a Markov process, issues

$$\begin{aligned}
\mathbb{E}\left[\prod_{i=0}^{n} f_i(X_{t_i})\right] =& \mathbb{E}\left[\prod_{i=0}^{n-1} f_i(X_{t_i}) \mathbb{E}[f_n(X_{t_n})|\mathscr{F}_{t_{n-1}}]\right] \\
=& \mathbb{E}\left[\prod_{i=0}^{n-1} f_i(X_{t_i}) P_{t_{n-1},t_n} f_n(X_{t_{n-1}})\right] \\
=& \mathbb{E}\left[\prod_{i=0}^{n-2} f_i(X_{t_i}) \widetilde{f}_{n-1}(X_{t_{n-1}})\right],
\end{aligned}$$

where $\widetilde{f}_{n-1}(x) := f_{n-1}(x) P_{t_{n-1},t_n} f_n(x)$. Clearly (why?) $\widetilde{f}_{n-1}$ is bounded and nonnegative. Therefore we can apply the induction assumption to conclude the proof. □

2.4.7 Corollary. *Under the assumptions of Theorem 2.4.6,*

$$\mathbb{P}[X_0 \in B_0, \dots, X_{t_n} \in B_n] = \int_{B_0} \mu(\mathrm{d}x_0) \int_{B_1} P_{0,t_1}(x_0, \mathrm{d}x_1) \int_{B_2} \cdots \int_{B_n} P_{t_{n-1},t_n}(x_{n-1}, \mathrm{d}x_n)$$

for every $B_0, \dots, B_n \in \mathscr{E}$ and every $0 = t_0 < t_1 < \ldots < t_n$, $n \in \mathbb{N}$.

Proof. Let $f_i(x) = 1_{B_i}(x)$, $i = 0, \dots, n$ and apply Theorem 2.4.6. □

Theorem 2.4.6 shows that the transition probability P and the initial distribution μ of a Markov process X, completely determine the finite dimensional distribution of X, which are given in Corollary 2.4.7.

Existence of Markov processes. As a next step we want to show the existence of Markov processes. More precisely, let P be a transition function and μ a distribution both on $(E, \mathscr{E})$, we prove the existence of a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, of a stochastic process X on it such X is a Markov process with transition function P and initial distribution μ with respect to $\tilde{\mathbb{F}}^X$ (see Example 2.3.2). We apply Kolmogorov Extension Theorem (Theorem 2.1.7 and Theorem 2.1.8).

As a first step we associate with a transition function P and an initial distribution μ on $(E, \mathscr{E})$ a system of finite dimensional distribution. This will be the system of finite dimensional distribution associated with a Markov process X. According to Corollary 2.4.7, the unique possible definition is the following: For every $B_0, \ldots, B_n \in \mathscr{E}$ and every $0 = t_0 < t_1 < \ldots < t_n$, $n \in \mathbb{N}$, we define

$$\begin{aligned} P^{\mu}_{t_1,\ldots,t_n}&(B_0 \times B_1 \times \cdots \times B_n) := \\ &\int_{B_0} \mu(\mathrm{d}\, x_0) \int_{B_1} P_{0,t_1}(x_0, \mathrm{d}\, x_1) \int_{B_2} \cdots \int_{B_n} P_{t_{n-1},t_n}(x_{n-1}, \mathrm{d}\, x_n). \end{aligned} \tag{2.15}$$

The system $P^{\mu} := \{P^{\mu}_{t_1,\ldots,t_n},\ 0 = t_0 < t_1 < \ldots < t_n, n \in \mathbb{N}\}$ is the system of finite dimensional distributions generated by P. The aim is to associate with P^{μ} a Markov process.

2.4.8 Remark. (i) At this point P and μ are only a transition function and a probability distribution on $(E, \mathscr{E})$: They are not yet related with a Markov process.

(ii) For every $0 = t_0 < t_1 < \ldots < t_n$, $P^{\mu}_{t_1,\ldots,t_n}$ is a probability measure on $(E^{n+1}, \mathscr{E}^{n+1})$: For $B_i = E$, $i = 0, \ldots, n$, then $P^{\mu}_{t_1,\ldots,t_n}(E^{n+1}) = 1$.

(iii) Let f be a measurable bounded function. Then

$$\begin{aligned} \int_{E^{n+1}} & P^{\mu}_{t_1,\ldots,t_n}(\mathrm{d}x_0, \ldots, \mathrm{d}x_n) f(x_0, \ldots, x_n) = \\ & \int_{E^{n+1}} \mu(dx_0) P_{0,t_1}(x_0, \mathrm{d}\, x_1) \cdots P_{t_{n-1},t_n}(x_{n-1}, \mathrm{d}\, x_n) f(x_0, \ldots, x_n). \end{aligned} \tag{2.16}$$

To see (2.16), consider $f_i = 1_{B_i}$, $B_i \in \mathscr{E}$, $i = 0, \ldots, n$ and $f(x_0, \ldots, x_n) = \prod_{i=0}^n f_i(x_i)$. In this special case (2.16) clearly follows from (2.15). The system of bounded measurable functions $\mathscr{C} := \{1_{B_0 \times \cdots \times B_n},\ B_i \in \mathscr{E},\ i = 0, \ldots, n,\ n \in \mathbb{N}\}$ is a multiplicative class which generates $\mathscr{E}^{n+1}$. Using dominated convergence, it is easy to show that the family of functions $\mathscr{K} := \{f$ measurable and bounded : (2.16) holds$\}$ is a monotone vector space. Furthermore $\mathscr{K}$ contains $\mathscr{C}$. An application of Theorem 1.2.2 yields (2.16) for every f measurable and bounded function. In particular, if $f(x_0, \ldots, x_n) = \prod_{i=0}^n f_i(x_i)$, with $f_i \geq 0$ measurable and bounded, $i = 0, \ldots, n$, then (2.16) becomes

$$\begin{aligned} \int_{E^{n+1}} & P^{\mu}_{t_1,\ldots,t_n}(\mathrm{d}x_0, \ldots, \mathrm{d}x_n) \prod_{i=0}^n f_i(x_i) = \\ & \int_E \mu(\mathrm{d}\, x_0) f_0(x_0) \int_E P_{0,t_1}(x_0, \mathrm{d}\, x_1) f_1(x_1) \int_E \cdots \int_E P_{t_{n-1},t_n}(x_{n-1}, \mathrm{d}\, x_n) f_n(x_n). \end{aligned} \tag{2.17}$$

2.4.9 Lemma. *The system P^{μ} of finite dimensional distributions in* (2.15)*, is consistent, that is, it satisfies condition* (2.7)*.*

Proof. We know that $P^{\mu}_{t_1,\dots,t_n}$ is a probability measure on $(E^{n+1}, \mathscr{E}^{n+1})$, for every $0 = t_0 < t_1 < \ldots < t_n$, $n \in \mathbb{N}$. We now show that it fulfils (2.7). Let $B_0, \dots, B_n \in \mathscr{E}$ and $0 = t_0 < t_1 < \ldots < t_n$, $n \in \mathbb{N}$. Then

$$\begin{aligned} P^{\mu}_{t_1,\dots,t_{k-1},t_k,t_{k+1},\dots,t_n}&(B_0 \times B_1 \times \cdots \times B_{k-1} \times E \times B_{k+1} \times \cdots \times B_n) = \\ &\int_{B_0} \mu(\mathrm{d}x_0) \int_{B_1} P_{0,t_1}(x_0, \mathrm{d}x_1) \cdots \int_{B_{k-1}} P_{t_{k-2},t_{k-1}}(x_{k-2}, \mathrm{d}x_{k-1}) \\ &\times \int_E P_{t_{k-1},t_k}(x_{k-1}, \mathrm{d}x_k) \int_{B_{k+1}} P_{t_k,t_{k+1}}(x_k, \mathrm{d}x_{k+1}) \cdots \\ &\times \int_{B_n} P_{t_{n-1},t_n}(x_{n-1}, \mathrm{d}x_n). \end{aligned} \tag{2.18}$$

Applying Fubini's theorem, we deduce the relation

$$\begin{aligned} \int_E P_{t_{k-1},t_k}(x_{k-1},\mathrm{d}x_k) \int_{B_{k+1}} P_{t_k,t_{k+1}}(x_k, \mathrm{d}x_{k+1}) &= \\ \int_{B_{k+1}} \int_E P_{t_{k-1},t_k}(x_{k-1}, \mathrm{d}x_k) P_{t_k,t_{k+1}}(x_k, \mathrm{d}x_{k+1}) &= \\ \int_{B_{k+1}} P_{t_{k-1},t_{k+1}}(x_{k-1}, \mathrm{d}x_{k+1}) & \end{aligned}$$

where in the last equality we used Chapman–Kolmogorov equation (CK). Inserting the latter computation in (2.18), yields

$$\begin{aligned} P^{\mu}_{t_1,\dots,t_{k-1},t_k,t_{k+1},\dots,t_n}&(B_0 \times B_1 \times \cdots \times B_{k-1} \times E \times B_{k+1} \times \cdots \times B_n) = \\ &P^{\mu}_{t_1,\dots,t_{k-1},t_{k+1},\dots,t_n}(B_0 \times B_1 \times \cdots \times B_{k-1} \times B_{k+1} \times \cdots \times B_n) \end{aligned}$$

which is the consistency condition (2.7) and the proof is concluded. □

Now we are ready to formulate the theorem about the existence of Markov processes. We recall that we use the notation $(E, \mathscr{E}) := (\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

2.4.10 Theorem. *Let P be a transition function and μ a distribution on $(E, \mathscr{E})$.*

(i) *Then there exists a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a stochastic process X on $(\Omega, \mathscr{F}, \mathbb{P})$ with state space $(E, \mathscr{E})$ such that X_0 is μ-distributed and P^{μ} given by* (2.15) *is the system of the finite dimensional distribution associated with X.*

(ii) *Furthermore the process X is a Markov process with respect to the filtration $\tilde{\mathbb{F}}^X$.*

Proof. The measurable space $(E, \mathscr{E})$ (with the Euclidean norm) is clearly a Polish space. Therefore it fulfils the assumptions of Theorem 2.1.8. Because of Lemma 2.4.9 P^{μ} is a consistent system of finite dimensional distributions. Hence (i) follows from Theorem 2.1.8. (Exercise: give the explicit definition of $(\Omega, \mathscr{F}, \mathbb{P})$ and of X). We have to prove claim (ii). Because of (i), $P^{\mu} = \{P^{\mu}_{t_1,\dots,t_n},\ 0 = t_0 < t_1 < \ldots < t_n\}$ is the system of finite dimensional distribution associated with X. Therefore, for every measurable nonnegative bounded f_i, $i = 0, \dots, n$, we have

$$\mathbb{E}\Big[\prod_{i=0}^{n} f_i(X_{t_i})\Big] = \int_{E^{n+1}} P^{\mu}_{t_1,\dots,t_n}(\mathrm{d}x_0, \dots, \mathrm{d}x_n) \prod_{i=0}^{n} f_i(x_i)$$
$$= \int_E \mu(\mathrm{d}x_0) f_0(x_0) \int_E P_{0,t_1}(x_0, \mathrm{d}x_1) f_1(x_1) \int_E \cdots \int_E P_{t_{n-1},t_n}(x_{n-1}, \mathrm{d}x_n) f_n(x_n)$$

as we have seen in Remark 2.4.8 (iii). To conclude, we apply Lemma 2.4.11 below. □

2.4.11 Lemma. *Let X be a stochastic process on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with value space $(E, \mathscr{E})$. If (2.14) holds for every measurable nonnegative bounded function f_i, $i = 0, \dots, n$, $0 = t_0 < \dots < t_n$, $n \in \mathbb{N}$, then X is a Markov process with respect to the filtration $\tilde{\mathbb{F}}^X$.*

Proof. Let $0 = t_0 < \dots < t_n \le s < t$ and let $g \ge 0$ be a measurable nonnegative bounded function. Let f_i, $i = 0, \dots, n$, be measurable nonnegative bounded functions and $f_{n+1} := g$. Then, because of (2.14),

$$\mathbb{E}\Big[\prod_{i=0}^{n} f_i(X_{t_i}) g(X_t)\Big] = \mathbb{E}\Big[\prod_{i=0}^{n+1} f_i(X_{t_i})\Big] = \int_E \mu(\mathrm{d}x_0) f_0(x_0) \int_E \cdots \int_E P_{t_{n-1},t_n}(x_{n-1}, \mathrm{d}x_n) f_n(x_n) \int_E P_{t_n,t}(x_n, \mathrm{d}y) g(y). \tag{2.19}$$

On the other side

$$\mathbb{E}\Big[\prod_{i=0}^{n} f_i(X_{t_i}) P_{s,t} g(X_s)\Big] = \int_E \mu(\mathrm{d}x_0) f_0(x_0) \int_E \cdots \int_E P_{t_{n-1},t_n}(x_{n-1}, \mathrm{d}x_n) f_n(x_n) \int_E P_{t_n,s}(x_n, \mathrm{d}z) \int_E P_{s,t}(z, \mathrm{d}y) g(y). \tag{2.20}$$

Applying Fubini theorem and Chapman–Kolmogorov equation yields

$$\int_E P_{t_n,s}(x_n, \mathrm{d}z) \int_E P_{s,t}(z, \mathrm{d}y) g(y) = \int_E P_{t_n,t}(x_n, \mathrm{d}y) g(y),$$

that is (2.19) and (2.20) are equal for every bounded f_i, $i = 0, \dots, n$. In particular,

$$\mathbb{E}[1_{B_0 \times \dots \times B_n}(X_0, \dots, X_{t_n}) g(X_t)] = \mathbb{E}[1_{B_0 \times \dots \times B_n}(X_0, \dots, X_{t_n}) P_{s,t} g(X_s)],$$

for every $B_i \in \mathscr{E}$, $i = 0, \dots, n$, $0 = t_0 < t_1 < \dots < t_n \le s$. These functions clearly build a multiplicative class $\mathscr{C}$ which generates $\tilde{\mathscr{F}}^X_s$. From Theorem 1.2.2 it follows

$$\mathbb{E}[W g(X_t)] = \mathbb{E}[W P_{s,t} g(X_s)]$$

for every bounded $\tilde{\mathscr{F}}^X_s$-measurable random variable W. But this means

$$\mathbb{E}\big[g(X_t) | \tilde{\mathscr{F}}^X_s\big] = P_{s,t} g(X_s), \quad 0 \le s < t,$$

for every measurable bounded nonnegative g and the proof is complete. □

Markov property and Markov processes. We complete this section about Markov processes discussing the converse of Theorem 2.4.5. More precisely, we have seen that any Markov process X possesses the Markov property (MP). We now consider a process X satisfying the Markov property (MP) and ask if it is a Markov process with respect to some filtration $\mathbb{F}$. The problem is here that one need to construct a transition probability starting from (MP). We shall see that this problem is strictly connected with the existence of a regular version of the conditional probabilities $\mathbb{P}[X_t \in B|X_s]$ and $\mathbb{P}[X_t \in B|\mathscr{F}_s]$.

We first give a result concerning the existence of a regular version of the conditional distribution $\mathbb{P}[X_t \in B|X_s]$ which involves transition probabilities. We refer to Kallenberg (1997), Theorem 5.3.

2.4.12 Theorem. *Let X be a stochastic process with state space $(E, \mathscr{E})$ satisfying the Markov property with respect to the filtration $\mathbb{F}$. Then there exists a transition probability P on $E \times \mathscr{E}$, such that*

$$P_{s,t}(X_s, B) = \mathbb{P}[X_t \in B|X_s], \quad B \in \mathscr{E}, \quad 0 \le s \le t$$

$\mathbb{P}_{X_s}$*-a.s.*

Because of Markov property we have

$$\mathbb{P}[X_t \in B|\mathscr{F}_s] = \mathbb{P}[X_t \in B|X_s] = P_{s,t}(X_s, B),$$

where P is the transition probability of Theorem 2.4.12. Notice that the previous equalities hold a.s. At this point it is important to use the structure of $(E, \mathscr{E})$: Because this is a Borel space, we can choose a common exceptional set for every $B \in \mathscr{E}$.

Let g be a measurable nonnegative bounded function and $0 \le s < t$. From Theorem 2.1.9, we have

$$\begin{aligned} \mathbb{E}[g(X_t)|\mathscr{F}_s] &= \int_E \mathbb{P}[X_s \in \mathrm{d}\, y|\mathscr{F}_s] g(y) = \int_E \mathbb{P}[X_s \in \mathrm{d}\, y|X_s] g(y) \\ &= \int_E P_{s,t}(X_s, \mathrm{d}\, y) g(y) = P_{s,t} g(X_s)\,, \end{aligned} \tag{2.21}$$

where in the second equality we used (MP). The previous relation shows that, if P is a transition function, then X is a Markov process in the sense of Definition 2.4.4. Therefore we have to show that P satisfies (CK). Let $0 \le r \le s \le t$ and $B \in \mathscr{E}$. Then

$$\begin{aligned} P_{r,t}(X_r, B) &= \mathbb{P}[X_t \in B|X_r] \overset{\text{(MP)}}{=} \mathbb{P}[X_t \in B|\mathscr{F}_r] = \mathbb{E}[\mathbb{P}[X_t \in B|\mathscr{F}_s]|\mathscr{F}_r] \\ &= \mathbb{E}[\mathbb{P}[X_t \in B|X_s]|\mathscr{F}_r] = \mathbb{E}[P_{s,t}(X_s, B)|\mathscr{F}_r]. \end{aligned}$$

where in the last but one equality we used (2.21). Setting $g(X_s) := P_{s,t}(X_s, B)$ in the previous formula, we deduce from (2.21)

$$P_{r,t}(X_r, B) = \mathbb{E}[g(X_t)|\mathscr{F}_r] = P_{r,t} g(X_r) = \int_E P_{r,s}(X_r, \mathrm{d}y) P_{s,t}(y, B)$$

which is (CK). We summarize the previous discussion in the following theorem:

2.4.13 Theorem. *Let X be an $\mathbb{F}$-adapted process with values in $(E, \mathscr{E})$ fulfilling* (MP). *Then X is a Markov process with respect to the filtration $\mathbb{F}$ with transition probability P given by $P_{s,t}(x, B) := \mathbb{P}[X_t \in B|X_s = x]$, for every $B \in \mathscr{E}$, $0 \le s \le t$.*

2.5 Processes with independent increments

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space and let $X = (X_t)_{t\geq 0}$ be a stochastic process. The random variable $X_t - X_s$, $0 \leq s \leq t$, is called *an increment* of the process X over $[s,t]$. We furthermore consider a filtration $\mathbb{F} = (\mathscr{F}_t)_{t\geq 0}$ to which the process X is adapted.

2.5.1 Definition. (i) We say that X has independent increments if the random vector $(X_{t_0}, X_{t_1} - X_{t_0}, \ldots X_{t_n} - X_{t_{n-1}})$ is independent, for every $0 \leq t_0 < t_1 < \ldots < t_n$, $n \in \mathbb{N}$.

(ii) If there exists $n \in \mathbb{N}$ such that the random vector $(X_{t_0}, X_{t_1} - X_{t_0}, \ldots X_{t_k} - X_{t_{k-1}})$ is distributed as the random vector $(X_{t_0}, X_{t_1 - t_0}, \ldots, X_{t_k - t_{k-1}})$, for every $0 \leq t_0 < t_1 < \ldots < t_k$, $k = 0, \ldots, n$ then we say that X has *homogeneous n-dimensional increments*. We say that X has *homogeneous increments* if it has homogeneous n-dimensional increments for every $n \in \mathbb{N}$.

(iii) If X is $\mathbb{F}$-adapted and $X_t - X_s$ is independent of $\mathscr{F}_s$, $0 \leq s \leq t$, then we say that X has *independent increments with respect to* $\mathbb{F}$ (or $\mathbb{F}$*-independent increments*).

2.5.2 Exercise. Let X be a process with independent and 1-dimensional homogeneous increments. Then it has homogeneous increments.

The next theorem establishes the relation between property (i) and (ii) in Definition 2.5.1. The proof of the theorem is left as an exercise. It can be found in Bauer (1996), §45, or in WTHM, Satz 11.3.

2.5.3 Theorem. (i) *If X has independent increments with respect to $\mathbb{F}$, then it has independent increments.*

(ii) *If $\mathbb{F} = \tilde{\mathbb{F}}^X$, then a process with independent increments has also independent* $\mathbb{F}$-independent increments.

Processes with independent increments are a special case of processes possessing the Markov property (MP). To see it, we need the following lemma.

2.5.4 Lemma. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and $(E, \mathscr{E})$ a measurable space. Let $\mathscr{G}$ and $\mathscr{H}$ independent sub-σ-algebras of $\mathscr{F}$. Let $\varphi : (E \times \Omega, \mathscr{E} \otimes \mathscr{H}) \longrightarrow (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ a bounded measurable function. If $X : (\Omega, \mathscr{G}) \longrightarrow (E, \mathscr{E})$ is a $(\mathscr{G}, \mathscr{E})$-measurable random variable, then*

$$\mathbb{E}[\varphi(X, \cdot)|\mathscr{G}] = \Big(\mathbb{E}[\varphi(x, \cdot)]\Big)\Big|_{x=X}$$

Proof. First we deduce the result for the simple case of φ with separated variables, that is $\varphi(x, \omega) = \varphi_1(x)\varphi_2(\omega)$ with φ_1, φ_2 bounded, φ_1 $(\mathscr{E}, \mathscr{B}(\mathbb{R}))$-measurable and ϕ_2 $(\mathscr{H}, \mathscr{B}(\mathbb{R}))$-measurable. In this special case, the claim follows using first the $\mathscr{G}$-measurability of X and then the $\mathscr{G}$-independence of φ_2:

$$\begin{aligned}\mathbb{E}[\varphi_1(X)\varphi_2(\cdot)|\mathscr{G}] &= \varphi_1(X)\mathbb{E}[\varphi_2(\cdot)|\mathscr{G}] = \varphi_1(X)\mathbb{E}[\varphi_2(\cdot)] \\ &= \Big(\varphi_1(x)\mathbb{E}[\varphi_2(\cdot)]\Big)\Big|_{x=X} = \Big(\mathbb{E}[\varphi(x, \cdot)]\Big)\Big|_{x=X}.\end{aligned}$$

As a next step, we consider the general case. Let $\mathscr{K}$ be the class of the functions φ as in the assumptions of the lemma, such that the claim holds. Then $\mathscr{K}$ clearly contains the multiplicative class $\mathscr{C}$ of functions with separated variables, which generates $\mathscr{E} \otimes \mathscr{H}$.

Because of the monotone convergence theorem, we can conclude that $\mathscr{K}$ is a monotone vector space and from Theorem 1.2.2, we deduce that $\mathscr{K}$ contains all bounded functions $\varphi(\cdot,\cdot)$ and the proof is complete. □

We remark that the notation $(\mathbb{E}[\varphi(x,\cdot)])\Big|_{x=X}$ is a computation rule and it says: First fix the $\mathscr{G}$-measurable part of the random variable $\varphi(X,\cdot)$ and then integrate the independent part. After this integration, replace the fixed x with X, to get a $\mathscr{G}$-measurable quantity.

2.5.5 Theorem. *Let X be a process with $\mathbb{F}$-independent increments.* (i) *Then it possesses the Markov property* (MP). *In particular, every process with independent increments possesses the Markov property with respect to $\tilde{\mathbb{F}}^X$.*

(ii) *X is a Markov process with respect to $\mathbb{F}$. In particular, every process with independent increments is a Markov process with respect to $\tilde{\mathbb{F}}^X$.*

Proof. We start proving (i). We define $\mathscr{H} := \sigma(X_{t+s} - X_t)$ and $\mathscr{G} := \mathscr{F}_t$. Because of the $\mathbb{F}$-independence of the increments of X, these two σ-algebras are independent. Because of the adaptedness of X, is X_t $\mathscr{F}_t$-measurable. If we show that $\mathbb{E}[\psi(X_{t+s})|\mathscr{F}_t]$ is $\sigma(X_t)$-measurable, then follows

$$\mathbb{E}[\psi(X_{t+s})|\mathscr{F}_t] = \mathbb{E}[\psi(X_{t+s})|X_t]$$

because of the properties of the conditional expectation. Let

$$\varphi(x,\omega) := \psi(X_{t+s}(\omega) - X_t(\omega) + x).$$

Then φ is $\mathscr{E}\otimes\mathscr{H}$-measurable and hence an application of Lemma 2.5.4 yields

$$\mathbb{E}[\psi(X_{t+s})|\mathscr{F}_t]\mathbb{E}[\varphi(X_t,\cdot)|\mathscr{F}_t] = \Big(\mathbb{E}[\varphi(x,\cdot)]\Big)\Big|_{x=X_t}$$

which is clearly $\sigma(X_t)$-measurable. If now $\mathbb{F} = \tilde{\mathbb{F}}^X$, the claim follows from Theorem 2.5.3. Now (ii) follows from (i) and Theorem 2.4.13. □

It is easy to check if a process with independent increments is a martingale.

2.5.6 Lemma. *Let X be an adapted process with $\mathbb{F}$-independent increments such that $X_0 = 0$. Then X is a martingale (not necessarily càdlàg) if and only if the random variable X_t is integrable and $\mathbb{E}[X_t] = 0$, for every $t \geq 0$.*

Proof. If X is a process with $\mathbb{F}$-independent increments such that $X_0 = 0$ and a (not necessarily càdlàg) martingale then $\mathbb{E}[X_t] = \mathbb{E}[X_0] = 0$, $t \geq 0$. Conversely, if X is a process with $\mathbb{F}$-independent increments such that $X_0 = 0$ and that $\mathbb{E}[X_t] = 0$, $t \geq 0$, we get

$$\mathbb{E}[X_t|\mathscr{F}_s] = \mathbb{E}[X_t - X_s|\mathscr{F}_s] + X_s = \mathbb{E}[X_t - X_s] + X_s = X_s,$$

proving that X is an $\mathbb{F}$-martingale. □

The $\mathbb{F}$-independence of the increments and the homogeneity of the one-dimensional increments are stable under convergence in probability, as the following lemma shows.

2.5.7 Lemma. *Let X be an $\mathbb{F}$-adapted process. If $(X^n)_{n\geq 1}$ is a sequence of processes with $\mathbb{F}$-independent increments (resp., homogeneous one-dimensional increments) such that X^n_t converges to X_t in probability, for every $t \geq 0$, as $n \to +\infty$, then X has $\mathbb{F}$-independent increments (resp., homogeneous one-dimensional increments).*

Proof. We assume that the sequence $(X^n)_{n\geq 1}$ has $\mathbb{F}$-independent increments (resp., homogeneous one-dimensional increments). For every $0 \leq s \leq t$, we have

$$\mathbb{E}[\mathrm{e}^{iu(X^n_t - X^n_s)}|\mathscr{F}_s] = \mathbb{E}[\mathrm{e}^{iu(X^n_t - X^n_s)}] \quad (\text{resp.,} \quad \mathbb{E}[\mathrm{e}^{iu(X^n_t - X^n_s)}] = \mathbb{E}[\mathrm{e}^{iuX^n_{t-s}}]), \quad u \in \mathbb{R}.$$

Letting n converge to $+\infty$ in the previous formula and applying the theorem of Lebesgue on dominated convergence we get

$$\mathbb{E}[\mathrm{e}^{iu(X_t - X_s)}|\mathscr{F}_s] = \mathbb{E}[\mathrm{e}^{iu(X_t - X_s)}], \quad (\text{ resp.,} \quad \mathbb{E}[\mathrm{e}^{iu(X_t - X_s)}] = \mathbb{E}[\mathrm{e}^{iuX_{t-s}}]), \quad u \in \mathbb{R},$$

which concludes the proof. □

We only consider processes X with $\mathbb{F}$-independent increments which are also *stochastically continuous* (cf. Definition 2.2.1 (iv)). In this special case we speak of *additive processes.*

2.5.8 Definition. Let X be an *adapted* and *stochastically continuous* process such that $X_0 = 0$.

(i) We say that X is an *additive process in law* if it has independent increments.

(ii) If X is a *càdlàg* additive process in law, we simply call it an *additive process.*

(iii) We say that X is an *additive process in law relative to the filtration* $\mathbb{F}$ if it has $\mathbb{F}$-independent increments. If X is also *càdlàg*, we simply call it an *additive process relative to* $\mathbb{F}$.

The notation $(X, \mathbb{F})$ emphasizes the filtration relative to which X is an additive process (resp., an additive process in law) and sometimes we simply say that $(X, \mathbb{F})$ is an additive process (resp., an additive process in law) to mean that X is an additive process (resp., an additive process in law) relative to the filtration $\mathbb{F}$.

2.5.9 Proposition. *Let X be an adapted stochastically continuous process such that $X_0 = 0$.*

(i) *If $(X, \mathbb{F})$ is an additive process in law, then X is an additive process in law.*

(ii) *The following statements are equivalent:*

(a) *X is an additive process in law.*

(b) *$(X, \tilde{\mathbb{F}}^X)$ is an additive process in law.*

(c) *$(X, \tilde{\mathbb{F}}^X_+)$ is an additive process in law.*

(d) *$(X, \mathbb{F}^X)$ is an additive process in law.*

(e) *$(X, \mathbb{F}^X_+)$ is an additive process in law.*

Proof. The statement (i) follows from Theorem 2.5.3 (i). Now we come to the equivalences in (iii). The statements (a) and (b) are equivalent from Theorem 2.5.3 (ii). Furthermore (b) and (d) are equivalent because $\tilde{\mathscr{F}}^X_t$ and $\mathscr{F}^X_t$ differs on null sets: In general if Y is a random variable which is independent of a sub-σ-algebra $\mathscr{G}$, then Y is also independent of $\overline{\mathscr{G}} := \mathscr{G} \vee \mathscr{N}(\mathbb{P})$. Indeed, if $A \in \overline{\mathscr{G}}$, then $A = A' \cup N$, with

$A' \in \mathscr{G}$ and $N \in \mathscr{N}(\mathbb{P})$. This implies that the disjoint union $A\Delta A' = N \cap (A')^c$ is also a null set (recall: $B\Delta C := B \setminus C \cup C \setminus B$). Hence $A \setminus A' \subseteq A\Delta A$ is a null set. Because the probability space is complete, then all these null sets are measurable and hence $\mathbb{P}[A] = \mathbb{P}[A \setminus A'] + \mathbb{P}[A \cap A'] = \mathbb{P}[A']$. More generally, for every $F \in \mathscr{G}$, we have $\mathbb{P}[A \cap F] = \mathbb{P}[A' \cap F]$. If we now take $F := \{Y \in B\}$, $B \in \mathscr{E}$, we get

$$\mathbb{P}[A, F] = \mathbb{P}[A', F] = \mathbb{P}[A']\mathbb{P}[F] = \mathbb{P}[A]\mathbb{P}[F]$$

for every $A \in \overline{\mathscr{G}}$, meaning that $\overline{\mathscr{G}}$ and Y are independent. This explains (b)$\Longrightarrow$(d). The converse implication follows because $\mathscr{F}_t^X \supseteq \tilde{\mathscr{F}}_t^X$, for every $t \geq 0$. In the same way one proofs the equivalence between (c) and (e). We have to show that (b) and (c) are equivalent. It is clear that from (c) follows (b). We now show the converse implication. Let $A \in \tilde{\mathscr{F}}_{t+}^X$ and $\xi := 1_A$. Because of the definition of $\tilde{\mathscr{F}}_{t+}^X$, ξ is $\tilde{\mathscr{F}}_{t+\frac{1}{n}}^X$-measurable, for every n. Furthermore, since the $\tilde{\mathbb{F}}^X$-independence of the increments of X, ξ is independent of $\eta_n := X_{t+\frac{1}{n}+s} - X_{t+\frac{1}{n}}$. The joint characteristic function of ξ and η_n, say $\varphi^{\eta_n,\xi}(u, v)$, equals $\varphi^{\eta_n}(u)\varphi^{\xi}(v)$, where φ^{η_n} and φ^{ξ} are the characteristic functions of η_n and ξ respectively. Let now $\eta := X_{t+s} - X_t$. Because of the stochastic continuity of X, η_n converges to η in law, as $n \to +\infty$. Hence

$$\varphi^{\eta}(u)\varphi^{\xi}(v) = \lim_{n\to\infty} \varphi^{\eta_n}(u)\varphi^{\xi}(v) = \lim_{n\to\infty} \varphi^{\eta_n,\xi}(u, v) = \varphi^{\eta,\xi}(u, v)$$

which implies the independence of η of $\tilde{\mathscr{F}}_{t+\frac{1}{n}}^X$, for every $n \in \mathbb{N}$. The prove is now complete because $\tilde{\mathscr{F}}_{t+\frac{1}{n}}^X \supseteq \tilde{\mathscr{F}}_{t+}^X$, for every $n \in \mathbb{N}$. □

We remark that Wang (1981) proved that for an additive process X, the filtration $\mathbb{F}^X$ coincides with $\mathbb{F}_+^X$.

2.5.10 Lemma. *Let X be a stochastic process with one-dimensional homogeneous increments such that $X_0 = 0$ a.s. and that $X_t \longrightarrow 0$ in probability as $t \downarrow 0$. Then X is stochastically continuous. In particular, any càdlàg process with one-dimensional homogeneous increments which starts at zero is stochastically continuous.*

2.5.11 Exercise. Prove Lemma 2.5.10.

For the next result we refer to He, Wang & Yan (1992), Theorem 2.68.

2.5.12 Theorem. *If $\mathbb{F}$ is a filtration satisfying the usual conditions and $(X, \mathbb{F})$ is an additive process in law, then there exists a modification, again denoted by X, such that $(X, \mathbb{F})$ is an additive process.*

We shall always assume that the filtration $\mathbb{F}$ with respect to which we consider an additive process X satisfies the usual conditions and therefore we always consider the càdlàg version of X. In other words, the notation $(X, \mathbb{F})$ has to be understood as: X is an additive process with respect to the filtration $\mathbb{F}$ satisfying the usual conditions. Notice that in Proposition 2.5.9 (ii), if we remove (a), the chain of equivalences holds also if we replace $\tilde{\mathbb{F}}^X$ with a general filtration $\tilde{\mathbb{F}}$. This means that, if $\tilde{\mathbb{F}}$ does not satisfies the usual conditions, then we can consider $\mathbb{F}_+$, that is the smallest filtration satisfying the usual conditions and containing $\tilde{\mathbb{F}}$. Then $(X, \mathbb{F}_+)$ is an additive process and we can pass to one of its càdlàg versions, which is again an additive process with respect to $\mathbb{F}_+$.

