PART I

The fundamental principles

1

Preliminaries

In order to fix the terminology, let us start with some well-established basic facts from functional analysis and measure theory. The reader is referred to the books of Reed and Simon [98] or Rudin [101] and Ash [5], Halmos [43] or Billingsley [13] for details.

We study Fréchet spaces, because the archetype of an abstract Wiener space is the space of real sequences, endowed with the topology of pointwise convergence. It is a Fréchet space and an abstract Wiener space over 'little' l^2.

Let $\mathbb{N}$ be the set of positive integers and let $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. We identify $n \in \mathbb{N}_0$ with the set $\{1, \ldots, n\}$, thus $0 = \emptyset$. For elements a, b of a totally ordered set, we set

$$a \wedge b := \min\{a, b\}, \quad a \vee b := \max\{a, b\}.$$

The following notation is important: if f is a binary relation, then $f[A]$ is the set of all second components of pairs in f, where the first components run through A, i.e.,

$$f[A] := \{y \mid \exists x \in A\,((x, y) \in f)\}.$$

In particular, if f is a function, then $f[A] := \{f(x) \mid x \in A\}$ and $f^{-1}[B] := \{x \mid f(x) \in B\}$. Note that, in contrast to $f[A]$, $f(A)$ denotes the value of A if A is an element of the domain of f (see the notion 'image measure' below).

A semi-metric d fulfils the same conditions as a metric, except that from $d(a, b) = 0$ it need not follow that $a = b$. Let $(|\cdot|_i) := (|\cdot|_i)_{i \in \mathbb{N}}$ be a separating sequence of semi-norms on a real vector space $\mathbb{B}$; **separating** means: if $x \neq 0$, then there exists an $i \in \mathbb{N}$ with $|x|_i \neq 0$. A neighbourhood base of an element a in $\mathbb{B}$ in the locally convex topology $\mathcal{T}_{(|\cdot|_i)}$, given by $(|\cdot|_i)$, is the family of sets of the form

$$U_{\frac{1}{m}}(a) := \left\{ x \in \mathbb{B} \mid \max_{j \leq m} |x - a|_j < \frac{1}{m} \right\},$$

with $m \in \mathbb{N}$. It follows that a sequence $(a_k)_{k\in\mathbb{N}}$ converges to a in the topology $\mathcal{T}_{(|\cdot|_i)}$ if and only if $\lim_{k\to\infty} |a_k - a|_i = 0$ for all $i \in \mathbb{N}$. The topology $\mathcal{T}_{(|\cdot|_i)}$ is generated by a **translation invariant** metric d, i.e., $d(a,b) = d(a+c, b+c)$, where

$$d(a,b) = \sum_{i=1}^{\infty} \frac{1}{2^i} \cdot \frac{|a-b|_i}{1+|a-b|_i}.$$

This metric d is called the **metric generated by** $(|\cdot|_i)_{i\in\mathbb{N}}$. Note that we have $d(a+b,0) \le d(a,0) + d(b,0)$. The space $(\mathbb{B}, d) = (\mathbb{B}, (|\cdot|_i))$ is called a **Pre-Fréchet space.** It is called a **Fréchet space**, provided $\mathbb{B}$ is complete. We always assume that Fréchet spaces are separable. The space $(\mathbb{B}, d)$ is complete as a metric space if and only if it is **complete as a locally convex space** $(\mathbb{B}, (|\cdot|_i))$, i.e., if (a_n) is a Cauchy sequence for each semi-norm $|\cdot|_i$, then there exists an $a \in \mathbb{B}$ such that $\lim_{n\to\infty} |a_n - a|_i = 0$ for all $i \in \mathbb{N}$.

The **topological dual** of a locally convex space $\mathbb{B}$ over $\mathbb{R}$ is denoted by $\mathbb{B}'$. It is the space of all linear and continuous functions $\varphi : \mathbb{B} \to \mathbb{R}$.

Fix a Hilbert space $\mathbb{H}$ over $\mathbb{R}$ with scalar product $\langle \cdot, \cdot \rangle : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$ and norm $\|\cdot\|$ given by $\|a\| := \sqrt{\langle a, a \rangle}$. For subsets $A, B \subseteq \mathbb{H}$, we set $A + B := \{a + b \mid a \in A \text{ and } b \in B\}$. The **Cauchy–Schwarz inequality** says that $|\langle a, b \rangle| \le \|a\| \cdot \|b\|$ for all $a, b \in \mathbb{H}$.

