ulm university universität uulm

Stochastics II

Lecture notes
(Working version)

Prof. Dr. Evgeny Spodarev

Ulm
2010

Contents

1 General theory of random functions

1.1 Random functions

Let $(\Omega, \mathcal{A}, \mathsf{P})$ be a probability space and $(\mathcal{S}, \mathcal{B})$ a measurable space, $\Omega, \mathcal{S} \neq \emptyset$.

Definition 1.1.1
A *random element* $X : \Omega \to \mathcal{S}$ is a $\mathcal{A}|\mathcal{B}$-measurable mapping (Notation: $X \in \mathcal{A}|\mathcal{B}$), i.e.,

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}, \quad B \in \mathcal{B}.$$

If X is a random element, then $X(\omega)$ is a *realization of* X for arbitrary $\omega \in \Omega$.

The σ-algebra $\mathcal{B}$ of subsets of $\mathcal{S}$ is induced by the set system $\mathcal{M}$ (Elements of $\mathcal{M}$ are also subsets of $\mathcal{S}$), if

$$\mathcal{B} = \bigcap_{\substack{\mathcal{F} \supset \mathcal{M} \\ \mathcal{F}\text{-}\sigma\text{-algebra on } \mathcal{S}}} \mathcal{F}$$

(Notation: $\mathcal{B} = \sigma(\mathcal{M})$).

If $\mathcal{S}$ is a toplological or metric space, then most of the times $\mathcal{M}$ is chosen as class of all open sets of $\mathcal{S}$ and $\sigma(\mathcal{M})$ is called Borel σ-algebra (Notation: $\mathcal{B} = \mathcal{B}(\mathcal{S})$).

Example 1.1.1
1. If $\mathcal{S} = \mathbb{R}$, $\mathcal{B} = \mathcal{B}(\mathbb{R})$, then a random element X is called a *random variable*.
2. If $\mathcal{S} = \mathbb{R}^m$, $\mathcal{B} = \mathcal{B}(\mathbb{R}^m)$, $m > 1$, then X is called *random vector*. Random variables and random vectors are considered in the lectures „Elementare Wahrscheinlichkeitsrechnung und Statistik“ and „Stochastik I“.
3. Let $\mathcal{S}$ be the class of all closed sets of $\mathbb{R}^m$. Let

   $$\mathcal{M} = \{\{A \in \mathcal{S} : A \cap B \neq \emptyset\}, \quad B - \text{arbitrary compactum of } \mathbb{R}^m\}.$$

   Then $X : \Omega \to \mathcal{S}$ is a *random closed set*.

As an example we consider n independent equally distributed points $Y_1, \ldots, Y_n \in [0,1]^m$ and $R_1, \ldots, R_n > 0$ almost surely independent random variables, which are defined on the same probability space $(\Omega, \mathcal{A}, \mathsf{P})$ as $Y_1, \ldots, Y_n$. Consider $X = \cup_{i=1}^n B_{R_i}(Y_i)$. Obviously, this is a random set. An example of a realization is provided in Figure 1.1.

Exercise 1.1.1
Let $(\Omega, \mathcal{A})$ and $(\mathcal{S}, \mathcal{B})$ be measurable spaces, $\mathcal{B} = \sigma(\mathcal{M})$, where $\mathcal{M}$ is a class of subsets of $\mathcal{S}$. Proof that $X : \Omega \to \mathcal{S}$ is $\mathcal{A}|\mathcal{B}$-measurable, if and only if $X^{-1}(C) \in \mathcal{A}$, $C \in \mathcal{M}$.

Definition 1.1.2
Let T be an arbitrary index set and $(\mathcal{S}_t, \mathcal{B}_t)_{t \in T}$ a family of measurable spaces. A family $X = \{X(t), t \in T\}$ of random elements $X(t) : \Omega \to \mathcal{S}_t$ defined on $(\Omega, \mathcal{A}, \mathsf{P})$ and $\mathcal{A}|\mathcal{B}_t$-measurable for all $t \in T$ is called *random function* (associated with $(\mathcal{S}_t, \mathcal{B}_t)_{t \in T}$).

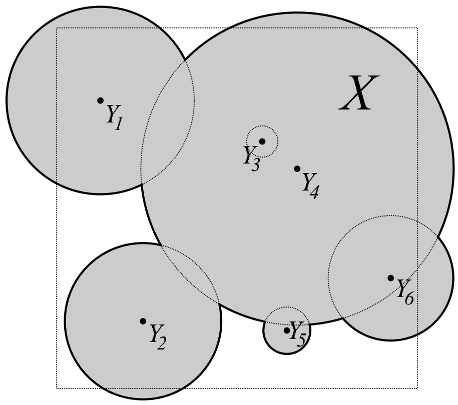

Abb. 1.1: Example of a random set $X = \cup_{i=1}^{6} B_{R_i}(Y_i)$

Therefore it holds $X : \Omega \times T \to (\mathcal{S}_t,\ t \in T)$, i.e. $X(\omega, t) \in \mathcal{S}_t$ for all $\omega \in \Omega$, $t \in T$ and $X(\cdot, t) \in \mathcal{A}|\mathcal{B}_t$, $t \in T$. A lot of times ω is discarded in the notation and we write $X(t)$ instead of $X(\omega, t)$. In most of the cases $(\mathcal{S}_t, \mathcal{B}_t)$ does not depend on $t \in T$ as well: $(\mathcal{S}_t, \mathcal{B}_t) = (\mathcal{S}, \mathcal{B})$ for all $t \in T$.

Special cases of random functions:

1. $T \subseteq \mathbb{Z} : X$ is called *random sequence* or *stochastic process in discrete time*.
   Example: $T = \mathbb{Z}$, $\mathbb{N}$.

2. $T \subseteq \mathbb{R} : X$ is called *stochastic process in continuous time*.
   Example: $T = \mathbb{R}_+$, $[a, b]$, $-\infty < a < b < \infty$, $\mathbb{R}$.

3. $T \subseteq \mathbb{R}^d, d \geq 2 : X$ is called *random field*.
   Example: $T = \mathbb{Z}^d$, $\mathbb{R}_+^d$, $\mathbb{R}^d$, $[a, b]^d$.

4. $T \subseteq \mathcal{B}(\mathbb{R}^d) : X$ is called *set-induced process*.
   If $X(t)$ is almost surely non-negative and σ-additive on the σ-algebra T, then X is called *random measure*.

The tradion of denoting the index set with T, arises from the interpretation of $t \in T$ for the cases 1 and 2 as *time parameter*.

For every $\omega \in \Omega$, $\{X(\omega, t),\ t \in T\}$ is called a *trajectory* or *path* of the random function X.