Strong Markov property. Let $(X,\mathbb{F})$ be an additive process. For any $u\in\mathbb{R}$, $0\le s<t$, we introduce

$$\varphi_{s,t}(u):=\mathbb{E}\big[\exp(\mathrm{i}\,u(X_t-X_s))\big] \tag{2.22}$$

that is, $\varphi_{s,t}$ is the characteristic function of the random variable X_t-X_s. The $\mathbb{F}$-independence of the increments implies

$$\varphi_{s,t}(u)=\varphi_{s,r}(u)\varphi_{r,t}(u),\quad 0\le s<v<t. \tag{2.23}$$

Because of the stochastic continuity of X, X_r converges to X_v *in probability*, as $r\to v$. This implies that X_r converges to X_v *in law*, as $u\to v$ (this means that $\mathbb{P}_{X_r}$ weakly converges to $\mathbb{P}_{X_v}$, $u\to v$). We can therefore conclude that $\varphi_{\cdot,\cdot}(\cdot)$ is continuous is all its arguments (a characteristic function is always continuous in u). Furthermore, the following lemma holds:

2.5.13 Lemma. *Let $(X,\mathbb{F})$ be an additive process. Then, for all $u\in\mathbb{R}$ and all $0\le s<t$, we have $\varphi_{s,t}(u)\neq 0$.*

Proof. Let $t_0:=\inf\{t\ge s:\varphi_{s,t}(u)=0\}$. Since $\varphi_{s,s}(u)=1$, then $t_0>s$. Let us assume $t_0<+\infty$. By definition of t_0, it holds $\varphi_{s,t_0}(u)=0$. Let us take $t\in(s,t_0)$. From (2.23), we have $0=\varphi_{s,t_0}(u)=\varphi_{s,t}(u)\varphi_{t,t_0}(u)$. But $\varphi_{s,t}(u)\neq 0$, hence it must be $\varphi_{t,t_0}(u)=0$. Letting $t\uparrow t_0$, by continuity of the characteristic function of an additive process, we deduce function $\varphi_{t_0,t_0}(u)=\lim_{t\uparrow t_0}\varphi_{t,t_0}(u)=0$. But this is a contradiction, because $\varphi_{t_0,t_0}(u)=1$ by definition. Therefore $t_0=+\infty$ and the proof is concluded. □

For an additive process relative to $\mathbb{F}$ we can introduce

$$Z_{s,t}(u):=\frac{1}{\varphi_{s,t}(u)}\exp(\mathrm{i}u(X_t-X_s))\,. \tag{2.24}$$

2.5.14 Proposition. *The process $(Z_{s,t}(u))_{t\ge s}$ defined by* (2.24) *is a martingale with respect to $(\mathscr{F}_t)_{t\ge s}$, for every $u\in\mathbb{R}$.*

Proof. Let $0\le s\le r<t$. Because of the $\mathbb{F}$-independence of the increments and (2.23), we get

$$\begin{aligned}\mathbb{E}[Z_{s,t}(u)|\mathscr{F}_r]&=\frac{1}{\varphi_{s,t}(u)}\exp(\mathrm{i}u(X_r-X_s))\mathbb{E}[\exp(\mathrm{i}u(X_t-X_r))|\mathscr{F}_r]\\&=\frac{\varphi_{r,t}(u)}{\varphi_{s,t}(u)}\exp(\mathrm{i}u(X_r-X_s))=Z_{s,r}(u)\,.\end{aligned}$$

□

We shall write $Z_t:=Z_{0,t}(u)$ and $\varphi_t:=\varphi_{0,t}$. The next theorem shows that for an additive process the *strong* Markov property holds:

2.5.15 Theorem (Strong Markov property)**.** *Let $(X,\mathbb{F})$ be an additive process and τ be a finite valued stopping time. Let $Y=(Y_t)_{t\ge0}$ be defined by $Y_t:=X_{\tau+t}-X_\tau$, $t\ge 0$. Then*

(i) *Y_t is independent of $\mathscr{F}_\tau$;*

(ii) *$(Y_t)_{t\ge0}$ has $(\mathscr{F}_{t+\tau})_{t\ge0}$ independent increments;*

(iii) *Y is independent of $\mathscr{F}_\tau$.*

If furthermore X has homogeneous 1-dimensional increments, then Y_t has the same law of X_t.

Proof. Let us first assume that τ is a bounded stopping time. Then because $Z = (Z_t)_{t\geq 0}$ defined is an $\mathbb{F}$-martingale, from Theorem 2.3.13 we deduce $\mathbb{E}[Z_{\tau+t}|\mathscr{F}_\tau] = Z_\tau$, a.s. Therefore

$$\mathbb{E}[\exp(\mathrm{i}u(X_{\tau+t} - X_\tau))|\mathscr{F}_\tau] = \frac{\varphi_{\tau+t}(u)}{\varphi_\tau(u)} = \varphi_{\tau,\tau+t}(u), \tag{2.25}$$

meaning that $X_{\tau+t} - X_\tau$ is independent of $\mathscr{F}_\tau$. If furthermore, X has homogeneous one dimensional increments, then $\varphi_{\tau,\tau+t} = \varphi_t$. We now assume that τ is an a.s. finite stopping time. Then $\tau_n := \tau \wedge n$ is a bounded stopping time and for every $A \in \mathscr{F}_\tau$, we have $A \cap \{\tau \leq n\} \in \mathscr{F}_{\tau_n} (= \mathscr{F}_\tau \cap \mathscr{F}_n)$. Because of the first step we get

$$\mathbb{E}[\exp(\mathrm{i}v1_{A\cap\{\tau\leq n\}} + \mathrm{i}u(X_{\tau_n+t} - X_{\tau_n}))] = \mathbb{E}[\exp(\mathrm{i}v1_{A\cap\{\tau\leq n\}})]\mathbb{E}[\exp(\mathrm{i}u(X_{\tau_n+t} - X_{\tau_n}))]$$

and dominated convergence yields

$$\mathbb{E}[\exp(\mathrm{i}v1_A + \mathrm{i}u(X_{\tau+t} - X_\tau))] = \mathbb{E}[\exp(\mathrm{i}v1_A)]\mathbb{E}[\exp(\mathrm{i}u(X_{\tau+t} - X_\tau))]$$

meaning that Y_t is independent of $\mathscr{F}_\tau$. We now see (ii). Clearly for $0 \leq s < t$, setting $u := t-s$ and $\sigma := \tau+s$, for every a.s. finite stopping time τ, we have $Y_t - Y_s = X_{\sigma+u} - X_\sigma$ which is independent of $\mathscr{F}_\sigma = \mathscr{F}_{\tau+s}$. We have to prove that Y is $(\mathscr{F}_{\tau+t})_{t\geq 0}$-adapted. To see it, because X is adapted and càdlàg, then it is a progressively measurable process and hence X_ζ is an $\mathscr{F}_\zeta$-measurable random variable, for every a.s. finite valued stopping time ζ (cf. Definition 2.3.3, Theorem 2.3.5 and the discussion after Definition 2.3.6). Therefore, X_τ and $X_{\tau+t}$ are respectively $\mathscr{F}_\tau$ and $\mathscr{F}_{\tau+t}$-measurable. We can now conclude that Y_t is $\mathscr{F}_{\tau+t}$measurable because $\mathscr{F}_\tau \subseteq \mathscr{F}_{\tau+t}$. To see (iii) it is enough to show that for every $t_1 < \ldots < t_n$, the vector $(Y_{t_1}, \ldots, Y_{t_n})$ is independent of $\mathscr{F}_\tau$. But this is clear because Y_{t_1} is independent of $\mathscr{F}_\tau$ by (i) and we can proceed by induction over n.

If we now assume that X has homogeneous 1-dimensional increments, for a bounded stopping time τ, (2.25) yields

$$\mathbb{E}[\exp(\mathrm{i}u(X_{\tau+t} - X_\tau))] = \varphi_t(u),$$

meaning that Y_t is distributed as X_t in this special case. To pass to the general case in which τ is an a.s. finite valued stopping time, we use dominated convergence as in the proof of (i) starting from this latter relation. The proof of the theorem is now complete. □

We remark that, because of Theorem 2.5.15, for every a.s. finite valued stopping time τ and for an additive process $(X, \mathbb{F})$, from Lemma 2.5.4, we have, for every $B \in \mathscr{E}$,

$$\mathbb{P}[X_{\tau+t} \in B|\mathscr{F}_\tau] = \mathbb{P}[X_{\tau+t} - X_\tau + X_\tau \in B|\mathscr{F}_\tau] = \Big(\mathbb{P}[X_{\tau+t} - X_\tau + x \in B]\Big)\Big|_{x=X_\tau}$$

and the right-hand side is $\sigma(X_\tau)$-measurable. Therefore, we have

$$\mathbb{P}[X_{\tau+t} \in B|\mathscr{F}_\tau] = \mathbb{P}[X_{\tau+t} \in B|X_\tau]$$

and this justifies why we called Theorem 2.5.15 strong Markov property.

A relevant subclass of additive processes, which we are going to introduce, are *Lévy processes*.

2.5.16 Definition. (i) We say that an additive process (resp., an additive process in law) is a *Lévy process (resp., a Lévy process in law)* if it has also homogeneous increments.

(ii) We say that an additive process (resp., an additive process in law) relative to $\mathbb{F}$ is a *Lévy process (resp., a Lévy process in law) relative to* $\mathbb{F}$ if it has also homogeneous increments.

Let X be a Lévy process (resp., a Lévy process in law) relative to $\mathbb{F}$. The notation $(X, \mathbb{F})$ emphasizes the filtration with respect to which X is a Lévy process (resp., a Lévy process in law) and sometimes we simply say that $(X, \mathbb{F})$ is a Lévy process (resp., a Lévy process in law) to mean that X is a Lévy process (resp., a Lévy process in law) relative to $\mathbb{F}$. As for additive processes, we shall always consider the càdlàg version of a Lévy process.

A Lévy process with bounded jumps has a finite moment of every order, as the following proposition states.

2.5.17 Theorem. *Let $(X, \mathbb{F})$ be a Lévy process with bounded jumps, i.e., such that $|\Delta X_t| \leq c$, for every $t \geq 0$, a.s. Then for every $t \geq 0$, the random variable X_t^m belongs to $L^1(\mathbb{P})$, for every $m \in \mathbb{N}$, thus,*

$$\mathbb{E}[|X_t|^m] < +\infty, \quad t \geq 0, \quad m \in \mathbb{N}.$$

Proof. We define the sequence of stopping times $(\tau_n)_{n \geq 0}$ by setting $\tau_0 := 0$ and then for $n \geq 1$, $\tau_{n+1} := \inf\{t > \tau_n : |X_t - X_{\tau_n}| > c\}$. Then it is clear that $(\tau_n)_{n \geq 0}$ is increasing by definition and that $|\Delta X_{\tau_n}| \leq c$. We have $\tau_0 < \tau_1 < \ldots < \tau_n < \ldots$. Without loss of generality, we can assume that each τ_n is an a.s. finite-valued stopping time. Indeed, if $\tau_n = +\infty$, by definition, this means that the process X will never exceed the level c after τ_{n-1}. In other words, this means that the process is bounded because there exist a level which will be never crossed (notice that if $\tau_n = +\infty$, then $\tau_m = +\infty$ for every $m \geq n$) and in this case the theorem clearly holds. We notice that

$$\sup_{0 < s < +\infty} |X_{s \wedge \tau_n}| \leq 2cn. \tag{2.26}$$

This can be easily proven by induction. Indeed, if $n = 0$, there is nothing to prove. Assuming (2.26) for $n = k$ we have, by the induction hypothesis,

$$\sup_{0 < s < +\infty} |X_{s \wedge \tau_{k+1}}| = \sup_{0 < s < \tau_k} |X_{s \wedge \tau_k}| \vee \sup_{\tau_k < s < +\infty} |X_{s \wedge \tau_k}| \leq 2kc \vee \sup_{\tau_k < s < +\infty} |X_{s \wedge \tau_k}|.$$

But then, again by the induction hypothesis we deduce

$$\begin{aligned}\sup_{\tau_k < s < +\infty} |X_{s \wedge \tau_k}| &\leq \sup_{0 < u < \tau_{k+1} - \tau_k} |X_{\tau_k + u} - X_{\tau_k}| + |X_{\tau_k}| \\ &\leq \sup_{0 < u < \tau_{k+1} - \tau_k} |X_{\tau_k + u} - X_{\tau_k}| + 2kc\end{aligned}$$

and we have to estimate the first summand on the right-hand side in the previous inequality. We have two possibilities: $\tau_{k+1}(\omega) = +\infty$ or $\tau_{k+1}(\omega) < +\infty$. The first possibility can be excluded because in this case $\sup_{0 < u < \tau_{k+1} - \tau_k} |X_{\tau_k + u} - X_{\tau_k}| \leq c$ and the prove is finished. In the second case we have

$$\sup_{0 < u < \tau_{k+1} - \tau_k} |X_{\tau_k + u} - X_{\tau_k}| = |X_{\tau_{k+1}} - X_{\tau_k}| \leq |\Delta X_{\tau_k}| + |X_{\tau_{k+1}-} - X_{\tau_k}|$$

and the first term on the right-hand side of this inequality satisfies $|\Delta X_{\tau_k}| \leq c$ by the assumptions on X, while $|X_{\tau_{k+1}-} - X_{\tau_k}| \leq c$ because, by definition of τ_k, $|X_t - X_{\tau_k}| \leq c$ on $[\tau_k, \tau_{k+1})$ (notice that $X_{\tau_{k+1}-}$ denotes the process X_- evaluated at τ_{k+1}). Therefore, we deduce $\sup_{\tau_k < s < +\infty} |X_{s\wedge\tau_k}| \leq 2(k+1)c$ and the claim holds. We now apply the strong Markov property for Lèvy processes, that is Theorem 2.5.15. The time gap between τ_n and τ_{n+1} is the smallest time $t > 0$ such that $|X_{\tau_n+t} - X_{\tau_n}| > c$, that is $\tau_{n+1} - \tau_n = \inf\{t > 0 : |X_{\tau_n+t} - X_{\tau_n}| > c\}$. By Theorem 2.5.15, $\tau_{n+1} - \tau_n$ is therefore independent of $\mathscr{F}_{\tau_n}$ and is distributed as $\tau_1 - \tau_0 = \tau_1$. Hence $(\tau_{k+1} - \tau_k)_{k=0,\dots,n}$ is an iid sequence of random variables. Because of $\mathbb{P}[\tau_1 > 0] = 1$, $\mathbb{E}[\mathrm{e}^{-\tau_1}] =: a \in (0,1)$ holds. Let us compute

$$\mathbb{E}[\mathrm{e}^{-\tau_n}] = \mathbb{E}[\mathrm{e}^{-(\tau_n - \tau_{n-1})}\mathrm{e}^{-\tau_{n-1}}] = \mathbb{E}[\mathrm{e}^{-(\tau_n-\tau_{n-1})}]\mathbb{E}[\mathrm{e}^{-\tau_{n-1}}] = a\mathbb{E}[\mathrm{e}^{-\tau_{n-1}}] = \ldots = a^n\,.$$

We now are ready to show that the process has finite moments of every order. Let $m \in \mathbb{N}$ be arbitrarily *fixed*.

$$\mathbb{E}[|X_t|^m] = \sum_{n=0}^{\infty} \mathbb{E}\left[|X_t|^m 1_{\{2nc < |X_t| \leq 2(n+1)c\}}\right] \leq \sum_{n=0}^{\infty} (2(n+1)c)^m \mathbb{P}\left[|X_t| > 2nc\right]. \quad (2.27)$$

Because of (2.26), we have that if $|X_t(\omega)| > 2nc$, then $\tau_n(\omega) < t$. Therefore $\{|X_t| > 2nc\} \subseteq \{\tau_n < t\}$. Using Chebychev inequality we can estimate

$$\mathbb{P}[\tau_n < t] = \mathbb{P}\left[\mathrm{e}^{-\tau_n} > \mathrm{e}^{-t}\right] \leq \frac{1}{\mathrm{e}^{-t}}\mathbb{E}[\mathrm{e}^{-\tau_n}] = \mathrm{e}^t a^n.$$

With this latter estimation, we can estimate in (2.27)

$$\begin{aligned}
\mathbb{E}[|X_t|^m] &\leq \sum_{n=0}^{\infty} (2(n+1)c)^m \mathbb{P}\left[|X_t| > 2nc\right] \\
&\leq \sum_{n=0}^{\infty} (2(n+1)c)^m \mathbb{P}\left[\tau_n < t\right] \\
&\leq \mathrm{e}^t c^m 2^m \sum_{n=0}^{\infty} (n+1)^m a^n < +\infty
\end{aligned}$$

because $\frac{(n+2)^m}{(n+1)^m}\, a \longrightarrow a < 1$ as $n \to \infty$ and the proof is complete. □

CHAPTER 3

Brownian Motion and stochastic integration

In this chapter we are going to introduce Brownian motion and stochastic integration with respect to it. We shall see that the pats of the Brownian motion are not of finite variation. Therefore, it is not possible to give a pathwise definition of the integral with respect to the Brownian motion.

We use the notation $(E, \mathscr{E}) := (\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

3.1 Definition, existence and continuity

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space and let $\mathbb{F}$ be a filtration of events of $\mathscr{F}$ which satisfies the usual conditions.

3.1.1 Definition of the Brownian motion

3.1.1 Definition. Let $B = (B_t)_{t \geq 0}$ be a process taking values in $(E, \mathscr{E})$.

(i) B is a Brownian motion if

(a) $B_0 = 0$ a.s.

(b) For $0 \leq s \leq$, the distribution of the increment $B_t - B_s$ is a central normal with variance $t - s$, in symbols $B_t - B_s \sim \mathscr{N}(0, t - s)$.

(c) B has independent increments.

(ii) B is a Brownian motion with respect to $\mathbb{F}$ if (c) is replaced by: "B has $\mathbb{F}$-independent increments". In this case we say that $(B, \mathbb{F})$ is a Brownian motion.

(iii) We say that B is a continuous Brownian motion (resp., a continuous Brownian motion with respect to $\mathbb{F}$) if the paths are continuous.

Clearly, if $(B, \mathbb{F})$ is a Brownian motion, then B is a Brownian motion. Furthermore, if B is a Brownian motion, then $(B, \tilde{\mathbb{F}}^B)$ and $(B, \mathbb{F}^B)$ are also Brownian motions. If furthermore B is continuous, then $(B, \tilde{\mathbb{F}}^B_+)$ and hence $(B, \mathbb{F}^B_+)$ are Brownian motions (cf. Proposition 2.5.9).

We recall that a one-dimensional stochastic process is called a *Gaussian process* if for every $n \geq 0$ and $0 = t_0 < t_1 < \ldots < t_n$ the vector $(X_{t_0}, \ldots, X_{t_n})$ is normal distributed with mean $b_{t_0,\dots,t_n}$ and variance $\Sigma_{t_0,\dots,t_n}$. In other words, a Gaussian process X is such that the family $\mu := \{\mu_{t_1,\dots,t_n},\ 0 = t_0 < t_1 < \dots < t_n,\ n \geq 1\}$ of the finite-dimensional distributions of X is such that $\mu_{t_0,\dots,t_n}$ is an n-dimensional normal distribution on $(E^n, \mathscr{E}^n)$.

From WTHM it is known that a Brownian motion is a special case of Gaussian process, as the next theorem shows. For the proof we refer to WTHM, Theorem 10.6 and Theorem 10.7. The proof is left as an exercise.

3.1.2 Theorem. *Let* $0 \leq t_1 < \dots < t_n$ *and* $n \in \mathbb{N}$.

(i) *If B is a Brownian motion, then the distribution of the random vector* $(B_{t_1}, \dots, B_{t_n})$ *is a n-dimensional central normal with covariance matrix* $(\Sigma_{t_1,\dots,t_n})_{ij} = t_i \wedge t_j,\ i,j = 1,\dots,n$.

(ii) *Let X be a Gaussian process such that $X_0 = 0$ a.s. and the distribution of* $(X_{t_1}, \dots, X_{t_n})$ *is an n-dimensional central normal with covariance matrix* $(\Sigma_{t_1,\dots,t_n})_{ij} = t_i \wedge t_j,\ i,j = 1,\dots,n,\ n \in \mathbb{N}$. *Then X is a Brownian motion.*

Let $\mu := \{\mu_{t_1,\dots,t_n},\ 0 \geq< t_1 < \dots < t_n,\ n \geq 1\}$ be the family of the finite dimensional distributions of a Brownian motion. From Theorem 3.1.2, we know that

$$\mu_{t_1,\dots,t_n} = \mathscr{N}\big(0, (t_i \wedge t_j)_{i,j=1,\dots,n}\big).$$

Let now $Y = (Y_1, \dots, Y_n)$ be a n-dimensional Gaussian vector with mean b and variance Σ. Because of the uniqueness of the characteristic function, this means that (cf. WTHM, Appendix to Chapter 9), that the characteristic function of Y is

$$\mathbb{E}\big[\exp(\mathrm{i}\langle \xi, Y\rangle)\big] = \exp\Big(\mathrm{i}\langle \xi, b\rangle - \tfrac{1}{2}\,\xi^\top \Sigma \xi\Big), \quad \xi \in \mathbb{R}^n,$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^n$. If Σ is non-degenerate, we can associate to this distribution a density which is

$$f_Y(x) := \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\Big(-\tfrac{1}{2}\,(x-b)^\top \Sigma^{-1}(x-b)\Big), \quad x \in \mathbb{R}^n. \tag{3.1}$$

Coming back to the Brownian motion B, we want to apply these general results to determine the density of $\mu_{t_1,\dots,t_n}$, for every $0 = t_0 < t_1 < \dots < t_n$, $n \geq 1$, of the system of finite dimensional distributions of B.

For $0 = t_0 < t_1 < \dots < t_n$, we define

$$\Gamma := (B_{t_1}, \dots, B_{t_n})^\top;\ \Delta := (B_{t_1} - B_{t_0}, \dots, B_{t_n} - B_{t_{n-1}})^\top;\ D := \mathrm{diag}(t_1 - t_0, \dots, t_n - t_{n-1}). \tag{3.2}$$

Then, because of the property of B, $\Delta \sim \mathscr{N}(0, D)$. Let now M be a lower diagonal matrix in $\mathbb{R}^{n\times n}$ with entries equal to one on and below the diagonal. Then we clearly have $\Gamma = M\Delta$. We now set for simplicity $\Sigma := \Sigma_{t_1,\dots,t_n}$. Applying the properties of the characteristic function we deduce

$$\begin{aligned}\exp\Big(-\tfrac{1}{2}\,\langle \xi, \Sigma\xi\rangle\Big) &= \mathbb{E}\big[\exp(\mathrm{i}\langle \xi, \Gamma\rangle)\big] = \mathbb{E}\big[\exp(\mathrm{i}\langle \xi, M\Delta\rangle)\big] \\ &= \mathbb{E}\big[\exp(\mathrm{i}\langle M^\top\xi, \Delta\rangle)\big] = \exp\Big(-\tfrac{1}{2}\,\langle M^\top \xi, DM^\top\xi\rangle\Big)\end{aligned}$$

meaning that $\Sigma = MDM^\top$ hence $\Sigma^{-1} = (M^\top)^{-1}D^{-1}M^{-1}$. Since M^{-1} is a two-band matrix with entries equal to 1 on the first diagonal and equal to -1 on the first sub-diagonal, we see

$$\langle x, \Sigma^{-1}x\rangle = \langle M^{-1}x, D^{-1}M^{-1}x\rangle = \sum_{j=1}^{n} \frac{(x_j - x_{j-1})^2}{t_j - t_{j-1}}$$

and $\det(\Sigma) = \prod_{j=1}^{n}(t_j - t_{j-1})$. From (3.1), setting $x_0 := 0$, we can now conclude that $\mu_{t_1,\dots,t_n}$ has a density given by

$$p_{t_1,\dots,t_n}(x_1,\dots,x_n) := \frac{1}{\sqrt{(2\pi)^n \prod_{j=1}^{n}(t_j - t_{j-1})}} \exp\left(-\frac{1}{2}\sum_{j=1}^{n} \frac{(x_j - x_{j-1})^2}{t_j - t_{j-1}}\right). \tag{3.3}$$

3.1.2 Existence of a Brownian motion

Let now $\mu = \{\mu_{t_1,\dots,t_n},\ 0 \le t_1 < \dots < t_n,\ n \in \mathbb{N}\}$ be the system of finite dimensional distributions given by

$$\mu_{t_1,\dots,t_n}(B_1 \times \dots \times B_n) := \int_{B_1\times\dots\times B_n} p_{t_1,\dots,t_n}(x_1,\dots,x_n)\mathrm{d}x_n \cdots \mathrm{d}x_1, \quad t_1 > 0\,, \tag{3.4}$$

$$\mu_{0,t_2,\dots,t_n}(B_1 \times \dots \times B_n) := \delta_0(B_1) \int_{B_2\times\dots\times B_n} p_{t_2,\dots,t_n}(x_2,\dots,x_n)\mathrm{d}x_n \cdots \mathrm{d}x_2, \tag{3.5}$$

with $p_{t_1,\dots,t_n}$ as in (3.3). It is clear that this family satisfies the consistency conditions (2.1) and (2.2). Therefore, Theorem 2.1.7 ensures the existence of a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and of a stochastic process X having μ as associated system of finite dimensional distributions. So $\mu_{t_1,\dots,t_n}$ is an n-variate normal distribution, because $p_{t_1,\dots,t_n}$ are Gaussian densities. We can then claim that X is a Gaussian process. Furthermore, from the special form of these densities, we get that the covariance matrix of $\mu_{t_1,\dots,t_n}$ is $(t_i \wedge t_j)_{i,j=1,\dots n}$. Since Theorem 3.1.2 (ii), we can conclude that the process X must be a Brownian motion.

3.1.3 Existence of a continuous modification

With the help of Kolmogorov Extension Theorem we were able to prove the existence of a Brownian motion. However, the process constructed in this way need not be continuous. The question then arise if it is possible at least to find a continuous version of a Brownian motion. The positive answer to this question can be given thanks to Kolmogorov Continuity Theorem (cf. Theorem 2.2.7). The proof is left as an exercise.

3.1.3 Theorem. *Let B be a Brownian motion. Then there exists a version X of B which is a Brownian motion and such that the paths of X are α-Hölder continuous, for every $\alpha \in (0, \frac{1}{2})$, that is for every $T > 0$ and $\alpha \in (0, \frac{1}{2})$, there exists a function $C_{T,\alpha} : \Omega \longrightarrow (0, +\infty)$ such that*

$$|X_t(\omega) - X_s(\omega)| \le C_{T,\alpha}|t - s|^\alpha\,, \quad t, s \in [0, T], \quad \forall \omega \in \Omega\,.$$

Proof. Let B be a Brownian motion. Then $Z := B_t - B_s/\sqrt{t-s}$ is a standard normal random variable. Hence, for $n > 2$, there exists $c_n > 0$ such that $0 < \mathbb{E}[|Z|^n] = c_n < +\infty$, which yields

$$\mathbb{E}[|B_t - B_s|^n] = c_n(t-s)^{\frac{n}{2}} = c_n(t-s)^{1+\frac{n-2}{2}}.$$

Because of Theorem 2.2.7, there exist a version X of B which is α-Hölder continuous with $\beta = n$, $\alpha = \frac{n-2}{2}$ and $c = c_n$. Clearly, X satisfies all the properties of the definition of the Brownian motion. In particular $(X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}})$ has the same distribution (hence characteristic function) of $(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}})$ (cf. Exercise 2.2.5). Notice that to ensure that X is again a stochastic process (see Definition 2.1.1), we need the completeness of the probability space. □

3.1.4 Exercise. ($\star$) Prove Theorem 3.1.3.

In the sequel we shall always consider the continuous version of the Brownian motion: For us a Brownian motion will always be a continuous Brownian motion. Notice that, if we work with a Brownian motion with respect to a filtration, we need to require that the filtration is complete to get that the continuous version is still an adapted process and hence a Brownian motion with respect to the same filtration.

3.1.4 (Ir)Regularity of paths

In Theorem 3.1.3 we have seen that we can always choose a version of the Brownian motion which is Hölder-continuous. It is then natural to investigate if it is possible to get better regularity of the paths of the Brownian motion. For example, it is of interest to understand if the paths of a Brownian motion are differentiable. However, this is not the case because the paths of a Brownian motion are not of finite variation.

3.1.5 Definition. Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a function and $[a, b] \subseteq \mathbb{R}$ be an interval.
(i) The quantity

$$V_b^a f := \sup_{\pi \in \Pi} \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|,$$

where π is a partition of $[a, b]$ with $a = x_0 < \ldots < x_n = b$ and Π is the set of all partition of $[a, b]$ as a such, is called *total variation* of f over $[a, b]$.
(ii) A function f is of *finite (total) variation* if $V_b^a f < +\infty$, for every $a, b \in \mathbb{R}$.

Let now B be a Brownian motion. In WTHM, Theorem 12.1 in Chapter 10, we gave a proof of the following result:

3.1.6 Theorem. *Let $\pi := \{t_0, \dots, t_m\}$ be a partition of $[s, t]$ with $s = t_0 < \ldots < t_m = t$, set $|\pi| := \max_{0 \le k \le m-1} |t_{k+1} - t_k|$ and*

$$S_\pi := \sum_{k=0}^{m-1} (B_{t_{k+1}} - B_{t_k})^2.$$

Then S_π converges to $t - s$ in $L^2(\mathbb{P})$, as $|\pi| \to 0^+$.

3.1.7 Exercise. Prove Theorem 3.1.6.

As a consequence of Theorem 3.1.6 we deduce

3.1.8 Corollary. *The paths of a Brownian motion B are almost surely of infinite total variation.*

Proof. Let S_π be defined as in Theorem 3.1.6. Then

$$0 \le S_\pi \le \max_{0\le k\le m-1} |B_{t_{k+1}} - B_{t_k}| \sum_{k=0}^{m-1} |B_{t_{k+1}} - B_{t_k}|. \tag{3.6}$$

Since the paths of B are continuous, $\max_{0\le k\le m-1} |B_{t_{k+1}} - B_{t_k}|$ goes to zero as $|\pi| \to 0^+$. If we now assume that the paths of B are of finite variation over a set $A \in \mathscr{F}$ with $\mathbb{P}[A] > 0$, we deduce from (3.6), $S_\pi(\omega) \longrightarrow 0$, as $|\pi| \to 0^+$ for $\omega \in A$ and this is a contradiction to Theorem 3.1.6. □

3.2 P. Lèvy Characterization of Brownian Motion

Theorem 3.1.2 is a characterization of the Brownian motion among Gaussian processes. The aim of this section is to characterize the Brownian motion as a continuous martingale. This characterization is due to Paul Lèvy (1948). We refer to Schilling & Partsch (2014), §9.4.

We start with two lemmas: The first one characterize the Brownian motion in term of characteristic functions and the second one is a statement about the existence of the fourth moment for some continuous local martingales.

3.1 Lemma. *Let X be an $\mathbb{F}$-adapted (continuous) stochastic process starting at zero. Then it is a (continuous) Brownian motion with respect to $\mathbb{F}$ if and only if*

$$\mathbb{E}\Big[\exp\big(\mathrm{i}u(X_t - X_s)\big)\Big|\mathscr{F}_s\Big] = \exp\left(-\frac{1}{2}\,u^2(t-s)\right), \quad 0 \le s \le t \tag{3.7}$$

for every $u \in \mathbb{R}$.

Proof. It is clear that, if $(X, \mathbb{F})$ is a Brownian motion, then (3.7) holds. We now assume that if (3.7) holds, then X is a Brownian motion. Let $0 = t_0 < \ldots < t_n$; $X_{-1} := 0$. Let Γ, Δ, D and M be defined as in (3.2) and the following lines starting from X. It is enough to show that

$$\mathbb{E}\big[\exp(\mathrm{i}\langle\xi, \Delta\rangle)\big] = \exp\Big(-\tfrac{1}{2}\,\xi^\top D\xi\Big), \quad \xi \in \mathbb{R}^n, \tag{3.8}$$

which means $\Delta \sim \mathscr{N}(0, D)$. Indeed, because $\Gamma = M\Delta$, we have

$$\mathbb{E}\big[\exp(\mathrm{i}\langle\xi, \Gamma\rangle)\big] = \exp\Big(-\tfrac{1}{2}\,\xi^\top MDM^\top\xi\Big), \quad \xi \in \mathbb{R}^n,$$

and $(MDM^\top)_{ij} = t_i \wedge t_j$, $i, j = 1, \ldots, n$, which from Theorem 3.1.2 (ii), implies that X is a Brownian motion. To show (3.8) we use induction. For $n = 1$ there is nothing to prove. Now we assume (3.8) for $1 \le m \le n-1$ and prove it for n.

$$
\begin{aligned}
\mathbb{E}\big[\exp(\mathrm{i}\langle \xi, \Delta\rangle)\big] &= \mathbb{E}\left[\exp\left(\mathrm{i}\sum_{j=0}^{n}\xi_j(X_{t_j}-X_{t_{j-1}})\right)\right]\\
&= \mathbb{E}\left[\exp\left(\mathrm{i}\sum_{j=1}^{n-1}\xi_j(X_{t_j}-X_{t_{j-1}})\right)\mathbb{E}\Big[\exp\big(\mathrm{i}\,\xi_n(X_{t_n}-X_{t_{n-1}})\big)\Big|\mathscr{F}_{t_{n-1}}\Big]\right]\\
&= \exp\Big(-\frac{1}{2}\,\xi_n^2(t_n-t_{n-1})\Big)\exp\left(-\frac{1}{2}\sum_{j=1}^{n-1}\xi_j^2(t_j-t_{j-1})\right)\\
&= \exp\left(-\frac{1}{2}\sum_{j=1}^{n}\xi_j^2(t_j-t_{j-1})\right)
\end{aligned}
$$

where in the last but one equality we used the induction hypothesis and (3.7). □

The proof of the next result can be found in Schilling & Partsch (2014), Lemma 9.10. We omit the proof and only stress that the proof given in Schilling & Partsch (2014) does not use stochastic integration.

3.2 Lemma. *Let X be an $\mathbb{F}$-continuous martingale such that $(X_t^2-t)_{t\ge 0}$ is an $\mathbb{F}$-martingale. Then $\mathbb{E}[X_t^4]<+\infty$ and, for every $0\le s\le t$, the estimate*

$$\mathbb{E}[(X_t-X_s)^4|\mathscr{F}_s]\le 4(t-s)^2$$

holds.

We are now ready to state and prove the P. Lèvy-characterization of the Brownian motion.

3.2.1 Theorem (P. Lèvy 1948)**.** *Let X be a continuous $\mathbb{F}$-martingale starting at zero. Then $(X,\mathbb{F})$ is a Brownian motion if and only if $(X_t^2-t)_{t\ge 0}$ is an $\mathbb{F}$-martingale.*

Proof. Clearly, if $(X,\mathbb{F})$ is a continuous Brownian motion then X and $(X_t^2-t)_{t\ge0}$ are continuous $\mathbb{F}$-martingale. Conversely, let X be a continuous $\mathbb{F}$-martingale such that $(X_t^2-t)_{t\ge0}$ is an $\mathbb{F}$-martingale as well. We are going to show that $(X,\mathbb{F})$ is a Brownian motion. Because of Lemma 3.7, it suffices to verify (3.7). We set $Y := X_t - X_s$ and $h := t-s$. We consider the Taylor expansion of $\exp(\mathrm{i}\,\xi Y)$ in ξ around zero up to the second order with a remainder of the third order and get

$$\exp(\mathrm{i}\xi Y) = 1+\mathrm{i}\,\xi Y-\frac{\xi^2}{2}Y^2-\mathrm{i}\,\frac{\xi^3}{6}(\eta Y)^3$$

where $\eta=\eta(\omega)$ is a complex valued random variable such that $|\eta|\in(0,1)$. The Taylor expansion of $\exp(-\frac{1}{2}\xi h)$ up to the fifth order with a remainder of the sixth order is

$$\exp(-\frac{1}{2}\xi h) = 1-\frac{\xi^2}{2}h+\frac{\xi^4}{8}h^2-\frac{\xi^6}{48}(\theta h)^3$$

where $\theta \in (0,1)$. Subtracting these two expressions we deduce:

$$\exp(\mathrm{i}\,\xi Y) - \exp(-\frac{1}{2}\xi h) = \mathrm{i}\xi Y + \frac{\xi^2}{2}h - \frac{\xi^2}{2}Y^2 - \frac{\xi^4}{8}h^2 - \mathrm{i}\frac{\xi^3}{6}(\eta Y)^3 + \frac{\xi^6}{48}(\theta h)^3.$$

Using that $\mathbb{E}[Y|\mathscr{F}_s] = 0$ because X is a martingale starting at zero, that $\mathbb{E}[Y^2|\mathscr{F}_s] = h$ because $(X_t^2 - t)_{t\geq 0}$ is an $\mathbb{F}$-martingale and taking conditional expectation in the previous expression, yields

$$\mathbb{E}\Big[\exp(\mathrm{i}\,\xi Y) - \exp(-\frac{1}{2}\xi h)|\mathscr{F}_s\Big] = -\frac{\xi^4}{8}h^2 + \frac{\xi^6}{48}(\theta h)^3 - \mathrm{i}\mathbb{E}[(\eta Y)^3|\mathscr{F}_s].$$

We now estimate $\mathbb{E}[(\eta Y)^3|\mathscr{F}_s]$. By Hölder inequality with $p = 4/3$ and $q = 4$ we get

$$|\mathbb{E}[(\eta Y)^3|\mathscr{F}_s]| \leq \mathbb{E}[|Y|^3|\mathscr{F}_s] \leq \mathbb{E}[|Y|^4|\mathscr{F}_s]^{4/3} \leq 4^{4/3}h^{3/2},$$

where in the last estimate we used Lemma 3.2. This means that for every $s < t$ with $h = t - s < 1$ there exists a constant c_ξ such that for every ξ in a neighbourhood of zero the estimate

$$\Big|\mathbb{E}\Big[\exp\big(\mathrm{i}\xi(X_t - X_s)\big) - \exp\big(-\frac{1}{2}\xi^2(t-s)\big)|\mathscr{F}_s\Big]\Big| \leq c_\xi(t-s)^{3/2}.$$

We now have to remove the restriction $t - s < 1$. Let $s < t$ be fixed and set $t_j := s + (t-s)j/n$, $j = 0, \dots, n$, $n \geq 1$. Then

$$\exp\big(\mathrm{i}\,\xi(X_t - X_s)\big) = \prod_{j=1}^{n}\exp\big(\mathrm{i}\,\xi(X_{t_j} - X_{t_{j-1}})\big), \quad \exp\big(-\tfrac{1}{2}\xi^2(t-s)\big) = \prod_{j=1}^{n}\exp\big(-\tfrac{1}{2}\xi(t_j - t_{j-1})\big).$$

Using the estimate

$$|\mathbb{E}[aA - bB]| \leq \mathbb{E}[|a-b||A||\mathscr{F}_s] + |\mathbb{E}[b(A-B)|\mathscr{F}_s]|$$

with $a = \mathbb{E}[\exp\big(\mathrm{i}\,\xi(X_t - X_{t_{n-1}})\big)|\mathscr{F}_{t_{n-1}}]$, $b = \exp\big(-\frac{1}{2}\xi(t - t_{n-1})\big)$, $A = \exp\big(\mathrm{i}\,\xi(X_{t_{n-1}} - X_s)\big)$, $B = \exp\big(-\frac{1}{2}\xi(t_{n-1} - s)\big)$, we get

$$\begin{aligned}
&\Big|\mathbb{E}\Big[\exp\big(\mathrm{i}\xi(X_t - X_s)\big) - \exp\big(-\tfrac{1}{2}\xi^2(t-s)\big)\Big|\mathscr{F}_s\Big]\Big| \leq \\
&\qquad \mathbb{E}\Big[\Big|\mathbb{E}\Big[\exp\big(\mathrm{i}\,\xi(X_t - X_{t_{n-1}})\big) - \exp\big(-\tfrac{1}{2}\xi(t - t_{n-1})\big)\Big|\mathscr{F}_{t_{n-1}}\Big]\Big|\Big|\mathscr{F}_s\Big] + \\
&\qquad\qquad + \Big|\mathbb{E}\Big[\exp\big(\mathrm{i}\,\xi(X_{t_{n-1}} - X_s)\big) - \exp\big(-\tfrac{1}{2}\xi(t_{n-1} - s)\big)\Big]\Big| \\
&\qquad \leq \ldots \leq \sum_{j=1}^{n}\mathbb{E}\Big[\Big|\mathbb{E}\Big[\exp\big(\mathrm{i}\,\xi(X_{t_j} - X_{t_{j-1}})\big) - \exp\big(-\tfrac{1}{2}\xi(t_j - t_{j-1})\big)\Big|\mathscr{F}_{t_{j-1}}\Big]\Big|\Big|\mathscr{F}_s\Big].
\end{aligned}$$

Let now $n \geq 1$ be big enough, so that $t_j - t_{j-1} < 1$. Because of the first step we get

$$\begin{aligned}
\Big|\mathbb{E}\Big[\exp\big(\mathrm{i}\,\xi(X_t - X_s)\big) - \exp\big(-\tfrac{1}{2}\xi^2(t-s)\big)\Big|\mathscr{F}_s\Big]\Big| &\leq c_\xi\sum_{j=1}^{n}(t_j - t_{j-1})^{3/2} \\
&= c_\xi\sum_{j=1}^{n}\left(\frac{(t-s)}{n}\right)^{3/2} \\
&= c_\xi\left(\frac{(t-s)}{\sqrt{n}}\right)^{3/2}
\end{aligned}$$

which converges to zero as $n \to +\infty$ and the proof is complete. □

3.3 Stochastic integration

We start this section recalling Riemann–Stiltjes integration.