By the Cauchy–Schwarz inequality, we have $(\langle a, \cdot \rangle : \mathbb{H} \to \mathbb{R}) \in \mathbb{H}'$ for each $a \in \mathbb{H}$. Vice versa, by the **Riesz lemma**, for each $\varphi \in \mathbb{H}'$ there exists a (unique) $a_\varphi \in \mathbb{H}$ such that $\varphi = \langle a_\varphi, \cdot \rangle$ (see Theorem II.4 in [98]). It is a common practice to identify φ and a_φ, thus $\mathbb{H}' = \mathbb{H}$.

We say that $a \in \mathbb{H}$ is **orthogonal to** $b \in \mathbb{H}$ if the distance between a and b is equal to the distance between a and $-b$; in this case we shall write $a \perp b$. Note that $a \perp b$ if and only if $\langle a, b \rangle = 0$. Let $S \subseteq \mathbb{H}$. We will write $a \perp S$ if $a \perp b$ for all $b \in S$ and set $S^\perp := \{x \in \mathbb{H} \mid x \perp S\}$. Since $\langle \cdot, b \rangle$ is linear and continuous for all $b \in \mathbb{H}$, $S^\perp$ is a closed linear subspace of $\mathbb{H}$.

Let $\mathcal{E}(\mathbb{H})$ denote the set of all finite-dimensional subspaces of $\mathbb{H}$. If $A \subseteq \mathbb{H}$, then **span**A denotes the linear subspace of $\mathbb{H}$, generated by A.

A subset $S \subseteq \mathbb{H}$ is called an **orthonormal set** (ONS) in $\mathbb{H}$ if $\|a\| = 1$ and $a \perp b$ for all $a, b \in S$ with $a \ne b$. An ONS $\mathfrak{B}$ in $\mathbb{H}$ is called an **orthonormal basis** (ONB) of $\mathbb{H}$ if $\mathfrak{B}$ is **maximal**, i.e., there does not exist an ONS S which is a strict extension of $\mathfrak{B}$. Each Hilbert space $\mathbb{H}$ has an ONB (see Theorem II.5 in [98]).

We will often use the so-called **projection theorem** (see Theorem II.3 in [98]), which tells us that, if G is a closed subspace of $\mathbb{H}$, then each $x \in \mathbb{H}$ can be composed of a sum $x = a + b$ with $a \in G$ and $b \in G^\perp$; the pair (a, b) is uniquely

Cambridge University Press
978-1-107-01614-9 - Malliavin Calculus for Lévy Processes and Infinite-Dimensional Brownian Motion: An Introduction
Horst Osswald
Excerpt
More information

determined by x; the mapping $f:\mathbb{H}\to G$ that assigns to x the element a is called the **orthogonal projection** from $\mathbb{H}$ onto G and is denoted by $\pi_{\mathbb{H}G}$.

In what follows let us assume that $\mathbb{H}$ is an infinite-dimensional Hilbert space and **separable**. This means there exists a countable **dense** subset $D\subseteq\mathbb{H}$, i.e., for each $a\in\mathbb{H}$ and each $\varepsilon>0$ there exists a $d\in D$ with $\|d-a\|<\varepsilon$. It follows that each separable $\mathbb{H}$ has a countable ONB $\mathfrak{E}:=(\mathfrak{e}_i)_{i\in\mathbb{N}}$ (see Theorem II.5 in [98]). Now each a is an infinite linear combination of elements in $\mathfrak{E}$, i.e., $a=\sum_{i=1}^{\infty}\langle a,\mathfrak{e}_i\rangle\,\mathfrak{e}_i$ (see Theorem II.6 in [98]). The $\langle a,\mathfrak{e}_i\rangle$ are called the **Fourier coefficients** of a with respect to $\mathfrak{E}$.