We want to proof that the random function $X = \{X(t),\ t \in T\}$ is a random element within the corresponding function space, which is equipped with a σ-algebra that now is specified.

Let $\mathcal{S}_T = \prod_{t \in T} \mathcal{S}_t$ be the cartesian product of $\mathcal{S}_t$, $t \in T$, i.e., $X \in \mathcal{S}_T$ if $X(t) \in \mathcal{S}_t$, $t \in T$. The *elementary cylindric set* in $\mathcal{S}_T$ is defined as

$$C_T(B_t) = \{X \in \mathcal{S}_t : X(t) \in B_t\},$$

where $t \in T$ is a selected point from T and $B_t \in \mathcal{B}_t$ a subset in $\mathcal{B}_t$. $C_T(B_t)$ therefore contains all trajectories X, which go through the „gate“ B_t, see Figure 1.2.

Definition 1.1.3
The *cylindric σ-algebra* $\mathcal{B}_T$ is introduced as a σ-algebra induced in $\mathcal{S}_T$ by the family of all

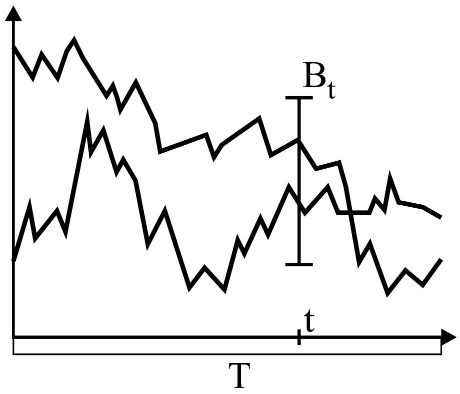

Abb. 1.2: Trajectories, which pass a „gate“ B_t.

elementary cylinders. They are labeled by $\mathcal{B}_T = \otimes_{t\in T}\mathcal{B}_t$. If $\mathcal{B}_t = \mathcal{B}$ for all $t \in T$, then $\mathcal{B}^T$ is written istead of $\mathcal{B}_T$.

Lemma 1.1.1
The family $\{X = X(t),\ t \in T\}$ is a random function on $(\Omega, \mathcal{A}, \mathsf{P})$ with phase spaces $(\mathcal{S}_t, \mathcal{B}_t)_{t\in T}$ if and only if for every $\omega \in \Omega$ the mapping $\omega \mapsto X(\omega, \cdot)$ is $\mathcal{A}|\mathcal{B}_T$-measurable.

Exercise 1.1.2
Proof lemma 1.1.1.

Definition 1.1.4
Let X be a random element: $X : \Omega \to \mathcal{S}$, i.e. X be $\mathcal{A}|\mathcal{B}$-measurable. The *distribution of* X is the probability measure P_X on $(\mathcal{S}, \mathcal{B})$, such that $\mathsf{P}_X(B) = \mathsf{P}(X^{-1}(B))$, $B \in \mathcal{B}$.

Lemma 1.1.2
An arbitrary probability measure μ on $(\mathcal{S}, \mathcal{B})$ can be considered as the distribution of a random element X.

Proof Take $\Omega = \mathcal{S}$, $\mathcal{A} = \mathcal{B}$, $\mathsf{P} = \mu$ and $X(\omega) = \omega$, $\omega \in \Omega$. □

When does a random function with given properties exist? A random function, which consists of independent random elements always exists. This assertion is known.

Theorem 1.1.1 (Lomnicki, Ulam):
Let $(\mathcal{S}_t, \mathcal{B}_t, \mu_t)_{t\in T}$ be a sequence of probability spaces. It exists a random sequence $X = \{X(t),\ t \in T\}$ on a probability space $(\Omega, \mathcal{A}, \mathsf{P})$ (associated with $(\mathcal{S}_t, \mathcal{B}_t)_{t\in T}$), such that

1. $X(t),\ t \in T$ are independent random elements.
2. $\mathsf{P}_{X(t)} = \mu_t$ on $(\mathcal{S}_t, \mathcal{B}_t)$, $t \in T$.

A lot of important random processes are built on the basis of independent random elements; cf. examples in section 1.2.

Definition 1.1.5
Let $X = \{X(t), t \in T\}$ be a random function on $(\Omega, \mathcal{A}, \mathsf{P})$ with phase space $(\mathcal{S}_t, \mathcal{B}_t)_{t\in T}$. The *finite-dimensional distributions of* X are defined as the distribution law $\mathsf{P}_{t_1,\ldots,t_n}$ of $(X(t_1), \ldots, X(t_n))^T$ on $(\mathcal{S}_{t_1,\ldots,t_n}, \mathcal{B}_{t_1,\ldots,t_n})$, for arbitrary $n \in \mathbb{N}$, $t_1, \ldots, t_n \in T$, where $\mathcal{S}_{t_1,\ldots,t_n} = \mathcal{S}_{t_1} \times \ldots \times \mathcal{S}_{t_n}$ and

$\mathcal{B}_{t_1,\dots,t_n} = \mathcal{B}_{t_1} \otimes \ldots \otimes \mathcal{B}_{t_n}$ is the σ-algebra in $\mathcal{S}_{t_1,\dots,t_n}$, which is induced by all sets $B_{t_1} \times \ldots \times B_{t_n}$, $B_{t_i} \in \mathcal{B}_{t_i}$, $i = 1, \dots, n$, i.e., $\mathsf{P}_{t_1,\dots,t_n}(C) = \mathsf{P}((X(t_1), \dots, X(t_n))^T \in C)$, $C \in \mathcal{B}_{t_1,\dots,t_n}$. In particular for $C = B_1 \times \ldots \times B_n$, $B_k \in \mathcal{B}_{t_k}$:

$$\mathsf{P}_{t_1,\dots,t_n}(B_1 \times \ldots \times B_n) = \mathsf{P}(X(t_1) \in B_1, \dots, X(t_n) \in B_n).$$

Exercise 1.1.3
Proof that $X_{t_1,\dots,t_n} = (X(t_1), \dots, X(t_n))^T$ is a $\mathcal{A}|\mathcal{B}_{t_1,\dots,t_n}$-measurable random element.

Definition 1.1.6
Let $\mathcal{S}_t = \mathbb{R}$ for all $t \in T$. The random function $X = \{X(t), t \in T\}$ is called *symmetric*, if all of its finite-dimensional distributions are symmetric probability measures, i.e., $\mathsf{P}_{t_1,\dots,t_n}(A) = \mathsf{P}_{t_1,\dots,t_n}(-A)$ for $A \in \mathcal{B}_{t_1,\dots,t_n}$ and all $n \in \mathbb{N}$, $t_1, \dots, t_n \in T$, whereby $\mathsf{P}_{t_1,\dots,t_n}(-A) = \mathsf{P}((-X(t_1), \dots, -X(t_n))^T$ $A)$.