Let f be a function over an interval $[a,b]$, and $\pi := \{x_0, \ldots, x_n\}$, $a = x_0 < \ldots < x_n = b$ a partition of $[a,b]$, setting $|\pi| := \max_{1\le i\le n} |x_i - x_{i-1}|$, we define

$$F_\pi := \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}), \qquad \xi_i \in [x_{i-1}, x_i].$$

If f is continuous then there exists $A \in \mathbb{R}$ such that $A = \lim_{|\pi|\to 0} F_\pi$. Furthermore A does not depend on the choice of π and ξ_i. In this case A is called Riemann integral of f over $[a,b]$ and $A =: \int_a^b f(x)\mathrm{d}x$.

Riemann–Stiltjes integration is a generalization of Riemann integration: For an increasing function g we set

$$F_\pi^g := \sum_{i=1}^{n} f(\xi_i)(g(x_i) - g(x_{i-1})), \qquad \xi_i \in [x_{i-1}, x_i].$$

If f is a continuous function then there exists $A \in \mathbb{R}$ such that $A = \lim_{|\pi|\to 0} F_\pi^g$. Furthermore A does not depend on the choice of π and ξ_i. In this case A is called Riemann–Stiltjes integral of f over $[a,b]$ and $A =: \int_a^b f(x)g(\mathrm{d}x)$.

A first generalization of Riemann–Stiltjes integral can be given if the function g is of finite variation. Indeed, in this case there exists two increasing functions h and k such that $g = k - h$ and $V_x^a g = k(x) + h(x)$. If f is a continuous function we then set

$$\int_a^b f(x)g(\mathrm{d}x) = \int_a^b f(x)k(\mathrm{d}x) - \int_a^b f(x)h(\mathrm{d}x).$$

The question is if it is possible to generalize the definition of the Riemann–Stiltjes integral to integrators which are not of finite variation. The negative answer to this question is given by the following theorem:

3.3.1 Theorem. *Let $g : [0,T] \longrightarrow \mathbb{R}$ be a finite-valued function of infinite total variation over $[0,T]$. Then there exist a continuous bounded function f over $[0,T]$ such that the Riemann–Stiltjes integral of f with respect to g on $[0,T]$ does not exists.*

To prove Theorem 3.3.1, we need Banach–Steinhaus theorem:

3.3.2 Theorem. *Let X be a Banach space and Y a normed linear space. Let $(T_\alpha)_{\alpha\in I}$ be a family of continuous linear (and hence bounded) operators on X in Y. If for every $x \in X$, $(T_\alpha x)_{\alpha\in I}$ is bounded (that is $\sup_\alpha \|T_\alpha x\|_Y < +\infty$), then $(\|T_\alpha\|)_{\alpha\in I}$ is bounded (that is $\sup_\alpha \|T_\alpha\| < +\infty$), where $\|T_\alpha\| := \sup_{x\in X} \|T_\alpha x\|_Y / \|x\|_X$.*

Proof of Theorem 3.3.1. Set $(X, \|\cdot\|_X) = (C([0,T]), \|\cdot\|_\infty)$, where $C([0,T])$ is the space of continuous functions over $[0,T]$ and $\|\cdot\|_\infty$ the uniform norm, and $(Y, \|\cdot\|_Y) = (\mathbb{R}, |\cdot|)$. Let π^n be a sequence of partitions of $[0,T]$ such that $|\pi^n| \longrightarrow 0$ as $n \to +\infty$. For every partition π of $[0,T]$ with N points,

$$T_\pi f := \sum_{i=0}^{N-1} f(t_i)(g(t_{i+1}) - g(t_i)), \quad f \in X$$

is a linear bounded (hence continuous) operator on $(X, \|\cdot\|_X)$ into $(Y, \|\cdot\|_Y)$. For every partition π, we can chose a function $f^*_\pi \in X$ such that $\|f^*_\pi\|_X = 1$ and that $f^*_\pi(t_i) = \mathrm{sign}(g(t_{i+1}) - g(t_i))$. Then

$$T_\pi f^* := \sum_{i=0}^{N-1} |g(t_{i+1}) - g(t_i)|, \quad f \in X.$$

The function g is of unbounded total variation over $[0, T]$ and therefore $(\|T_{\pi^n}\|)_{n\in\mathbb{N}}$ is unbounded. Because of Banach–Steinhaus's Theorem (cf. Theorem 3.3.2) there exists a function $f \in X$ such that $(|T_{\pi^n} f|)_{n\in\mathbb{N}}$ is unbounded. Therefore $\lim_{n\to+\infty} F_{\pi^n}$ cannot be finite, meaning that $\int_0^T f(t)g(\mathrm{d}t)$ does not exist. □

We now apply these results to the Brownian motion. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space and $\mathbb{F}$ a filtration of events of $\mathscr{F}$ satisfying the usual conditions. Let $T > 0$ be a fixed time horizon. All stochastic processes in this section are meant to be $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$-valued. Let B be an $\mathbb{F}$-Brownian motion. The aim of this section is to give sense to the random variable

$$\int_0^T H_t \mathrm{d}B_t.$$

A nïve (and WRONG!!!) way to define the integral with respect to the Brownian motion is to set

$$\left(\int_0^T H_t \mathrm{d}B_t\right)(\omega) := \int_0^T H_t(\omega)\mathrm{d}B_t(\omega)\,, \quad \omega \in \Omega$$

for every measurable and bounded continuous process H. However, we know that $t \to B_t(\omega)$ is a function of unbounded variation over $[0, T]$ (cf. §3.1.4) and, from Theorem 3.3.1, a pathwise definition of the integral with respect to the Brownian motion is not possible.

Our unique hope to give sense to $\int_0^T H_t \mathrm{d}\, B_t$ is to put into play the probability measure $\mathbb{P}$: If a pathwise definition of the integral is not possible, it may be possible to give a definition of the integral *in mean*. This intuition is due to Itô, who developed the theory of stochastic integration which we are going to present.

3.3.1 Stochastic integral of elementary processes

As a firs step we define the stochastic integral for *elementary* processes. We want to define the integral in such a way that it fulfils some elementary properties, that is it should be linear and boundeb continuous functions have to be integrable.

3.3.3 Definition. Let H be a stochastic process with time parameter set $[a, b] \subseteq \mathbb{R}_+$ and values in $\mathbb{R}$. We say that it is an *elementary process* if there exist a partition $a = t_0 < t_1 < \ldots < t_n = b$ of $[a, b]$ and a sequence $e_0, \ldots, e_{n-1}$ of random variable such that e_i is $\mathscr{F}_{t_i}$-measurable and the decomposition

$$H_t = \sum_{i=0}^{n-1} e_i 1_{[t_i, t_{i+1})}(t)\,, \quad t \in [a, b] \tag{3.9}$$

holds.

3.3.4 Exercise. (i) Show that $\alpha H + \beta K$ is an elementary process, whenever H and K are elementary and $\alpha, \beta \in \mathbb{R}$.

(ii) Show that if H is an elementary process with representation (3.9), then H_t is square integrable if and only if $e_0, \ldots, e_{n-1}$ are square integrable.

(iii) Let H be an elementary process. Is it adapted? Is it progressively measurable?

3.3.5 Definition. Let H be an elementary process with representation (3.9). The *elementary stochastic integral* (over $[a,b]$) of H with respect to B is defined by

$$\int_a^b H_s \mathrm{d}B_s := \sum_{i=0}^{n-1} e_i (B_{t_{i+1}} - B_{t_i}) \,. \tag{3.10}$$

We also use the notation $I_a^b(H) := \int_a^b H_s \mathrm{d}B_s$.

Let us see some first properties of the elementary stochastic integral.

3.3.6 Proposition. *Let H and K be two elementary processes.*

(i) *The elementary stochastic integral is linear.*

(ii) *The random variable $I_a^b(H)$ is $\mathscr{F}_b$-measurable.*

(iii) *The mapping $t \mapsto I_0^t(H)$ is continuous and adapted.*

(iv) *For $a \le c \le b$ the identity $I_a^b(H) = I_a^c(H) + I_c^b(H)$ holds.*

Proof. (i) is clear, (ii) follows because of the definition of $I_a^b(H)$ and B is adapted. The adaptedness in (iii) follows from (ii) whenever $a = 0$ and $b = t$. For the continuity we observe that, for $t \in [t_k, t_{k+1})$

$$I_0^t(H) = \sum_{i=0}^{k-1} e_i (B_{t_{i+1}} - B_{t_i}) + e_k (B_t - B_{t_k})$$

and $t \mapsto B_t$ is continuous. Property (iv) is left as an exercise. □

We now come to some important properties of the elementary stochastic integral.

3.3.7 Theorem. *Let H be an elementary process such that $\mathbb{E}[H_t^2] < +\infty$, for every $t \ge 0$. Then the following relation holds*

(i) $I_a^b(H) \in L^2(\mathbb{P})$;

(ii) $\mathbb{E}\big[I_a^b(H) | \mathscr{F}_a\big] = 0$;

(iii) $\mathbb{E}\Big[\big(I_a^b(H)\big)^2\Big] = \int_a^b \mathbb{E}\big[H_t^2\big] \mathrm{d}\, t$.

Proof. We show (i). Because of Proposition 3.3.6, the stochastic integral $I_a^b(H)$ is $\mathscr{F}_b$-measurable. Furthermore, $\sum_{i=0}^{n-1} e_i (B_{t_{i+1}} - B_{t_i})$ is square integrable because e_i and $(B_{t_{i+1}} - B_{t_i})$ are square integrable and *independent*, for $i = 1, \ldots, n$. To see (iii) we notice that

$$\big(I_a^b(H)\big)^2 = \left(\sum_{i=0}^{n-1} e_i (B_{t_{i+1}} - B_{t_i})\right)^2 = \sum_{i=0}^{n-1} e_i^2 (B_{t_{i+1}} - B_{t_i})^2 + 2 \sum_{\substack{i,j=0 \\ i>j}}^{n-1} e_i e_j (B_{t_{j+1}} - B_{t_j})(B_{t_{i+1}} - B_{t_i}) \,.$$

For $i > j$, $e_i e_j (B_{t_{j+1}} - B_{t_j})$ is independent of $(B_{t_{i+1}} - B_{t_i})$. Furthermore, because $e_i(B_{t_{i+1}} - B_{t_i})$ and $e_j(B_{t_{j+1}} - B_{t_j})$ are square integrable, the right-hand side in the previous equality is integrable. We can then compute the expectation, which yields

$$
\begin{aligned}
&\mathbb{E}\Big[\big(I_a^b(H)\big)^2\Big] \\
&= \sum_{i=0}^{n-1} \mathbb{E}\left[e_i^2 (B_{t_{i+1}} - B_{t_i})^2\right] + 2 \sum_{\substack{i,j=0 \\ i>j}}^{n-1} \mathbb{E}\left[e_i e_j (B_{t_{j+1}} - B_{t_j})(B_{t_{i+1}} - B_{t_i})\right] \\
&= \sum_{i=0}^{n-1} \mathbb{E}\left[e_i^2\right] \mathbb{E}\left[(B_{t_{i+1}} - B_{t_i})^2\right] + 2 \sum_{\substack{i,j=0 \\ i>j}}^{n-1} \mathbb{E}\left[e_i e_j (B_{t_{j+1}} - B_{t_j})\right] \mathbb{E}\left[(B_{t_{i+1}} - B_{t_i})\right] \\
&= \sum_{i=0}^{n-1} \mathbb{E}\left[e_i^2\right] (t_{i+1} - t_i) = \int_a^b \mathbb{E}[H_t^2] \mathrm{d}\, t \,,
\end{aligned}
$$

where in the second line of the previous computation we used the independence of $e_i e_j (B_{t_{j+1}} - B_{t_j})$ and $(B_{t_{i+1}} - B_{t_i})$ for $i > 0$ and the properties of the Brownian motion to set the second sum equal to zero. We now show (ii). We have $\mathscr{F}_a \subseteq \mathscr{F}_{t_i}$, $i = 1, \ldots, n$. Hence

$$
\begin{aligned}
\mathbb{E}\Big[I_a^b(H)\big|\mathscr{F}_a\Big] &= \sum_{i=0}^{n-1} \mathbb{E}\left[e_i (B_{t_{i+1}} - B_{t_i})\Big|\mathscr{F}_a\right] \\
&= \sum_{i=0}^{n-1} \mathbb{E}\left[\mathbb{E}\big[e_i (B_{t_{i+1}} - B_{t_i})|\mathscr{F}_{t_i}\big]\Big|\mathscr{F}_a\right] \\
&= \sum_{i=0}^{n-1} \mathbb{E}\left[e_i \mathbb{E}\big[(B_{t_{i+1}} - B_{t_i})\big]\Big|\mathscr{F}_a\right] = 0 \,.
\end{aligned}
$$

□

3.3.8 Exercise. Let H be a square integrable elementary process. Show that the definition of the stochastic integral is well posed.

3.3.2 Extension of the stochastic integral

In §3.3.1 we defined the stochastic integral for elementary integrand and called it elementary stochastic integral. Aim of this section is to extend the definition of the stochastic integral to more general integrands. The idea is to define a certain space of stochastic processes containing the square integrable elementary processes and in which they are dense. By approximation, we will then extend the definition of the stochastic integral to random variables which are limit of sequences elementary iterated integrals. The main point here is to prove that if we approximate consider the limit of a sequence of square integrable elementary integrands, then the corresponding sequence of elementary integrals converge. At this point property (iii) in Theorem 3.3.7 will play a crucial role.

We anticipate that we are going to develop an L^2-theory: that is, the limit of square integrable elementary process will be taken in some L^2-sense. Therefore, the processes for which we shall define the stochastic integral will be progressively measurable processes, the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ being complete and the elementary processes being progressively measurable (they are adapted and right-continuous, cf. Theorem 2.3.5).

Let $[0,T]$, $T>0$, be the time horizon and $\mathscr{S}_T^2$ the space of elementary processes H over $[0,T]$ which are square integrable, that is such that $\mathbb{E}[H_t^2] < +\infty$, $t \in [0,T]$. Notice that, because of the definition of elementary process, if H is square integrable, we have

$$\int_0^T \mathbb{E}\big[H_t^2\big]\mathrm{d}\,t < +\infty$$

which, applying the theorem of Fubini becomes

$$\mathbb{E}\left[\int_0^T H_t^2 \mathrm{d}\,t\right] < +\infty$$

meaning that $H \in L^2([0,T]\times\Omega, \mathscr{B}([0,T])\otimes\mathscr{F}, \lambda\otimes\mathbb{P}, \|\cdot\|_{\lambda\otimes\mathbb{P}}) =: L_T^2(\lambda\otimes\mathbb{P})$, where λ denotes the Lebesgue measure on $[0,T]$ and $\|\cdot\|_{\lambda\otimes\mathbb{P}}$ the norm of the scalar product, that is

$$\|X\|_{\lambda\otimes\mathbb{P}}^2 := \mathbb{E}\left[\int_0^T X_t^2 \mathrm{d}\,t\right],$$

for every measurable process X. More precisely $\|\cdot\|_{\lambda\otimes\mathbb{P}}$ is only a semi norm, to get a norm we have to identify, as we do, measurable processes which are equivalent with respect to the following equivalence relation

$$X \sim Y \Longleftrightarrow \mathbb{P}\left[\int_0^T |X_t - Y_t| \mathrm{d}\,t = 0\right] = 1.$$

We want now to define a space of progressively measurable processes containing $\mathscr{S}_T^2$. Because we want to develop an L^2-theory, the unique possible choice is

$$\mathrm{L}^2(B) := \left\{H \text{ progressively measurable}, H \in L_T^2(\lambda\otimes\mathbb{P})\right\}. \tag{3.11}$$

The space $\mathrm{L}^2(B)$ in (3.11) will be the space of the integrands for the Brownian motion, that is, it is the space of processes for which we are going to define the stochastic integral with respect to B. Obviously the inclusions

$$\mathscr{S}_T^2 \subseteq \mathrm{L}^2(B) \subseteq L_T^2(\lambda\otimes\mathbb{P})$$

hold. We define

$$I(H) := I_0^T(H), \quad H \in \mathscr{S}_T^2. \tag{3.12}$$

If we now regard the elementary stochastic integral $I(H)$ as a mapping on $\mathscr{S}_T^2$, Theorem 3.3.7 (i) shows that it maps $\mathscr{S}_T^2$ to $L^2(\mathbb{P})$, that is $I : \mathscr{S}_T^2 \longrightarrow L^2(\mathbb{P})$. Since Theorem 3.3.7 (iii), I is an isometry on $\mathscr{S}_T^2 \subseteq L_T^2(\lambda\otimes\mathbb{P})$ into $L^2(\mathbb{P})$:

$$\|I(H)\|_{L^2(\mathbb{P})} = \|H\|_{\lambda\otimes\mathbb{P}}. \tag{3.13}$$

The isometry relation (3.13) is known as *Itô isometry.*

Let now H^n be a sequence in $\mathrm{L}^2(B)$ converging in $L_T^2(\lambda\otimes\mathbb{P})$ to H. Because our probability space is complete and $\mathbb{F}$ satisfies the usual condition, this yields that also H is a progressive process. In other words, $\mathrm{L}^2(B)$ is a closed subspace of $L_T^2(\lambda\otimes\mathbb{P})$ and hence it is an Hilbert space with respect to the scalar product of $L_T^2(\lambda\otimes\mathbb{P})$.

We now summarize the situation:

1. We defined the mapping I, that is the stochastic integral for elements in $\mathscr{S}_T^2$.
2. We introduced the subspace $\mathrm{L}^2(B)$ of $L_T^2(\lambda \otimes \mathbb{P})$ of progressively measurable processes and we have the inclusion $\mathscr{S}_T^2 \subseteq \mathrm{L}^2(B)$.
3. $I : \mathscr{S}_T^2 \longrightarrow L^2(\mathbb{P})$ is an isometry of $L_T^2(\lambda \otimes \mathbb{P})$ into $L^2(\mathbb{P})$.

Our aim is to extend the isometry I to processes in $\mathrm{L}^2(B)$. For this we shall proceed as follow: We first show that we can approximate elements in $\mathrm{L}^2(B)$ with elements in $\mathscr{S}_T^2$ in $L_T^2(\lambda \otimes \mathbb{P})$. We then show that, if $H \in \mathrm{L}^2(B)$ and $H^n \in \mathscr{S}_T^2$ is the sequence of elementary integrands approximating H, the sequence of the elementary stochastic integrals associated to H^n converges in $L^2(\mathbb{P})$ to a random variable X. Then we show that the limit X does not depend on the approximating sequence but only on H. Hence we are allowed to call X the stochastic integral of H with respect to B.

We start with the following proposition, about the possibility of approximating elements of $\mathrm{L}^2(B)$ with elementary integrands.

3.3.9 Proposition. *Let $H \in \mathrm{L}^2(B)$. Then there exists a sequence $(H^n)_{n\in\mathbb{N}} \subseteq \mathscr{S}_T^2$ such that $\|H - H^n\|_{\lambda\otimes\mathbb{P}} \longrightarrow 0$, as $n \to +\infty$.*

To prove Proposition 3.3.9 we need two lemmas, the first of which is a purely deterministic one. The prove is elementary but a little bit technical. Therefore we postpone it to the end of this section.

3.3.10 Lemma. *Let $(H^n)_{n\in\mathbb{N}} \subseteq \mathscr{S}_T^2$ converge to H in $\mathrm{L}^2(B)$. Then the sequence $I(H^n)$ is a Cauchy sequence in $L^2(\mathbb{P})$.*

Proof. This is an exercise. □

3.3.11 Theorem. *There exists a unique linear and isometric mapping J on $\mathrm{L}^2(B)$ into $L^2(\mathbb{P})$ such that $J(H) = I(H)$, for $H \in \mathscr{S}_T^2$.*

Proof. For $H \in \mathscr{S}_T^2$, we set $J(H) := I(H)$, where I has been defined in (3.12). Then, J is a linear isometric mapping on $\mathscr{S}_T^2$ into $L^2(\mathbb{P})$ (cf. Proposition 3.3.6 and Theorem 3.3.7). We are going to extend this isometry to elements in $\mathrm{L}^2(B)$. Let $H \in \mathrm{L}^2(B)$. Then, from Proposition 3.3.9, there exists $(H^n)_{n\in\mathbb{N}} \subseteq \mathscr{S}_T^2$ converging to H in $\mathrm{L}^2(B)$. Because of Lemma 3.3.10, the sequence $I(H^n)$ of the elementary iterated integrals is then a Cauchy sequence in $L^2(\mathbb{P})$, which is an Hilbert space. Therefore, there exists $X \in L^2(\mathbb{P})$ such that

$$X = L^2(\mathbb{P})\text{-}\lim_{n\to+\infty} I(H^n). \tag{3.14}$$

So, for $H \in \mathrm{L}^2(B)$, we define

$$J(H) := X. \tag{3.15}$$

We have to show that this definition is well-posed, that is it does not depend on the approximating sequence $(H^n)_{n\in\mathbb{N}} \subseteq \mathscr{S}_T^2$. Let $(\tilde{H}^n)_{n\in\mathbb{N}} \subseteq \mathscr{S}_T^2$ be another approximating sequence for H in $\mathrm{L}^2(B)$. Then, because of the linearity of the elementary stochastic integral (cf. Proposition 3.3.6 (i)) and Itô isometry (3.13) we get

$$\|I(H^n) - I(\tilde{H}^n)\|_{L^2(\mathbb{P})} = \|H^n - \tilde{H}^n\|_{\lambda\otimes\mathbb{P}} \longrightarrow 0, \quad \text{as } n \to +\infty,$$

showing that (3.15) is a well-posed definition. We now show that J is an isometry. Let $H \in \mathscr{L}^2(B)$ and $(H^n)_{n\in\mathbb{N}} \subseteq \mathscr{S}_T^2$ an approximating sequence for H in $\mathscr{L}^2(B)$. Then, using the continuity of the norm and Itô isometry for elementary integrals, yields

$$\|J(H)\|_{L^2(\mathbb{P})} = \lim_{n\to+\infty} \|J(H^n)\|_{L^2(\mathbb{P})} = \lim_{n\to+\infty} \|H^n\|_{\lambda\otimes\mathbb{P}} = \|H\|_{\lambda\otimes\mathbb{P}}$$

which is the extension of Itô isometry to elements of $\mathrm{L}^2(B)$. We now show that J is linear: Let $K \in \mathrm{L}^2(B)$ and $(K^n)_{n\in\mathbb{N}} \subseteq \mathscr{S}_T^2$ an approximating sequence for K in $\mathscr{L}^2(B)$. Then, for every $a, b \in \mathbb{R}$

$$J(aH+bK) = \lim_{n\to+\infty} J(aH^n+bK^n) = a \lim_{n\to+\infty} J(H^n)+b \lim_{n\to+\infty} J(K^n) = aJ(H)+bJ(K).$$

To see the uniqueness of the linear isometric mapping J, we assume that there exists another linear isometric mapping $\tilde{J}$ on $\mathrm{L}^2(B)$ into $L^2(\mathbb{P})$ such that $\tilde{J}(H) = I(H)$ for every $H \in \mathscr{S}_T^2$. Then, for every $H \in \mathrm{L}^2(B)$ and for every $(H^n)_{n\in\mathbb{N}} \subseteq \mathscr{S}_T^2$ converging to H in $L_T^2(\lambda \otimes \mathbb{P})$, we have

$$\|\tilde{J}(H) - \tilde{J}(H^n)\|_{L^2(\mathbb{P})} = \|\tilde{J}(H - H^n)\|_{L^2(\mathbb{P})} = \|H - H^n\|_{\lambda\otimes\mathbb{P}} \longrightarrow 0, \quad \text{as } n \to +\infty.$$

Therefore,

$$\begin{aligned}\|\tilde{J}(H) - J(H)\|_{L^2(\mathbb{P})} &= \|\tilde{J}(H) - \tilde{J}(H^n) + J(H^n) - J(H)\|_{L^2(\mathbb{P})} \\ &\le \|\tilde{J}(H) - \tilde{J}(H^n)\|_{L^2(\mathbb{P})} + \|J(H^n) - J(H)\|_{L^2(\mathbb{P})}\end{aligned}$$

and the right hand side converges to zero as $n \to +\infty$. So $\tilde{J}(H) = J(H)$ a.s., for every $H \in \mathrm{L}^2(B)$ and the proof is complete. □

3.3.12 Definition. Let $H \in \mathrm{L}^2(B)$. Then the unique linear and isometric mapping J of Theorem 3.3.11 is called *the stochastic integral* of H with respect to B. We use the notation

$$J(H) =: \int_0^T H_t \mathrm{d}B_t =: H \cdot B_T\,.$$

3.3.13 Exercise. (i) Let B be a Brownian motion. Show that the following processes belong to $\mathrm{L}^2(B)$: B, B^n, $\mathrm{e}^{\lambda B}$, $\int_0^\cdot |B_t|^p \mathrm{d}t$, $p > 1$, $\sup_{s\in[0,\cdot]} B_s$. Does e^{B^3} belong to $L^2(B)$?
(ii) Compute $B \cdot B_T$.

3.3.14 Exercise. Let B be a Brownian motion over $\mathbb{R}_+$. Show that it is a martingale. Is it a uniformly integrable martingale?

We conclude this chapter stating some properties of the stochastic integral. The proof is an exercise.

3.3.15 Proposition (Moments)**.** *Let $H, K \in \mathrm{L}^2(B)$, $f \in L^2([0,T])$. Then*
(i) $\mathbb{E}[H \cdot B_T] = 0$*;*
(ii) $\mathrm{Var}[H \cdot B_T] = \int_0^T \mathbb{E}[H_t^2]\mathrm{d}\,t$*;*
(iii) $\mathrm{Cov}[H \cdot B_T, K \cdot B_T] = \int_0^T \mathbb{E}[H_t K_t]\mathrm{d}\,t$*;*
(iv) $H \cdot B_T = K \cdot B_T$ *a.s. if and only if* $\int_0^T \mathbb{E}[(H_t - K_t)^2]\mathrm{d}\,t = 0$*;*
(v) $f \cdot B_T$ *is a Gaussian random variable.*

Proof of Proposition 3.3.9. Let $L^2([0,T]) := L^2([0,T],\mathscr{B}([0,T]),\lambda)$. By $\|\cdot\|_2$ we denote the norm of the scalar product in $L^2([0,T])$. Let $P_n : L^2([0,T]) \longrightarrow L^2([0,T])$ be the linear mapping defined by

$$P_n f := \sum_{n=1}^{\lfloor nT\rfloor-1} c_{n,i}(f)1_{[\frac{i}{n},\frac{i+1}{n})}(t), \quad c_{n,i}(f) := \frac{1}{(\frac{1}{n})}\int_{\frac{i-1}{n}}^{\frac{i}{n}} f(s)\mathrm{d}\, s\,. \tag{3.16}$$

Notice that $c_{n,i}(f) < +\infty$, because $f \in L^2([0,T])$. The mapping P_n is defined in such a way to associate to f an approximating sequence $P_n f$ which is a step function over interval of length $\frac{1}{n}$. The value of $P_n f$ over each of these intervals, is given by the mean of f over the preceding interval. The function $P_n f$ is right continuous.

3.3.16 Lemma. *Let $f \in L^2([0,T])$ and P_n be the mapping defined in* (3.16). *Then*
(i) $\|P_n f\|_2 \le \|f\|_2$;
(ii) $\|P_n f - f\|_2 \longrightarrow 0$, *as* $n \to +\infty$.

Proof. We start proving (i). Because of Jensen's inequality applied to the probability measure $n\lambda$ over an interval of length $\frac{1}{n}$, we get

$$(c_{n,i}(f))^2 \le \frac{1}{(\frac{1}{n})}\int_{\frac{i-1}{n}}^{\frac{i}{n}} f(s)^2\mathrm{d}s$$

which yields

$$\begin{aligned}\|P_n f\|_2^2 &= \int_0^T (P_n f)(s)^2\mathrm{d}s = \sum_{n=1}^{\lfloor nT\rfloor-1} c_{n,i}(f)^2\frac{1}{n}\\ &\le \sum_{n=1}^{\lfloor nT\rfloor-1}\int_{\frac{i-1}{n}}^{\frac{i}{n}} f(s)^2\mathrm{d}\, s \le \int_0^T f(s)^2\mathrm{d}\, s = \|f\|_2^2\end{aligned}$$

which proves (i). To see (ii) we first consider a continuous function $g \in L^2([0,T])$. Notice that in this case $P_n g(t)$ converges to $g(t)$, a.e. as $n \to +\infty$. Indeed, by continuity of g, for every $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon)$ such that $g(t) - \varepsilon < g(s) < g(t) + \varepsilon$, for every $s \in (t-\delta, t+\delta)$, for every fixed $t \in [0,T]$. For a fixed t, there exists a unique i^* such that $t \in [\frac{i^*}{n}, \frac{i^*+1}{n})$. Therefore,

$$P_n g(t) = n\int_{\frac{i^*-1}{n}}^{\frac{i^*}{n}} g(s)\mathrm{d}s.$$

Hence for $n > 1/\delta$, $(\frac{i^*-1}{n}, \frac{i^*}{n}) \subseteq (t-\delta, t+\delta)$ and therefore

$$|P_n g(t) - g(t)| \le n\int_{\frac{i^*-1}{n}}^{\frac{i^*}{n}} |g(s) - g(t)|\mathrm{d}s < \varepsilon$$

which shows that $P_n g(t)$ converges to $g(t)$, a.e. as $n \to +\infty$ for a continuous function g. By dominated convergence, because of

$$\sup_{t\in[0,T]} |P_n g(t)| \le \sum_{n=1}^{\lfloor nT\rfloor-1} n\int_{\frac{i-1}{n}}^{\frac{i}{n}} |g(s)|\mathrm{d}\, s \le \int_0^T |g(s)|\mathrm{d}s < +\infty,$$

we also get $P_n g \longrightarrow g$ in $L^2([0,T])$ as $n \to +\infty$. For the general case, we proceed by approximation: Continuous functions are dense in $L^2([0,T])$. Hence, if $f \in L^2([0,T])$, for every $\varepsilon > 0$, then there exists a continuous g such that $\|f-g\|_2 \le \varepsilon/3$, which yields $\|P_n f - P_n g\|_2 \le \varepsilon/3$ because of step (i). We know that $P_n g \longrightarrow g$ in $L^2([0,T])$, as $n \to +\infty$. Therefore, we can fix $n_0 \ge 0$ such that $\|P_n g - g\|_2 \le \varepsilon/3$ for $n \ge n_0$. Using triangular inequality we get

$$\|P_n f - f\|_2 \le \|P_n f - P_n g\|_2 + \|P_n g - g\|_2 + \|f-g\|_2 \le \varepsilon$$

and the proof is complete. □

We now want to approximate processes. Let $H \in \mathrm{L}^2(B)$. We denote by $H_\cdot(\omega)$ the mapping $u \mapsto X_u(\omega)$. Then, from the definition of $\mathrm{L}^2(B)$, the mapping $H_\cdot(\omega)$ belongs to $L^2([0,T])$ for almost all $\omega \in \Omega$. This means that we can define the mapping $\hat{P}_n : \mathrm{L}^2(B) \longrightarrow \mathrm{L}^2(B)$ setting $(\hat{P}_n H)_t(\omega) := 0$, if $H_\cdot(\omega) \notin L^2([0,T])$, $(\hat{P}_n H)_t(\omega) := (P_n H_\cdot(\omega))(t)$ otherwise. Observe that $(\hat{P}_n H) \in \mathscr{S}^2_T$. Indeed,

$$(\hat{P}_n H)_t := \sum_{n=1}^{\lfloor nT \rfloor -1} C_{n,i} 1_{[\frac{i}{n}, \frac{i+1}{n})}(t), \quad C_{n,i} := \frac{1}{(\frac{1}{n})} \int_{\frac{i-1}{n}}^{\frac{i}{n}} X_s \mathrm{d}\, s\,.$$

Then $\hat{P}_n X$ is an elementary process as in Definition 3.3.3, (the $\mathscr{F}_{\frac{i}{n}}$-measurability of $C_{i,n}$ follows because X is progressively measurable). Furthermore

$$\mathbb{E}[C^2_{n,i}] \le \frac{1}{(\frac{1}{n})} \int_{\frac{i-1}{n}}^{\frac{i}{n}} \mathbb{E}[H_s^2] \mathrm{d}\, s \le n \|H\|_{\lambda \otimes \mathbb{P}} < +\infty$$

and hence $(\hat{P}_n H) \in \mathscr{S}^2_T$.

The proof of Proposition 3.3.9 is completed by the following lemma.