One form of the **Hahn–Banach theorem for locally convex spaces** $\mathbb{B}$, given by a separating sequence $(|\cdot|_i)$ of semi-norms, is (see Theorem 3.3 in Rudin [101]): Fix $i\in\mathbb{N}$ and a subspace M of $\mathbb{B}$. Then each linear map $f:M\to\mathbb{R}$ with $|f(\cdot)|\le|\cdot|_i$ on M can be extended to a linear mapping $g:\mathbb{B}\to\mathbb{R}$ such that $|g(\cdot)|\le|\cdot|_i$ on $\mathbb{B}$. In the special case of normed spaces we obtain the following result. Let $(\mathbb{B},|\cdot|)$ be a normed space and fix $a\in\mathbb{B}$ with $a\neq 0$. Then there exists a $\varphi\in\mathbb{B}'$ such that $|\varphi|:=\sup_{|x|\le 1}|\varphi(x)|\le 1$ and $\varphi(a)=|a|$.

We also need some elementary facts from measure theory (see Ash [5] or Billingsley [13] for details). The **power set** of a set Λ, i.e., the set of all subsets of Λ, is denoted by $\mathcal{P}(\Lambda)$. The symmetric difference of sets A,B is denoted by

$$A\mathbin{\triangle}B:=(A\setminus B)\cup(B\setminus A).$$

An **algebra** on a set Λ is a subset of $\mathcal{P}(\Lambda)$, closed under finite unions and under complements, and containing Λ. An algebra on Λ is called a **σ-algebra** on Λ if it is closed under countable unions. If $\mathcal{D}\subseteq\mathcal{P}(\Lambda)$, then the intersection $\sigma(\mathcal{D})$ of all σ-algebras $\mathcal{S}\supseteq\mathcal{D}$ is again a σ-algebra, which is called the **σ-algebra**, **generated by** $\mathcal{D}$. For two subsets $\mathcal{X},\mathcal{Y}\subseteq\mathcal{P}(\Lambda)$ let us set

$$\mathcal{X}\vee\mathcal{Y}=\sigma(\mathcal{X}\cup\mathcal{Y}).$$

The **Borel σ-algebra** on a topological space Λ is generated by the set of open sets in Λ and is denoted by $\mathcal{B}(\Lambda)$. The elements of $\mathcal{B}(\Lambda)$ are called the **Borel sets** of Λ.

Fix $a,b\in\mathbb{R}\cup\{-\infty\}$. **Right intervals** in $\mathbb{R}$ are sets of the form $]a,b]=\{x\in\mathbb{R}\mid a<x\le b\}$ or of the form $]a,\infty[=\{x\in\mathbb{R}\mid a<x<\infty\}$. Fix $n\in\mathbb{N}$. **right rectangles** in $\mathbb{R}^n$ are sets of the form $J_1\times\ldots\times J_n$, where $J_1,\ldots,J_n$ are right intervals in $\mathbb{R}$. Note that the set $\mathcal{R}(\mathbb{R}^n)$ of finite unions of pairwise disjoint right rectangles is an algebra on $\mathbb{R}^n$. This set also generates the Borel σ-algebra on $\mathbb{R}^n$.

Let $\mathcal{A}$ be an algebra on a set Λ. A function $\mu:\mathcal{A}\to[0,\infty]$ is called a **finitely additive measure on** $\mathcal{A}$ if $\mu(\emptyset)=0$ and $\mu(A\cup B)=\mu(A)+\mu(B)$ for all $A,B\in\mathcal{A}$

with $A \cap B = \emptyset$. A finitely additive measure μ on $\mathcal{A}$ is called a **measure** if it is σ**-additive**, i.e., $\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$ for all pairwise disjoint $A_n \in \mathcal{A}$ such that $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$. In this case $(\Lambda, \mathcal{A}, \mu)$ is called a **measure space**, provided that $\mathcal{A}$ is a σ-algebra. A measure μ on $\mathcal{A}$ is called **finite** if $\mu(\Lambda) < \infty$, and μ is called a **probability measure** if $\mu(\Lambda) = 1$, in which case $(\Lambda, \mathcal{A}, \mu)$ is called a **probability space**. The measure μ is called a **Borel measure** if $\mathcal{A}$ is a Borel σ-algebra.

Since we are working only with finite measures with the exception of the σ-finite Lebesgue measure on $]0,\infty]$, the Lévy measure on $\mathbb{R}$ and the counting measure on $\mathbb{N}$, it is always assumed that measures μ are σ**-finite**, i.e., there exists a sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}$ with $\Lambda = \bigcup_{n\in\mathbb{N}} A_n$ and $\mu(A_n) < \infty$. Here we have to be careful, because σ-finite measures, restricted to a σ-subalgebra of $\mathcal{A}$, and image measures (see below) of a σ-finite measure need not to be σ-finite.