Exercise 1.1.4
Proof that the finite-dimensional distributions of a random function X have the following properties: for arbitrary $n \in \mathbb{N}$, $n \geq 2$, $\{t_1, \dots, t_n\} \subset T$, $B_k \in \mathcal{S}_{t_k}$, $k = 1, \dots, n$ and an arbitrary permutation $(i_1, \dots, i_n)$ of $(1, \dots, n)$ it holds:

1. *Symmetry*: $\mathsf{P}_{t_1,\dots,t_n}(B_1 \times \ldots \times B_n) = \mathsf{P}_{t_{i_1},\dots,t_{i_n}}(B_{i_1} \times \ldots \times B_{i_n})$
2. *Consistency*: $\mathsf{P}_{t_1,\dots,t_n}(B_1 \times \ldots \times B_{n-1} \times \mathcal{S}_{t_n}) = \mathsf{P}_{t_1,\dots,t_{n-1}}(B_1 \times \ldots \times B_{n-1})$

The following theorem evidences that these properties are sufficient to proof the existence of a random function X with given finite-dimensional distributions.

Theorem 1.1.2 (Kolmogorov):
Let $\{\mathsf{P}_{t_1,\dots,t_n}, n \in \mathbb{N}, \{t_1, \dots, t_n\} \subset T\}$ be a family of probability measures on $(\mathbb{R}^m \times \ldots \times \mathbb{R}^m, \mathcal{B}(\mathbb{R}^m) \otimes \ldots \otimes \mathcal{B}(\mathbb{R}^m))$, which fulfill conditions 1 and 2 of exercise 1.1.4. Then there exists a random function $X = \{X(t), t \in T\}$ defined on a probability space $(\Omega, \mathcal{A}, \mathsf{P})$ with finite-dimensional distributions $\mathsf{P}_{t_1,\dots,t_n}$.

Proof See [13], section II.9. □

This theorem also holds on more general (however not arbitrary!) spaces than $\mathbb{R}^m$, on so-called *Borel spaces*, which in a sense are isomorphic to $([0,1], \mathcal{B}[0,1])$ or a subspace of that.

Definition 1.1.7
Let $X = \{X(t), t \in T\}$ be a random function with values in $(\mathcal{S}, \mathcal{B})$, i.e., $X(t) \in \mathcal{S}$ almost surely for arbitrary $t \in T$. X is called *measurable* if the mapping $X : (\omega, t) \mapsto X(\omega, t) \in \mathcal{S}$, $(\omega, t) \in \Omega \times T$, is $\mathcal{A} \otimes C|\mathcal{B}$-measurable.

Thus, definition 1.1.7 not only provides the measurability of X with respect to $\omega \in \Omega$: $X(\cdot, t) \in \mathcal{A}|\mathcal{B}$ for all $t \in T$, but $X(\cdot, \cdot) \in \mathcal{A} \otimes C|\mathcal{B}$ as a function of (ω, t). The measurability of X is of significance if $X(\omega, t)$ is considered at random moments $\tau : \Omega \to T$: $X(\omega, \tau(\omega))$. This in particular is the case in the theory of martingals if τ is a so-called stop time for X. Since the distribution of $X(\omega, \tau(\omega))$ might differ considerably from the distribution of $X(\omega, t)$, $t \in T$.

1.2 Elementary examples

The theorem of Kolmogorov can be used directly for the explicit construction only in few cases, since for a lot of random function their finite-dimensional distributions are not given explicitly. In this cases a new random function $X = \{X(t), t \in T\}$ is built as $X(t) = g(t, Y_1, Y_2, \ldots)$, $t \in T$, where g is a measurable function and $\{Y_n\}$ a sequence of random elements (also random functions), whose existence has already been ensured. For that we give several examples.

Let $X = \{X(t), t \in T\}$ be a real-valued random function with a probability space $(\Omega, \mathcal{A}, \mathsf{P})$.

1. *White noise*:

   Definition 1.2.1
   The random function $X = \{X(t), t \in T\}$ is called *white noise*, if all $X(t)$, $t \in T$, are independent and identically distributed (i.i.d.) random variables.

   White noise exists according to the theorem 1.1.1. It is used to depict the noise in (electromagnetic or acoustical) signals. If $X(t) \sim \mathrm{Ber}(p)$, $p \in (0,1)$, $t \in T$, one means *Salt-and-pepper noise*, the binary noise, which occurs at the transfer of binary data in computer-networks. If $X(t) \sim \mathcal{N}(0, \sigma^2)$, $\sigma^2 > 0$, $t \in T$, then X is called *Gaussian white noise*. It occures e.g. in acoustical signals.

2. *Gaussian random function*:

   Definition 1.2.2
   The random function $X = \{X(t), t \in T\}$ is called *Gaussian*, if all of its finite-dimensional distributions are Gaussian, i.e. for all $n \in \mathbb{N}$, $t_1, \ldots, t_n \subset T$ it holds

   $$X_{t_1,\ldots,t_n} = ((X(t_1), \ldots, X(t_n))^\top \sim \mathcal{N}(\mu_{t_1,\ldots,t_n}, \sum_{t_1,\ldots,t_n}),$$

   where the mean is given by $\mu_{t_1,\ldots,t_n} = (\mathsf{E}X(t_1), \ldots, \mathsf{E}X(t_n))^\top$ and the covariance matrix is given by $\sum_{t_1,\ldots,t_n} = ((\mathsf{cov}(X(t_i), X(t_j))_{i,j=1}^n$.

   Exercise 1.2.1
   Proof that the distribution of an Gaussian random function X is uniquely determined by its mean value function $\mu(t) = \mathsf{E}X(t), t \in T$, and covariance function $C(s,t) = \mathsf{E}[X(s)X(t)]$, $s, t \in T$, respectively.

   An example for a Gaussian process is the so-called *Wiener process* (or *Brownian motion*) $X = \{X(t), t \geq 0\}$, which has the expected value zero ($\mu(t) \equiv 0, t \geq 0$) and the covariance function $C(s,t) = \min\{s,t\}$, $s, t \geq 0$. Usually it is required addionally that the paths of X are continuous functions.

   We will investigate the regularity properties of the paths of random functions in more detail in section 1.3. Now we can say that such a process exists with probability one (with almost surely continuous trajectories.

   Exercise 1.2.2
   Proof that the Gaussian white noise is a Gaussian random function.