3.3.17 Lemma. *Let $H \in \mathrm{L}^2(B)$. Then $\hat{P}_n H$ converges to H in $\mathrm{L}^2(B)$ as $n \to +\infty$.*

Proof. First we set

$$A_n(\omega) := \int_0^T |(\hat{P}_n H)_t(\omega) - H_\cdot(\omega)|^2 \mathrm{d}t = \|P_n H_\cdot(\omega) - H_\cdot(\omega)\|_2^2,$$

$\|\cdot\|_2$ denoting the norm in $L^2([0,T])$. So with this notation we are going to show

$$\lim_{n\to+\infty} \|\hat{P}_n H - H\|^2_{\lambda\otimes\mathbb{P}} = \lim_{n\to+\infty} \mathbb{E}[A_n(\omega)] = 0. \tag{3.17}$$

Because of Lemma 3.3.16 (ii), we know that $P_n H_\cdot(\omega)$ converges to $H_\cdot(\omega)$ in $L^2([0,T])$ for almost all $\omega \in \Omega$. This means $A_n(\omega) \longrightarrow 0$ a.s. as $n \to +\infty$. To conclude we apply the theorem of Lebesgue on dominated convergence: Because $\|a-b\|_2^2 \le 2(\|a\|_2^2 + \|b\|_2^2)$, from Lemma 3.3.16 (i)

$$A_n(\omega) \le 2\big(\|P_n H_\cdot(\omega)\|_2^2 + \|H_\cdot(\omega)\|_2^2\big) \le 4\|H_\cdot(\omega)\|_2^2$$

and the last term in the previous inequalities is integrable because

$$\mathbb{E}[\|H_\cdot(\omega)\|_2^2] = \|H\|^2_{\lambda\otimes\mathbb{P}} < +\infty$$

since $H \in \mathrm{L}^2(B)$. We can therefore apply the theorem on dominated convergence to deduce $A_n(\omega) \longrightarrow 0$ in $L^2(\mathbb{P})$, as $n \to +\infty$ and the proof is complete. □

3.4 Miscellanea

In this section we collect some important properties and consequences of the stochastic integral which we cannot develop in this lecture notes.

3.4.1 Stochastic integral as a process

First of all we discuss some properties of the stochastic integral as a function with respect to the integration extremes. In other words, for every $t \in [0,T]$ and $H \in \mathrm{L}^2_t(B) = \{H = (H_s)_{s\in[0,t]}$ progressively measurable: $\mathbb{E}[\int_0^t H_s^2 \mathrm{d}s] < +\infty\}$, we consider the random variable

$$X_t := \int_0^t H_s \mathrm{d}B_s, \quad t \in [0,T] \tag{3.18}$$

and then the stochastic process $X = (X_t)_{t\in[0,T]}$. Notice that X_t is a random variable for every $t \in [0,T]$ (the definition can be given starting from $\mathscr{S}^2_t$ exactly as we have done before). Therefore X is a stochastic process. However, from the construction, we cannot conclude that it is a measurable process.

3.4.1 Theorem. *Let $(B,\mathbb{F})$ be a Brownian motion. The stochastic-integral process X defined in* (3.18) *is a square integrable $\mathbb{F}$-martingale. Furthermore there exists a continuous modification of X, which is in particular a measurable process.*

Proof. First we show that X is adapted and has the martingale property, that is X is a non-càdlàg martingale. Let $H^n \in \mathscr{S}^2_t$ be a sequence converging to $H = (H_s)_{s\in[0,t]}$ in $\mathrm{L}^2_t(B)$. Then $H^n \cdot B_t$ converges to X_t in $L^2(\mathbb{P})$ as $n \to +\infty$. Therefore, because of the $\mathscr{F}_t$-measurability of $H^n \cdot B_t$ (cf. Proposition 3.3.6 (ii)) and the completeness of the probability space, X_t is $\mathscr{F}_t$-measurable and hence X is adapted. Because of

$$H^n \cdot B_t - H^n \cdot B_s = \int_s^t H_u \mathrm{d}B_u$$

for $s = t_0^n < \ldots < t_k^n = t$, we get

$$\mathbb{E}\Big[\int_s^t H_u^n \mathrm{d}B_u \Big| \mathscr{F}_s\Big] = \sum_{j=0}^{k-1} \mathbb{E}\Big[e_j^n (B_{t_{j+1}^n} - B_{t_j^n}) \Big| \mathscr{F}_s\Big] = \sum_{j=0}^{k-1} \mathbb{E}\Big[e_j^n \mathbb{E}\big[(B_{t_{j+1}^n} - B_{t_j^n}) \big| \mathscr{F}_{t_j^n}\big] \Big| \mathscr{F}_s\Big] = 0.$$

Therefore, $(H^n \cdot B_t)_{t\in[0,T]}$ is a sequence of $\mathbb{F}$-martingales converging in $L^2(\mathbb{P})$ to X_t. Because of Lemma 2.3.11, X is a non-càdlàg martingale. Now we show the existence of a continuous modification of X. We can assume, passing if necessary to a subsequence, that $\|H^n - H\|_{\lambda\otimes\mathbb{P}} \leq \frac{1}{2}\frac{1}{n^3}$. This implies, by triangular inequality, $\|H^n - H^{n+1}\|_{\lambda\otimes\mathbb{P}} \leq \frac{1}{n^3}$. By Doob's maximal inequality and Itô's isometry we get

$$\mathbb{P}\Big[\sup_{t\in[0,T]} \big|H^n \cdot B_t - H^{n+1} \cdot B_t\big| \geq \frac{1}{n^2}\Big] = \mathbb{P}\Big[\sup_{t\in[0,T]} \big|H^n \cdot B_t - H^{n+1} \cdot B_t\big|^2 \geq \frac{1}{n^4}\Big]$$
$$\leq n^4 \|H^n - H^{n+1}\|^2_{\lambda\otimes\mathbb{P}} \leq \frac{1}{n^2}.$$

Because $\sum \frac{1}{n^2} < +\infty$. by Borel–Cantelli Lemma we deduce the existence of a measurable set A of probability one and of $n_0 = n_0(\omega)$ such that

$$\sup_{t\in[0,T]} \left| H^n \cdot B_t(\omega) - H^{n+1} \cdot B_t(\omega) \right| \le \frac{1}{n^2},$$

for every $n \ge n_0$ and $\omega \in A$. But then, by triangular inequality

$$\sup_{t\in[0,T]} \left| H^n \cdot B_t(\omega) - H^m(\omega) \cdot B_t \right| \le \sum_{k=n}^{m-1} \sup_{t\in[0,T]} \left| H^k \cdot B_t(\omega) - H^{k+1} \cdot B_t(\omega) \right| \le \sum_{k=n}^{\infty} \frac{1}{k^2} = \frac{\text{const}}{n}$$

that is $H^n \cdot B_t(\omega)$ is uniformly Cauchy for every $\omega \in A$, $\mathbb{P}[A] = 1$. Therefore, we can define the random variable $X_t^\infty(\omega)$ as the uniform limit of $H^n \cdot B_t(\omega)$ for $\omega \in A$ and zero else. By Proposition 3.3.6 (iii), $t \to H^n \cdot B_t(\omega)$ is a continuous mapping. Hence $t \to X_t^\infty(\omega)$ is continuous for $\omega \in A$, that is $t \to X_t^\infty(\omega)$ is continuous for every $\omega \in \Omega$. We now have to show that X^∞ is a modification of X. We know that $H^n \cdot B_t(\omega)$ converges to X_t^∞ a.s. and $H^n \cdot B_t$ converges to X_t in $L^2(\mathbb{P})$. Therefore there exists a subsequence $H^{n_k} \cdot B_t(\omega)$ converging a.s. to $X_t(\omega)$. By uniqueness of the a.s. limit we get $X_t(\omega) = X_t^\infty(\omega)$ a.s. for every $t \in [0,T]$ and the proof is complete. □

3.4.2 A further extension of the stochastic integral

We now introduce the set

$$\mathrm{L}^2_{\mathrm{loc}}(B) := \left\{ H \text{ progressively measurable: } \int_0^T H_t^2 \mathrm{d}t < +\infty \text{ a.s.} \right\}.$$

We notice that the inclusion $\mathrm{L}^2(B) \subseteq \mathrm{L}^2_{\mathrm{loc}}(B)$ holds. We are going to introduce the stochastic integral with respect to the Brownian motion B for processes in $\mathrm{L}^2_{\mathrm{loc}}(B)$. Notice that every locally bounded progressively measurable process, and in particular any continuous adapted process, belongs to $\mathrm{L}^2_{\mathrm{loc}}(B)$. Therefore, if f is a continuous function, then $f(B) = (f(B_t))_{t\in[0,T]} \in \mathrm{L}^2_{\mathrm{loc}}(B)$. These assertions are not true for $\mathrm{L}^2(B)$.

Let $H \in \mathrm{L}^2_{\mathrm{loc}}(B)$ and define

$$\tau_n := \inf \left\{ t \ge 0 : \int_0^t H_s^2 \mathrm{d}s \ge n \right\}.$$

Then $(\tau_n)_{n\ge1}$ is an increasing sequence of stopping times. Indeed, by the progressive measurability of H, the process $t \to \int_0^t H_s^2 \mathrm{d}s$ is adapted (cf. He, Wang & Yan (1992), Theorem 3.46). Furthermore this is a continuous process. Hence, by Theorem 2.3.5, it is progressively measurable. By Theorem 2.3.8 (iii), τ_n is a stopping time for every $n \ge 1$. It is an increasing sequence because of the continuity of $t \to \int_0^t H_s^2 \mathrm{d}s$.

We define the stochastic interval

$$[0, \tau_n] := \{(\omega, t) \in \Omega \otimes [0,T] : 0 \le t \le \tau_n(\omega)\}$$

and the stochastic process $1_{[0,\tau_n]}$. Notice that $1_{[0,\tau_n]}(\omega, t) = 1$ on $\tau_n(\omega) \ge t$ and $\{\tau_n \ge t\}$ is $\mathscr{F}_t$-measurable, because τ_n is a stopping time. Hence $1_{[0,\tau_n]}$ is adapted and left-continuous. By Theorem 2.3.5 it is a progressively measurable process and the process $H^n := 1_{[0,\tau_n]} H$ belongs to $\mathrm{L}^2(B)$. Indeed

$$\mathbb{E}\left[\int_0^T 1_{[0,\tau_n]}(t) H_t^2 \right] \le \mathbb{E}\left[\int_0^{\tau_n} H_t^2 \mathrm{d}t \right] \le n.$$

3.4.2 Theorem. *Let $H \in \mathrm{L}^2_{\mathrm{loc}}(B)$ and $H^n := H1_{[0,\tau_n]}$. Then there exists a continuous adapted stochastic process X such that*

$$\mathbb{P}\left[X_t = \lim_{n\to+\infty} H^n \cdot B_t, \quad t \in [0,T]\right] = 1.$$

Furthermore X is a (locally bounded) local martingale.

The stochastic process X of Theorem 3.4.2 is again called stochastic integral of $H \in \mathrm{L}^2_{\mathrm{loc}}(B)$ with respect to B. The notation is again

$$X =: \int_0^{\cdot} H_s \mathrm{d}B_s = H \cdot B.$$

We stress that the space $\mathrm{L}^2_{\mathrm{loc}}(B)$ is not a subspace of $L^2_T(\lambda \otimes \mathbb{P})$ and for $H \in \mathrm{L}^2_{\mathrm{loc}}(B)$, $H \cdot B_T$ is not, in general, a square integrable random variable. In particular Itô's isometry does not extend to all stochatic integrals $H \cdot B$ for $H \in \mathrm{L}^2_{\mathrm{loc}}(B)$. For a more detailed discussion of this topic we refer to Schilling & Partsch (2014), §16.

3.4.3 Itô's formula

In this subsection we state Itô's formula in its simpler formulation. For a complete presentation of this topic we refer to Schilling & Partsch (2014), §17.

Itô's formula extends the well known chain-rule for the derivative of the composition $f \circ g$ of two differentiable functions f and g, that is $[f(g(t))]' = f'(g(t))g'(t)$, which in integral form becomes

$$f(g(t)) - f(g(0)) = \int_0^t f'(g(s))g'(s)\mathrm{d}s = \int_0^t f'(g(s))\mathrm{d}g(s).$$

We know that the Brownian motion is a process of unbounded variation. However, if we look back to Theorem 3.1.6, we immediately see that the *quadratic variation* of the Brownian motion is finite, converges in $L^2(\mathbb{P})$ and is equal to the increasing process $A_t = t$. By Itô's formula we can compute the differential of $f(B)$, where f is a twice continuously differentiable function and $(B, \mathbb{F})$ a Brownian motion. However, the chain rule has a correction term involving an integral of the second derivative of f with respect to the quadratic variation of the Brownian motion.

Let $H \in \mathrm{L}^2_{\mathrm{loc}}(B)$ and let K be a progressively measurable process such that $K(\omega, \cdot) \in L^1([0,T])$ for every $\omega \in \Omega$. Then we say that X is an Itô's process if

$$X_t = X_0 + \int_0^t K_s \mathrm{d}s + \int_0^t H_s \mathrm{d}B_s, \quad t \in [0,T]$$

which, in differential, form becomes

$$\mathrm{d}X_t = K_t \mathrm{d}t + H_t \mathrm{d}B_t, \quad t \in [0,T].$$

Let now f be a twice continuously differentiable function on $\mathbb{R}$ into itself. Then Itô's formula is

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(X_s)H_s^2 \mathrm{d}s$$

that is

$$f(X_t) = f(X_0) + \int_0^t \left(f'(X_s)K_s + \frac{1}{2} f''(X_s)H_s^2 \right) \mathrm{d}s + \int_0^t f'(X_s)\mathrm{d}B_s.$$

We notice that to write $f(X_t)$ we have to consider Taylor's expansion of f up to the second order:

$$\mathrm{d}f(X_t) = f'(X_t)\mathrm{d}X_t + \frac{1}{2} f''(X_t)(\mathrm{d}X_t)^2.$$

To determine $(\mathrm{d}X_t)^2$, we use the computational rule given by the so-called Itô's table:

$\times$	$\mathrm{d}t$	$\mathrm{d}B_t$
$\mathrm{d}t$	0	0
$\mathrm{d}B_t$	0	$\mathrm{d}t$

Hence

$$(\mathrm{d}X_t)^2 = \big(K_t \mathrm{d}t + H_t \mathrm{d}B_t\big)^2 = K_t^2(\mathrm{d}t)^2 + H_t^2(\mathrm{d}B_t)^2 + 2K_tH_t\mathrm{d}t\mathrm{d}B_t = H_t^2\mathrm{d}t.$$

This show that Itô's calculus is differs from usual differential calculus. On can say that stochastic calculus is a differential calculus of "second order".

3.4.3 Exercise. (i) Let $\lambda \in \mathbb{R}$. Determine the stochastic differential of $Y = \exp(\lambda B)$. Is Y a martingale? If not, define a process Z such that $X = ZY$ is a martingale.

(ii) Let $f \in L^2([0,T])$. Show that $f \cdot B_T$ is a centred normal random variable and compute explicitely its variance.

CHAPTER 4

Poisson Process

In this chapter we introduce the Poisson process and give its characterization among processes of finite variation. To this end, it is necessary to introduce processes of finite variation and stochastic integration with respect such processes. We notice that in this case, we can define the integral pathwise. We shall also discuss some problem concerning with the independence of Poisson processes.

4.1 Processes of finite variation and stochastic integration

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space and $T > 0$ a fixed time horizon. We consider a filtration $\mathbb{F} = (\mathscr{F}_t)_{t\in[0,T]}$ which satisfies the usual conditions. We shal only consider processes with values in $(E, \mathscr{E}) := (\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

4.1.1 Definition. (i) Let A be a stochastic process. By $\mathrm{Var}(A) := (\mathrm{Var}(A)_t)_{t\in[0,T]}$, we denote the variation process of A defined as a such: $\mathrm{Var}(A)_t(\omega)$ is the variation of $u \mapsto A_u(\omega)$ over the interval $[0, t]$, $t \le T$.

(ii) We say that A is of *finite variation* if $\mathrm{Var}(A)_t < +\infty$, for every $t \le T$.

(iii) We denote by $\mathscr{V}$ the set of processes of finite variation which are $\mathbb{F}$-adapted, càdlàg and starts at zero.

(iv) We say that A is an *increasing* process if its paths are increasing. We denote by $\mathscr{V}^+$ the subspace of $\mathscr{V}$ consisting of increasing processes.

4.1.2 Proposition. *Let $A \in \mathscr{V}$. Then $\mathrm{Var}(A)$ is adapted and there exists a unique pair of processes $B, C \in \mathscr{V}^+$ such that $A = B - C$ and $\mathrm{Var}(A) = B + C$.*

Proof. Set $B := (A + \mathrm{Var}(A))/2$ and $C := B - A$. Clearly the paths of $\mathrm{Var}(A)$ are increasing, hence B and C are increasing. Furthermore we have

$$\mathrm{Var}(A)_t(\omega) := \lim_{n\to+\infty} \sum_{k=1}^{n} |A_{tk/n}(\omega) - A_{t(k-1)/n}(\omega)|$$

implying that $\mathrm{Var}(A)$ is adapted. Hence B and C are adapted and so they are in $\mathscr{V}^+$. The uniqueness of the pair B, C is clear. □

For the next theorem we refer to Appelbaum (2009), Theorem 2.3.14.

4.1.3 Theorem. *Let $A \in \mathscr{V}$. Then*

$$\sum_{0 \le s \le t} |\Delta A_s| \le \mathrm{Var}(A)_t < +\infty$$

for every $t \ge 0$.

Notice that the process $A^d := \sum_{0 \le s \le \cdot} \Delta A_s$ is adapted: Indeed A is càdlàg and therefore the sum on the right-hand side consists of at most countably many addends. The process $A^c := A - A^d$ is therefore adapted. Furthermore it is continuous, because we deleted from A all its jumps. Hence we see that every process of finite variation A can be decomposed in its continuous and purely discontinuous part:

$$A = A^c + A^d. \tag{4.1}$$

We stress that for this decomposition, it is important to know that A is of finite variation, otherwise the process A^d is, in general, not finite-valued.

Let $A \in \mathscr{V}^+$. We want now define the stochastic integral for this processes. Clearly $t \mapsto A_t(\omega)$ is a measure generating function (it is right-continuous and increasing). Therefore (cf. Bauer (2001), Theorem 6.5), there exists a unique measure μ_A^ω over $([0,T], \mathscr{B}([0,T]))$ such that

$$\mu_A^\omega((s,t]) := A_t(\omega) - A_s(\omega), \qquad 0 \le s \le t \le T. \tag{4.2}$$

Notice that μ_A^ω is a finite measure on $[0,T]$. The stochastic integral with respect to $A \in \mathscr{V}^+$ of a measurable process H is defined as a *Stiltjes–Lebesgue integral*, i.e., by fixing $\omega \in \Omega$ and defining the integral pathwise with respect to the trajectory $t \mapsto A_t(\omega)$.

4.1.4 Definition. Let $A \in \mathscr{V}^+$.

(i) We say that a measurable process H is integrable with respect to A if

$$\int_0^t |H_s(\omega)| \mu_\omega^A(\mathrm{d}s) < +\infty$$

for every $t \ge 0$ and for every $\omega \in \Omega$.

(ii) If H is integrable with respect to A by

$$\int_0^t H_s(\omega)\,\mathrm{d}A_s(\omega) := \int_0^t H_s(\omega)\,\mu_\omega^A(\mathrm{d}s), \quad t \ge 0,$$

we denote the integral of H with respect to A up to time t. We introduce the *integral process* $H \cdot A = (H \cdot A_t)_{t \ge 0}$ by

$$H \cdot A_t(\omega) := \begin{cases} \int_0^t H_s(\omega)\,\mathrm{d}A_s(\omega), & \text{if } \int_0^t |H_s(\omega)|\,\mathrm{d}A_s(\omega) < +\infty \\ \\ +\infty, & \text{otherwise.} \end{cases}$$

We now consider the general case $A \in \mathscr{V}$. The stochastic integral with respect to A of a measurable process H can be introduced in a similar way. Indeed, because of Proposition 4.1.2 there exist a unique pair of processes $B, C \in \mathscr{V}^+$ such that $A = B - C$ and $\mathrm{Var}(A) = B + C$.

4.1.5 Definition. Let $A \in \mathscr{V}$.

(i) We say that a measurable process H is integrable with respect to A if it is integrable with respect to $\mathrm{Var}(A)$.

(ii) If H is integrable with respect to A, we introduce the integral process $H \cdot A = (H \cdot A_t)_{t \geq 0}$ by

$$H \cdot A_t(\omega) := \begin{cases} \int_0^t H_s(\omega)\, \mathrm{d}A_s(\omega), & \text{if } \int_0^t |H_s(\omega)|\, \mathrm{dVar}(A)_s(\omega) < +\infty \\ \\ +\infty, & \text{otherwise.} \end{cases}$$

Notice that now the process A induces a *signed measure*: If B and C is the unique pair given in Proposition 4.1.2, then $\mu_A^\omega := \mu_B^\omega - \mu_C^\omega$, where μ_B^ω and μ_C^ω are defined in (4.2). If H is integrable with respect to A, then by definition,

$$\int_0^t |H_s| \mathrm{dVar}(A)_s < +\infty$$

and the formula

$$\mathrm{Var}(H \cdot A) = |A| \cdot \mathrm{Var}(A) \tag{4.3}$$

holds.

The next lemma gives a class of measurable processes which are integrable with respect to processes in $\mathscr{V}$.

4.1.6 Lemma. *Let H be a locally bounded measurable processes. Then H is integrable with respect to $A \in \mathscr{V}$.*

Proof. If H is a locally bounded measurable process and $(\tau_n)_{n\in\mathbb{N}}$ is a localizing sequence, say such that $|H^\tau| \leq c_n$, we have, for every $n \geq 1$ and every fixed $t \geq 0$, $|H^{T_n}| \cdot \mathrm{Var}(A)_t \leq c_n \mathrm{Var}(A)_t < +\infty$. On the other side, $\tau \uparrow +\infty$. Hence, for every fixed $t \geq 0$, there exists $n(t) \in \mathbb{N}$ such that $\tau_n \geq t$ for every $n \geq n(t)$ and so $|H| \cdot \mathrm{Var}(A)_t \leq c_n \mathrm{Var}(A)_t < +\infty$, $n \geq n(t)$. □

From (4.3), we know that if a process H is integrable with respect to A, then $H \cdot A$ is a process of finite variation. Furthermore, the following result, for which we refer to He, Wang & Yan (1992), Theorem 3.46.

4.1.7 Proposition. *Let $A \in \mathscr{V}$ and H be a measurable process which is integrable with respect to A. If H is progressively measurable, then $H \cdot A$ is $\mathbb{F}$-adapted.*

We can now state a result of partial integration for processes in $\mathscr{V}$. For the proof, we refer to Lipster & Shiryaev (2001), Lemma 18.7.

4.1.8 Theorem. *Let $A, B \in \mathscr{V}$. Then*

$$A_t B_t = A_- \cdot B_t + B_- \cdot A_t + \sum_{0 \leq s \leq t} \Delta A_s \Delta B_s \quad t \in [0, T]. \tag{4.4}$$

The formula of integration by part given in Theorem 4.1.8 is well defined. Indeed, A_- and B_- are locally bounded processes, as we remarked after the solution of Exercise 2.3.18, and hence the two integrals on the right-hand side of (4.4) make sense because of Lemma 4.1.6. Furthermore, the sum appearing in (4.4) is finite valued because of Theorem 4.1.3.

If $A \in \mathscr{V}^+$, it may non be integrable. This leads to the next definition.

4.1.9 Definition. (i) A process $A \in \mathscr{V}^+$ is called *integrable* if $\mathbb{E}[A_T] < +\infty$. We denote by $\mathscr{A}^+$ the set of integrable processes.

(ii) A process $A \in \mathscr{V}$ is called of *integrable variation* if $\mathrm{Var}(A) \in \mathscr{A}^+$. We denote by $\mathscr{A}$ the set of processes of integrable variation.

We now want to investigate the situation in which a process $A \in \mathscr{V}$ is a martingale and to see under which conditions the integral $H \cdot A$ is a gain a martingale. Before we consider this problem, we need to introduce *stochastic intervals* and the σ-algebra of *predictable sets.*

4.1.10 Definition. (i) Let τ be a stopping time. The subset $[0, \tau] \subseteq \Omega \times [0, T]$ defined by $[0, \tau] := \{(\omega, t) \in \Omega \times [0, T] : 0 \leq t \leq \tau(\omega)\}$ is called stochastic interval.

(ii) Let $\mathscr{C}^* := \{[0, \tau],\ \tau \text{ stopping time}\}$. Then the σ-algebra $\mathscr{P} := \sigma(\mathscr{C}^*)$ is called *predictable* σ-algebra.

(iii) A stochastic process $X : \Omega \times [0, T] \longrightarrow \mathbb{R}$ is called *predictable*, if it is $(\mathscr{P}, \mathscr{E})$ measurable.

(iv) Let $B \in \mathscr{F}_T \otimes \mathscr{B}([0, T])$. Then we say that B is a progressively measurable set if $\{B \cap [0, t] \times \Omega\}$ belongs to $\mathscr{F}_t \otimes \mathscr{B}([0, t])$, for every $t \in [0, T]$. We denote by $\mathscr{D}$ the σ-algebra generated by progressively measurable sets and call it *progressive* σ-algebra.

The next theorem explains the relation between the σ-algebras $\mathscr{P}$ and $\mathscr{D}$ introduced in Definition 4.1.10.

4.1.11 Theorem. (i) *A stochastic process X is progressively measurable if and only if the mapping $X : \Omega \times [0, T] \longrightarrow \mathbb{R}$ is measurable with respect to $\mathscr{D}$.*

(ii) *Let τ be a stopping time. Then the stochastic interval $[0, \tau]$ is a progressively measurable set. In particular, $\mathscr{P} \subseteq \mathscr{D}$ and predictable processes are adapted and measurable.*

(iii) *Let $\mathscr{C}'$ be the system of all adapted processes which are left-continuous. Then $\sigma(\mathscr{C}') = \mathscr{P}$.*

Proof. We start proving (i). If X is $\mathscr{D}$-measurable, then it is clear that it is progressively measurable. Let now X be progressively measurable. If $X_t(\omega) = 1_B(\omega, t)$, with $B \subseteq \Omega \times [0, T]$, then X is progressively measurable if and only if $B \in \mathscr{D}$. If we now denote by $\mathscr{C}$ the class of all these progressively measurable processes and by $\mathscr{K}$ the class of all bounded and progressively measurable processes which are $\mathscr{D}$-measurable, then $\mathscr{C} \subseteq \mathscr{K}$ and $\mathscr{K}$ is a monotone class. Because of Theorem 1.2.2, we can conclude that every bounded progressively measurable process is $\mathscr{D}$-measurable. If now X is positive $X^n := X \wedge n$ is bounded and converges pointwise to X, therefore we get the statement for progressively measurable and nonnegative processes. For a general progressively measurable process, $X = X^+ - X^-$ and the prove of (i) is complete. We now show (ii). Let τ be a stopping time. Then $1_{[0,\tau]}$ is left-continuous and adapted. By Theorem 2.3.5, $1_{[0,\tau]}$ is progressively measurable and, by (i) it is $\mathscr{D}$-measurable. Hence $[0, \tau]$ is $\mathscr{D}$-measurable. To see (iii), we mention that Dellacherie (1972), Theorem IV.22, it is proven that left-continuous adapted processes are predictable, hence $\sigma(\mathscr{C}') \subseteq \mathscr{P}$. On the other side, $1_{[0,\tau]}$ is left-continuous and adapted, for every stopping time τ. This proves the converse inclusion and the proof of the theorem is complete. □

The next lemma, explain the way in which the stopping procedure acts on the integral with respect to processes in $\mathscr{V}$.

4.1.12 Lemma. *Let τ be a stopping time and H a measurable process which is integrable with respect to A. Then $(H \cdot A)^\tau = (1_{[0,\tau]}H) \cdot A$.*

Proof. The proof is given as an application of the monotone class theorem (cf. Theorem 1.2.2): First take $H = 1_{(u,v]\times B}$, where $B \in \mathscr{F}_T$. For this class of processes the claim holds and they are stable under multiplication and generate $\mathscr{F}_T \otimes \mathscr{B}([0,T])$. Therefore, we get the claim for bounded H and then, by approximation for nonnegative and hence for integrable and measurable H. □

More generally the following result holds:

4.1.13 Lemma. *Let $A \in \mathscr{A}$ and K, H be locally bounded measurable processes. Then $K \cdot A$ and $H \cdot (K \cdot A)$ are of finite variation and*

$$H \cdot (K \cdot A) = HK \cdot A$$

Proof. The proof is also consequence of the monotone class theorem. First we consider H bounded and $K = 1_{(u,v]\times B}$, with $B \in \mathscr{F}$. The the statement clearly hold and an application of the monotone class theorem yields the result for every bounded and measurable H and K. To pass to the general case it is enough to observe that, if H and K are locally bounded and $(\tau_n)_{n\geq 0}$ and $(\sigma_n)_{n\geq 0}$ are localizing sequences for H and K respectively, then $(\rho_n)_{n\geq 0}$, $\rho_n := \tau_n \wedge \sigma_n$, is a localizing sequence for HK and use Lemma 4.1.12. □

Let X be a martingale. X is of finite variation if $X \in \mathscr{V}$; X is of integrable variation if $X \in \mathscr{A}$.

It is clear that increasing martingales must be evanescent: this is a consequence of the monotonicity of the conditional expectation. Furthermore the following result holds (cf. Revuz & Yor (1999), Proposition IV.1.2).

4.1.14 Proposition. *Let X be a martingale in $\mathscr{V}$. If furthermore X is continuous, then it is constant.*

In the next theorem, we now give sufficient conditions for $H \cdot X$ to be a martingale, if X is a martingale in $\mathscr{A}$.

4.1.15 Theorem. *Let H be a predictable and bounded process and $X \in \mathscr{A}$ be a martingale. Then $H \cdot X \in \mathscr{A}$ is a martingale.*

Proof. Let τ be a stopping time taking values in $[0,T]$ and $H = 1_{[0,\tau]}$. Then, because of Lemma 4.1.12 $X^\tau = H \cdot X$ and, from Doob Stopping Theorem (cf. Theorem 2.3.13), X^τ is a martingale. From (4.3) and Lemma 4.1.12 we get

$$\mathrm{Var}(X^\tau)_T = \mathrm{Var}(H \cdot X)_T = H \cdot \mathrm{Var}(X)_T = \mathrm{Var}(X)_T^\tau \leq \mathrm{Var}(X)_T \in L^1(\mathbb{P})$$

meaning that $H \cdot X$ is a martingale in $\mathscr{A}$. Let $\mathscr{C}$ be the class of all such processes. Then, by the definition of predictable σ-algebra, $\mathscr{C}$ generates $\mathscr{P}$. We now denote by $\mathscr{K}$ the class of bounded predictable processes H such that $H \cdot X$ is a martingale. Then, $\mathscr{K}$ is a linear space containing 1 and $\mathscr{C} \subseteq \mathscr{K}$. We now show that $\mathscr{K}$ is a monotone vector space. Let $(H^n)_n \subseteq \mathscr{K}$ be a uniformly bounded nonnegative and increasing sequence converging pointwise to H. An application of Lebesgue theorem on dominated convergence yields $H^n \cdot \mathrm{Var}(X)_t \longrightarrow H \cdot \mathrm{Var}(X)_t$ ω-wise as $n \to +\infty$. Furthermore, H^n is uniformly bounded. Let $c > 0$ be such that $H^n \leq c$. Then $H^n \cdot \mathrm{Var}(X)_t \leq c\mathrm{Var}(X)_T$ and we can conclude that $H^n \cdot \mathrm{Var}(X)_t \longrightarrow H \cdot \mathrm{Var}(X)_t$ in $L^1(\mathbb{P})$ as $n \to +\infty$. But then, because $H - H^n \geq 0$

$$\mathbb{E}\big[|H \cdot X_t - H^n \cdot X_t|\big] \leq \mathbb{E}\big[(H - H^n) \cdot \mathrm{Var}(X)_T\big] \longrightarrow 0 \quad \text{as } n \to +\infty.$$

Because of Lemma 2.3.11, $H \cdot X$ is a martingale. An application of Theorem 1.2.2, yields the claim for all bounded predictable H. Notice that $H \cdot X \in \mathscr{A}$, H being bounded. □

4.2 Poisson Process: definition and characterization

In this part we want to discuss an important example of increasing process: Poisson process. As a first step, we introduce *simple point processes*.

4.2.1 Definition. Let $N \in \mathscr{V}^+$ (hence, in particular, adapted). We say that it is a *simple point process* relative to $\mathbb{F}$ if
(i) N takes values in $\mathbb{N}$ a.s.;
(ii) ΔN takes values in $\{0, 1\}$ a.s.

Let N be a simple point process and introduce the increasing sequence of stopping times $(\tau_n)_{n\geq 0}$ setting $\tau_0 := 0$ and $\tau_{n+1} := \inf\{t > \tau_n : N_t = n + 1\}$. Then N is equal to zero on $[0, \tau_1)$. In τ_1 it makes a jump of size one and stays constant over $[\tau_1, \tau_2)$. In τ_2 it makes a jump of size one, its value becomes two and it stays constant over $[\tau_2, \tau_3)$ and so on. In conclusion, the paths of N are a right-continuous increasing step functions taking values $0, 1, \ldots$ and the values of N changes only because of its jumps. Therefore the continuous part N^c in the decomposition (4.1) of N is identically equal to zero.

4.2.2 Definition. Let N be a simple point process relative to $\mathbb{F}$. We say that it is a *Poisson process* relative to $\mathbb{F}$ if:
(i) $\mathbb{E}[N_t] < +\infty$, for every $t \in \mathbb{R}_+$;
(ii) The function $a(t) := \mathbb{E}[N_t]$, called *intensity* function of N, is continuous;
(iii) $N_t - N_s$ is independent of $\mathscr{F}_s$, for every $0 \leq s < t$.
If N is a Poisson process relative to $\mathbb{F}$ with intensity function $a(\cdot)$ of the form $a(t) = \gamma\, t$, $t \geq 0$, $\gamma > 0$, we say that N is a *homogeneous* Poisson process (with parameter γ) relative to $\mathbb{F}$.

We notice that a Poisson process N is stochastically continuous. Indeed, by definition it is a càdlàg process end hence it is right-continuous. To show that it is stochastically continuous it is enough to prove that it is stochastically continuous from the left. But this is an immediate consequence of Markov's inequality: For any $\varepsilon > 0$ and $0 \le s \le t$ we have $N_t \ge N_s$ and

$$\lim_{s\uparrow t} \mathbb{P}[(N_t - N_s) \ge \varepsilon] \le \frac{1}{\varepsilon} \lim_{s\uparrow t} \mathbb{E}[N_t - N_s] = \frac{1}{\varepsilon} \lim_{s\uparrow t}(a(t) - a(s)) = 0,$$

where in the first passage we applied Markov's inequality and in the last equality we used the continuity of the intensity function $a(\cdot)$.

We now give the characterization of the Poisson process.