A set $N \in \mathcal{A}$ is called a μ**-nullset** if $\mu(N) = 0$. The set of all μ-nullsets is denoted by $\mathcal{N}_\mu$. A measure space $(\Lambda, \mathcal{A}, \mu)$ is called **complete** if each subset of a μ-nullset belongs to $\mathcal{A}$, whence it is a μ-nullset.

Fix a measure space $(\Lambda, \mathcal{A}, \mu)$. The **Borel–Cantelli lemma** will be often used (see Ash [5] 2.2.4): Let $(U_i)_{i\in\mathbb{N}}$ be a sequence of sets in $\mathcal{A}$ such that $\sum_{i\in\mathbb{N}} \mu(U_i) < \infty$. Then

$$\mu\left(\bigcap_{n\in\mathbb{N}} \bigcup_{n\le i\in\mathbb{N}} U_i\right) = 0.$$

Two functions F, G, defined on Λ, are identified if $G = F$ μ-a.e. (μ-almost everywhere), i.e., $\mu\{X \mid F(X) \neq G(X)\} = 0$. Sometimes let us write μ-a.s. (μ-almost surely) instead of μ-a.e., in case μ is a probability measure. We are interested in L^p-spaces with $p \in [1,\infty]$. For $p \in\ [1,\infty[$ let $L^p(\mu)$ be the space of all real random variables F such that $|F|^p$ is μ-integrable and set $\|F\|_p := (\mathbb{E}|F|^p)^{\frac{1}{p}}$. Let $L^\infty(\mu)$ be the set of all μ-a.e. bounded random variables and for $F \in L^\infty(\mu)$ set $\|F\|_\infty := \inf\{c > 0 \mid |F| \le c\ \mu\text{-a.e.}\}$. All these L^p-spaces are Banach spaces with norm $\|\cdot\|_p$. We will write L^p instead of $L^p(\mu)$ if it is clear which measure we mean.

Let $(\Lambda, \mathcal{A}, \mu)$ be a probability space. The elements in $\mathcal{A}$ are called **events.** The **expected value** of a real-valued random variable f, i.e., the integral $\int_\Lambda f d\mu$, if it exists, is denoted by $\mathbb{E}_\mu(f)$ or simply $\mathbb{E}(f)$ or $\mathbb{E}f$ if it is clear which measure μ we mean. The **conditional expectation** of a μ-integrable random variable f with respect to a sub-σ-algebra $\mathcal{C}$ of $\mathcal{A}$ is denoted by $\mathbb{E}^{\mathcal{C}} f$ or $\mathbb{E}_\mu^{\mathcal{C}} f$ if it is not clear which measure we mean. It is the μ-a.s. uniquely determined

$\mathcal{C}$-measurable function g such that for all $C \in \mathcal{C}$

$$\int_C g\,d\mu = \int_C f\,d\mu.$$

In Ash's book [5] Section 6.5 the reader can find an arrangement of all properties of the conditional expectation we need; in particular, we use Jensen's inequality

$$\left(\mathbb{E}^{\mathcal{C}} |f|\right)^p \leq \mathbb{E}^{\mathcal{C}} \left(|f|^p\right) \ \mu\text{-a.s.}$$

over and over again, where $p \in [1, \infty[$ and $|f|^p$ is integrable.

Let $(\Lambda, \mathcal{A}, \mu)$ be a measure space, let $\mathcal{A}'$ be a σ-algebra on a set Λ' and let $f : \Lambda \to \Lambda'$ be $(\mathcal{A}, \mathcal{A}')$**-measurable**, i.e., $f^{-1}[B] \in \mathcal{A}$ for all $B \in \mathcal{A}'$. The measure μ_f, defined on $\mathcal{A}'$ by $\mu_f(B) := \mu\left(f^{-1}[B]\right)$, is called the **image measure of** μ **by** f. If $\mathcal{A}'$ is a Borel σ-algebra, then f is simply called $\mathcal{A}$**-measurable**.