3. *Lognormal- and χ^2-functions*:
   The random function $X = \{X(t),\ t \in T\}$ is called *lognormal*, if $X(t) = e^{Y(t)}$, where $Y = \{Y(t), t \in T\}$ is a Gaussian random function. X is called *χ^2-function*, if $X(t) = \|Y(t)\|^2$,

where $Y = \{Y(t), t \in T\}$ is a Gaussian random function with values in $\mathbb{R}^n$, for which $Y(t) \sim \mathcal{N}(0, \mathrm{I})$, $t \in T$; here I is the $(n \times n)$-unit matrix. Then it holds that $X(t) \sim \chi^2_n$, $t \in T$.

4. *Cosine wave*:
   $X = \{X(t), t \in \mathbb{R}\}$ is defined by $X(t) = \sqrt{2}\cos(2\pi Y + tZ)$, where $Y \sim \mathcal{U}([0,1])$ and Z is a random variable, which is independent of Y.

   Exercise 1.2.3
   Let $X_1, X_2, \ldots$ be i.i.d. cosine waves. Determine the weak limit of the finite-dimensional distributions of the random function $\left\{\frac{1}{\sqrt{n}}\sum_{k=1}^n X_k(t),\ t \in \mathbb{R}\right\}$ for $n \to \infty$.

5. *Poisson process*:
   Let $\{Y_n\}_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables $Y_n \sim \mathrm{Exp}(\lambda), \lambda > 0$. The stochastic process $X = \{X(t), t \geq 0\}$ defined as $X(t) = \max\{n \in \mathbb{N} : \sum_{k=1}^n Y_k \leq t\}$ is called *Poisson process* with intensity $\lambda > 0$. $X(t)$ counts the number of certain events until the time $t > 0$, where the typical interval between two of these events is $\mathrm{Exp}(\lambda)$-distributed. These events can be a notification of claim, the record of an elementary particle in the Geiger counter, etc. Then $X(t)$ represents the number of damages or particles within the time interval $[0, t]$.

1.3 Regularity properties of trajectories

The theorem of Kolmogorov provides the existence of the distribution of a random function with given finite-dimensional distributions. However, it does not provide a statement about the properties of the paths of X. This is understandable since all random objects are defined in the almost surely sense(a.s.) in probability theory, with the exception of a set $A \subset \Omega$ with $\mathsf{P}(A) = 0$.

Example 1.3.1
Let $(\Omega, \mathcal{A}, \mathsf{P}) = ([0,1], \mathcal{B}([0,1]), \nu_1)$, where ν_1 is the Lebesgue measure on $[0,1]$. We define $\{X = X(t),\ t \in [0,1]\}$ by $X(t) \equiv 0$, $t \in [0,1]$ and $Y = \{Y(t),\ t \in [0,1]\}$ by

$$Y(t) = \begin{cases} 1, & t = U, \\ 0, & \text{sonst}, \end{cases}$$

where $U(\omega) = \omega$, $\omega \in [0,1]$, is a $\mathcal{U}([0,1])$-distributed random variable defined on $(\Omega, \mathcal{A}, \mathsf{P})$. Since $\mathsf{P}(Y(t) = 0) = 1$, $t \in T$, since $\mathsf{P}(U = t) = 0$, $t \in T$, it is clear that $X \stackrel{d}{=} Y$. Nevertheless, X and Y have different path properties since X has continuous and Y has volatile trajectories, and $\mathsf{P}(X(t) = 0,\ \forall t \in T) = 1$, where $\mathsf{P}(Y(t) = 0,\ \forall t \in T) = 0$.

It may be that the „set of exceptions“ A (see above) is very different for $X(t)$ for every $t \in T$. Therefore, we require that all $X(t)$, $t \in T$, are defined simultaneously on a subset $\Omega_0 \subseteq \mathcal{N}$ with $\mathsf{P}(\Omega_0) = 1$. The defined random function $\tilde{X} : \Omega_0 \times T \to \mathbb{R}$ is called *modification* of $X : \Omega \times T \to \mathbb{R}$. X and $\tilde{X}$ differ on a set Ω/Ω_0 with probability zero. Therefore we indicate later when stating that „random function X possesses a property C “ that it exists a modification of X with this property C.

Definition 1.3.1
The random functions $X = \{X(t),\ t \in T\}$ and $Y = \{Y(t),\ t \in T\}$ defined on the same proba-

bility space $(\Omega, \mathcal{A}, \mathsf{P})$ are called *(stochastically) equivalent*, if

$$B_t = \{\omega \in \Omega : X(\omega, t) \neq Y(\omega, t)\} \in A,\ \ t \in T,$$

and $\mathsf{P}(B_t) = 0,\ \ t \in T$.

We also can say, that X and Y are versions of one and the same random function. It is clear, that all modifications (or versions) of X are equivalent to Y.

Exercise 1.3.1
Proof that the random functions X and Y in example 1.3.1 are stochastically equivalent.

Definition 1.3.2
The random functions $X = \{X(t),\ t \in T\}$ and $Y = \{Y(t),\ t \in T\}$ (not necessarily defined on the same probability space) are called *equivalent in distribution*, if $\mathsf{P}_X = \mathsf{P}_Y$ on $(\mathcal{S}_t, \mathcal{B}_t)$. Notation: $X \stackrel{d}{=} Y$.

According to the theorem 1.1.2 it is sufficient for the equivalence in distribution of X and Y, if they possess the same finite-dimensional distributions. It is clear that stochastic equivalence implies equivalence in distribution, but not the other way around.

Definition 1.3.3
The random functions $X = \{X(t),\ t \in T\}$ and $Y = \{Y(t),\ t \in T\}$ defined on the same probability space $(\Omega, \mathcal{A}, \mathsf{P})$ associated with $(\mathcal{S}_t, \mathcal{B}_t)_{t\in T}$ have *equivalent trajectories* (or are called *stochastically indistinguishable*), if

$$A = \{\omega \in \Omega : X(\omega, t) \neq Y(\omega, t) \text{ for a } t \in T\} \in \mathcal{A}$$

and $\mathsf{P}(A) = 0$.

This term implies that X and Y have paths, which coincide with probability one. If the space $(\Omega, \mathcal{A}, \mathsf{P})$ is *complete* (i.e. the implication of $A \in \mathcal{A} : \mathsf{P}(A) = 0$ is for all $B \subset A$: $B \in \mathcal{A}$ (and then $\mathsf{P}(B) = 0$)), then the indistinguishable processes are stochastically equivalent.

Now, let T and $\mathcal{S}$ be *Banach spaces* with norms $|\cdot|_T$ and $|\cdot|_{\mathcal{S}}$, respectively. The random function $X = \{X(t),\ t \in T\}$ is now defined on $(\Omega, \mathcal{A}, \mathsf{P})$ with values in $(\mathcal{S}, \mathcal{B})$.