4.2.3 Theorem. *Let N be a simple point process with respect to $\mathbb{F}$. Then N is a Poisson process with respect to $\mathbb{F}$ if and only if there exists a continuous increasing function $t \to a(t)$ such that $N - a(\cdot)$ is an $\mathbb{F}$-local martingale. In this case, the function $a(\cdot)$ is the intensity function of N and $N - a(\cdot)$ is a true $\mathbb{F}$-martingale.*

Proof. If N is a Poisson process with intensity $a(\cdot)$, then $(N - a(\cdot), \mathbb{F})$ is a centred process with independent increments. Hence, because of Lemma 2.5.6, it is an $\mathbb{F}$-martingale. We now prove the converse implication, that is, we assume that N is a simple point process such that there exists a continuous increasing function $a(\cdot)$ such that $N - a(\cdot)$ is an $\mathbb{F}$-local martingale. Then we show that $N - a(\cdot)$ is an $\mathbb{F}$ martingale and that N is a Poisson process. Let $N - a(\cdot)$ be an $\mathbb{F}$-local martingale and $(\tau_n)_{n\ge1}$ a localizing sequence for it. Then $(N - a(\cdot))^{\tau_n}$ is a martingale starting at zero and therefore

$$\mathbb{E}\big[N_T^{\tau_n}\big] = \mathbb{E}\big[a(T)^{\tau_n}\big].$$

Furthermore applying B. Levi's theorem on monotone convergence we get

$$\mathbb{E}\big[N_T\big] = \mathbb{E}\Big[\lim_{n\to+\infty} N_{T\wedge\tau_n}\Big] = \lim_{n\to+\infty} \mathbb{E}\big[N_{T\wedge\tau_n}\big] = \lim_{n\to+\infty} \mathbb{E}\big[a(T\wedge\tau_n)\big] = \mathbb{E}\Big[\lim_{n\to+\infty} a(T\wedge\tau_n)\Big] = a(T) < +\infty.$$

Hence $N_T \in L^1(\mathbb{P})$. Furthermore, the estimate $\sup_{t\in[0,T]} |N_t - a(t)| \le N_T + a(T)$ holds, N and $a(\cdot)$ being increasing. Hence, $N - a(\cdot)$ is a local martingale which is bounded in $L^1(\mathbb{P})$ and therefore it is a true martingale (cf. Exercise 2.3.18 (ii)). Furthermore, $\mathrm{Var}(N - a(\cdot))_T \le N_T + a(T)$, hence $N - a(\cdot) \in \mathscr{A}$. We now show that N is a Poisson process with respect to $\mathbb{F}$. We only need to verify that $(N, \mathbb{F})$ is a process with independent increments. We consider the (complex-valued) process $Y = (\mathrm{e}^{\mathrm{i}uN_t})_{t\in[0,T]}$, $u \in \mathbb{R}$, which changes only because of its jumps. Therefore, we can write

$$\begin{aligned}
\mathrm{e}^{\mathrm{i}uN_t} &= 1 + \sum_{0\le s\le t} \Delta \mathrm{e}^{\mathrm{i}uN_s} = 1 + \sum_{0\le s\le t} \Big(\mathrm{e}^{\mathrm{i}uN_s} - \mathrm{e}^{\mathrm{i}uN_{s-}}\Big) = 1 + \sum_{0\le s\le t} \mathrm{e}^{\mathrm{i}uN_{s-}} \Big(\mathrm{e}^{\mathrm{i}u\Delta N_s} - 1\Big) \\
&= 1 + \sum_{0\le s\le t} \mathrm{e}^{\mathrm{i}uN_{s-}} \Big(\mathrm{e}^{\mathrm{i}u} - 1\Big)\Delta N_s = 1 + \Big(\mathrm{e}^{\mathrm{i}u} - 1\Big) \sum_{0\le s\le t} \mathrm{e}^{\mathrm{i}uN_{s-}} \Delta N_s \\
&= 1 + \Big(\mathrm{e}^{\mathrm{i}u} - 1\Big) \int_0^t \mathrm{e}^{\mathrm{i}uN_{s-}} \mathrm{d}N_s
\end{aligned} \tag{4.5}$$

where in the last passage we used the definition of Stiltjes–Lebesgue integral with respect to N. We now introduce the process $Z = Y\mathrm{e}^{-a(\cdot)(\mathrm{e}^{\mathrm{i}u}-1)} = \mathrm{e}^{\mathrm{i}uN - a(\cdot)(\mathrm{e}^{\mathrm{i}u}-1)}$. By partial integration (Theorem 4.1.8), we get

$$Z_t = 1 + \int_0^t \mathrm{e}^{-a(s)(\mathrm{e}^{\mathrm{i}u}-1)} \mathrm{d}Y_s + \int_0^t Y_{s-} \mathrm{d}\mathrm{e}^{-a(s)(\mathrm{e}^{\mathrm{i}u}-1)} + \\ + \sum_{0\le s\le t} \Delta Y_s \Delta \mathrm{e}^{-a(s)(\mathrm{e}^{\mathrm{i}u}-1)}$$

but the last summand on the right-hand side is equal to zero because $a(\cdot)$ is continuous. Therefore, by (4.5), we deduce

$$\begin{aligned} Z_t &= 1 + \Big(\mathrm{e}^{\mathrm{i}u} - 1\Big) \int_0^t \mathrm{e}^{-a(s)(\mathrm{e}^{\mathrm{i}u}-1)} \mathrm{e}^{\mathrm{i}uN_{s-}} \mathrm{d}N_s - \Big(\mathrm{e}^{\mathrm{i}u} - 1\Big) \int_0^t \mathrm{e}^{-a(s)(\mathrm{e}^{\mathrm{i}u}-1)} \mathrm{e}^{\mathrm{i}uN_{s-}} \mathrm{d}a(s) \\ &= 1 + \Big(\mathrm{e}^{\mathrm{i}u} - 1\Big) \int_0^t \mathrm{e}^{-a(s)(\mathrm{e}^{\mathrm{i}u}-1)} \mathrm{e}^{\mathrm{i}uN_{s-}} \mathrm{d}(N_s - a(s)) \\ &= 1 + \Big(\mathrm{e}^{\mathrm{i}u} - 1\Big) \int_0^t Z_{s-} \mathrm{d}(N_s - a(s)) \end{aligned} \tag{4.6}$$

where in the firs passage we used the chain rule for the Stiltjes–Lebesgue integral to compute $\mathrm{d}\mathrm{e}^{-a(s)(\mathrm{e}^{\mathrm{i}u}-1)}$, in the last-but-one passage the linearity of the Stiltjes–Lebesgue integral with respect to the integrator and in the last passage the definition of Z. By Theorem 4.1.11 (iii), the process Z_- is predictable. Furthermore, because of the definition of Z and the continuity of $a(\cdot)$, $|Z_-|$ is bounded. Hence, we can apply Theorem 4.1.15, to conclude that Z is a martingale. This in particular yields

$$\mathbb{E}\Big[\exp\big(\mathrm{i}u(N_t - N_s)\big)\Big|\mathscr{F}_s\Big] = \exp\big((a(t) - a(s))(\mathrm{e}^{\mathrm{i}u} - 1)\big), \quad 0 \le s \le t \tag{4.7}$$

and hence $(N, \mathbb{F})$ is an integrable simple point process with independent increments and continuous intensity function, that is a Poisson process with respect to $\mathbb{F}$. □

Notice, if N is a Poisson process with respect to $\mathbb{F}$, then (4.7) holds. Choosing $s = 0$ and taking the expectation in (4.7) yields

$$\mathbb{E}\Big[\exp\big(\mathrm{i}uN_t\big)\Big] = \exp\big(a(t)(\mathrm{e}^{\mathrm{i}u} - 1)\big), \quad t \in [0, T], \tag{4.8}$$

meaning that N_t is Poisson distributed with parameter $a(t)$, for every $t \in [0, T]$.

4.3 Independence of Poisson processes

Let $N^1, \ldots, N^m$ be Poisson processes with respect to the same filtration $\mathbb{F}$ and let $a^1(\cdot), \ldots, a^m(\cdot)$ be the respective intensity function. Because $(N^j, \mathbb{F})$ is a process with independent increments, from (4.8) and Proposition 2.5.14, we know that the process $Z^j = (Z^j_t)_{t\in[0,T]}$,

$$Z^j_t := \exp\big(\mathrm{i}u^j N^j_t - a^j(t)(\mathrm{e}^{\mathrm{i}u} - 1)\big), \quad t \in [0, T] \tag{4.9}$$

is an $\mathbb{F}$-martingale, for every $j = 1, \dots m$ and $u^j \in \mathbb{R}$. We define the process $Z = (Z_t)_{t \geq 0}$ by

$$Z := \prod_{j=1}^{m} Z^j. \tag{4.10}$$

The following theorem give sufficient conditions for the process Z to be an $\mathbb{F}$-martingale.

4.3.1 Theorem. *Let $N^1, \dots, N^m$ be Poisson processes relative to $\mathbb{F}$ with intensity function $a^1(\cdot), \dots, a^m(\cdot)$, respectively. If $\Delta N^j \Delta N^k = 0$, $j, k = 1, \dots, m$, $k \neq j$, then Z defined in* (4.10) *is an $\mathbb{F}$-martingale with $Z_0 = 1$.*

To prove Theorem 4.3.1 we need two preliminary lemmas, the first of which is a purely algebraic result.

Let $x_1, \dots, x_m$ and $x_{1-}, \dots, x_{m-}$ be two sequence of real numbers. We define

$$\Delta x_j := x_j - x_{j-}, \quad j = 1, \dots, m; \qquad \Delta \prod_{j=1}^{m} x_j := \prod_{j=1}^{m} x_j - \prod_{j=1}^{m} x_{j-}.$$

4.3.2 Lemma. *For any two sequences $x_1, \dots, x_m$ and $x_{1-}, \dots, x_{m-}$ of real numbers such that $\Delta x_j \Delta x_k = 0$, $j, k = 1, \dots, m$, $j \neq k$, it follows:*

$$\Delta \prod_{j=1}^{m} x_j = \sum_{j=1}^{m} \Big(\prod_{k \neq j}^{m} x_{k-} \Big) \Delta x_j. \tag{4.11}$$

Proof. We proceed by induction on $m \in \mathbb{N}$. If $m = 1$ there is nothing to show. Now we assume (4.11) for $m = n$ and show it for $m = n + 1$.

$$\begin{aligned}
\Delta \textstyle\prod_{j=1}^{n+1} x_j &= \textstyle\prod_{j=1}^{n+1} x_j - \prod_{j=1}^{n+1} x_{j-} \\
&= x_{n+1} \textstyle\prod_{j=1}^{n} x_j - x_{(n+1)-} \prod_{j=1}^{n} x_j + x_{(n+1)-} \prod_{j=1}^{n} x_j - x_{(n+1)-} \prod_{j=1}^{n} x_{j-} \\
&= \Delta x_{n+1} \Big(\textstyle\prod_{j=1}^{n} x_{j-} + \prod_{j=1}^{n} x_j - \prod_{j=1}^{n} x_{j-} \Big) + x_{(n+1)-} \Big(\prod_{j=1}^{n} x_j - \prod_{j=1}^{n} x_{j-} \Big) \\
&= \Delta x_{n+1} \Big(\textstyle\prod_{j=1}^{n} x_{j-} + \sum_{j=1}^{n} \Big(\prod_{k \neq j} x_{j-} \Big) \Delta x_j \Big) + x_{(n+1)-} \sum_{j=1}^{n} \Big(\prod_{k \neq j} x_{k-} \Big) \Delta x_j \\
&= \textstyle\sum_{j=1}^{n+1} \Big(\prod_{k \neq j} x_{k-} \Big) \Delta x_j,
\end{aligned}$$

where in the last but one equality we used the induction hypothesis. □

4.3.3 Lemma. *Let $N^1, \dots, N^m$ be Poisson processes relative to $\mathbb{F}$. For the martingales $Z^1, \dots, Z^m$ associated with $N^1, \dots, N^m$ it follows that $\Delta Z^j \Delta Z^k = 0$, $j, k = 1, \dots, m$, $k \neq j$, if and only if $\Delta N^j \Delta N^k = 0$, $j, k = 1, \dots, m$, $k \neq j$.*

Proof. By the definition of Z^j and Z^k, for every $t \geq 0$ a.s., for $j \neq k$, we get

$$\begin{aligned}
|\Delta Z_t^j \Delta Z_t^k| = |\exp(iu^j \Delta N_t^j) - 1||\exp(iu^k \Delta N_t^k) - 1| \times \\
\times |\exp(-a^j(t)(\mathrm{e}^{\mathrm{i}u^j} - 1)||\exp(-a^k(t)(\mathrm{e}^{\mathrm{i}u^k} - 1)|.
\end{aligned}$$

The second factor on the right-hand side of the previous formula is different from zero, for every $t \in [0, T]$ and $u^j, u^k \in \mathbb{R}$. Therefore $|\Delta Z_t^j \Delta Z_t^k| = 0$ for every $t \in [0, T]$ a.s., for every $u^j, u^k \in \mathbb{R}$, if and only if $|\exp(iu^j \Delta N_t^j) - 1||\exp(iu^k \Delta N_t^k) - 1| = 0$ for every $t \in [0, T]$ a.s., for every $u^j, u^k \in \mathbb{R}$, which is verified if and only if $\Delta N^j \Delta N^k = 0$, $k \neq j$. □

Proof of Theorem 4.3.1. We consider the function $F(x_1, \dots, x_m) := \prod_{j=1}^m x^j$. Then

$$Z = F(Z^1, \dots, Z^m).$$

We write

$$F(Z^1, \dots, Z^m)_t := F(Z_t^1, \dots, Z_t^m), \quad t \in [0, T].$$

By induction, we show the identity

$$Z_t = 1 + \sum_{r=1}^m \frac{\partial}{\partial x_r} F(Z^1, \dots, Z^m)_- \cdot Z_t^r. \tag{4.12}$$

Let $F^m := F^m(x_1, \dots, x_m) := \prod_{j=1}^m x^j$ and $Z(m) := \prod_{j=1}^m Z^j$. With this notation, $Z(m)_t = F^m(Z^1, \dots, Z^m)_t$. For $m = 2$, partial integration (cf. Theorem 4.1.8 for complex-valued processes) yields

$$Z(2)_t = Z_t^1 Z_t^2 = 1 + Z_-^1 \cdot Z_t^2 + Z_-^2 \cdot Z_t^1 + \sum_{0 \le s \le t} \Delta Z_s^1 \Delta Z_s^2$$

and the last summand vanishes because of Lemma 4.3.3. Hence

$$Z(2)_t = 1 + Z_-^1 \cdot Z_t^2 + Z_-^2 \cdot Z_t^1 = 1 + \sum_{r=1}^2 \frac{\partial}{\partial x_r} F(Z^1, \dots, Z^m)_- \cdot Z_t^r.$$

We now assume (4.12) for $m-1$ and prove it for m. By partial integration, we have

$$\begin{aligned} Z(m)_t &= Z_t^m Z(m-1)_t \\ &= 1 + Z_-^m \cdot Z(m-1)_t + Z(m-1)_- \cdot Z_t^m + \sum_{0 \le s \le t} \Delta Z_s^m \Delta Z(m-1)_s. \end{aligned}$$

Because of Lemma 4.3.2

$$\Delta Z(m-1) = \Delta \prod_{j=1}^{m-1} Z^j = \sum_{j=1}^{m-1} \Big(\prod_{k \ne j}^{m-1} Z_-^k \Big) \Delta Z^j.$$

Hence, because of Lemma 4.3.3,

$$\sum_{0 \le s \le t} \Delta Z_s^m \Delta Z(m-1)_s = \sum_{j=1}^{m-1} \Big(\prod_{k \ne j}^{m-1} Z_{s-}^k \Big) \Delta Z_s^j \Delta Z_s^m = 0.$$

So, by induction and Lemma 4.1.13,

$$\begin{aligned} Z(m)_t &= 1 + Z_-^m \cdot Z(m-1)_t + Z(m-1)_- \cdot Z_t^m \\ &= 1 + Z(m-1)_- \cdot Z_t^m + \sum_{r=1}^{m-1} \frac{\partial}{\partial x_r} F^{m-1}(Z^1, \dots, Z^{m-1})_- Z_-^m \cdot Z_t^r \\ &= 1 + Z(m-1)_- \cdot Z_t^m + \sum_{r=1, r \ne m}^{m} \frac{\partial}{\partial x_r} F^m(Z^1, \dots, Z^m)_- \cdot Z_t^r \\ &= 1 + \sum_{r=1}^{m} \frac{\partial}{\partial x_r} F(Z^1, \dots, Z^m)_- \cdot Z_t^r \end{aligned}$$

and formula (4.12) is proven. From (4.12) and (4.6), we deduce

$$
\begin{aligned}
Z &= 1 + \sum_{r=1}^{m} \frac{\partial}{\partial x_r} F(Z^1, \dots, Z^m)_- \big(\mathrm{e}^{\mathrm{i}u^r} - 1\big) Z^r_- \cdot (N^r - a^r(\cdot))_t \\
&= 1 + \sum_{r=1}^{m} \big(\mathrm{e}^{\mathrm{i}u^r} - 1\big)\Big(Z_- \cdot (N^r - a^r(\cdot))_t\Big).
\end{aligned}
$$

But by definition $|Z_-|$ is predictable and bounded, as a product of bounded processes ($|Z^j_-|$ is bounded over $[0, T]$ by continuity of $a^j(\cdot)$) and $(N^r - a^r(\cdot)) \in \mathscr{A}$. Therefore, from Theorem 4.1.15, Z is a martingale and the proof is complete. □

We are now able to proof the following result:

4.3.4 Theorem. *Let $N^1, \dots, N^m$ be Poisson processes relative to $\mathbb{F}$. If*

$$
\Delta N^j \Delta N^k = 0, \quad j, k = 1, \dots, m;\ j \neq k,
$$

then the random vector $(N^1_t - N^1_s, \dots, N^m_t - N^m_s)$ is independent and independent of $\mathscr{F}_s$, for every $0 \le s \le t$. In particular the vector $(N^1, \dots, N^m)$ has $\mathbb{F}$-independent increments.

Proof. Let Z^j be the martingale associated with N^j, $j = 1, \dots, m$. By Theorem 4.3.1 the process $Z := \prod_{j=1}^m Z^j$ is an $\mathbb{F}$-martingale: for every $0 \le s \le t$, $\mathbb{E}[Z_t|\mathscr{F}_s] = Z_s$. Hence

$$
\mathbb{E}\left[\exp\left(\mathrm{i}\sum_{j=1}^{m} u^j (N^j_t - N^j_s)\right)\middle|\mathscr{F}_s\right] = \exp\left(\sum_{j=1}^{m} \big(\mathrm{e}^{\mathrm{i}u^j} - 1\big)(a^j(t) - a^j(s))\right), \tag{4.13}
$$

for every $0 \le s \le$ and $u^j \in \mathbb{R}$, which implies, in particular, that $(N^1_t - N^1_s, \dots, N^m_t - N^m_s)$ is independent of $\mathscr{F}_s$. Clearly, the identities

$$
\begin{aligned}
\exp\left(\sum_{j=1}^{m} \Big(\mathrm{e}^{\mathrm{i}u^j} - 1\Big)\left(a^j(t) - a^j(s)\right)\right) &= \prod_{j=1}^{m} \exp\Big(\big(\mathrm{e}^{\mathrm{i}u^j} - 1\big)(a^j(t) - a^j(s))\Big) \\
&= \prod_{j=1}^{m} \mathbb{E}\Big[\exp\Big(\mathrm{i}\,u^j (N^j_t - N^j_s)\Big)\Big]
\end{aligned} \tag{4.14}
$$

holds. Taking now the expectation in (4.13) and using (4.14), we deduce

$$
\begin{aligned}
\mathbb{E}\left[\exp\left(i\sum_{j=1}^{m} u^j (N^j_t - N^j_s)\right)\right] &= \prod_{j=1}^{m} \exp\Big(\big(\mathrm{e}^{\mathrm{i}u^j} - 1\big)(a^j(t) - a^j(s))\Big) \\
&= \prod_{j=1}^{m} \mathbb{E}\Big[\exp\Big(\mathrm{i}\,u^j (N^j_t - N^j_s)\Big)\Big]
\end{aligned}
$$

meaning that the vector $(N^1_t - N^1_s, \dots, N^m_t - N^m_s)$ is independent. The proof is now complete. □

As a consequence of Theorem 4.13, we get the following result about the independence of Poisson processes.

4.3.5 Theorem. *Let $N^1, \ldots, N^m$ be Poisson processes. If*

$$\Delta N^j \Delta N^k = 0, \quad j, k = 1, \ldots, m; \quad j \neq k,$$

then the vector $(N^1, \ldots, N^m)$ is independent.

Proof. To keep notation simpler, we only prove the result for $m = 2$. We show that for every $n \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots t_n$ the formula

$$\begin{aligned} &\mathbb{E}\left[\exp\left(\mathrm{i}\sum_{j=1}^{n} u^j (N^1_{t_j} - N^1_{t_{j-1}}) + \mathrm{i}\sum_{j=1}^{n} v^j (N^2_{t_j} - N^2_{t_{j-1}})\right)\right] \\ &\qquad = \mathbb{E}\left[\exp\left(\mathrm{i}\sum_{j=1}^{n} u^j (N^1_{t_j} - N^1_{t_{j-1}})\right)\right] \mathbb{E}\left[\exp\left(\mathrm{i}\sum_{j=1}^{n} v^j (N^2_{t_j} - N^2_{t_{j-1}})\right)\right], \end{aligned} \tag{4.15}$$

for every $u^j, v^j \in \mathbb{R}$, $j = 1, \ldots, n$. Notice that, because of the independence of the increments, we have

$$\begin{aligned} &\mathbb{E}\left[\exp\left(\mathrm{i}\sum_{j=1}^{n} u^j (N^1_{t_j} - N^1_{t_{j-1}})\right)\right] \mathbb{E}\left[\exp\left(\mathrm{i}\sum_{j=1}^{n} v^j (N^2_{t_j} - N^2_{t_{j-1}})\right)\right] \\ &\qquad = \prod_{j=1}^{n} \mathbb{E}\left[\exp\left(\mathrm{i}\, u^j (N^1_{t_j} - N^1_{t_{j-1}})\right)\right] \prod_{j=1}^{n} \mathbb{E}\left[\exp\left(\mathrm{i}\, v^j (N^2_{t_j} - N^2_{t_{j-1}})\right)\right]. \end{aligned} \tag{4.16}$$

Because of Theorem 4.13, $(N^1_t - N^1_s, N^2_t - N^2_s)$ is an independent vector which is independent of $\mathscr{F}_s$, $0 \leq s \leq t$. Assuming that the vector

$$A := ((N^1_{t_1} - N^1_{t_0}, N^2_{t_1} - N^2_{t_0}), \ldots, (N^1_{t_{n-1}} - N^1_{t_{n-2}}, N^2_{t_{n-1}} - N^2_{t_{n-2}}))$$

is independent, we get that the vector

$$B := ((N^1_{t_1} - N^1_{t_0}, N^2_{t_1} - N^2_{t_0}), \ldots, (N^1_{t_n} - N^1_{t_{n-1}}, N^2_{t_n} - N^2_{t_{n-1}}))$$

is independent. Indeed, A is $\mathscr{F}_{t_{n-1}}$-measurable and $(N^1_{t_n} - N^1_{t_{n-1}}, N^2_{t_n} - N^2_{t_{n-1}}))$ is an independent vector which is independent of $\mathscr{F}_{t_{n-1}}$. Therefore, by induction, we get that, for every $n \in \mathbb{N}$, the vector

$$B := ((N^1_{t_1} - N^1_{t_0}, N^2_{t_1} - N^2_{t_0}), \ldots, (N^1_{t_n} - N^1_{t_{n-1}}, N^2_{t_n} - N^2_{t_{n-1}}))$$

is independent. Hence, setting

$$(u^j, v^j)(N^1_{t_j} - N^1_{t_{j-1}}, N^2_{t_j} - N^2_{t_{j-1}}) = u^j (N^1_{t_j} - N^1_{t_{j-1}}) + v^j (N^2_{t_j} - N^2_{t_{j-1}}),$$

$$\begin{aligned}
&\mathbb{E}\Bigg[\exp\Bigg(\mathrm{i}\sum_{j=1}^{n}u^j(N^1_{t_j}-N^1_{t_{j-1}})+\mathrm{i}\sum_{j=1}^{n}v^j(N^2_{t_j}-N^2_{t_{j-1}})\Bigg)\Bigg]\\
&\quad=\mathbb{E}\Bigg[\exp\Bigg(\mathrm{i}(u^1,v^1)(N^1_{t_1}-N^1_{t_0},N^2_{t_1}-N^2_{t_0})+\ldots\\
&\qquad\qquad+\mathrm{i}(u^n,v^n)(N^1_{t_n}-N^1_{t_{n-1}},N^2_{t_n}-N^2_{t_{n-1}})\Bigg)\Bigg]\\
&\quad=\prod_{j=1}^{n}\mathbb{E}\Bigg[\exp\Bigg(\mathrm{i}(u^j,v^j)(N^1_{t_j}-N^1_{t_{j-1}},N^2_{t_j}-N^2_{t_{j-1}})\Bigg)\Bigg]\\
&\quad=\prod_{j=1}^{n}\mathbb{E}\Bigg[\exp\Bigg(\mathrm{i}u^j(N^1_{t_j}-N^1_{t_{j-1}})\Bigg)\Bigg]\mathbb{E}\Bigg[\exp\Bigg(\mathrm{i}v^j(N^2_{t_j}-N^2_{t_{j-1}})\Bigg)\Bigg]\\
&\quad=\prod_{j=1}^{n}\mathbb{E}\Bigg[\exp\Bigg(\mathrm{i}u^j(N^1_{t_j}-N^1_{t_{j-1}})\Bigg)\Bigg]\prod_{j=1}^{n}\mathbb{E}\Bigg[\exp\Bigg(\mathrm{i}v^j(N^2_{t_j}-N^2_{t_{j-1}})\Bigg)\Bigg],
\end{aligned}$$

where, to get the second equality, we used the independence of the vector B for every $n \in \mathbb{N}$. For the third equality we applied Theorem 4.13, which ensures the independence of $(N^1_{t_j}-N^1_{t_{j-1}}, N^2_{t_j}-N^2_{t_{j-1}})$. Therefore, (4.15) holds, because of (4.16). This means that the vectors $(N^1_{t_1}-N^1_{t_0},\ldots,N^1_{t_n}-N^1_{t_{n-1}})$ and $(N^2_{t_1}-N^2_{t_0},\ldots,N^2_{t_n}-N^2_{t_{n-1}})$ are independent, for every $0=t_0<t_1<\ldots<t_n$. Hence, the vectors $(N^1_{t_1},\ldots,N^1_{t_n})$ and $(N^2_{t_1},\ldots,N^2_{t_n})$ are independent because they can be obtained as linear combination of the previous two vectors respectively. Therefore (N^1,N^2) is an independent vector of processes and the proof is complete. □

The proof of the following theorem is very instructive for the understanding of the developed theory. However, it will not be part of the exam (contrarily, the formulation of the theorem will).

4.3.6 Theorem. *Let N^1 and N^2 be independent Poisson processes with respect to the same filtration $\mathbb{F}$. Then $\Delta N^1 \Delta N^2 = 0$.*

Proof. We first prove that the process $(N^1-a^1(\cdot))(N^2-a^2(\cdot))$ is a martingale with respect to $\mathbb{F}$. Because of the independence of the factors, the random variable $(N^1_t - a^1(t))(N^2_t-a^2(t))$ is integrable and. Furthermore, setting $\overline{N}^j := (N^j - a^j(\cdot))$, $j=1,2$, we get

$$\mathbb{E}\big[(\overline{N}^1_t-\overline{N}^1_s)(\overline{N}^2_t-\overline{N}^2_s)\big|\mathscr{F}_s\big]=\mathbb{E}\big[(\overline{N}^1_t-\overline{N}^1_s)\mathbb{E}\big[(\overline{N}^2_t-\overline{N}^2_s)\big|\mathscr{F}_s\vee\sigma(\overline{N}^1_t-\overline{N}^1_s)\big]\big|\mathscr{F}_s\big]=0,$$

where in the last passage we used that $(\overline{N}^2_t-\overline{N}^2_s)$ is independent of $\mathscr{F}_s\vee\sigma(\overline{N}^1_t-\overline{N}^1_s)$ and centred. Therefore, $\overline{N}^1\overline{N}^2$ is an $\mathbb{F}$-martingale. By partial integration and continuity of $a^1(\cdot)$ and $a^1(\cdot)$, we get

$$\begin{aligned}
(N^1_t-a^1(t))(N^2_t-a^2(t)) &=\\
&(N^1_- - a^1(\cdot))\cdot(N^2-a^2(\cdot))_t+(N^2_- - a^2(\cdot))\cdot(N^1-a^1(\cdot))_t+\\
&\qquad+\sum_{0\le s\le t}\Delta N^1_s\Delta N^2_s,
\end{aligned}$$

Because $(N^1 - a^1(\cdot))(N^2 - a^2(\cdot))$ is a martingale and $(N^j_- - a^j(\cdot)) \cdot (N^k - a^k(\cdot))$, is a local martingale $j, k = 1, 2$, $j \neq k$, (cf. Theorem 4.1.15), both starting at zero, we deduce that the process $\sum_{0 \leq s \leq \cdot} \Delta N^1_s \Delta N^2_s$ is a local martingale which starts at zero. Let now $(\tau_n)_{n \geq 1}$ be a localizing sequence. Then

$$\mathbb{E}\left[\sum_{0 \leq s \leq \tau_n \wedge t} \Delta N^1_s \Delta N^2_s\right] = 0$$

which implies $\sum_{0 \leq s \leq \tau_n \wedge t} \Delta N^1_s \Delta N^2_s = 0$, $t \in [0, T]$ a.s., because $\sum_{0 \leq s \leq \cdot} \Delta N^1_s \Delta N^2_s$ is a nonnegative process. Furthermore, for every $t \in [0, T]$ a.s. holds

$$0 = \Delta \sum_{0 \leq s \leq \tau_n \wedge t} \Delta N^1_s \Delta N^2_s = \sum_{0 \leq s \leq \tau_n \wedge t} \Delta N^1_s \Delta N^2_s - \sum_{0 \leq s < \tau_n \wedge t} \Delta N^1_s \Delta N^2_s = \Delta N^1_{\tau_n \wedge t} \Delta N^2_{\tau_n \wedge t}$$

hence

$$\Delta N^1_t \Delta N^2_t = \lim_{n \to +\infty} \Delta N^1_{\tau_n \wedge t} \Delta N^2_{\tau_n \wedge t} = 0, \quad t \in [0, T], \text{ a.s.}$$

and the proof is complete. □

The next result states that for a family of Poisson processes, independence of the family and pairwise independence are equivalent properties.

4.3.7 Corollary. *Let $N^1, \ldots, N^m$ be Poisson processes relative to the same filtration $\mathbb{F}$. Then $(N^1, \ldots, N^m)$ is independent if and only if (N^j, N^k), $j, k = 1, \ldots, m$, $j \neq k$, are independent.*

Proof. It is enough to show that the independence of (N^j, N^k), $j, k = 1, \ldots, m$, $j \neq k$ implies the independence of $(N^1, \ldots, N^m)$. The converse implication is clear. Because of Theorem 4.3.6, the independence of (N^j, N^k), $j, k = 1, \ldots, m$, $j \neq k$ implies $\Delta N^j \Delta N^k = 0$, $j, k = 1, \ldots, m$, $j \neq k$. Hence by Theorem 4.13 we get the independence of $(N^1, \ldots, N^m)$. □

CHAPTER 5

Poisson random measures

5.1 Random measures

We devote this section to Poisson random measures relative to a filtration. We do not consider general Poisson random measures. Rather we restrict our attention to random measures associated with the jumps of adapted càdlàg processes and consider only the *homogeneous* case. Before we need to introduce the notion of random measure and of integer-valued random measure. Of particular interest will be the part concerning the definition of the stochastic integral of deterministic functions with respect to a Poisson random measure and with respect to a compensated Poisson random measure. We fix a *complete* probability space $(\Omega, \mathscr{F}, \mathbb{P})$, a time horizon $[0,T]$, $T > 0$ and a filtration $\mathbb{F} = (\mathscr{F}_t)_{t\in[0,T]}$ satisfying the *usual conditions*. Then we assume $\mathscr{F} = \mathscr{F}_T$. For the sake of simplicity, we introduce the following notation:

$$(E, \mathscr{B}(E)) := ([0,T] \times \mathbb{R}, \mathscr{B}([0,T]) \otimes \mathscr{B}(\mathbb{R})). \tag{5.1}$$

All needed notions of measure theory (definition of ring, semiring...) are summarized in the appendix to Chapter 1.

The jump jeasure of a càdlàg process

5.1.1 Definition. A random measure μ on $(E, \mathscr{B}(E))$ is a mapping on $\Omega \times \mathscr{B}(E)$ in $[0,+\infty]$ such that:

(i) $\mu(\cdot, A)$ is a random variable for every $A \in \mathscr{B}(E)$.

(ii) $\mu(\omega, \cdot)$ is a measure on $(E, \mathscr{B}(E))$ such that $\mu(\omega; \{0\} \times \mathbb{R}) = 0$, $\omega \in \Omega$.

If μ is a random measure on $(E, \mathscr{B}(E))$, we write

$$\mu(A) := \mu(\omega, A), \quad A \in \mathscr{B}(E).$$

For any measurable set A, $\mu(A)$ is a nonnegative random variable on $(\Omega, \mathscr{F}, \mathbb{P})$. We can therefore introduce the expectation of $\mu(A)$ (note that, by definition, $\mu(A) \geq 0$). We call *intensity measure* of μ the mapping m on $\mathscr{B}(E)$ in $[0,+\infty]$ defined by

$$\mathrm{m}(A) := \mathbb{E}[\mu(A)]. \tag{5.2}$$

The intensity measure m is a (deterministic) measure on $(E, \mathscr{B}(E))$. Indeed, $\mathrm{m}(\emptyset) = 0$ because $\mu(\omega, \emptyset) = 0$, for every ω, $\mu(\omega, \cdot)$ being a measure. The σ-additivity of m follows from the theorem of B. Levi on monotone convergence (cf. Theorem 1.A.4).

5.1.2 Definition. We say that a random measure μ on $(E, \mathscr{B}(E))$ is an *integer-valued* random measure if $\mu(A)$ takes values in $\mathbb{N} \cup \{+\infty\}$, for every $A \in \mathscr{B}(E)$.

Integer-valued random measures are of special importance because of the relation that they have with càdlàg adapted processes. Let X be a càdlàg adapted process. For every $A \in \mathscr{B}(E)$ we define on $(E, \mathscr{B}(E))$ the random measure μ by

$$\mu(\omega; A) = \sum_{s\geq 0} 1_{\{\Delta X_s(\omega)\neq 0\}} 1_A(s, \Delta X_s(\omega)), \quad \omega \in \Omega, \quad A \in \mathscr{B}(E). \tag{5.3}$$

5.1.3 Proposition. *Let X be an adapted càdlàg process with values in $\mathbb{R}$. Then the random measure μ defined on $(E, \mathscr{B}(E))$ by* (5.3) *is an integer-valued random measure.*

Proof. We show that (5.3) defines an integer-valued random measure. The process X is adapted and càdlàg . Therefore it is progressively measurable (cf. Theorem 2.3.5) and the set $\{t > 0 : \Delta X_t \neq 0\}$ is at most countable (cf. Theorem 2.2.6). Therefore, from this latter property, if (5.3) defines a random measure, it is an integer-valued random measure. It is clear that for every $\omega \in \Omega$ $\mu(\omega, \cdot)$ is a measure on $(E, \mathscr{B}(E))$. We have to show that $\mu(\cdot, A)$ is a random variable, for every $A \in \mathscr{B}(E)$. This is consequence of the provgressive measurability of X. Indeed, X is a measurable process, being progressively measurable and adapted. Therefore $\omega \to 1_{\{\Delta X_t(\omega)\neq 0\}} 1_A(t, \Delta X_t(\omega))$ is $\mathscr{F}$-measurable for every $t \in [0, T]$. Therefore $\omega \to \mu(\omega, A)$ is a random variable for every $A \in \mathscr{B}(E)$, as a countable sum of random variables. □

5.1.4 Definition. We call the integer-valued random measure μ defined in (5.3) the *jump measure* of X.

We shall only consider integer-valued random measures which are jump measure of some adapted càdlàg process.