We have the following **transformation rule** (see Bauer [6], 19.1 and 19.2): Let $f : \Lambda \to \mathbb{R}^n$ be $\mathcal{A}$-measurable. Then we have, for all Borel-measurable $g : \mathbb{R}^n \to \mathbb{R}^d$,

$$\int_{\mathbb{R}^n} g\,d\mu_f = \int_\Lambda g \circ f\,d\mu,$$

provided that at least one integral exists.

Assume that $f : \Lambda \to \mathbb{R}_0^+$ is $\mathcal{A}$-measurable. The set function $f\mu : \mathcal{A} \to [0, \infty]$ is defined by setting

$$f\mu(B) := \int_B f\,d\mu = \int_\Lambda \mathbf{1}_B \cdot f\,d\mu.$$

Here $\mathbf{1}_B$ is the **indicator function** of B, i.e., $\mathbf{1}_B(x) = 1$ for $x \in B$ and $\mathbf{1}_B(x) = 0$ for $x \notin B$. By the monotone convergence theorem, $f\mu$ is a measure on $\mathcal{A}$. We say that the measure $f\mu$ has **density** f **with respect to** μ. The previous equality can be extended from $\mathbf{1}_B$ to all $\mathcal{A}$-measurable functions $g : \Lambda \to \mathbb{R}$:

$$\int_\Lambda g\,df\mu = \int_\Lambda g \cdot f\,d\mu,$$

provided that at least one integral exists.

We call a measure μ **absolutely continuous to** a measure ν, where μ and ν are defined on the same domain, if each ν-nullset is also a μ-nullset. Both measures are called **equivalent** if μ is absolutely continuous to ν and ν is absolutely continuous to μ.

The product measure of measures $\mu_1, \ldots, \mu_k$ is denoted by $\mu_1 \otimes \ldots \otimes \mu_k$.

In the following chapter we will use extensions of the previously mentioned **Riesz lemma** for σ-finite measures. The dual space of $L^p(\mu)$ with $1 \leq p < \infty$ is $L^q(\mu)$, where $\frac{1}{p} + \frac{1}{q} = 1$ and $q = \infty$ for $p = 1$, in the following sense: for each $\varphi \in (L^p(\mu))'$ there exists an $\psi \in L^q(\mu)$ such that $\varphi(f) = \int_\Lambda \psi \cdot f d\mu$ for all $f \in L^p(\mu)$.

2

Martingales

In this chapter a detailed introduction to martingale theory is presented. In particular, we study important Banach spaces of martingales with regard to the supremum norm and the quadratic variation norm. The main results show that the martingales in the associated dual spaces are of bounded mean oscillation. The Burkholder–Davis–Gandy (B–D–G) inequalities for L^p-bounded martingales are very useful applications. All results in this chapter are well known; I learned the proofs from Imkeller's lecture notes [47]. We also need the B–D–G inequalities for special Orlicz spaces of martingales.

In this chapter we study martingales, defined on standard finite timelines. Later on the notion 'finite' is extended and the results in this chapter are transferred to a finite timeline, finite in the extended sense. We obtain all established results also for the new finite timeline. Then we shall outline some techniques to convert processes defined on this new finite timeline to processes defined on the continuous timeline $[0,\infty[$ and vice versa. The reader is referred to the fundamental articles of Keisler [53], Hoover and Perkins [46] and Lindstrøm [64].

From what we have now said it follows that we only need to study martingales defined on a discrete, even finite, timeline.

2.1 Martingales and examples

Fix a countable set I, totally ordered by $<$ with smallest element $\square$. If not otherwise determined, we assume that I is finite and $H = \max I$. The set I can be viewed as a **timeline**. Choose an arbitrary object $H^+ \notin I$, define $t < H^+$ for all $t \in I$ and set $\inf \emptyset := H^+$. For each $t \in I$ we set $I_t := \{s \in I \mid s \leq t\}$.

Fix a complete probability space $(\Lambda, \mathcal{C}, \mu)$. Then $\left(\Lambda, \mathcal{C}, \mu, (\mathcal{C}_t)_{t\in I}\right)$ is called an **adapted probability space** if $(\mathcal{C}_t)_{t\in I}$ is a **filtration on** $\mathcal{C}$, i.e., $\mathcal{C}_t$ is a

σ-subalgebra of $\mathcal{C}$ and $\mathcal{C}_s \subseteq \mathcal{C}_t$ for $s \leq t$. The events in $\mathcal{C}_t$ represent the **state of information at time** $t \in I$. We tacitly assume that each $\mathcal{C}_t$ contains all μ-nullsets.