Definition 1.3.4
The random function $X = \{X(t),\ t \in T\}$ is called

a) *stochastically continuous on* T, if $X(s) \xrightarrow[s\to t]{\mathsf{P}} X(t)$, for arbitrary $t \in T$, i.e.

$$\mathsf{P}(|X(s) - X(t)|_{\mathcal{S}} > \varepsilon) \xrightarrow[s\to t]{} 0,\ \text{ for all } \varepsilon > 0.$$

b) L^p*-continuous on* T, $p \geq 1$, if $X(s) \xrightarrow[s\to t]{L^p} X(t)$, $t \in T$, i.e. $\mathsf{E}|X(s) - X(t)|^p \xrightarrow[s\to t]{} 0$. For $p = 2$ the specific notation „continuity in the squared mean "is used.

c) *a.s. continuous on* T, if $X(s) \xrightarrow[s\to t]{f.s.} X(t)$, $t \in T$, i.e., $\mathsf{P}(X(s) \xrightarrow[s\to t]{} X(t)) = 1$, $t \in T$.

d) *continuous*, if all trajectories of X are continuous functions.

In applications one is interested in the cases c) and d), although the weakest form of continuity is the stochastic continuity.

$$\boxed{L^p\text{-continuity}} \Longrightarrow \boxed{\text{stochastic continuity}} \Longleftarrow \boxed{\text{a.s. continuity}} \Longleftarrow \boxed{\text{continuity of all paths}}$$

Why are cases c) and d) important? Let's consider an example.

Example 1.3.2
Let $T = [0,1]$ and $(\Omega, \mathcal{A}, \mathsf{P})$ be an canonical probability space with $\Omega = \mathbb{R}^{[0,1]}$, i.e. $\Omega = \prod_{t\in[0,1]} \mathbb{R}$. Let $X = \{X(t),\ t \in [0,1]\}$ be a stochastic process on $(\Omega, \mathcal{A}, \mathsf{P})$. Not all events are elements of A, like e.g. $A = \{\omega \in \Omega : X(\omega,t) = 0 \text{ for all } t \in [0,1]\} = \cap_{t\in[0,1]} \{X(\omega,t) = 0\}$, since this is an intersection of measurable events from $\mathcal{A}$ in uncountable number. If however X is continuous, then all of its paths are continuous functions and one can depict $A = \cap_{t\in D} \{X(\omega,t) = 0\}$, where D is a dense countable subset of $[0,1]$, e.g., $D = \mathbb{Q} \cap [0,1]$. Then it holds that $A \in \mathcal{A}$.

However, in many applications (like e.g. in financial mathematics) it is not realistic to consider stochastic processes with continuous paths as models for real phenomena. Therefore, a bigger class of possible trajectories of X is allowed: the so-called *càdlàg-class* (càdlàg = continue à droite, limitée à gonche (fr.)).

Definition 1.3.5
A stochastic process $X = \{X(t),\ t \in \mathbb{R}\}$ is called *càdlàg*, if all of its trajectories are right-sided continuous functions, which have left-sided limits.

Now, we want to consider the properties of the notations of continuity introduced above in more detail. One can note e.g., that the stochastic continuity is a property of the two-dimensional distribution $\mathsf{P}_{s,t}$ of X. This is shown by the following lemma.

Lemma 1.3.1
Let $X = \{X(t),\ t \in T\}$ be a random function associated with $(\mathcal{S}, \mathcal{B})$, where $\mathcal{S}$ and T are Banach spaces. The following statements are equivalent:

a) $X(s) \xrightarrow[s\to t_0]{\mathsf{P}} Y$,

b) $\mathsf{P}_{s,t} \xrightarrow[s,t\to t_0]{d} \mathsf{P}_{(Y,Y)}$,

where $t_0 \in T$ and Y is a $\mathcal{B}$-random element. For the stochastic continuity of X, one should choose $t_0 \in T$ arbitrarily and $Y = X(t_0)$.

Proof *a) ⇒ b)*
$X(s) \xrightarrow[s\to t_0]{\mathsf{P}} Y$ means $(X(s), X(t))^\top \xrightarrow[s,t\to t_0]{\mathsf{P}} (Y,Y)^\top$. This results in $\mathsf{P}_{s,t} \overset{d}{\to} \mathsf{P}_{(Y,Y)}$, since $\overset{\mathsf{P}}{\to}$-convergence is stricter than $\overset{d}{\to}$-convergence.

b) ⇒ a)
For arbitrary $\varepsilon > 0$ we consider a continuous function $g_\varepsilon : \mathbb{R} \to [0,1]$ with $g_\varepsilon(0) = 0$, $g_\varepsilon(x) = 1$, $x \notin B_\varepsilon(0)$. It holds for all $s,t \in T$ that

$$\mathsf{E}g_\varepsilon(|X(s) - X(t)|_{\mathcal{S}}) = \mathsf{P}(|X(s) - X(t)|_{\mathcal{S}} > \varepsilon) + \mathsf{E}(g_\varepsilon(|X(s) - X(t)|_{\mathcal{S}})\mathsf{E}(|X(s) - X(t)|_{\mathcal{S}} \leq \varepsilon)),$$

hence $\mathsf{P}(|X(s)-X(t)|_{\mathcal{S}} > \varepsilon) \leq \mathsf{E}g_\varepsilon(|X(s)-X(t)|_{\mathcal{S}}) = \int_{\mathcal{S}}\int_{\mathcal{S}} g_\varepsilon(|x-y|_{\mathcal{S}})\mathsf{P}_{s,t}(d(x,y)) \underset{\substack{s\to t_0\\ t\to t_0}}{\longrightarrow}$ $\int_{\mathcal{S}}\int_{\mathcal{S}} g_\varepsilon(|x-y|_{\mathcal{S}})\mathsf{P}_{(Y,Y)}(d(x,y)) = 0$, since $\mathsf{P}_{(Y,Y)}$ is concentrated on $\{(x,y)\in\mathcal{S}^2 : x=y\}$ and $g_\varepsilon(0)=0$. Thus $\{X(s)\}_{s\to t_0}$ is a fundamental sequence (in probability), therefore $X(s) \xrightarrow[s\to t_0]{\mathsf{P}} Y$. □

It may be that X is continuous, although all of the paths of X have jumps, i.e. X cannot possess any a.s. continuous modification. The descriptive explanation for that is that such X may have a jump for concrete $t\in T$ with probability zero. Therefore jumps of the paths of X always occur at different locations.

Exercise 1.3.2
Proof that the Poisson process is stochastically continuous, although it does not possess any a.s. continuous modifications.