Let X be an $\mathbb{F}$-adapted càdlàg process and let μ be its jump measure. It is easy to see that $\mu(\{t\} \times \mathbb{R}) \in \{0, 1\}$. Indeed, from the definition of μ, we get

$$\begin{aligned}\mu(\{t\} \times \mathbb{R}) &= \textstyle\sum_{s\geq 0} 1_{\{\Delta X_s\neq 0\}} 1_{\{t\}\times\mathbb{R}}(s, \Delta X_s)\\ &= 1_{\{\Delta X_t\neq 0\}} 1_{\{t\}\times\mathbb{R}}(t, \Delta X_t)\\ &= 1_{\{\Delta X_t\neq 0\}} \in \{0, 1\}.\end{aligned}$$

If $A \in \mathscr{B}(E)$, we define the process $N^A = (N^A_t)_{t\geq 0}$ by

$$N^A_t := \mu(A \cap [0, t] \times \mathbb{R}). \tag{5.4}$$

5.1.5 Lemma. *Let X be an $\mathbb{F}$-adapted càdlàg process and let μ be its jump measure with intensity measure* m. *If $A \in \mathscr{B}(E)$ is such that $\mu(A \cap [0, t] \times \mathbb{R}) < +\infty$ for every $t \geq 0$, then the process N^A introduced in* (5.4) *is a simple point process.*

Proof. By definition, the process N^A is increasing. Proposition 5.1.3 implies that μ is an integer-valued random measure and so $N_t^A \in \mathbb{N}$, $t \geq 0$. Furthermore, we have $\mu(A \cap [0,t] \times \mathbb{R}) < +\infty$ for every $t \geq 0$ which yields $\mu(A \cap [0, t + \frac{1}{n}] \times \mathbb{R}) < +\infty$ for every $n \geq 1$ and $\mu(A \cap [0, t + \frac{1}{n}] \times \mathbb{R}) \downarrow \mu(A \cap [0,t] \times \mathbb{R})$ as $n \to +\infty$. Therefore N^A is right-continuous. The left-limit of N_t^A is given by $\mu(A \cap [0,t) \times \mathbb{R}) < +\infty$. Thus, N^A is a càdlàg increasing process. For every $t \geq 0$, $N_{t-}^A = \mu(A \cap [0,t) \times \mathbb{R})$ and from this it it follows that

$$\Delta N_t^A = \mu(A \cap \{t\} \times \mathbb{R}) \leq \mu(\{t\} \times \mathbb{R}) \in \{0,1\}.$$

We know that $\mu(A \cap \{t\} \times \mathbb{R}) \in \mathbb{N}$, because μ is integer-valued and so the previous formula yields $\Delta N^A \in \{0,1\}$. It remains to prove that N^A is an $\mathbb{F}$-adapted process. But this easily follow from the progressive measurability of X. Indeed, $\omega \to 1_{\{\Delta X_s(\omega) \neq 0\}} 1_A(s, \Delta X_s(\omega))$ is $\mathscr{F}_t$-measurable, for every $s \leq t$ and hence N_t^A is $\mathscr{F}_t$-measurable, as a countable sum of $\mathscr{F}_t$-measurable random variables. □

The assumption of Lemma 5.1.5 can be weakened requiring that $\mu(A \cap [0,t] \times \mathbb{R}) < +\infty$ *a.s.* for every $t \geq 0$. In this case the process N^A introduced in (5.4) is defined only a.s. and therefore the statement of Lemma 5.1.5 holds a.s. To extend the definition of N^A everywhere we can set $N_t^A(\omega) = 0$ on the exceptional set on which it is not defined by (5.4). The filtration $\mathbb{F}$ satisfies the usual conditions and this version of the process N^A is adapted and càdlàg. We denote this process again by N^A and obviously the statement of Lemma 5.1.5 holds also for such a modification.

5.2 Definition of Poisson random measure

We consider an $\mathbb{F}$-adapted càdlàg process X with jump measure μ. Let m be the intensity measure of μ. Thanks to Proposition 5.1.3, we know that the random measure μ is an integer-valued random measure. Now we are going to discuss the case in which the jump measure of X is an *homogeneous* Poisson random measure relative to the filtration $\mathbb{F}$. The definition of Poisson random measure relative to a filtration can be given in full generality, without relating it to the jump measure of an adapted càdlàg process. Such a general definition requires some technical preparation which would exceed the purpose of these notes. For complete treatment of the topic we refer to Jacod & Shiryaev (2000), Chapter II.

5.2.1 Definition. Let X be an $\mathbb{F}$-adapted process and let μ be its jump measure with intensity measure m (cf. (5.3) and (5.2), respectively). We say that μ is a *Poisson random measure relative to the filtration* $\mathbb{F}$ if:

(i) The intensity measure m is of the form $\mathrm{m} = \lambda_+ \otimes \rho$, where λ_+ is the Lebesgue measure on $([0,T], \mathscr{B}([0,T]))$ and ρ is a σ-finite measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

(ii) For every fixed $s \in (0,T]$ and every $A \in \mathscr{B}(E)$ such that $A \subseteq (s,T] \times \mathbb{R}$, $\mathrm{m}(A) < +\infty$, the random variable $\mu(A)$ is independent of $\mathscr{F}_s$.

We remark that if X is a càdlàg process and its jump measure μ is a Poisson random measure relative to $\mathbb{F}$, then the process X as no fixed-time discontinuities a.s. Indeed, we have $\mathrm{m}(\{t\} \times \mathbb{R}) = \lambda_+(\{t\})\rho(\mathbb{R})$. Because ρ is σ-finite, this implies $\mathrm{m}(\{t\} \times \mathbb{R}) = 0$. Therefore $\mathbb{E}[\mu(\{t\} \times \mathbb{R})] = 0$ and so $\mu(\{t\} \times \mathbb{R}) = 0$ a.s., $t \geq 0$.

We now show that if μ is a Poisson random measure relative to $\mathbb{F}$ and $A \in \mathscr{B}(E)$ is such that $\mathrm{m}(A) < +\infty$, then the process N^A introduced in (5.4) is a Poisson process relative to $\mathbb{F}$ (cf. Definition 4.2.2).

5.2.2 Lemma. *Let μ be the jump measure of a càdlàg adapted process X with intensity measure* m. *If μ is a Poisson random measure relative to $\mathbb{F}$, then for every set $A \in \mathscr{B}(E)$ such that $\mathrm{m}(A \cap [0,t] \times \mathbb{R}) < +\infty$, $t \geq 0$, the process N^A introduced in* (5.4) *is an $\mathbb{F}$-adapted Poisson process relative to the filtration $\mathbb{F}$ with intensity function $a^A(\cdot) := \mathrm{m}(A \cap [0,\cdot] \times \mathbb{R})$.*

Proof. We are going to prove that under the stated assumptions the process N^A fulfils all the properties of Definition 4.2.2. Because of $\mathrm{m}(A \cap [0,t] \times \mathbb{R}) < +\infty$ for every $t \geq 0$, it follows that $\mu(A \cap [0,t] \times \mathbb{R}) < +\infty$ a.s. for every $t \geq 0$. From Lemma 5.1.5 and the comment following it, the process N^A is an $\mathbb{F}$-adapted simple point process. Moreover, $\mathrm{m}(A \cap [0,t] \times \mathbb{R}) < +\infty$, $t \geq 0$, yields $N^A_t \in L^1(\mathbb{P})$, $t \geq 0$. Thus,

$$a^A(t) := \mathbb{E}[N^A_t] = \mathrm{m}(A \cap [0,t] \times \mathbb{R}_+) = (\lambda_+ \otimes \mu)(A \cap [0,t] \times \mathbb{R}_+) < +\infty, \quad t \geq 0. \tag{5.5}$$

The previous equalities imply that $a^A(\cdot)$ is a continuous function because λ_+ is the Lebesgue measure on $[0,T]$. It remains to prove that N^A has $\mathbb{F}$-independent increments. But this is immediate from the properties of μ. Indeed,

$$N^A_t - N^A_s = \mu(A \cap (s,t] \times \mathbb{R}), \quad 0 \leq s \leq t,$$

and $A \cap (s,t] \times \mathbb{R} \subseteq (s,T] \times \mathbb{R}$. Because μ is a Poisson random measure relative to the filtration $\mathbb{F}$, it follows that $N^A_t - N^A_s$ is independent of $\mathscr{F}_s$, $0 \leq s \leq t$. Therefore $(N^A, \mathbb{F})$ is a Poisson process relative to $\mathbb{F}$ with intensity function $a^A(\cdot)$. □

We now state some properties of Poisson random measures relative to a filtration.

5.2.3 Theorem. *Let μ be the jump measure of an $\mathbb{F}$-adapted càdlàg process X with intensity measure* m. *If μ is a Poisson random measure relative to the filtration $\mathbb{F}$, then it has the following properties:*

(i) *For every $A \in \mathscr{B}(E)$ such that $\mathrm{m}(A) < +\infty$, the random variable $\mu(A)$ is Poisson distributed with parameter $\mathrm{m}(A)$.*

(ii) *If $A_1, \dots, A_m$, $m \geq 1$, are $\mathscr{B}(E)$-measurable pairwise disjoint subsets such that $\mathrm{m}(A_j) < +\infty$, $j = 1, \dots, m$, then the vector $(\mu(A_1), \cdots, \mu(A_m))$ of random variable is independent.*

(iii) *If $A_1, \dots, A_m$, $m \geq 1$, are $\mathscr{B}(E)$-measurable subsets such that $\mathrm{m}(A_j) < +\infty$ and that $A_j \subseteq (s,T] \times \mathbb{R}$, $s > 0$, $j = 1, \dots, m$, then $(\mu(A_1), \cdots, \mu(A_m))$ is a random vector independent of $\mathscr{F}_s$.*

Proof. Let $A \in \mathscr{B}(E)$ be such that $\mathrm{m}(A) < +\infty$. We show that $\mu(A)$ is Poisson distributed with parameter $\mathrm{m}(A)$. By Lemma 5.2.2, the process $(N^A, \mathbb{F})$ is a Poisson process with intensity function $a(\cdot) = \mathrm{m}(A \cap [0,\cdot] \times \mathbb{R})$ and therefore, by (4.8)

$$\mathbb{E}\big[\mathrm{e}^{iu\mu(A)}\big] = \mathbb{E}\big[\mathrm{e}^{iuN^A_T}\big] = \exp\big((\mathrm{e}^{iu} - 1)\mathrm{m}(A)\big)\,.$$

In particular, this implies that $M(A)$ is Poisson distributed with parameter $\mathrm{m}(A)$ and (i) is proven. Let $A_1,\ldots,A_m\in\mathscr{B}(E)$ be pairwise disjoint and such that $\mathrm{m}(A_j)<+\infty$. By Lemma 5.2.2, we know that the process $(N^{A_j},\mathbb{F})$ is a Poisson process. Furthermore,

$$\Delta N^{A_j}\Delta N_t^{A_k}=\mu(A_j\cap\{t\}\times\mathbb{R})\mu(A_k\cap\{t\}\times\mathbb{R})=1_{\{\Delta X_t\neq 0\}}1_{A_j\cap A_k}(t,\Delta X_t)=0, \quad (5.6)$$

for every $j,k=1,\ldots,m,\ j\neq k$.

Now we discuss (ii). If $A_1,\ldots,A_m$ are pairwise disjoint subsets of E such that $\mathrm{m}(A_j)<+\infty$, because of (5.6) and Theorem 4.3.5, the vector $(N^{A_1},\ldots N^{A_m})$ is an independent vector of Poisson processes and therefore $(\mu(A_1),\ldots,\mu(A_m))=(N_T^{A_1},\ldots,N_T^{A_m})$ is an independent vector of random variables. To see (iii) it is sufficient to prove it for pairwise disjoint sets. Indeed, we can always reduce the general situation to this particular case by considering an appropriate partition of the union of the A_js. If $A_1,\ldots,A_m$ are pairwise disjoint subsets of E such that $\mathrm{m}(A_j)<+\infty$ and that $A_j\subseteq(s,+\infty)\times\mathbb{R}$, $j=1,\ldots,m$, because of (5.6) and of Theorem 4.3.4, the vector $(N_t^{A_1}-N_s^{A_1},\ldots,N_t^{A_m}-N_s^{A_m})$ is independent of $\mathscr{F}_s$, $0\leq s\leq t$. On the other side, $A_j\subseteq(s,+\infty)\times\mathbb{R}$ implies $N_s^{A_j}=0$, $j=1,\ldots,m$. This yields the vector $(N_t^{A_1},\ldots,N_t^{A_m})$ is independent of $\mathscr{F}_s$, $t\geq 0$. Therefore $(\mu(A_1),\ldots,\mu(A_m))=(N_T^{A_1},\ldots,N_T^{A_m})$ is independent of $\mathscr{F}_s$. □

5.3 Stochastic integration for Poisson random measures

Let X be an $\mathbb{F}$-adapted càdlàg process. We assume that that jump measure of X is a Poisson random measure relative to the filtration $\mathbb{F}$ with intensity measure $\mathrm{m}=\lambda_+\otimes\rho$. In this section we define the integral of deterministic measurable functions with respect to μ. To simplify the terminology, in this section we call a Poisson random measure relative to a filtration simply a *Poisson random measure*. We recall that the definition of $(E,\mathscr{B}(E))$ was given in (5.1). For a deterministic *numerical* function f which is $\mathscr{B}(E)$-measurable we have introduce the notation

$$f*\mathrm{m}:=\int_E f(t,x)\,\mathrm{m}(\mathrm{d}t,\mathrm{d}x)$$

if the integral on the right-hand side exists. In particular $f*\mathrm{m}$ is well defined if f is nonnegative. We define the integral of f with respect to μ *ω-wise* in an analogous way, because $\mu(\omega,\cdot)$ is a (nonnegative) measure on $(E,\mathscr{B}(E))$ for every $\omega\in\Omega$. If f is a nonnegative measurable function, then the integral $\int_E f(t,x)\,\mu(\omega,\mathrm{d}t,\mathrm{d}x)$ always exists. We shall use the notation $f*\mu$ for this random variable with values in $[0,+\infty]$. This definition extends to functions f of arbitrary sign. More precisely, for *any* measurable function f on $(E,\mathscr{B}(E))$, by Ω_f we denote the set of all $\omega\in\Omega$ such that $\int_E f(t,x)\,\mu(\omega,\mathrm{d}t,\mathrm{d}x)$ exists and is finite a.s. Obviously $\Omega_f\in\mathscr{F}$. We say that the integral of f with respect to μ exists and is finite a.s. if $\mathbb{P}[\Omega_f]=1$. In this case the random variable $f*\mu$ defined by

$$f*\mu(\omega):=\begin{cases}\int_E f(t,x)\,\mu(\omega,\mathrm{d}t,\mathrm{d}x), & \text{if } \omega\in\Omega_f;\\[2ex] 0, & \text{otherwise};\end{cases} \quad (5.7)$$

is called the *stochastic integral* of f with respect to the Poisson random measure μ. Note that the stochastic integral $f*\mu$ exists and is finite a.s. if and only if $|f|*\mu<+\infty$ a.s. We now state the so-called *exponential formula* (cf. Kallenberg (1997), Lemma 10.2).

5.3.1 Lemma (Exponential Formula)**.** *Let f be a function on $(E, \mathscr{B}(E))$. If $f \geq 0$, then*

$$\mathbb{E}[\mathrm{e}^{-f*\mu}] = \exp((\mathrm{e}^{-f} - 1) * \mathrm{m}). \tag{5.8}$$

Proof. First we take $f \geq 0$ of the form $f = u1_A$, where $u \geq 0$ and $A \in (E, \mathscr{B}(E))$ is such that $\mathrm{m}(A) < +\infty$. From Theorem 5.2.3 (i), we know that $\mu(A)$ is Poisson distributed with parameter $\mathrm{m}(A)$. This implies

$$\mathbb{E}[\mathrm{e}^{-f*\mu}] = \mathbb{E}[\mathrm{e}^{-u\mu(A)}] = \exp((\mathrm{e}^{-u} - 1)\mathrm{m}(A)) = \exp((\mathrm{e}^{-f} - 1) * \mathrm{m}).$$

Let now f be a simple function of the form $f = \sum_{j=1}^m c_j 1_{A_j}$ where $c_j \geq 0$ for every $j = 1, \ldots, m$ and $A_1, \ldots, A_m \in (E, \mathscr{B}(E))$ are *pairwise disjoint* sets such that $\mathrm{m}(A_j) < +\infty$, for every $j = 1, \ldots, m$. From Theorem 5.2.3 (ii), the random vector $(\mu(A_1), \ldots, \mu(A_m))$ is independent and so, from the previous step, we get

$$\begin{aligned}\mathbb{E}[\mathrm{e}^{-f*\mu}] &= \mathbb{E}\Big[\exp\Big(-\textstyle\sum_{j=1}^m c_j \mu(A_j)\Big)\Big] \\ &= \textstyle\prod_{j=1}^m \exp((\mathrm{e}^{-c_j} - 1)\mathrm{m}(A_j)) \\ &= \exp\Big(-\textstyle\sum_{j=1}^m (\mathrm{e}^{-c_j} - 1)\mathrm{m}(A_j)\Big) = \exp((\mathrm{e}^{-f} - 1) * \mathrm{m}).\end{aligned}$$

If f is an arbitrary nonnegative function, it can be approximated by an *increasing* sequence $(f_n)_{n\in\mathbb{N}}$ of simple functions and from the previous step formula (5.8) is satisfied for every $n \in \mathbb{N}$. We conclude the proof passing to the limit and applying the theorem of B. Levi on monotone convergence (cf. Theorem 1.A.4). □

Now we characterize, in terms of the intensity measure m, under which conditions the integral of a deterministic function f with respect to μ exists and is a.s. finite (cf. Kallenberg (1997), Lemma 10.2).

5.3.2 Proposition. *Let μ be a Poisson random measure on $(E, \mathscr{B}(E))$ with intensity measure* m. *Then $f * \mu$ exists and is finite a.s. if and only if $(|f| \wedge 1) * \mathrm{m} < +\infty$.*

Proof. If $f * \mu$ exists and is finite a.s., then, by definition, $|f| * \mu < +\infty$ a.s. Applying Lemma 5.3.1 to $|f|$ and using the estimate $(x \wedge 1) \leq c(1 - \mathrm{e}^{-x})$ for some constant $c > 0$ and $x > 0$, we deduce $(|f| \wedge 1) * \mathrm{m} < +\infty$. Indeed, because of $|f| * \mu < +\infty$, we have

$$0 < \mathbb{E}[\mathrm{e}^{-f*\mu}] = \exp((\mathrm{e}^{-f} - 1) * \mathrm{m}),$$

meaning $(1 - \mathrm{e}^{-f}) * \mathrm{m} < +\infty$. Hence $(|f| \wedge 1) * \mathrm{m} \leq c(1 - \mathrm{e}^{-|f|}) * \mu < +\infty$. Conversely, we now assume that $(|f| \wedge 1) * \mathrm{m} < +\infty$. We apply Lemma 5.3.1 to $c|f|$, where $c > 0$:

$$\mathbb{E}[\mathrm{e}^{-c|f|*\mu}] = \exp((\mathrm{e}^{-c|f|} - 1) * \mathrm{m}). \tag{5.9}$$

Letting c converge to zero on the left-hand side of (5.9) and using dominated convergence, we get $\mathbb{E}[\mathrm{e}^{-c|f|*\mu}] \longrightarrow \mathbb{P}[|f| * \mu < +\infty]$. For the right-hand side of (5.9) we observe that the inequality

$$|\mathrm{e}^{-c|f|} - 1| \leq c|f| \wedge 2$$

holds. Because of $c \downarrow 0$, we can assume $c \in (0, 2]$ and so we get $|\mathrm{e}^{-c|f|} - 1| \leq 2(|f| \wedge 1)$. We apply again the theorem of Lebesgue on dominated convergence to get $(|\mathrm{e}^{-c|f|} - 1|) * \mathrm{m} \longrightarrow 0$. In conclusion we have

$$\mathbb{P}[|f| * \mu < +\infty] = 1.$$

□

The following lemma can be shown in a similar way.

5.3.3 Lemma. *Let f be a function on $(E, \mathscr{B}(E))$. If $(|f| \wedge 1) * \mathrm{m} < +\infty$, then*

$$\mathbb{E}[\mathrm{e}^{-\mathrm{i}f*\mu}] = \exp((\mathrm{e}^{-\mathrm{i}f} - 1) * \mathrm{m}). \tag{5.10}$$

We now show how to compute the expectation of the random variable $f * \mu$, where f is a function which belongs to $L^1(\mathrm{m})$.

5.3.4 Lemma. *Let $f \in L^1(\mathrm{m})$. Then*

$$\mathbb{E}[f * \mu] = f * \mathrm{m}. \tag{5.11}$$

Moreover, the stochastic integral with respect to μ is a continuous operator on $L^1(\mathrm{m})$ into $L^1(\mathbb{P})$.

Proof. For every nonnegative function $f \in L^1(\mathrm{m})$ formula (5.11) holds. Indeed, this is true for indicator functions of the form 1_A, $A \in \mathscr{B}(E)$, $\mathrm{m}(A) < +\infty$, and hence for nonnegative simple functions f. For an arbitrary nonnegative function f we can find a sequence $(f_n)_{n\geq 1}$ of nonnegative simple functions such that $f_n \uparrow f$ pointwise as $n \to +\infty$. The result follows by monotone convergence. Clearly, formula (5.11) extends to functions f such that $|f| * \mathrm{m} < +\infty$. The statement on the continuity follows from

$$\mathbb{E}[|f * \mu|] \leq \mathbb{E}[|f| * \mu] = |f| * \mathrm{m} < +\infty.$$

□

5.4 Compensated Poisson random measures

Let μ be the jump measure of a càdlàg adapted process and a Poisson random measure relative to $\mathbb{F}$ with intensity measure $\mathrm{m} = \lambda_+ \otimes \rho$, where λ_+ is the Lebesgue measure on $(\mathbb{R}_+, \mathscr{B}(\mathbb{R}_+))$ and ρ a σ-finite measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. In this section we define the *compensated* Poisson random measure and, using the theory of *orthogonal measures* developed in Gihman & Skorohod (1974), we introduce the stochastic integral of deterministic functions with respect to a compensated Poisson random measure. First, we introduce the following ring of sets of E:

$$\mathscr{A} := \{A \in \mathscr{B}(E) : \mathrm{m}(A) < +\infty\}. \tag{5.12}$$

To each set $A \in \mathscr{A}$ we associate the random variable $\overline{\mu}(A)$ by

$$\overline{\mu}(A) := \mu(A) - \mathrm{m}(A), \quad A \in \mathscr{A}. \tag{5.13}$$

We remark that $\overline{\mu}$ is not defined on $\mathscr{B}(E)$ but only on $\mathscr{A}$. Indeed, on $\mathscr{B}(E)$ expressions of the type "$+\infty - \infty$" could appear. Notice that $\overline{\mu}$ defines a mapping on $\Omega \times \mathscr{A}$ into $(-\infty, +\infty]$ and that $\mathbb{E}[\overline{\mu}(A)] = 0$ for every $A \in \mathscr{A}$. Now we recall the notion of orthogonal measures (cf. Gihman & Skorohod (1974), IV§4).

5.4.1 Definition. Let $\mathscr{K}$ be a *semiring* of sets of E. We assume that to each $A \in \mathscr{K}$ there corresponds a real-valued random variable $\zeta(A)$ with the following properties.
(i) $\zeta(A) \in L^2(\mathbb{P})$ and $\zeta(\emptyset) = 0$.
(ii) $\zeta(A \cup B) = \zeta(A) + \zeta(B)$ a.s. for disjoint A and B in $\mathscr{K}$.
(iii) $\mathbb{E}[\zeta(A)\zeta(B)] = \alpha(A \cap B)$, where α is a set function on $\mathscr{K}$.
The family $\zeta := \{\zeta(A),\ A \in \mathscr{K}\}$ of random variables satisfying the previous three conditions is called an *orthogonal random measure* and $\alpha(\cdot)$ is called *structural function* (of ζ).

5.4.2 Lemma. *The family* $\overline{\mu} := \{\overline{\mu}(A),\ A \in \mathscr{A}\}$ *of random variables defined by* (5.13) *is an orthogonal random measure on the ring* $\mathscr{A}$ *with structural function* m.

Proof. It is clear that $\overline{\mu}(\emptyset) = 0$ and that $\overline{\mu}(A \cup B) = \overline{\mu}(A) + \overline{\mu}(B)$ for disjoint A and B in $\mathscr{A}$. Because of Theorem 5.2.3, the random variable $\mu(A)$ is Poisson distributed with parameter $\mathrm{m}(A)$, for every A in $\mathscr{A}$. Therefore, $\mathbb{E}[\overline{\mu}(A)^2]$ is the variance of a Poisson-distributed random variable with parameter $\mathrm{m}(A)$, i.e.,

$$\mathbb{E}[\overline{\mu}(A)^2] = \mathbb{E}[\mu(A)] = \mathrm{m}(A), \quad A \in \mathscr{A}. \tag{5.14}$$

Because of $\mathrm{m}(A) < +\infty$ and of (5.14), we obtain $\overline{\mu}(A) \in L^2(\mathbb{P})$, $A \in \mathscr{A}$. It remains to show that m is the structural function of $\overline{\mu}$. For this aim, we notice that the relation

$$\overline{\mu}(A)\overline{\mu}(B) = [\overline{\mu}(A \cap B) + \overline{\mu}(A \setminus B)] \times [\overline{\mu}(A \cap B) + \overline{\mu}(B \setminus A)], \quad A, B \in \mathscr{A},$$

holds. This and (5.14), together with the fact that $\overline{\mu}(C)$ and $\overline{\mu}(D)$ are independent if C and D are pairwise disjoint sets in $\mathscr{A}$ (cf. Theorem 5.2.3), yield

$$\begin{aligned}\mathbb{E}[\overline{\mu}(A)\overline{\mu}(B)] &= \mathbb{E}[\overline{\mu}(A \cap B)^2] + \mathbb{E}[\overline{\mu}(A \cap B)\overline{\mu}(A \setminus B)] \\ &\quad + \mathbb{E}[\overline{\mu}(A \cap B)\overline{\mu}(B \setminus A)] + \mathbb{E}[\overline{\mu}(A \setminus B)\overline{\mu}(B \setminus A)] \\ &= \mathbb{E}[\overline{\mu}(A \cap B)^2] = \mathrm{m}(A \cap B), \quad A, B \in \mathscr{A},\end{aligned}$$

and the proof is concluded. □

We call the orthogonal random measure $\overline{\mu}$ defined by (5.13) on the ring $\mathscr{A}$ the *compensated Poisson random measure* associated to the random measure μ or simply *compensated Poisson random measure.*

5.4.1 Stochastic integration for compensated Poisson random measures

We are going to define the stochastic integral with respect to $\overline{\mu}$ for functions in $L^2(\mathrm{m})$. We start defining the stochastic integral for functions belonging to the set $\mathscr{D} \subseteq L^2(\mathrm{m})$ of simple functions:

$$\mathscr{D} := \left\{ f = \sum_{k=1}^{m} a_k 1_{A_k}, \quad (a_k)_{k=1}^m \subseteq \mathbb{R}; \quad (A_k)_{k=1}^m \subseteq \mathscr{A} \text{ pairwise disjoint} \right\}. \tag{5.15}$$

For $A \in \mathscr{A}$, we define

$$1_A * \overline{\mu} = \overline{\mu}(A)\,.$$

Notice that $\mathbb{E}[(1_A * \overline{\mu})^2] = 1_A * \mathrm{m}$, that is the stochastic integral with respect to $\overline{\mu}$ is an isometry from $\{1_A,\ A \in \mathscr{A}\}$ into $L^2(\mathbb{P})$. Therefore, the integral with respect to $\overline{\mu}$ can be uniquely extended to a linear isometry on $\mathscr{D}$ into $L^2(\mathbb{P})$: If $f \in \mathscr{D}$, the *elementary stochastic integral* with respect to $\overline{\mu}$ is again denoted by $\overline{\mu}$ and

$$f * \overline{\mu} := \sum_{k=1}^{m} a_k\, \overline{\mu}(A_k), \quad f \in \mathscr{D}. \tag{5.16}$$

Moreover,

$$\mathbb{E}[(f * \overline{\mu})(g * \overline{\mu})] = (fg) * \mathrm{m}, \quad f, g \in \mathscr{D}. \tag{5.17}$$

Now we extend the definition of elementary stochastic integral to every function $f \in L^2(\mathrm{m})$ (cf. Gihman & Skorohod (1974), IV.§4 Theorem 1).

5.4.3 Lemma. *The system $\mathscr{D}$ of simple functions is dense in $L^2(\mathrm{m})$.*

Proof. We only need to verify that $\mathscr{D}$ satisfies the conditions of Lemma 1.2.3. The set of simple functions is clearly such that $\sigma(\mathscr{D}) = \mathscr{B}(E)$ (notice that m is a σ-finite measure on E). Because $\mathscr{A}$ is a ring, $\mathscr{D}$ is $\cap$-stable and so $\mathscr{D}$ is stable under multiplication. The measure m is σ-finite on E, hence there exists a sequence $(A_n)_{n\in\mathbb{N}} \subseteq \mathscr{A}$ such that $A_n \uparrow E$ as $n \to +\infty$. Therefore, we can construct a sequence $(h_n)_{n\in\mathbb{N}} \subseteq \mathscr{D}$ converging *pointwise* to 1. □

5.4.4 Theorem. *There exists a unique (continuous) isometric mapping on $L^2(\mathrm{m})$ into $L^2(\mathbb{P})$, again denoted by $\overline{\mu}$, such that*

$$1_A * \overline{\mu} = \overline{\mu}(A), \quad A \in \mathscr{A}. \tag{5.18}$$

Proof. For a function $f \in \mathscr{D}$, we have defined the elementary stochastic integral and we know that it is an isometric mapping on $\mathscr{D}$ into $L^2(\mathbb{P})$. Moreover, by definition, the elementary stochastic integral satisfies relation (5.18). The linear space $\mathscr{D}$ is dense in $L^2(\mathrm{m})$, hence the elementary stochastic integral with respect to $\overline{\mu}$, regarded as a mapping from $\mathscr{D}$ into $L^2(\mathbb{P})$, has a unique *isometric* extension on $L^2(\mathrm{m})$. We denote this extension again by $\overline{\mu}$. We need to show the uniqueness of $\overline{\mu}$ satisfying (5.18). But for this it is enough to observe that any isometric mapping between two Hilbert spaces is continuous. Relation (5.18) defines $\overline{\mu}$ uniquely on $\mathscr{D}$ and, by the density of $\mathscr{D}$ in $L^2(\mathrm{m})$ and by continuity, uniquely on $L^2(\mathrm{m})$. □

Let $\overline{\mu}$ be the unique isometric mapping of Theorem 5.4.4 and $f \in L^2(\mathrm{m})$. We call $f * \overline{\mu}$ the *stochastic integral of f* with respect to the compensated Poisson random measure $\overline{\mu}$. We can extend the definition of the stochastic integral with respect to $\overline{\mu}$ also to functions in $L^1(\mathrm{m})$: Let $f \in L^1(\mathrm{m})$. We define the stochastic integral of f with respect to $\overline{\mu}$ by

$$f * \overline{\mu} := f * \mu - f * \mathrm{m}. \tag{5.19}$$

Of course, from Proposition 5.3.2, the right-hand side of (5.19) is well-defined and finite-valued a.s. It remains to show that it is consistent with the definition of the stochastic integral with respect to $\overline{\mu}$.

5.4.5 Proposition. *If $f \in L^1(\mathrm{m}) \cap L^2(\mathrm{m})$, then $f * \mu$ and $f * \overline{\mu}$ are both well-defined and*

$$f * \overline{\mu} = f * \mu - f * \mathrm{m} \quad \text{a.s.} \tag{5.20}$$

Proof. By Proposition 5.3.2 and Theorem 5.4.4, the stochastic integrals $f*\mu$ and $f*\overline{\mu}$ are both well-defined and finite-valued a.s. for any $f \in L^1(\mathrm{m}) \cap L^2(\mathrm{m})$. For proving (5.20), in a first step we assume that $f \in L^1(\mathrm{m}) \cap L^2(\mathrm{m})$ is such that $\mathrm{m}(\{f \neq 0\}) < +\infty$. Because $\mathscr{D} \subseteq L^2(\mathrm{m})$ is a dense set, there exists a sequence $(f_n)_{n\geq 1} \subseteq \mathscr{D}$ converging to f in $L^2(\mathrm{m})$. Replacing, if necessary, f_n with $f_n\, 1_{\{f\neq 0\}}$, without loss of generality we can assume that f_n vanishes outside of $\{f \neq 0\}$. This implies that $(f_n)_{n\geq 1}$ converges in $L^1(\mathrm{m})$ as well. Relation (5.20), being obviously true for every f_n, now extends to f by the $L^2(\mathrm{m})$-continuity of $\overline{\mu}$ and the $L^1(\mathrm{m})$-continuity of μ and m. (cf. Theorem 5.4.4 and Lemma 5.3.4). In case that $f \in L^1(\mathrm{m}) \cap L^2(\mathrm{m})$ is chosen arbitrarily, we define $f_n := f\, 1_{B_n}$ where $B_n \in \mathscr{A}$ is such that $B_n \uparrow E$ (such a sequence exists because m is a σ-finite measure on $(E, \mathscr{B}(E))$). Using the theorem of Lebesgue on dominated convergence, we observe that $(f_n)_{n\geq 1}$ converges to f in $L^1(\mathrm{m})$ and $L^2(\mathrm{m})$. Since $\mathrm{m}(\{f_n \neq 0\}) < +\infty$, we can apply the first step and obtain (5.20) for every f_n. Again by the continuity property of μ, $\overline{\mu}$ and m, we conclude that (5.20) remains valid for f. □

We conclude this section by giving necessary and sufficient conditions for a $\mathscr{B}(E)$-measurable function f to be *integrable* with respect to $\overline{\mu}$, i.e., to be such that $f * \overline{\mu}$ is well-defined and finite-valued a.s.

5.4.6 Theorem. *Let f be a measurable function on $(E, \mathscr{B}(E))$. The integral $f*\overline{\mu}$ exists and is a.s. finite if and only if $(f^2 \wedge |f|) * \mathrm{m} < +\infty$.*

Proof. We first assume that $(f^2 \wedge |f|) * \mathrm{m} < +\infty$. We have

$$f = f1_{\{|f|\leq 1\}} + f1_{\{|f|>1\}}$$

with $f1_{\{|f|\leq 1\}} \in L^2(\mathrm{m})$ and $f1_{\{|f|>1\}} \in L^1(\mathrm{m})$. By Theorem 5.4.4, $(f1_{\{|f|\leq 1\}}) * \overline{\mu}$ exists and is a.s. finite. By Proposition 5.3.2 and Proposition 5.4.5, $(f1_{\{|f|>1\}}) * \overline{\mu}$ exists, is a.s. finite and consistent. By linearity we put

$$f * \overline{\mu} := (f1_{\{|f|\leq 1\}}) * \overline{\mu} + (f1_{\{|f|>1\}}) * \overline{\mu}.$$

Hence, $f*\overline{\mu}$ exists, is a.s. finite and consistent. We do not verify the converse implication and refer to Kallenberg (1997), Theorem 10.15 for a complete proof. □

5.5 Construction of Lévy processes

In §5.3 and §5.4 we have introduced the stochastic integral of deterministic functions with respect to a Poisson random measure and the associated compensated Poisson random measure, respectively. Now we want to apply the developed theory to construct Lévy processes by integration of deterministic functions. We consider a càdlàg $\mathbb{F}$-adapted process X with jump measure μ. We assume that μ is a Poisson random measure *relative to the filtration* $\mathbb{F}$. The intensity measure of μ is $\mathrm{m} = \lambda_+ \otimes \nu$, where λ_+ is the Lebesgue measure on $([0,T], \mathscr{B}([0,T]))$, while ν is a σ-finite measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. We stress that this assumption implies that the process X has no fixed time discontinuities a.s. Furthermore, we require that ν is a Lévy measure, i.e., $\nu(\{0\}) = 0$ and $(x^2 \wedge 1) \in L^1(\nu)$. If $A \in E$ is of the form $A = [0,1] \times B$, then

$$\mathbb{E}[\mu(A)] = \nu(B)$$

that is the Lévy measure describes the mean of the jumps of X with size in B over a time interval $[0,1]$. Clearly

$$\mathbb{E}[\mu([0,t] \times B)] = t\nu(B), \quad t \in [0,T],$$

meaning that the expected values of the number of jumps in B does not depend on time t. If $B = \{|x| > 1\}$, we know that X will only have finitely many jumps in B (cf. Theorem 2.2.6) and, because of the integrability conditions on the Lévy measure, we also get $\nu(B) < +\infty$. Notice that also the stronger integrability condition $|x| \wedge 1 \in L^1(\nu)$ yields $\nu(B) < +\infty$. However, this condition also yields $\nu(\{|x| \leq 1\}) < +\infty$, which does not correspond to the property of a càdlàg process X: From Theorem 2.2.6 we only know that the "big jumps" of X are finitely many over compact time interval. However, small jumps could be countably many. This is related with the fact that for any càdlàg process the sum $\sum_{0\leq s\leq t}(\Delta X_s)^2$ is always finite, while the sum $\sum_{0\leq s\leq t}|\Delta X_s|$ is finite if and only if X is a càdlàg process of *finite variation*. In conclusion, one can see that, if X is a càdlàg process with jump measure μ which is a Poisson random measure of the form $\lambda_+ \otimes \rho$, then ρ is necessarily a Lévy measure.

Notice that the function h defined by $h(t,x) := 1_{[0,t]}f(x)$ belongs to $L^q(\lambda_+ \otimes \nu)$ if and only if the function f belongs to $L^q(\nu)$, $q \geq 1$. We recall that, because $\mathbb{F}$ satisfies the usual conditions, every adapted Lévy process in law relative to $\mathbb{F}$ has an adapted càdlàg modification which is a Lévy process relative to $\mathbb{F}$ (cf. Theorem 2.5.12). In the sequel we do not distinguish a Lévy process in law from such a càdlàg modification.