A property $P(X)$ about elements $X \in \Lambda$ holds **almost surely** if the set $\{P \text{ fails}\} := \{X \in \Lambda \mid P(X) \text{ fails}\}$ is a μ-nullset. As it is a common practice, we will write $\{P\}$ instead of $\{X \in \Lambda \mid P(X) \text{ holds}\}$ if P is a property on elements of Λ.

For $F : I \to \mathbb{R}$ and $t \in I$ the difference $\Delta F_t := F(t) - F(t^-)$ is called the **increment of** F **to** t. Here t^- is the immediate predecessor of t if $t > \square$. Set $F(\square^-) := 0$ and assume that $\square^- < t$ for all $t \in I$. So $\Delta F_\square = F_\square$ is the first 'jump' of F. As usual, we write F_t instead of $F(t)$. Set $\mathcal{C}_{\square^-} := \{\emptyset, \Lambda\} \vee \mathcal{N}_\mu$. Then $\mathbb{E}^{\mathcal{C}_{\square^-}} F = \mathbb{E}F$ μ-a.s. for all $F \in L^1$.

A process $F : \Lambda \times I \to \mathbb{R}$ is called μ-p-**integrable** if $F_t \in L^p(\mu)$ for all $t \in I$. If F is μ-1-integrable, then we simply say F is μ-**integrable** or simply **integrable**. A process $M := (M_t)_{t\in I}$ is called a $(\mathcal{C}_t)_{t\in I}$-μ-**martingale** if the following conditions are fulfilled.

(a) $(M_t)_{t\in I}$ is $(\mathcal{C}_t)_{t\in I}$-**adapted**, i.e., M_t is $\mathcal{C}_t$-measurable for all $t \in I$.
(b) M is integrable.
(c) $\mathbb{E}^{\mathcal{C}_s} M_{s^+} = M_s$ μ-a.s. if $s^+ \in I$ is the immediate successor of $s \in I$.

If under (c) we have "$\geq$" instead of "$=$", then M is called an $(\mathcal{C}_t)_{t\in I}$-μ-**submartingale**. By Jensen's inequality, $|M|^p$ with $1 \leq p < \infty$ is a $(\mathcal{C}_t)_{t\in I}$-μ-submartingale if M is a $(\mathcal{C}_t)_{t\in I}$-μ-martingale and M is p-integrable.

If we understand $M_t(X)$ as the result of the chance X at time $t \in I$, then condition (a) means that the result at time t does not depend on what will happen after time t. Condition (c) means that, under the present state $\mathcal{C}_t$ of information, the expected result at the future time t^+ is identical to the achieved result at the present time t.

Let us drop $(\mathcal{C}_t)_{t\in I}$ or μ in the phrases martingale or submartingale if it is clear which filtration or measure we mean. We call $F : \Lambda \times I \to \mathbb{R}$ a **canonical martingale** if F is a $\left(\mathcal{C}_t^F\right)_{t\in I}$-martingale, where $\left(\mathcal{C}_t^F\right)_{t\in I}$ is the **filtration generated by** F, i.e.,

$$\mathcal{C}_t^F = \left\{ (\Delta F_s)_{s\in I_t}^{-1} [B] \mid B \text{ is a Borel set in } \mathbb{R}^{I_t} \right\}.$$

Examples 2.1.1 Let $N : \Lambda \to \mathbb{R}$ be μ-integrable.

(i) $\left(\mathbb{E}^{\mathcal{C}_t} N\right)_{t\in I}$ is a martingale.
(ii) $\mathbf{1}_A \cdot M$ is a martingale if $A \in \mathcal{C}_\square$ and M is a martingale.
(iii) $M : (\cdot, t) \mapsto \mathbf{1}_A(\cdot) \cdot \mathbf{1}_{[s,H]}(t) \cdot \left(\mathbb{E}^{\mathcal{C}_t} N - \mathbb{E}^{\mathcal{C}_s} N\right)$ is a martingale if $s \in I$ and $A \in \mathcal{C}_s$.