Exercise 1.3.3
Let T be compact. Proof that if X is stochastically continuous on T, then it also is uniformly stochastically continuous, i.e., for all $\varepsilon,\eta>0$ $\exists\delta>0$, such that for all $s,t\in T$ with $|s-t|_T<\delta$ it holds that $\mathsf{P}(|X(s)-X(t)|_{\mathcal{S}}>\varepsilon)<\eta$.

Now let $\mathcal{S}=\mathbb{R}$, $\mathsf{E}X^2(t)<\infty$, $t\in T$, $\mathsf{E}X(t)=0$, $t\in T$. Let $C(s,t)=\mathsf{E}\left[X(s)X(t)\right]$ be the covariance function of X.

Lemma 1.3.2
For all $t_0\in T$ and a random variable Y with $\mathsf{E}Y^2<\infty$ the following assertions are equivalent:

a) $X(s) \xrightarrow[s\to t_0]{L^2} Y$

b) $C(s,t) \xrightarrow[s,t\to t_0]{} \mathsf{E}Y^2$

Proof *a) ⇒ b)*
The assertion results from the Cauchy-Schwarz inequality:

$$\begin{aligned}
|C(s,t)-\mathsf{E}Y^2| &= |\mathsf{E}(X(s)X(t))-\mathsf{E}Y^2)| = |\mathsf{E}\left[(X(s)-Y+Y)(X(t)-Y+Y)\right]-\mathsf{E}Y^2| \\
&\leq \mathsf{E}|(X(s)-Y)(X(t)-Y)| + \mathsf{E}|(X(s)-Y)Y| + \mathsf{E}|(X(t)-Y)Y| \\
&\leq \sqrt{\underbrace{\mathsf{E}(X(s)-Y)^2 \quad \mathsf{E}(X(t)-Y)^2}_{||X(s)-Y||^2_{L^2}\cdot||X(t)-Y||^2_{L^2}}} \\
&\quad + \sqrt{\mathsf{E}Y^2\,\underbrace{\mathsf{E}(X(s)-Y)^2}_{||X(s)-Y||^2_{L^2}}} + \sqrt{\mathsf{E}Y^2\,\underbrace{\mathsf{E}(X(t)-Y)^2}_{||X(t)-Y||^2_{L^2}}} \xrightarrow[s,t\to t_0]{} 0
\end{aligned}$$

with assumption a).

b) ⇒ a)

$$\begin{aligned}
\mathsf{E}(X(s)-X(t))^2 &= \mathsf{E}(X(s))^2 - 2\mathsf{E}[X(s)X(t)] + \mathsf{E}(X(t))^2 \\
&= C(s,s)+C(t,t)-2C(s,t) \xrightarrow[s,t\to t_0]{} 2\mathsf{E}Y^2 - 2\mathsf{E}Y^2 = 0.
\end{aligned}$$

Thus, $\{X(s),\ s\to t_0\}$ is a fundamental sequence in the L^2-sense, and we get $X(s) \xrightarrow[s\to t_0]{L^2} Y$. □

A random function X, which is continuous in the mean-squared sense, may still have uncontinuous trajectories. In most of the cases which are practically relevant, X however has an a.s. continuous modification. Later on this will become more precise by stating a theorem.

Conclusion 1.3.1
The random function X, which satisfies the conditions of the lamma 1.3.2, is continuous on T in the mean-squared sense if and only if its covariance function $C : T^2 \to \mathbb{R}$ is continuous on the diagonal $\operatorname{diag} T^2 = \{(s,t) \in T^2 : s = t\}$, i.e., $\lim_{s,t\to t_0} C(s,t) = C(t)$ for all $t_0 \in T$.

Proof Choose $Y = X(t_0)$ in lemma 1.3.2. □

Remark 1.3.1
If X is not centered, then the continuity of $\mu(\cdot)$ together with the continuity of C on $\operatorname{diag} T^2$ is required to ensure the L^2-continuity of X on T.

Exercise 1.3.4
Give an example of a stochastic process with a.s. uncontinuous trajectories, which is L^2-continuous.

Now we consider the property of (a.s.) continuity in more detail. Like mentioned before, we merely can talk about continuous modification or version of a process. The possibility to possess such a version also depends on the properties of the two-dimensional distributions of the process. This is proven by the following theorem (originally proven by A. Kolmogorov).

Theorem 1.3.1
Let $X = \{X(t),\ t \in [a,b]\}$, $-\infty < a < b \le +\infty$. A real-valued stochastic process X has a continuous version, if there exist constants $\alpha, c, \delta > 0$ such that

$$\mathsf{E}|X(t+h) - X(t)|^\alpha < c|h|^{1+\delta},\ t \in (a,b), \tag{1.3.1}$$

for sufficiently small $|h|$.

Proof See, e.g. [7], theorem 2.23. □

Now we turn to processes with càdlàg-trajectories. Let $(\Omega, \mathcal{A}, \mathsf{P})$ be a complete probability space.

Theorem 1.3.2
Let $X = \{X(t),\ t \ge 0\}$ be a real-valued stochastic process and D a countable dense subset of $[0,\infty)$. If

a) X is stochastically right-sided continuous, i.e., $X(t+h) \xrightarrow[h\to+0]{\mathsf{P}} X(t)$, $t \in [0,+\infty)$,

b) the trajectories of X for every $t \in D$ have finite right- and left-sided limits, i.e., $\exists \lim_{h\to\pm 0} X(t+h)$, $t \in D$ a.s.,

then X has a version with a.s. càdlàg-paths.

Without proof.

Lemma 1.3.3
Let $X = \{X(t),\ t \ge 0\}$ and $\{Y = Y(t),\ t \ge 0\}$ be two versions of a random function, both defined on the probability space $(\Omega, \mathcal{A}, \mathsf{P})$, with property that X and Y have a.s. right-sided continuous trajectories. Then X and Y are indistinguishable.