We introduce the system of simple functions in $L^q(\nu)$, $q \geq 1$, by

$$\mathscr{D} := \left\{ f = \textstyle\sum_{j=1}^m a_j\, 1_{C_j},\ a_j \in \mathbb{R}; \quad C_j \in \mathscr{B}(\mathbb{R}) \text{ p.d.}, \quad \nu(C_j) < +\infty \right\}, \tag{5.21}$$

where the acronym *p.d.* stands for *pairwise disjoint*.

5.5.1 Lemma. *Let X be a càdlàg adapted process with jump measure μ. If μ is a Poisson random measure relative to the filtration $\mathbb{F}$ with intensity measure $\lambda_+ \otimes \nu$, then for every $f \in \mathscr{D}$ the process $(1_{[0,\cdot]}f) * \mu = ((1_{[0,t]}f) * \mu)_{t\geq 0}$ is a Lévy process relative to $\mathbb{F}$.*

Proof. Let $f \in \mathscr{D}$ have the representation $f = \sum_{j=1}^m a_j\, 1_{C_j}$. We put $A_j := [0,T] \times C_j$. For every $t \geq 0$ we have $\mu(A_j \cap [0,t] \times \mathbb{R}) < +\infty$ a.s. because $\mathrm{m}(A_j \cap [0,t] \times \mathbb{R}) = \mathrm{m}([0,t] \times C_j) = t\nu(C_j) < +\infty$ and the identity $\mu(A_j \cap [0,t] \times \mathbb{R}) = \mu([0,t] \times C_j)$, $t \geq 0$, holds. Lemma 5.2.2 ensures that the process $N^{A_j} = (\mu([0,t] \times C_j))_{t\geq 0}$ is a Poisson process relative to $\mathbb{F}$. Therefore the process $\mu(1_{[0,\cdot]}f)$ is càdlàg and adapted because

$$\mu(1_{[0,t]}f) = \sum_{j=1}^m a_j \mu([0,t] \times C_j).$$

Furthermore, $\mu(1_{\{0\}}f) = 0$. For $0 \leq s \leq t$, the function $1_{(s,t]}f$ belongs to $L^1(\lambda_+ \otimes \nu)$ and we can apply Lemma 5.3.3 to get

$$\mathbb{E}\left[\exp\bigl(\mathrm{i}u((1_{[0,t]}f) * \mu - (1_{[0,s]}f) * \mu)\bigr)\right] = \mathbb{E}\left[\exp\bigl(\mathrm{i}u(1_{(s,t]}f) * \mu\bigr)\right] = \exp((t-s)(\mathrm{e}^{\mathrm{i}uf}-1)*\nu).$$

This means, in particular, that $(1_{[0,\cdot]}f)*\mu$ has homogeneous one-dimensional increments. We show the $\mathbb{F}$-independence of the increments. Obviously,

$$(1_{[0,t]}f) * \mu - (1_{[0,s]}f) * \mu = (1_{(s,t]}f) * \mu = \sum_{j=1}^{m} a_j \mu((s,t] \times C_j).$$

The sets $(s,t] \times C_1, \dots, (s,t] \times C_m$ are pairwise disjoint because $C_1, \dots, C_m$ are. Furthermore, $(s,t] \times C_j \subseteq (s,T] \times \mathbb{R}$ for every $j = 1, \dots, m$. Hence, the vector $(\mu((s,t] \times B_1), \dots, \mu((s,t] \times B_m))$ is an independent random vector independent of $\mathscr{F}_s$ (cf. Theorem 5.2.3) and so $(1_{(s,t]}f) * \mu$ is independent of $\mathscr{F}_s$. We have that $(1_{[0,\cdot]}f) * \mu$ starts at zero, is càdlàg and has homogeneous increments. This is sufficient to assert that it is a stochastically continuous process. Therefore $(1_{[0,\cdot]}f) * \mu$ is a Lévy process relative to $\mathbb{F}$ for every $f \in \mathscr{D}$. □

As a consequence of Lemma 1.2.3, the system $\mathscr{D}$ introduced in (5.21) is total in $L^q(\nu)$, $q \geq 1$. This allows to extend Lemma 5.5.1.

5.5.2 Proposition. *Let X be a càdlàg adapted process with jump measure μ. If μ is a Poisson random measure relative to the filtration $\mathbb{F}$ with intensity measure $\lambda_+ \otimes \nu$, then for every f such that $|f| \wedge 1 \in L^1(\nu)$ the process $(1_{[0,\cdot]}f) * \mu = ((1_{[0,t]}f) * \mu)_{t \geq 0}$ is a Lévy process relative to $\mathbb{F}$.*

Proof. First we assume $f \in L^1(\nu)$. The system $\mathscr{D}$ is dense in $L^1(\nu)$ and so there exists a sequence $(f_n)_{n\in\mathbb{N}} \subseteq \mathscr{D}$ such that $f_n \longrightarrow f$ in $L^1(\nu)$ as $n \to +\infty$. By Lemma 5.5.1, the sequence $(1_{[0,\cdot]}f_n) * \mu$ is a sequence of Lévy processes relative to $\mathbb{F}$. From the linearity of the stochastic integral and Lemma 5.3.4 we get

$$\begin{aligned}0 \leq \mathbb{E}[|(1_{[0,t]}f_n) * \mu - (1_{[0,t]}f) * \mu|] &\leq \mathbb{E}[(1_{[0,t]}|f_n - f|) * \mu] \\ &= t(|f_n - f|) * \nu \longrightarrow 0, \quad n \to +\infty.\end{aligned}$$

So we can conclude that $(1_{[0,t]}f_n) * \mu$ converges to $(1_{[0,t]}f) * \mu$ in $L^1(\mathbb{P})$. Because the filtration $\mathbb{F}$ satisfies the usual conditions, this implies that the process $(1_{[0,\cdot]}f) * \mu$ is $\mathbb{F}$-adapted. Furthermore, the $L^1(\mathbb{P})$-convergence of $(1_{[0,t]}f_n) * \mu$ to $(1_{[0,t]}f) * \mu$ implies convergence in probability. By Lemma 2.5.7, $(1_{[0,\cdot]}f) * \mu$ is a process with $\mathbb{F}$-independent and homogeneous increments. Obviously, $(1_{[0,t]}f) * \mu \longrightarrow 0$ a.s. as $t \downarrow 0$ and from the homogeneity of the increments, we can assert that $(1_{[0,\cdot]}f)*\mu$ is stochastically continuous (cf. Lemma 2.5.10). In conclusion, $((1_{[0,\cdot]}f) * \mu, \mathbb{F})$ is a Lévy process in law. Because the filtration $\mathbb{F}$ satisfies the usual conditions, we can find a version of $(1_{[0,\cdot]}f) * \mu$ which is in fact a Lévy process relative to $\mathbb{F}$, i.e., also càdlàg. We do not distinguish these two processes and denote the càdlàg version again by $((1_{[0,\cdot]}f) * \mu, \mathbb{F})$. Hence, for every $f \in L^1(\mathbb{P})$, the process $((1_{[0,\cdot]}f) * \mu, \mathbb{F})$ is a Lévy process.

The remain of the proof is not part of the exam:

We now weaken the assumptions and consider f such that $|f| \wedge 1 \in L^1(\nu)$. Because of Proposition 5.3.2, the stochastic integral $\mu(1_{[0,t]}f)$ exists and is finite a.s. for every

$t \geq 0$. Denoting by f^+ and f^- the positive and the negative part of f, respectively, we have $f^{\pm} \leq |f|$ so that $f^{\pm} \wedge 1 \in L^1(\nu)$ and the stochastic integrals $\mu(1_{[0,t]}f^{\pm})$ are well defined and the relation $\mu(1_{[0,t]}f) = \mu(1_{[0,t]}f^+) - \mu(1_{[0,t]}f^-)$ holds a.s. Let us introduce the functions $f_n^{\pm} := f^{\pm}1_{\{|f^{\pm}|<n\}}$ and $f_n := f_n^+ - f_n^-$, $n \geq 1$. Then $f_n^{\pm} \geq 0$ and $f_n^{\pm} \uparrow f^{\pm}$ pointwise as $n \to +\infty$. Furthermore, $f_n^{\pm} \in L^1(\nu)$ because $f_n^{\pm} \leq (f^{\pm} \wedge n) \leq n(f^{\pm} \wedge 1) \in L^1(\nu)$. Therefore we also have $f_n \in L^1(\nu)$. Hence $\mu(1_{[0,t]}f_n^{\pm}) < +\infty$ a.s. $t \geq 0$, $n \geq 1$, and the theorem of B. Levi on monotone convergence implies that $\mu(1_{[0,t]}f_n^{\pm}) \uparrow \mu(1_{[0,t]}f^{\pm})$ a.s. as $n \to +\infty$, $t \geq 0$. Because of the previous step $((\mu(1_{[0,t]}f_n))_{t\geq 0}, \mathbb{F})$ is a Lévy process for every $n \geq 1$. Moreover,

$$\begin{aligned}\lim_{n\to+\infty} \mu(1_{[0,t]}f_n) &= \lim_{n\to+\infty} \mu(1_{[0,t]}f_n^+) - \lim_{n\to+\infty} \mu(1_{[0,t]}f_n^-)\\ &= \mu(1_{[0,t]}f^+) - \mu(1_{[0,t]}f^+)\\ &= \mu(1_{[0,t]}f), \quad t \geq 0.\end{aligned}$$

The previous convergence takes place a.s. and so $\mu(1_{[0,\cdot]}f)$ is $\mathbb{F}$-adapted (the filtration satisfies the usual conditions). Furthermore, Lemma 2.5.7 yields $\mu(1_{[0,\cdot]}f)$ has $\mathbb{F}$-independent and homogeneous increments. Clearly, $\mu(1_{[0,t]}f)$ converges to 0 a.s. as $t \downarrow 0$ and we know that this fact, together with the homogeneity of the increments implies that $\mu(1_{[0,\cdot]}f)$ is stochastically continuous. Therefore $(\mu(1_{[0,\cdot]}f), \mathbb{F})$ is a Lévy process in law. We do not distinguish such process from its càdlàg modification and so we assert that $(\mu(1_{[0,\cdot]}f), \mathbb{F})$ is a Lévy process. □

Let now $\overline{\mu}$ be the compensated Poisson random measure of L. From Theorem 5.4.4, every deterministic function in $L^2(\lambda_+ \otimes \nu)$ can be integrated with respect to $\overline{\mu}$. For any $f \in L^2(\nu)$ we introduce the process $X^{(f)} = (X_t^{(f)})_{t\geq 0}$ by

$$X_t^{(f)} = (1_{[0,t]}f) * \overline{\mu}, \quad t \geq 0. \tag{5.22}$$

5.5.3 Theorem. *Let X be a càdlàg adapted process with jump measure μ. If μ is a Poisson random measure relative to the filtration $\mathbb{F}$ with intensity measure $\lambda_+ \otimes \nu$ and $\overline{\mu}$ is the associated compensated Poisson random measure, then for every $f \in L^2(\nu)$ the process $X^{(f)}$ defined by* (5.22) *has the following properties:*

(i) $\mathbb{E}[(X_t^{(f)})^2] = t\,(f^2 * \nu)$ *and, in particular, the random variable $X_t^{(f)}$ is square integrable, $t \geq 0$;*

(ii) $(X^{(f)}, \mathbb{F})$ *is a Lévy process;*

(iii) $X^{(f)}$ *is a square integrable martingale;*

(iv) $\Delta X^{(f)} = f(\Delta X)1_{\{\Delta X \neq 0\}}$ *a.s.*

Proof. (i) is consequence of the isometry of the stochastic integral with respect to $\overline{\mu}$ for functions in $L^2(\nu)$. We show (ii) and (iii) together. If $f \in L^2(\nu)$, then the sequence $f_n := f1_{\{|f|>\frac{1}{n}\}}$ belongs to $L^1(\nu) \cap L^2(\nu)$ and converges to f in $L^2(\nu)$ as $n \to +\infty$. Indeed, $f_n^2 \leq f^2$ and

$$|f_n| \leq f^2 1_{\{|f|>1\}} + |f| 1_{\{\frac{1}{n}<|f|\leq 1\}} \leq f^2 1_{\{|f|>1\}} + n f^2 1_{\{\frac{1}{n}<|f|\leq 1\}} \leq (1+n)f^2 \in L^1(\nu)\,.$$

For every $n \geq 1$, Proposition 5.4.5 yields

$$X_t^{(f_n)} = (1_{[0,t]}f_n) * \mu - t(f_n * \nu) = \textstyle\sum_{0<s\leq t} f_n(\Delta X_s)1_{\{\Delta X_s \neq 0\}} - t(f_n * \nu), \quad t \geq 0. \tag{5.23}$$

From Proposition 5.5.2 we know that $(X^{(f_n)}, \mathbb{F})$ is a Lévy process for every $n \geq 1$. Moreover, from (5.23) and Lemma 5.3.4 it follows that $\mathbb{E}[X_t^{(f_n)}] = 0$, $t \geq 0$. Lemma 2.5.6 yields $X^{(f_n)}$ is an $\mathbb{F}$-martingale. We can apply the linearity of the stochastic integral with respect to $\overline{\mu}$ and its isometry property for functions in $L^2(\lambda_+ \otimes \nu)$ to obtain

$$\mathbb{E}[(X_t^{(f_n)} - X_t^{(f)})^2] = \mathbb{E}[(X_t^{(f_n - f)})^2] = t((f_n - f)^2 * \nu) \longrightarrow 0 \quad \text{as } n \to +\infty. \tag{5.24}$$

Therefore $X_t^{(f_n)}$ converges to $X_t^{(f)}$ in $L^2(\mathbb{P})$, hence in probability, as $n \to +\infty$. By Lemma 2.5.7 we obtain that $X^{(f)}$ has $\mathbb{F}$-independent and homogeneous increments. The process $X^{(f_n)}$ is $\mathbb{F}$-adapted because it is a Lévy process relative to $\mathbb{F}$, so (5.24) shows that also $X^{(f)}$ is $\mathbb{F}$-adapted, because the filtration $\mathbb{F}$ satisfies the usual conditions. The stochastic continuity of $X^{(f)}$ is clear. Indeed, because of the isometry of the stochastic integral with respect to $\overline{\mu}$ for functions in $L^2(\nu)$, for every $t \in [0, T]$ and $\epsilon > 0$, by Chebyschev inequality,

$$\mathbb{P}[|X_t^{(f)} - X_s^{(f)}| > \varepsilon] \leq \frac{1}{\varepsilon^2}\mathbb{E}[(X_t^{(f)} - X_s^{(f)})^2] = \frac{1}{\varepsilon^2}(t - s)(f^2 * \nu) \longrightarrow 0, \text{ as } s \to t.$$

Formula (5.24) implies that $X_t^{(f_n)}$ converges in $L^1(\mathbb{P})$ to $X_t^{(f)}$ and thanks to Lemma 2.3.11 we can conclude that $X^{(f)}$ is an $\mathbb{F}$-martingale. The process $(X^{(f)}, \mathbb{F})$ is a Lévy process in law and we consider a version of $X^{(f)}$ which is also càdlàg, i.e., which is in fact a Lévy process relative to $\mathbb{F}$. We denote this modification again by $X^{(f)}$: $(X^{(f)}, \mathbb{F})$ is a Lévy process and an $\mathbb{F}$-martingale. Because of (i), the martingale $X^{(f)}$ is square integrable and this concludes the proof of (ii) and (iii). By (5.23) we have

$$\Delta X_t^{(f_n)} = f_n(\Delta X_t) 1_{\{\Delta X_t \neq 0\}}, \quad t \geq 0, \quad \text{a.s.} \tag{5.25}$$

Because $X^{(f)}$ is càdlàg we can define the process $\Delta X^{(f)}$ and because

$$\lim_{n \to +\infty} \Delta X_t^{(f_n)} = \Delta X_t^{(f)}, \quad t \leq T, \quad \text{a.s.}$$

we can conclude that $\Delta X^{(f)} = f(\Delta X) 1_{\{\Delta X \neq 0\}}$ a.s. because $T > 0$ was chosen arbitrarily. The proof of the theorem is now complete. □

5.6 The jump measure of a Lévy processes

In this section we prove that the jump measure of a Lévy process $(X, \mathbb{F})$ is a Poisson random measure relative to the filtration $\mathbb{F}$.

Because the filtration $\mathbb{F}$ satisfies the usual conditions, we can assume that a Lévy process $(X, \mathbb{F})$ is a càdlàg process. For this reason, the jump measure μ of X, given on $(E, \mathscr{B}(E))$ (cf. (5.1)) by (5.3), i.e.,

$$\mu(\omega, A) := \sum_{s \geq 0} 1_{\{\Delta X(\omega)_s \neq 0\}} 1_A(s, \Delta X_s(\omega)), \quad A \in \mathscr{B}(E),$$

is an integer-valued random measure (cf. Proposition 5.1.3). In this section we show that the jump measure of a Lévy process $(X, \mathbb{F})$ is a Poisson random measure relative to the filtration $\mathbb{F}$ with intensity measure $\mathrm{m} := \lambda_+ \otimes \nu$, where λ_+ is the Lebesgue measure on $([0,T], \mathscr{B}([0,T]))$ and ν is a Lévy measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, i.e., ν is σ-finite, such that $\nu(\{0\}) = 0$ and $(x^2 \wedge 1) \in L^1(\nu)$. For any Borel subset B of $\mathbb{R}$ contained in $\{|x| > \varepsilon\}$, where $\varepsilon > 0$ is arbitrary but fixed, we introduce the process $\xi^B = (\xi^B_t)_{t \geq 0}$ by

$$\xi^B_t := \mu([0,t] \times B), \quad t \geq 0. \tag{5.26}$$

Notice that ξ^B can be obtained from (5.4) by choosing $A = [0,T] \times B$. Because of $B \subseteq \{|x| > \varepsilon\}$ and of the càdlàg property of the paths of X, we have that for every $t \geq 0$, $\mu([0,t] \times B) < +\infty$. Therefore, from Lemma 5.1.5, ξ^B is an $\mathbb{F}$-adapted simple point process. We are going to show that the process ξ^B is a homogeneous Poisson process relative to $\mathbb{F}$. Now the situation is different from the one of Lemma 5.2.2 because we only know that μ is the jump measure of a Lévy process and not (yet) that it is a Poisson random measure relative to $\mathbb{F}$.

With the Lévy process $(X, \mathbb{F})$ we associate the process ${}^tX$ by

$${}^tX_s := X_{t+s} - X_t, \quad s \geq 0. \tag{5.27}$$

Because of the Strong Markov Property (Theorem 2.5.15) the process ${}^tX$ is independent of $\mathscr{F}_t$. If we can show that the increment $\xi^B_{t+s} - \xi^B_t$ is a measurable time-homogeneous functional of ${}^tX$, then we can conclude that $\xi^B_{t+s} - \xi^B_t$ is independent of $\mathscr{F}_t$ and has the same distribution of ξ^B_s, that is $(\xi^B, \mathbb{F})$ has homogeneous and independent increments. However, we have to explain what we mean with "measurable functional".

In general, if Y is a random variable taking values in some general measurable space $(Z, \mathscr{Z})$ and it is independent of a σ-algebra $\mathscr{G}$, then $F(Y)$ is also independent of $\mathscr{G}$, if $F : Z \longrightarrow \mathbb{R}$ is a $(\mathscr{Z}, \mathscr{B}(\mathbb{R}))$-measurable functional. To apply this result to our case we have to understand a Lévy process X as a random variable in some measurable space.

We observe that for every $\omega \in \Omega$, the mapping $\omega \to \{X_t(\omega)\}_{t \in [0,T]}$ is a càdlàg function. This suggest that we can look regard a càdlàg process X as a random variable taking value in the space of the paths. More precisely, we introduce the set

$$D := \{z : [0,T] \longmapsto \mathbb{R} \text{ such that the mapping } t \mapsto z(t) \text{ is càdlàg}\} \tag{5.28}$$

of càdlàg functions on $[0,T]$ into $\mathbb{R}$ is called the *Skorohod space* over $\mathbb{R}$. We consider the application $Z_t : D \longmapsto \mathbb{R}$ such that $Z_t(z) := z(t)$, called coordinate projection. The σ-algebra $\mathscr{F}_D := \sigma(Z_t,\ t \geq 0)$ is called the *Skorohod σ-algebra* on D.

One can see that $\omega \to \{X_t(\omega)\}_{t \in [0,T]}$ is $(\mathscr{F}, \mathscr{F}_D)$-measurable and therefore it is possible to understand X as a random variable taking values on the Skorohod space $(D, \mathscr{F}_D)$.

Notice that, if $\mathbb{P}$ is a probability measure on $(\Omega, \mathscr{F})$, then we can consider the law of the process X (as a random variable with values in the Skorohod space) $\mathbb{P}_X$ by setting

$$\mathbb{P}_X[A] := \mathbb{P}[\omega \in \Omega : X(\omega) \in A], \quad A \in \mathscr{F}_D.$$

Notice that, if for example X is a càdlàg Gaussian process and A is a cylindrical set, then $\mathbb{P}_X[A]$ is a more-dimensional normal distribution.

In the same spirit we can consider a Brownian motion as a random variable taking values in the measurable space $C([0,T])$ of continuous functions endowed with Skorohod σ-algebra generated by continuous functions.

We now show that the increment $\xi^B_{t+s} - \xi^B_t$ can be represented as a time-homogeneous $(\mathscr{F}_D, \mathscr{B}(\mathbb{R}))$-measurable functional of the process ${}^tX$. From this it follows, in particular, that ξ^B has $\mathbb{F}$-independent and homogeneous increments.

5.6.1 Lemma. *Let $B \in \mathscr{B}(\mathbb{R})$ be such that $B \subseteq \{|x| > \varepsilon\}$, where $\varepsilon > 0$ is arbitrary but fixed. We consider the process ξ^B introduced by* (5.26)*. For any $s \geq 0$ there exists an $(\mathscr{F}^{\mathbb{Q}}_D, \mathscr{B}(\mathbb{R}))$-measurable functional $F_{s,B}$ such that*

$$\xi^B_{t+s} - \xi^B_t = F_{s,B}({}^tX) \quad \text{a.s.,} \quad t \geq 0.$$

Moreover, there exists a constant, say $\nu^B \geq 0$, such that

$$a^B(t) := \mathbb{E}[\xi^B_t] = t\nu^B. \tag{5.29}$$

In particular, $(\xi^B, \mathbb{F})$ is a homogeneous Poisson process.

Proof. We define the functional $G_{s,B}$ by

$$G_{s,B}(z) := \sum_{0<u\leq s} 1_{\{\Delta Z_u(z)\neq 0\}} 1_B(\Delta Z_u(z)), \quad z \in D,$$

and the sequence $(\tau_k)_{k\geq 0}$ by

$$\tau_0 := 0, \qquad \tau_{k+1}(z) := \inf\{t > \tau_k(z) : \Delta Z_t(z) \in B\}, \quad k \geq 1.$$

Let $\mathbb{Q}$ be a probability measure on $(D, \mathscr{F}_D)$ and $\mathscr{F}^{\mathbb{Q}}_D$ the $\mathbb{Q}$-completion of $\mathscr{F}_D$. Let $\mathbb{G} = (\mathscr{G}_t)_{t\in[0,T]}$ be the right-continuous and complete filtration defined by $\mathscr{G}_t := \mathscr{F}^{\mathbb{Q}}_D$, $t \in [0,T]$. Then $(Z_t)_{t\in[0,T]}$ is a càdlàg $\mathbb{G}$-adapted process. Therefore from Theorem 2.3.5 it is progressively measurable. Since $\mathbb{G}$ satisfies the usual conditions, we can apply Theorem 2.3.8 (iii) to deduce that τ_k are stopping times with respect to $\mathbb{G}$. In particular τ_k is $(\mathscr{F}^{\mathbb{Q}}_D, \mathscr{B}(\mathbb{R}_+))$-measurable, because they are random variables on $(D, \mathscr{F}^{\mathbb{Q}}_D, \mathbb{Q})$ taking values in $[0, +\infty]$.

The functional $G_{s,B}$ is integer valued and because

$$\{z : G_{s,B}(z) = k\} = \{z : \tau_k(z) \leq t < \tau_{k+1}(z)\}$$

and $\{\tau_k \leq t < \tau_{k+1}\}$ is $\mathscr{F}^{\mathbb{Q}}_D$-measurable, the functional $G_{s,B}$ is $(\mathscr{F}^{\mathbb{Q}}_D, \mathscr{B}(\mathbb{R}))$-measurable, for every probability measure $\mathbb{Q}$ on $(D, \mathscr{F}_D)$. Furthermore,

$$\begin{aligned} \xi^B_{t+s} - \xi^B_t &= \textstyle\sum_{t<u\leq t+s} 1_{\{\Delta X_u \neq 0\}} 1_B(\Delta X_u) \\ &= \textstyle\sum_{0<u\leq s} 1_{\{\Delta^t X_u \neq 0\}} 1_B(\Delta^t X_u) = G_{s,B}({}^tX), \quad t, s \geq 0. \end{aligned}$$

If we now choose $\mathbb{Q} := \mathbb{P}_L = \mathbb{P}_{{}^tL}$ we find $F_{s,B}$ on $(D, \mathscr{F}_D)$ such that $F_{s,B} = G_{s,B}$ $\mathbb{P}_L$-a.s. (where $\mathbb{P}_L$ denotes the law of L) and because under $\mathbb{P}_L$ $F_{s,B}({}^tX)$ is distributed as $F_{s,B}(X) = \xi^B_s - \xi^B_0$, we can conclude that $\xi^B_{t+s} - \xi^B_t$ is distributed as ξ^B_s. From Lemma 5.1.5, we know that ξ^B is a simple point process relative to $\mathbb{F}$. We can assert that ξ^B is a simple point process with $\mathbb{F}$-independent and homogeneous increments. Because ξ^B is càdlàg and starts at 0, the homogeneity of the increments implies that ξ^B is also stochastically continuous (Lemma 2.5.10), i.e., $(\xi^B, \mathbb{F})$ is a Lévy process with bounded jumps. Because of Theorem 2.5.17, the process ξ^B has a finite moment of every order. We can introduce the process $\overline{\xi}^B = (\overline{\xi}_t^B)_{t\geq 0}$ by $\overline{\xi}_t^B := \xi_t^B - \mathbb{E}[\xi_t^B]$. Because of Lemma 2.5.6, $\overline{\xi}^B$ is an $\mathbb{F}$-martingale square integrable martingale. By Doob's inequality, we get that $\sup_{t\in[0,T]} |\overline{\xi}_t|$ belongs to $L^2(\mathbb{P})$, hence to $L^1(\mathbb{P})$. Therefore

$$\sup_{t\in[0,T]} |\xi_t^B|^2 \leq \sup_{t\in[0,T]} |\overline{\xi}_t|^2 + \mathbb{E}[(\xi_T^B)^2].$$

Therefore, by monotone convergence, because ξ_s^B converges to ξ_t^B in probability (convergence in probability is here enough, cf. Theorem 1.1.3), we can conclude that the function $t \mapsto \mathbb{E}[\xi_t^B]$ is continuous. Let $a(t) := \mathbb{E}[\xi_t^B]$. Then because of the homogeneity of the increments we get

$$a^B(t+s) = \mathbb{E}[\xi_{t+s}^B - \xi_t^B] + \mathbb{E}[\xi_t^B] = a^B(s) + a^B(t), \quad t, s \geq 0,$$

which is the Cauchy functional equality. This relation together with the continuity, implies that the function $a^B(\cdot)$ is linear and therefore that there exists a $\nu^B \geq 0$ such that $a^B(t) = t\nu^B$. In conclusion ξ^B is a stochastically continuous simple point process with $\mathbb{F}$-independent and homogeneous increments such that $\mathbb{E}[\xi_t^B] = t\nu^B$, $t \geq 0$, $\nu^B \geq 0$, i.e., $(\xi^B, \mathbb{F})$ is a homogeneous Poisson process (cf. Definition 4.2.2) and the proof of the lemma is complete. □

Thanks to Lemma 5.6.1, we can compute the explicit form of the intensity measure of the random measure μ. For any $C \in \mathscr{B}(\mathbb{R})$, we put

$$\nu(C) := \mathbb{E}[\mu([0,1] \times C)]. \tag{5.30}$$

Clearly, (5.30) defines a measure on the space $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. From the definition of μ, we have that $\nu(\{0\}) = 0$. Moreover, $\nu(\{|x| > \varepsilon\}) < +\infty$, for every $\varepsilon > 0$. Indeed, because of Lemma 5.6.1, the process $(\xi^B, \mathbb{F})$, where $B := \{|x| > \varepsilon\}$, is a homogeneous Poisson process and therefore

$$\nu(B) = \mathbb{E}[\mu([0,1] \times B)] = \mathbb{E}[\xi_1^B] = \nu^B < +\infty,$$

where the constant $\nu^B \geq 0$ was introduced in (5.29). This implies that ν is a σ-finite measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ because the sequence $(B_n)_{n\geq 1}$ defined by $B_n := \{|x| > \frac{1}{n}\}$, $n \geq 1$, is such that $B_n \uparrow \mathbb{R} \setminus \{0\}$ as $n \to +\infty$ and $\nu(\mathbb{R}) = \nu(\mathbb{R} \setminus \{0\})$. It can be proven that the function $(x^2 \wedge 1)$ is integrable with respect to ν. We do not verify this property and we refer to, e.g., Kallenberg (1997), Theorem 13.4. In conclusion, ν is a *Lévy measure.* We call ν *the Lévy measure of the process* X.

5.6.2 Lemma. *Let μ be the jump measure of the Lévy process $(X, \mathbb{F})$ with intensity measure* m*. Then* $\mathrm{m} = \lambda_+ \otimes \nu$*, where λ_+ is the Lebesgue measure on* $([0,T], \mathscr{B}([0,T]))$ *and ν is the Lévy measure of X.*

Proof. Because of Lemma 5.6.1, we have that $\mathrm{m}(A) = (\lambda_+ \otimes \nu)(A)$ if $A \in E$ is such that $A = [0,u] \times \{|x| > \varepsilon\}$, $\varepsilon > 0$, $u \geq 0$. To obtain that $\mathrm{m} = \lambda_+ \otimes \nu$ is now a standard procedure with the help of the uniqueness theorem for measures (cf. Theorem 1.A.2). We introduce the system of sets

$$\begin{aligned}\mathscr{C} := \{\{0\} \times C,\, C \in \mathscr{B}(\mathbb{R}) : \nu(C) < +\infty\} \cup \\ \cup \{(r,v] \times C : 0 \leq r \leq v,\ C \in \mathscr{B}(\mathbb{R}) : \nu(C) < +\infty\}\end{aligned}$$

which generates $\mathscr{B}(E)$ and is stable under intersections. If $A = (r, v] \times B$ with $0 < r < v$ and $B \in \mathscr{B}(\mathbb{R})$ is such that $B \subseteq \{|x| > \varepsilon\}$, $\varepsilon > 0$, then $A \in \mathscr{C}$ and $\mu(A) = \mu([0, v] \times B) - \mu([0, r] \times B) = \xi_v^B - \xi_r^B$. From the previous step, this yields $\mathrm{m}(A) = \mathrm{m}([0, v] \times B) - \mathrm{m}([0, r] \times B) = (v - r)\nu(B) = (\lambda_+ \otimes \nu)(A)$. The sets $A_n := [0, T] \times \left(\{|x| > \frac{1}{n}\} \cup \{0\}\right)$, $n \geq 1$, belong to $\mathscr{C}$ and are such that $\cup_{n=1}^{\infty} A_n = E$. Hence $(A_n)_{n \geq 1}$ is a sequence of sets of finite measure with respect to both m and $\lambda_+ \otimes \nu$. An application of the uniqueness theorem for measures (cf. Theorem 1.A.2) shows that $\mathrm{m} = \lambda_+ \otimes \nu$ on $\mathscr{B}(E)$. □

Now we show that, for every $A \in \mathscr{B}(E)$ such that $\mathrm{m}(A) < +\infty$, the process $(N^A, \mathbb{F})$ defined by (5.4) is a Poisson process.

5.6.3 Proposition. *Let μ be the jump measure of a Lévy process $(X, \mathbb{F})$ with intensity measure $\mathrm{m} = \lambda_+ \otimes \nu$ and let $A \in \mathscr{B}(E)$ be such that $(\lambda_+ \otimes \nu)(A) < +\infty$. Then the process $N^A := (N_t^A)_{t \geq 0}$ defined by $N_t^A := \mu(A \cap [0, t] \times \mathbb{R})$, $t \geq 0$, is a Poisson process relative to $\mathbb{F}$ and $a^A(\cdot) := (\lambda_+ \otimes \nu)(A \cap [0, \cdot] \times \mathbb{R})$ is its intensity function.*

Proof. We know that N^A is a simple point process for every $A \in \mathscr{B}(E)$ be such that $(\lambda_+ \otimes \nu)(A) < +\infty$. For the claim of the proposition it is sufficient to show that $(N^A - a^A(\cdot), \mathbb{F}^L)$ is a local martingale and then to apply Theorem 4.2.3. First we consider the case $A = (r, v] \times B$ with $B \subseteq \{|x| > \varepsilon\}$, $\varepsilon > 0$. Then, for every $t \geq 0$, we have

$$N_t^A - (\lambda_+ \otimes \nu)(A \cap [0, t] \times \mathbb{R}) = (\xi_t^B - t\nu(B))^v - (\xi_t^B - t\nu(B))^r,$$

where the superscripts v and r denote the stopping operation at the deterministic times v and r, respectively. By Lemma 5.6.1, $(\xi^B, \mathbb{F})$ is a Poisson process with intensity $a(t) = t\nu(B)$. Therefore, by Doob's Stopping Theorem (cf. Theorem 2.3.13) $N_t^A - (\lambda_+ \otimes \nu)(A \cap [0, t] \times \mathbb{R})$ is a martingale as a difference of martingales.