Proof Let Ω_X, Ω_Y be „sets of exception", for which the trajectories of X and Y, respectively are not right-sided continuous. It holds that $\mathsf{P}(\Omega_X) = \mathsf{P}(\Omega_Y) = 0$. Consider $A_t = \{\omega \in \Omega : X(\omega, t) \neq Y(\omega, t)\}$, $t \in [0, +\infty)$ and $A = \cup_{t \in \mathbb{Q}_+} A_t$, where $\mathbb{Q}_+ = \mathbb{Q} \cap [0, +\infty)$. Since X and Y are stochastically equivalent, it holds that $\mathsf{P}(A) = 0$ and therefore $P(\tilde{A}) = \mathsf{P}(A \cup \Omega_X \cup \Omega_Y) \leq \mathsf{P}(A) + \mathsf{P}(\Omega_X) + \mathsf{P}(\Omega_Y) = 0$, where $\tilde{A} = A \cup \Omega_X \cup \Omega_Y$. Therefore $X(\omega, t) = Y(\omega, t)$ holds for $t \in \mathbb{Q}_+$ and $\omega \in \Omega \setminus \tilde{A}$. Now, we proof this for all $t \geq 0$. For arbitrary $t \geq 0$ a sequence $\{t_n\} \subset \mathbb{Q}_+$ exists, such that $t_n \downarrow t$. Since $X(\omega, t_n) = Y(\omega, t_n)$ for all $n \in \mathbb{N}$ and $\omega \in \Omega \setminus \tilde{A}$, it holds that $X(\omega, t) = \lim_{n\to\infty} X(\omega, t_n) = \lim_{n\to\infty} Y(\ omega, t_n) = Y(\omega, t)$ for $t \geq 0$ and $\omega \in \Omega \setminus \tilde{A}$. Therefore X and Y are indistinguishable. □

Conclusion 1.3.2
If càdlàg-processes $X = \{X(t),\ t \geq 0\}$ and $Y = \{Y(t),\ t \geq 0\}$ are versions of a random function, then they are indistinguishable.

1.4 Differentiability of trajectories

Let T be a linear normed space.

Definition 1.4.1
A real-valued random function $X = \{X(t),\ t \in T\}$ is *differentiable on T in direction $h \in T$ stochastically, in the L^p-sense, $p \geq 1$*, or *a.s.*, if

$$\lim_{l \to 0} \frac{X(t+hl) - X(t)}{l} = X_h'(t),\ t \in T$$

exists in the corresponding sense, namely stochastically, in the L^p-space or a.s..

The lemmata 1.3.2 - 1.3.3 show, that the stochastic differentiability is a property that is determined by three-dimensional distributions of X (because the common distribution of $\frac{X(t+hl)-X(t)}{l}$ and $\frac{X(t+hl')-X(t)}{l'}$ should converge weakly), whereas the differentiability in the mean-squared sense is determined by the smoothness of the covariance function $C(s,t)$.

Exercise 1.4.1
Show that

1. the Wiener process is not stochastically differentiable on $[0, \infty)$.

2. the Poisson process is stochastically differentiable on $[0, \infty)$, however not in the L^p-mean, $p \geq 1$.

Lemma 1.4.1
A centered random function $X = \{X(t),\ t \in T\}$ (i.e., $\mathsf{E}X(t) \equiv 0$, $t \in T$) is L^2-differentiable in $t \in T$ in direction $h \in T$ if its covariance function C is differentiable twice in (t,t) in direction h, i.e., if $C''_{hh}(t,t) = \left.\frac{\partial^2 C(s,t)}{\partial s_h \partial t_h}\right|_{s=t}$. $X_h'(t)$ is L^2-continuous in $t \in T$ if $C''_{hh}(s,t) = \frac{\partial^2 C(s,t)}{\partial s_h \partial t_h}$ is continuous in $s = t$. Therefore $C''_{hh}(s,t)$ is the covariance function of $X_h' = \{X_h'(t),\ t \in T\}$.

Proof According to 1.3.3 it is enough to show that

$$I = \lim_{l,l' \to 0} \mathsf{E}\left(\frac{X(t+lh) - X(t)}{l} \cdot \frac{X(s+l'h) - X(s)}{l'}\right)$$

exists for $s = t$. Indeed we recieve

$$\begin{aligned} I &= \frac{1}{ll'}\Big(C(t+lh, s+l'h) - C(t+lh, s) - C(t, s+l'h) + C(t,s)\Big) \\ &= \frac{1}{l}\left(\frac{C(t+lh, s+l'h) - C(t+lh, s)}{l'} - \frac{C(t, s+l'h) - C(t,s)}{l'}\right) \xrightarrow[l,l'\to 0]{} C''_{hh}(s,t)\,. \end{aligned}$$

All other statements of the lemma result from this relation. □

Remark 1.4.1
The properties of the L^2-differentiability and a.s. differentiability of random functions are defined in the following sense: there are stochastic processes that have L^2-differentiable paths, although they are a.s. uncontinuous, and vice versa, processes with a.s. differentiable paths are not always L^2-differentiable, since e.g. the first derivation of their covariance function is not continuous.

Exercise 1.4.2
Give appropriate examples!

1.5 Moments und covariance

Let $X = \{X(t),\ t \in T\}$ be a random function that is real-valued, and let T be an arbitrary index space.

Definition 1.5.1
The *mixed moment* $\mu^{(j_1,\dots,j_n)}(t_1,\dots,t_n)$ of X of order $(j_1,\dots,j_n) \in \mathbb{N}^n$, $t_1,\dots,t_n \in T$ is given by $\mu^{(j_1,\dots,j_n)}(t_1,\dots,t_n) = \mathsf{E}\left[X^{j_1}(t_1)\cdot\ldots\cdot X^{j_n}(t_n)\right]$, where it is required that the expected value exists and that it is finite. Then it is sufficient to assume that $\mathsf{E}|X(t)|^j < \infty$ for all $t \in T$ and $j = j_1 + \ldots + j_n$.

Important special cases:

1. $\mu(t) = \mu^{(1)}(t) = \mathsf{E}X(t)$, $t \in T$ – *mean value function of* X.
2. $\mu^{(1,1)}(s,t) = \mathsf{E}\left[X(s)X(t)\right] = C(s,t)$ – *(non-centered) covariance function of* X. Whereas the *centered covariance function* is: $K(s,t) = \mathsf{cov}((X(s), X(t)) = \mu^{(1,1)}(s,t) - \mu(s)\mu(t)$, $s,t \in T$.

Exercise 1.5.1
Show that the centered covariance function of a real-valued random function X

1. is *symmetric*, i.e., $K(s,t) = K(t,s)$, $s,t \in T$.
2. is *positive semidefinite*, i.e., for $n \in \mathbb{N}$, $t_1,\dots,t_n \in T$, $z_1,\dots,z_n \in \mathbb{R}$ it holds that
$$\sum_{i,j=1}^{n} K(t_i,t_j) z_i z_j \geq 0.$$
3. satisfies $K(t,t) = \mathsf{var}\, X(t)$, $t \in T$.

Property 2) also holds for the non-centered covariance function $C(s,t)$.

The mean value function $\mu(t)$ shows a (non random) tendency. If $\mu(t)$ is known, the random function X can be centered by considering a random function $Y = \{Y(t),\ t \in T\}$ with $Y(t) = X(t) - \mu(t)$, $t \in T$.