Now we consider $\varepsilon > 0$ arbitrarily fixed and let $\mathscr{A}$ be the algebra of measurable subsets A of $(0, T] \times \{|x| > \varepsilon\}$ of the form $A = \bigcup_{j=1}^n A_j$ with pairwise disjoint rectangles $A_j := (r_j, v_j] \times C_j$, $0 \leq r_j \leq v_j \leq T$, $C_j \in \mathscr{B}(\mathbb{R})$ and $C_j \subseteq \{|x| > \varepsilon\}$. We notice that $\mathscr{A}$ generates the σ-field $\mathscr{B}((0, T]) \otimes \mathscr{B}(\{|x| > \varepsilon\})$. For any $A \in \mathscr{A}$ it follows

$$N^A - (\lambda_+ \otimes \nu)(A \cap [0, \cdot] \times \mathbb{R}) = \sum_{j=1}^n \left(N^{A_j} - (\lambda_+ \otimes \nu)(A_j \cap [0, \cdot] \times \mathbb{R})\right)$$

and hence $(N^A - (\lambda_+ \otimes \nu)(A \cap [0, \cdot] \times \mathbb{R}), \mathbb{F}^L)$ is a martingale by the previous step. The class $\mathscr{C} \subseteq \mathscr{B}(E)$ of all $A \subseteq (0, T] \times \{|x| > \varepsilon\}$ such that the process $(N^A - (\lambda_+ \otimes \nu)(A \cap [0, \cdot] \times \mathbb{R}), \mathbb{F}^L)$ is a martingale is a monotone class of subsets of $(0, T] \times \{|x| > \varepsilon\}$ (cf. §1.2). Indeed, if $A_n \in \mathscr{C}$ monotonically converges to a subset A of $(0, T] \times \{|x| > \varepsilon\}$ then it is easy to see that $N_t^{A_n} - (\lambda_+ \otimes \nu)(A_n \cap [0, t] \times \mathbb{R})$ converges to $N_t^A - (\lambda_+ \otimes \nu)(A \cap [0, t] \times \mathbb{R})$ in $L^1(\mathbb{Q})$, $t \geq 0$, and hence the process $(N^A - (\lambda_+ \otimes \nu)(A \cap [0, \cdot] \times \mathbb{R}), \mathbb{F}^L)$ is again a martingale (cf. Lemma 2.3.11). Because of the above-stated, we have $\mathscr{A} \subseteq \mathscr{C}$ and by the monotone class theorem for sets (cf. Theorem 1.2.1), $\mathscr{C} = \mathscr{B}((0, T]) \otimes \mathscr{B}(\{|x| > \varepsilon\})$. Let now A be an arbitrary Borel subset of $[0, T] \times \mathbb{R}$ such that $(\lambda_+ \otimes \nu)(A) < +\infty$. The sequence $(A_n)_{n \geq 1}$ defined by $A_n := A \cap (0, T] \times \{|x| > \frac{1}{n}\}$, $n \geq 1$, converges to $A \setminus (0, +\infty) \times \{0\}$ increasingly. From this it follows that $N_t^{A_n} - (\lambda_+ \otimes \nu)(A_n \cap [0, t] \times \mathbb{R})$ converges to $N_t^A - (\lambda_+ \otimes \nu)(A \cap [0, t] \times \mathbb{R})$ in $L^1(\mathbb{P})$ for every $t \geq 0$. From the previous step we know that $(N^{A_n} - (\lambda_+ \otimes \nu)(A_n \cap [0, \cdot] \times \mathbb{R}), \mathbb{F}^X)$ are martingales and hence

$(N^A - (\lambda_+ \otimes \nu)(A \cap [0, \cdot] \times \mathbb{R}), \mathbb{F}^X)$ is a martingale, too. This completes the proof of the proposition. □

It is now immediate to see that if μ is the jump measure of a Lévy process $(X, \mathbb{F})$, then it is a Poisson random measure relative to the filtration $\mathbb{F}$. Indeed, if we fix $s > 0$ and we consider a Borel subset A of $(s, T] \times \mathbb{R}$ such that $\mathrm{m}(A) < +\infty$, because of Proposition 5.6.3, the process $(N^A, \mathbb{F})$ is a Poisson process. Notice that $N_s^A = 0$. Therefore

$$\mu(A) = N_T^A = N_T^A - N_s^A,$$

which is independent of $\mathscr{F}_s$. Lemma 5.6.2 implies that the intensity measure m of μ is equal to $\lambda_+ \otimes \nu$. In conclusion we have shown the following result:

5.6.4 Theorem. *Let $(X, \mathbb{F})$ be a Lévy process and let μ be its jump measure. Then μ is a Poisson random measure relative to the filtration $\mathbb{F}$ and its intensity measure is given by $\mathrm{m} = \lambda_+ \otimes \nu$, where λ_+ is the Lebesgue measure on $([0, T], \mathscr{B}([0, T]))$ and ν is the Lévy measure of the process X.*

5.7 Lévy–Itô and Lévy–Khintchine decomposition

In this section we are going to establish two important decompositions which are valid for Lévy processes: Lévy–Itô Decomposition and Lévy–Khintchine Decomposition. The first one give as informations about the structure of a Lévy process and more precisely it states that each Lévy process $(X, \mathbb{F})$ can be decomposed as the sum of a deterministic drift, a Brownian motion and the sum of its jumps which can be decomposed as big jumps and small jumps. These terms in the decomposition are independent. Lévy–Khintchine decomposition, which thanks to the independence of the therms of the Lévy–Itô decomposition can be easily obtained by the latter one, gives informations about the distribution of a Lévy process $(X, \mathbb{F})$, that is, it explicitely gives the form of the characteristic function of X_t, for every $t \in [0, T]$.

In §5.5, we have shown that if X is a càdlàg adapted process whose jump measure μ is a Poisson random measure with respect to the filtration $\mathbb{F}$ with intensity measure $\lambda_+ \otimes \nu$, where λ_+ is the Lebesgue measure on $([0, T], \mathscr{B}([0, T]))$ and ν a Lévy measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, then for every f such that $|f| \wedge 1 \in L^1(\nu)$, the process $(\mu(1_{[0,\cdot]} f), \mathbb{F})$ is a Lévy process.

We now assume that $(X, \mathbb{F})$ is a Lévy process. Therefore, from Theorem 5.6.4, the jump measure μ of X is a Poisson random measure relative to the filtration $\mathbb{F}$ with intensity measure $\lambda_+ \otimes \nu$, where ν is the Lévy measure of the process X.

To prove the Lévy–Itô decomposition, we need to show that the increments of the process $\mu(1_{[0,\cdot]} f)$, where f is such that $|f| \wedge 1 \in L^1(\nu)$, can be represented as a time-homogeneous $(\mathscr{F}_D, \mathscr{B}(\mathbb{R}))$-measurable functional of the process ${}^tX$. In other words, our aim is to extend Lemma 5.6.1 to any deterministic function f such that $|f| \wedge 1 \in L^1(\nu)$.

We recall that ${}^tX$ we denotes the process defined by

$${}^tX_s := X_{t+s} - X_t, \quad s, t \geq 0 \tag{5.31}$$

while $(D, \mathscr{F}_D)$ denotes the Skorohod space (cf. §sec:gen.lev.pr). By $\mathscr{D}$ we disign as usual the system of simple functions in $L^q(\nu)$, for $q \in [1, +\infty]$, that is,

$$\mathscr{D} := \left\{ f = \sum_{j=1}^{m} a_j 1_{C_j}, \quad a_j \in \mathbb{R}, \quad C_j \in \mathscr{B}(\mathbb{R}) : \quad \nu(C_j) < +\infty \right\}. \tag{5.32}$$

5.7.1 Lemma. *Let $f \in \mathscr{D}$, $f \geq 0$. Then the increments of the process Y^f, $Y_t^f := (1_{[0,t]} f) * \mu$, can be represented as*

$$Y_{t+s}^f - Y_t^f = F_{s,f}({}^t X), \quad \text{a.s.}, \quad t, s \geq 0,$$

where $F_{s,f}$ is an $(\mathscr{F}_D, \mathscr{B}(\mathbb{R}))$-measurable functional.

Proof. Let $f \in \mathscr{D}$, $f \geq 0$, with representation $f = \sum_{j=1}^m a_j \, 1_{C_j}$, where $a_j \geq 0$ for every $j = 1, \dots, m$. For each C_j appearing in the representation of f, we introduce the sequence $B_j^n := C_j \cap \{|x| > \frac{1}{n}\}$ and the function

$$f_n := \sum_{j=1}^{m} a_j \, 1_{B_j^n}, \quad n \geq 1.$$

The process $\xi^{B_j^n}$ is given in (5.26). Because of Lemma 5.6.1, for every $n \geq 1$, we have

$$\mu(1_{(t,t+s]} f_n) = \sum_{j=1}^{m} a_j (\xi_{t+s}^{B_j^n} - \xi_t^{B_j^n}) = \sum_{j=1}^{m} a_j F_{s,B_j^n}({}^t L), \quad \text{a.s.}, \quad t, s \geq 0,$$

where $F_{s,B_j^n}(\cdot)$ is an $(\mathscr{F}_D, \mathscr{B}(\mathbb{R}))$-measurable functional. For every $t \geq 0$, $\mu(1_{[0,t]} f_n)$ increasingly converges to $\mu(1_{[0,t]} f)$ pointwise in ω as $n \to +\infty$. Therefore, if we put

$$F_{s,f}({}^t X) := \liminf_{n \to +\infty} \sum_{j=1}^{m} a_j F_{s,B_j^n}({}^t X),$$

we have $Y_{t+s}^f - Y_t^f = (1_{(t,t+s]} f) * \mu = F_{s,f}({}^t X)$ a.s. for every $t, s \geq 0$, $f \in \mathscr{D}$, $f \geq 0$. □

We observe that, from Lemma 1.2.3, the system $\mathscr{D} \subseteq L^q(\nu)$ of simple functions is dense in $L^q(\nu)$. As a consequence of this fact we can extend Lemma 5.7.1.

5.7.2 Proposition. *Let f be such that $|f| \wedge 1 \in L^1(\nu)$. The increments of the process Y^f, $Y_t^f := (1_{[0,t]} f) * \mu$ can be represented, as*

$$Y_{t+s}^f - Y_t^f = F_{s,f}({}^t L), \quad \text{a.s.}, \quad t, s \geq 0,$$

where $F_{s,f}$ is an $(\mathscr{F}_D, \mathscr{B}(\mathbb{R}))$-measurable functional.

Proof. First we consider the case $f \in L^1(\nu)$, $f \geq 0$. The set $\mathscr{D}$ of simple functions is dense in $L^1(\nu)$ and so there exists a sequence $(f_n)_{n \in \mathbb{N}} \subseteq \mathscr{D}$ such that $f_n \geq 0$, $n \geq 1$, and $f_n \uparrow f$ in $L^1(\nu)$ as $n \to +\infty$. By monotone convergence, $\mu(1_{[0,t]} f_n) \uparrow \mu(1_{[0,t]} f)$ pointwise in ω as $n \to +\infty$. By Lemma 5.7.1, we have

$$\mu(1_{(t,t+s]} f_n) = F_{s,f_n}({}^t X), \quad \text{a.s.}, \quad s, t \geq 0, \quad n \geq 1,$$

where F_{s,f_n} is a $(\mathscr{F}_D, \mathscr{B}(\mathbb{R}))$-measurable functional. If we put

$$F_{s,f}({}^tX) := \liminf_{n\to+\infty} F_{s,f_n}({}^tX)$$

we obtain, $Y^f_{t+s} - Y^f_t = (1_{(t,t+s]}f) * \mu = F_{s,f_n}({}^tX)$ and $F_{s,f}$ is an $(\mathscr{F}_D, \mathscr{B}(\mathbb{R}))$-measurable functional.

The remaing of this proof is non relevant for the exam. We now weaken the assumptions and consider f such that $|f| \wedge 1 \in L^1(\nu)$. Because of Proposition 5.3.2, the stochastic integral $(1_{[0,t]}f) * \mu$ exists and it is finite a.s. for every $t \geq 0$. If f^+ and f^- are the positive and the negative part of f, respectively, we have $f^{\pm} \leq |f|$ so that $f^{\pm} \wedge 1 \in L^1(\nu)$ and the stochastic integrals $(1_{[0,t]}f^{\pm}) * \mu$ are well defined and the relation $(1_{[0,t]}f) * \mu = (1_{[0,t]}f^+) * \mu - (1_{[0,t]}f^-) * \mu$ holds. Let us introduce the functions $f^{\pm}_n := f^{\pm}1_{\{|f^{\pm}|<n\}}$ and $f_n := f^+_n - f^-_n$. Then $f^{\pm}_n \uparrow f^{\pm}$ and $f^{\pm}_n \geq 0$. By monotone convergence we get $(1_{[0,t]}f^{\pm}_n)*\mu \uparrow (1_{[0,t]}f^{\pm})*\mu$ pointwise in ω as $n \to +\infty$. Furthermore, $f^{\pm}_n \in L^1(\nu)$, indeed $f^{\pm}_n \leq (f^{\pm} \wedge n) \leq n(f^{\pm} \wedge 1) \in L^1(\nu)$, for every $n \geq 1$. Because of the previous step, $(1_{(t,t+s]}f^{\pm}_n)*\mu = F_{s,f^{\pm}_n}({}^tX)$ a.s., where $F_{s,f^{\pm}_n}$ is an $(\mathscr{F}_D, \mathscr{B}(\mathbb{R}))$-measurable functional. We can put

$$F_{s,f^{\pm}}({}^tX) := \liminf_{n\to+\infty} F_{s,f_n}({}^tX^{\pm}),$$

which is an $(\mathscr{F}_D, \mathscr{B}(\mathbb{R}))$-measurable functional. Hence it follows that $\mu(1_{(t,t+s]}f^{\pm}) = F_{s,f^{\pm}}({}^tX)$ a.s. and from $(1_{[0,t]}f) * \mu = (1_{[0,t]}f^+) * \mu - (1_{[0,t]}f^-) * \mu$ we have

$$Y^f_{t+s} - Y^f_t = \mu(1_{(t,t+s]}f) = F_{s,f^+}({}^tL) - F_{s,f^-}({}^tX) =: F_{s,f}({}^tX), \quad \text{a.s.}, \quad s, t \geq 0.$$

which is an $(\mathscr{F}_D, \mathscr{B}(\mathbb{R}))$-measurable functional and the proof is complete. □

Let $(X, \mathbb{F})$ be a Lévy process and let μ be its jump measure. From Theorem 5.6.4, we know that μ is a Poisson random measure relative to $\mathbb{F}$ and that its intensity measure is $\lambda_+ \otimes \nu$, where λ_+ is the Lebesgue measure on $[0, T]$ and ν the Lévy measure of X. We define $f(x) := 1_{\{|x|>1\}}x$. The function f is such that $|f| \wedge 1 \in L^1(\nu)$. Clearly this implies that the function $h(t, x) := 1_{[0,t]}f$ is such that $|h| \wedge 1 \in L^1(\lambda_+ \otimes \nu)$. From Proposition 5.3.2, we can consider the stochastic integral of h with respect to μ. We introduce the process $L^2 = (L^2_t)_{t\geq 0}$ by

$$X^2_t := (1_{[0,t]}f) * \mu := (1_{[0,t]\times\{|x|>1\}}x) * \mu, \quad t \geq 0. \tag{5.33}$$

By Proposition 5.5.2, we know that $(X^2, \mathbb{F})$ is a Lévy process and furthermore, from Proposition 5.7.2, that its increments can be written as it follows:

$$X^2_{t+s} - X^2_t = F_{s,f}({}^tX), \quad \text{a.s.}, \quad t, s \geq 0,$$

where $F_{\cdot,f}({}^tL)$ is an $(\mathscr{F}_D, \mathscr{B}(\mathbb{R}))$-measurable homogeneous functional of the process ${}^tL$. Let now $Y = (Y_t)_{t\geq 0}$ be the process defined by

$$Y_t := X_t - X^2_t, \quad t \geq 0. \tag{5.34}$$

Clearly, $(Y, \mathbb{F})$ is a càdlàg process starting at 0 and its increments are $\mathbb{F}$-independent and homogeneous, because $Y_{t+s} - Y_t = {}^tL_s - F_{s,f}({}^tX)$ a.s., $s, t \geq 0$, which is clearly an $(\mathscr{F}_D, \mathscr{B}(\mathbb{R}))$-measurable homogeneous functional of ${}^tX$. These properties imply that

Y is also stochastically continuous and therefore $(Y, \mathbb{F})$ is a Lévy process (cf. Lemma 2.5.10).

We stress that to claim that $(Y, \mathbb{F})$ is a Lévy process, it is not sufficient that $(X, \mathbb{F})$ and $(X^2, \mathbb{F})$ are Lévy process: For the $\mathbb{F}$-independence of the increments of Y we need that the two dimensional process (X, X^2) has $\mathbb{F}$-independent increments. This is ensured by Proposition 5.34.

Notice that the relation

$$X_t^2 = \sum_{0<s\leq t} 1_{\{|\Delta X_s|>1\}} \Delta X_s, \quad t \geq 0, \quad a.s.,$$

holds. Therefore X_t^2 is the sum of all the jumps of the process X up to time t which are bigger than one. This implies that the process Y has only jumps of size smaller than or equal to one. Consequently, $(Y, \mathbb{F})$ is a Lévy process with bounded jumps and, because of Theorem 2.5.17, it has finite moments of every order. In particular, $\mathbb{E}[Y_t^2] < +\infty$ for every $t \geq 0$. Hence the Lévy process $\overline{Y} = (\overline{Y}_t)_{t\geq 0}$ defined by

$$\overline{Y}_t := Y_t - \mathbb{E}[Y_t], \quad t \geq 0, \tag{5.35}$$

is a square-integrable $\mathbb{F}$-martingale (cf. Lemma 2.5.6). Because of the square integrability of $\overline{Y}_t$, ≥ 0, an application of Doob's inequality shows that the process $(\overline{Y})_{0\leq t\leq T}$ is uniformly integrable. Consequently, the process $(Y)_{0\leq t\leq T}$ is uniformly integrable. An application of the generalization of theorem of Lebesgue on dominated convergence to uniformly integrable families of random variables (cf. Theorem 1.1.3) shows that the mapping $t \mapsto \mathbb{E}[Y_t]$ is continuous (see the proof of Lemma 5.6.1). We now define $a(t) := \mathbb{E}[Y_t]$, $t \geq 0$. Because of the homogeneity of the increments we have

$$a(t+s) = \mathbb{E}[Y_{t+s} - Y_t] + \mathbb{E}[Y_t] = a(t) + a(s), \quad s, t \geq 0,$$

which is a Cauchy functional equation. We know that $a(\cdot)$ is continuous and from the previous relation it follows that $a(\cdot)$ is a linear function. Therefore there exists $\beta \in \mathbb{R}$ such that

$$a(t) = \beta t, \quad t \geq 0. \tag{5.36}$$

Now we are ready to prove the Itô–Lévy decomposition.

5.7.3 Theorem (Itô–Lévy decomposition). *Let L be a càdlàg adapted process with jump measure μ. Then $(X, \mathbb{F})$ is a Lévy process if and only if μ is a Poisson random measure relative to $\mathbb{F}$ with intensity function $\lambda_+ \otimes \nu$, where λ_+ is the Lebesgue measure on $[0, T]$ and ν a Lévy measure, and there exists a Wiener process $(\mathrm{W}^\sigma, \mathbb{F})$ with variance function $\sigma^2(t) = \sigma^2 t$, $\sigma^2 \geq 0$, called Gaussian part of L, such that the following decomposition holds*

$$X_t = \beta t + \mathrm{W}_t^\sigma + (1_{[0,t]\times\{|x|>1\}}\, x) * \mu + (1_{[0,t]\times\{|x|\leq 1\}}\, x) * \overline{\mu}, \quad t \geq 0, \text{ a.s.}, \tag{5.37}$$

where $\beta \in \mathbb{R}$.

Proof. We assume first that $(X, \mathbb{F})$ is a Lévy process. From Theorem 5.6.4 we know that μ is a Poisson random measure relative to $\mathbb{F}$ and that its intensity measure is $\lambda_+ \otimes \nu$, where λ_+ is the Lebesgue measure on $[0, T]$ and ν the Lévy measure of X.

We introduce the function $g(x) := 1_{\{|x|\leq 1\}}x$ and the sequence $g^n := 1_{\{\frac{1}{n}<|x|\leq 1\}}x$. The function g belongs to $L^2(\nu)$, while $g^n \in L^1(\nu)\cap L^2(\nu)$, $n\geq 1$, and $g^n \longrightarrow g$ in $L^2(\nu)$ as $n\to +\infty$, by dominated convergence. We now define the processes $X^3 = (X^3_t)_{t\geq 0}$ and $L^{3,n} = (X^{3,n}_t)_{t\geq 0}$ by

$$X^3_t := (1_{[0,t]}g) * \overline{\mu}; \qquad X^{3,n}_t := (1_{[0,t]}g^n) * \overline{\mu}, \quad t\geq 0, \quad n\geq 1. \tag{5.38}$$

Because of Theorem 5.5.3, we have $L^{3,n}_t \longrightarrow X^3_t$ in $L^2(\mathbb{P})$ as $n\to +\infty$, for every $t\geq 0$ and the process L^3 is an $\mathbb{F}$-square integrable martingale. Moreover, $(X^3,\mathbb{F})$ is a Lévy process. The same statements holds for the process $(X^{3,n},\mathbb{F})$. Furthermore, from Proposition 5.7.2, we know that the increments of $X^{3,n}$ can be represented by a homogeneous $(\mathscr{F}_D,\mathscr{B}(\mathbb{R}))$-measurable functional of the process ${}^tX$ (cf. (5.31)). We denote such a functional by $G_{\cdot,g^n}$. We now introduce the processes $\tilde{X}^1 = (\tilde{X}^1_t)_{t\geq 0}$ and $\tilde{X}^{1,n} = (\tilde{X}^{1,n}_t)_{t\geq 0}$ by

$$\tilde{X}^1 := X - X^2 - X^3 = Y - X^3, \qquad \tilde{X}^{1,n} := X - X^2 - X^{3,n} = Y - X^{3,n}, \tag{5.39}$$

respectively, where the processes X^2 and Y were introduced by (5.33) and (5.34), respectively. The process $\tilde{X}^{1,n}$ has $\mathbb{F}$-independent and homogeneous increments because

$$\tilde{X}^{1,n}_{t+s} - \tilde{X}^{1,n}_s = {}^tX_s - F_{s,f}({}^tX) - G_{s,g^n}({}^tX), \quad s,t\geq 0.$$

Furthermore, $\tilde{X}^{1,n}_t$ converges in probability to $\tilde{X}^1_t$, for every $t\geq 0$, as $n\to+\infty$, because $X^{3,n}_t$ converges in $L^2(\mathbb{P})$, and hence in probability, to X^3_t, as $n\to+\infty$, for every $t\geq 0$. Hence Lemma 2.5.7 yields that $\tilde{X}^1$ has $\mathbb{F}$-independent and homogeneous increments. From Theorem 5.5.3.(iv), we know that the process X^3 has the following jumps:

$$\Delta X^3 = 1_{\{\Delta X\neq 0\}}1_{\{|\Delta X|\leq 1\}}\Delta X \quad \text{a. s.}$$

Therefore the process (X^2+X^3) has the same jumps of the process X a. s. But then we can claim that the process $\tilde{X}^1$ is continuous a.s. The filtration $\mathbb{F}$ satisfies the unusual conditions and so we can find an adapted version of the process $\tilde{X}^1$ which is in fact continuous and has $\mathbb{F}$-independent and homogeneous increments. We denote again by $\tilde{X}^1$ such a modification. Then the process $(\tilde{X}^1,\mathbb{F})$ is a continuous Lévy process. The process Y introduced by (5.34) is such that Y_t is square integrable, for every $t\geq 0$ and the same holds for X^3. Moreover, $\mathbb{E}[X^3_t]=0$. This implies that $\tilde{X}^1_t$ is square integrable and that $\mathbb{E}[\tilde{X}^1_t] = \mathbb{E}[Y_t]$, for every $t\geq 0$. We have seen that $\mathbb{E}[Y_t]=\beta t$, $t\geq 0$, where $\beta\in\mathbb{R}$. We introduce the process $X^1 = (X^1_t)_{t\geq 0}$ by

$$X^1_t = \tilde{X}^1_t - \beta t, \quad t\geq 0. \tag{5.40}$$

Clearly $(X^1,\mathbb{F})$ is a continuous Lévy process and because of $\mathbb{E}[X^1_t]=0$, Lemma 2.5.6 implies that it is an $\mathbb{F}$-martingale. Moreover, X^1_t is square integrable, for every $t\geq 0$ (its jumps are bounded by zero!) We can assert that X^1 is a continuous square integrable martingale. By Doob's inequality, we get that the family $((X^1_t)^2)_{0\leq t\leq T}$ is uniformly integrable. As a consequence of Theorem 1.1.3, the mapping $t\mapsto \mathbb{E}[(X^1_t)^2]$ is continuous. We put $\sigma^2(t) := \mathbb{E}[(X^1_t)^2]$, $t\geq 0$. This is a continuous function and moreover, because of the homogeneity and the $\mathbb{F}$-independence of the increments of X^1 and of $\mathbb{E}[X_t^1] = 0$, $t \geq 0$, we get

$$\sigma^2(t+s) = \mathbb{E}[(X^1_{t+s} - X^1_t + X^1_t)^2] = \mathbb{E}[(X^1_{t+s} - X^1_t)^2] + \mathbb{E}[(X^1_t)^2] = \sigma^2(t) + \sigma^2(s), \quad t, s \geq 0.$$

This relation and the continuity of $\sigma^2(\cdot)$, imply that $\sigma^2(\cdot)$ is a linear function, i.e., there exists $\sigma^2 \geq 0$ such that $\sigma^2(t) = \sigma^2 t$, for every $t \geq 0$. In conclusion, we have shown that $(X^1, \mathbb{F})$ is a continuous Lévy process and a square integrable martingale with a linear variance function. We now discuss two cases.

The first case is $\sigma^2 = 0$. In this case $X^1_t = 0$ a. s. for every t and by continuity we deduce that X^1 is indistinguishable from the zero-process. By (5.40) we the obtain $\tilde{X}^1_t = \beta t$ and from (5.39)

$$X_t = \beta t + X^2_t + X^3_t = \beta t + (1_{[0,t]} 1_{\{|x|>1\}} x) * \mu + (1_{[0,t]} 1_{\{|x| \leq 1\}} x) * \overline{\mu}$$

which is the Lévy–Itô decomposition in this special case.

We now consider the case $\sigma^2 > 0$ and introduce the process $W = (W_t)_{t \in [0,T]}$ by setting

$$W_t := \frac{1}{\sigma} X^1_t, \quad t \in [0, T].$$

Clearly, W is an $\mathbb{F}$-square integrable continuous martingale and for every $t \geq s$

$$\mathbb{E}[W_t^2 - t | \mathscr{F}_s] = \frac{1}{\sigma^2} \mathbb{E}[(X^1_t)^2 - t | \mathscr{F}_s] = \frac{1}{\sigma^2} \mathbb{E}[(X^1_t)^2 - (X^1_s)^2 | \mathscr{F}_s] + \frac{1}{\sigma^2} (X^1_s)^2 - t = W_s^2 - s,$$

where in the last passage we used the independence of the increments, the definition of W and the linearity of the variance-function of X^1. The previous computation shows that $(W_t^2 - t)_{t \in [0,T]}$ is a martingale. Hence, an application of P. Lévy characterization of the Brownian motion (cf. Theorem 3.2.1) yields that $(W, \mathbb{F})$ is a Brownian motion. Hence, by $X^1 = \sigma W$ and the definition of X^1 (cf. (5.40)), we deduce

$$X_t = \beta t + \sigma W_t + X^2_t + X^3_t = \beta t + \sigma W_t + (1_{[0,t]} 1_{\{|x|>1\}} x) * \mu + (1_{[0,t]} 1_{\{|x| \leq 1\}} x) * \overline{\mu}.$$

Conversely, we now assume that X is a càdlàg process with jump measure μ, which is a Poisson random measure with intensity $\lambda_+ \otimes \nu$, where ν is a Lévy measure. We assume that $(W, \mathbb{F})$ is a Brownian motion and $\sigma^2 \geq 0$, such that (5.37) holds. We verify that X is a Lévy process. We put

$$(X^1, X^2, X^3) := (\sigma W, (1_{[0,\cdot] \times \{|x|>1\}}\, x) * \mu, (1_{[0,\cdot] \times \{|x| \leq 1\}}\, x) * \overline{\mu}).$$

By assumption W is a Brownain motion process. Moreover by Proposition 5.5.2 $(X^2, \mathbb{F})$ is a Lévy process. Analogously we know that $(X^3, \mathbb{F})$ is a Lévy process and an $\mathbb{F}$-martingale. The processes X^2 and X^3 do not have common jumps. Because of the continuity of X^1, we have that $\Delta X^1 \Delta X^j = 0$, $j = 1, 2$. From these properties one can deduce that the vector (X^1, X^2, X^3) is independent and has $\mathbb{F}$-independent increments (see Remark 5.7.4 below). Because of (5.37), we can assert that also the process X has $\mathbb{F}$-independent increments. Moreover, from (5.37) and the independence of the vector (X^1, X^2, X^3) we have

$$\begin{aligned} \mathbb{E}\big[\mathrm{e}^{iu(X_t - X_s)}\big] &= \mathrm{e}^{iu\beta(t-s)} \textstyle\prod_{j=1}^3 \mathbb{E}\big[\mathrm{e}^{iu(X^j_t - X^j_s)}\big] \\ &= \mathrm{e}^{iu\beta(t-s)} \textstyle\prod_{j=1}^3 \mathbb{E}\big[\mathrm{e}^{iuX^j_{t-s}}\big] \\ &= \mathbb{E}\big[\mathrm{e}^{iuX_{t-s}}\big], \quad 0 \leq s \leq t, \quad u \in \mathbb{R}, \end{aligned}$$

where in the last but one equality we used that X^j has homogeneous increments, $j = 1, 2, 3$. Then X is a càdlàg adapted process with homogeneous and $\mathbb{F}$-independent increments such that $X_0 = 0$. Hence X is also stochastically continuous. In conclusion $(X, \mathbb{F})$ is a Lévy process and ν is its Lévy measure. This completes the proof of the theorem. □

5.7.4 Remark. In the last part of the proof of Theorem 5.7.3, we used that the vector-process $(\sigma W, (1_{[0,\cdot]\times\{|x|>1\}}\, x) * \mu, (1_{[0,\cdot]\times\{|x|\le 1\}}\, x) * \overline{\mu})$ is independent and has $\mathbb{F}$-independent increments. We do not show this statement but we try now to justify it.

We know that $(1_{[0,\cdot]\times\{|x|>1\}}\, x) * \mu$ and $(1_{[0,\cdot]\times\{|x|\le 1\}}\, x) * \overline{\mu}$ can be approximated by linear combinations of Poisson processes and of compensated Poisson Processes, respectively. Hence it is enough to show that if W is a Brownain motion and $N^1, \dots, N^n$ are Poisson processes with respect to $\mathbb{F}$ such that $\Delta Y^j \Delta Y^k = 0$, $j \neq k$, then the vector $(W, Y^1, \dots, Y^n)$ is independent and has $\mathbb{F}$-independent increments. To this aim it is sufficient to prove that the process $\tilde{Z} = (\tilde{Z}_t)_{t\in[0,T]}$,

$$\tilde{Z}_t := \exp\left(\mathrm{i}uW_t - \frac{u^2}{2}\,t\right) Z_t,$$

is a martingale, where $Z = (Z_t)_{t\in[0,T]}$ is the martingale of Theorem 4.3.1. For this goal we need a general Itô-formula (or partial integration), which is outside of the scope of this notes (also its formulation). The interested reader can find something about this problem in He, Wang & Yan (1992), Theorem 11.43.

An immediate but important consequence of the Lévy–Itô decomposition is the Lévy–Kintchine decomposition of the characteristic function of a Lévy process. We formulate the Lévy–Kintchine decomposition as a corollary of Theorem 5.7.3. However it is also possible to deduce Lévy–Itô decomposition from Lévy–Kintchine decomposition.

Before we give the following definition:

5.7.5 Definition. Let $(X, \mathbb{F})$ be a Lévy process. The triplet (β, σ^2, ν) of Theorem 5.7.3 (β is the coefficient of the drift, σ^2 is the variance-parameter of W and ν is the Lévy measure of X) is called *characteristic triplet* of the Lévy process. If $\sigma^2 = 0$, we say that X is a purely non-Gaussian Lévy process.

Notice that the Lévy processes which we constructed via stochastic integration with respect to a Poisson random measure and a compensated Poisson random measure in §5.5 are all purely non-Gaussian Lévy process.

5.7.6 Corollary. *Let X be a Lévy process with characteristics (β, σ^2, ν). Then for every $u \in \mathbb{R}$ and for every $t \ge 0$ we have*

$$\mathbb{E}[\mathrm{e}^{iuX_t}] = \exp\left(\left(iu\beta - \frac{1}{2}u^2\sigma^2 + \int_{\mathbb{R}} (\mathrm{e}^{iux} - 1 - iux1_{\{|x|\le 1\}})\, \nu(\mathrm{d}x)\right) t\right). \qquad (5.41)$$

Proof. We put $\varphi_t^X(u) := \mathbb{E}[\mathrm{e}^{iuX_t}]$, $t \ge 0$, $u \in \mathbb{R}$. We put

$$(X^1, X^2, X^3) := (\mathrm{W}^\sigma, \mu(1_{[0,\cdot]\times\{|x|>1\}}\, x), \overline{\mu}(1_{[0,\cdot]\times\{|x|\le 1\}}\, x)).$$

From Remark 5.5, the vector (X^1, X^2, X^3) is independent and has $\mathbb{F}$-independent increments. Therefore

$$\varphi_t^X(u) = \mathrm{e}^{iu\beta t}\varphi_t^{X^1}(u)\varphi_t^{X^2}(u)\varphi_t^{X^3}(u), \quad u \in \mathbb{R}, \quad t \geq 0. \tag{5.42}$$

We know that $\varphi_t^{X^1}(u) = \exp\left(-\frac{1}{2}u^2\sigma^2 t\right)$. Moreover, Lemma 5.3.3 implies that

$$\varphi_t^{L^2}(u) = \exp\left(t\int_{\{|x|>1\}}(\mathrm{e}^{iux}-1)\,\nu(\mathrm{d}x)\right).$$

To compute $\varphi^{L^3}(u)$ we have to proceed by approximation. If we define $f_n(x) := 1_{\{\frac{1}{n}<|x|\leq 1\}}\, x$, then we have $f_n \in L^1(\nu) \cap L^2(\nu)$ and $f_n \longrightarrow 1_{\{|x|\leq 1\}}x$ in $L^2(\nu)$ as $n \to +\infty$. By Proposition 5.4.5 and Lemma 5.3.3, we have

$$\mathbb{E}\left[\mathrm{e}^{iu(1_{[0,t]}f_n)*\overline{\mu}}\right] = \exp\left(iut\int_{\{\frac{1}{n}<|x|\leq 1\}} x\,\nu(\mathrm{d}x)\right)\exp\left(t\int_{\{\frac{1}{n}<|x|\leq 1\}}(\mathrm{e}^{iux}-1)\,\nu(\mathrm{d}x)\right).$$

Passing to the limit as $n \to +\infty$, it follows

$$\varphi_t^{L^3}(u) = \exp\left(t\int_{\{|x|\leq 1\}}(\mathrm{e}^{iux}-1-iux)\,\nu(\mathrm{d}x)\right)$$

and (5.42) implies (5.41). □

Bibliography

D. Appelbaum. *Lévy processes and stochastic calculus. Second edition.* Cambridge University Press, 2009.

P. Baldi. *Introduction to stochastic processes.* Scuola Matem. Interuniv. Cortona, 2007. http://160.80.10.11/ processi/cortonalevy.pdf

H. Bauer. *Probability Theory.* De Gruytier studies in mathematics; 23, 1996.

H. Bauer. *Measure and Integration Theory.* De Gruytier studies in mathematics; 26, 2001.

H. Brezis. *Analyse fonctionelle-Téorie et Application.* Masson Editeur; Paris, 1983.

F. Caravenna. Dispense del corso di analisi stocastica. `www.matapp.unimib.it/~fcaraven/did0910/mat/dispense-2.1-2pagine.pdf`

C. Dellacherie, P.-A. Meyer. *Probabilités et potential*I. Hermann: Paris, 1978.

C. Dellacherie. *Capacités et processus stocahstiques.* Springer, Berlin 1972.

I.I. Gihman, A.V. Skorohod. *The theory of stochastic processes I.* Springer-Verlag 1974.

S. He, J. Wang, J. Yan. *Semimartingale Theory and Stochastic Calculus.* Science Press, Beijing-New York, CRC Press INC, 1992.

F. Flandoli. Dispense del corso di istituzioni di Probabilitá 2013-2014, `users.dma.unipi.it/~flandoli/dispense_Istituzioni_2014.pdf`.

M. Keller-Ressel. Vorlesungskript: "Stochastische Analysis", Sommersemester 2014-2015 `www.math.tu-dresden.de/~mkeller/stochana/stochana_skript.pdf`

Statistics of Random Processes, Vol. 2. Springer Verlag, Berlin 2001.

J. Jacod. *Calcul Stochastique et Problèmes de Martingales.* Springer, Berlin Heidelberg 1979.

J. Jacod, A. Shiryaev. *Limit Theorems for Stochastic Processes 2ed.* Springer, 2000.

O. Kallenberg. *Foundations of Modern Probability.* Springer series in statistics. Probability and its applications. Springer-Verlag, 1997.

H. Kunita. *Representation of martingales with jumps and applications to finance.* Advanced Studies in Pure Mathematics 41, 2004 Stochastic Analysis and Related Topics pp. 209-232.

P. Protter. *Stochastic Integration and Differential Equations.* Springer, 2005; Edition 2.1.

D. Revuz, M. Yor. *Continuous Martingales and Brownian Motion.* Springer Verlag, Berlin 1999; Edition 3.

K. Sato. *Lévy processes and infinitely divisible distributions.* Cambridge University Press 1999.

R. Schilling, L. Partzsch. *Brownian motion: an introduction to stochastic processes.* De Gruyter, 2014.

M. Sharpe. *General theory of Markov provesses.* Pure and applied mathematics (Academic Press 1988): vol.133.

A. N. Shiryaev. *Probability.* Graduate Texts in Mathematics, vol. 95. Springer, 1996.

J. Wang. *Some remarks on processes with independent increments.* Séminaire de Probabilités (Strasbourg) XV(1981). pg. 627-631. Springer Verlag, 1981.