The covariance function $K(s,t)$ and $C(s,t)$, respectively contains informations about the dependence structer X. Sometimes the correlation function $R(s,t) = \frac{K(s,t)}{\sqrt{K(s,s)K(t,t)}}$ for all $s,t \in T$: $K(s,s) = \mathsf{var}\, X(s) > 0$, $K(t,t) = \mathsf{var}\, X(t) > 0$ is used instead of K and C, respectively. Because of the Cauchy-Schwarz inequality it holds that $|R(s,t)| \leq 1$, $s,t \in T$. The set of all mixed moments in general does not (uniquely) determine the distribution of a random function.

Exercise 1.5.2
Give an example of different random functions $X = \{X(t),\ t \in T\}$ und $Y = \{Y(t),\ t \in T\}$, for which holds that $\mathsf{E}X(t) = \mathsf{E}Y(t)$, $t \in T$ and $\mathsf{E}(X(s)X(t)) = \mathsf{E}(Y(s)Y(t))$, $s,t \in T$.

Exercise 1.5.3
Let $\mu : T \to \mathbb{R}$ be a random function and $K : T \times T \to \mathbb{R}$ be a positive semidefinite symmetric function. Proof that a random function $X = \{X(t),\ t \in T\}$ exists with $\mathsf{E}X(t) = \mu(t)$, $\mathsf{cov}(X(s), X(t)) = C(s,t)$, $s,t \in T$.

Lent now $X = \{X(t),\ t \in T\}$ be a real-valued random function with $\mathsf{E}\,|X(t)|^k < \infty$, $t \in T$, for a $k \in \mathbb{N}$.

Definition 1.5.2
The *mean increment of order k* of X is given by $\gamma_k(s,t) = \mathsf{E}(X(s) - X(t))^k$, $s,t \in T$.

Special attention is paid to the function $\gamma(s,t) = \frac{1}{2}\gamma_2(s,t) = \frac{1}{2}\mathsf{E}(X(s) - X(t))^2$, $s,t \in T$, which is called *variogram of X*. In geostatistics the variogram is often used instead of the covariance function. A lot of times we discard the condition $\mathsf{E}X^2(t) < \infty$, $t \in T$, instead we assume that $\gamma(s,t) < \infty$ for all $s,t \in T$.

Exercise 1.5.4
Proof that there exist random functions without finite second moments with $\gamma(s,t) < \infty$, $s,t \in T$ gibt.

Exercise 1.5.5
Show that for a random function $X = \{X(t),\ t \in T\}$ with mean value function μ and covariance function K it holds that:

$$\gamma(s,t) = \frac{K(s,s) + K(t,t)}{2} - K(s,t) + \frac{1}{2}(\mu(s) - \mu(t))^2, \quad s,t \in T.$$

If the random function X is *complex-valued*, i.e., $X : \Omega \times T \to \mathbb{C}$, with $\mathsf{E}\,|X(t)|^2 < \infty$, $t \in T$, then the covariance function of X is introduced as $K(s,t) = \mathsf{E}(X(s) - \mathsf{E}X(s))\overline{(X(t) - \mathsf{E}X(t))}$, $s,t \in T$, where $\overline{z}$ is the complex conjugate of $z \in \mathbb{C}$. Then it holds that $K(s,t) = \overline{K(t,s)}$, $s,t \in T$, and K is positive semidefinite, i.e, for all $n \in \mathbb{N}$, $t_1, \ldots, t_n \in T$, $z_1, \ldots, z_n \in \mathbb{C}$ it holds that $\sum_{i,j=1}^n K(t_i,t_j) z_i \overline{z_j} \geq 0$.

1.6 Stationarity and Independence

T be a subset of the linear vector space with operations $+$, $-$ over space $\mathbb{R}$.

Definition 1.6.1
The random function $X = \{X(t),\ t \in T\}$ is called *stationary* (strict sense stationary) if for all

$n \in \mathbb{N}$, $h, t_1, \ldots, t_n \in T$ with $t_1 + h, \ldots, t_n + h \in T$ it holds that:

$$\mathsf{P}_{(X(t_1),\ldots,X(t_n))} = \mathsf{P}_{(X(t_1+h),\ldots,X(t_n+h))},$$

i.e., all finite-dimensional distributions of X are invariant against translations in T.

Definition 1.6.2
A (complex-valued) random function $X = \{X(t),\ t \in T\}$ is called *second-order stationary* (or *wide sense stationary*) if $\mathsf{E}|X(t)|^2 < \infty$, $t \in T$, and $\mu(t) \equiv \mathsf{E}X(t) \equiv \mu$, $t \in T$, $K(s,t) = \mathsf{cov}(X(s), X(t)) = K(s+h, t+h)$ for all $h, s, t \in T : s+h, t+h \in T$.

If X is second-order stationary, it is beneficial to introduce a function $K(t) := K(0,t)$, $t \in T$ whereby $0 \in T$.

Strict sense stationarity and wide sense stationarity do not result from each other. However it is clear that if a complex-valued function is strict sense stationary and possesses finite second-order moments, then the function also is second-order stationary.

Definition 1.6.3
A real-valued random function $X = \{X(t),\ t \in T\}$ is *intrinsic second-order stationary* if $\gamma_k(s,t)$, $s, t \in T$ exist for $k \leq 2$, and for all $s, t, h \in T$, $s+h, t+h \in T$ it holds that $\gamma_1(s,t) = 0$, $\gamma_2(s,t) = \gamma_2(s+h, t+h)$.

For real-valued random functions, intrinsic second-order stationarity is more general than second-order stationarity since the existence of $\mathsf{E}|X(t)|^2$, $t \in T$ is not required.

The analogue of the stationarity of increments of X also exists in strict sense.

Definition 1.6.4
Let $X = \{X(t),\ t \in T\}$ be a real-valued stochastic process, $T \subset \mathbb{R}$. It is said that X

1. possesses *stationary increments* if for all $n \in \mathbb{N}$, $h, t_0, t_1, t_2, \ldots, t_n \in T$, with $t_0 < t_1 < t_2 < \ldots < t_n$, $t_i + h \in T$, $i = 0, \ldots, n$ the distribution of $(X(t_1+h) - X(t_0+h), \ldots, X(t_n+h) - X(t_{n-1}+h))^\top$ does not depend on h.

2. possesses *independent increments* if for all $n \in \mathbb{N}$, $t_0, t_1, \ldots, t_n \in T$ with $t_0 < t_1 < \ldots < t_n$ the random variables $X(t_0), X(t_1) - X(t_0), \ldots, X(t_n) - X(t_{n-1})$ are pairwise independent.

Let $(\mathcal{S}_1, \mathcal{B}_1)$ and $(\mathcal{S}_2, \mathcal{B}_2)$ be measurable spaces. In general it is said that two random elements $X : \Omega \to \mathcal{S}_1$ and $X : \Omega \to \mathcal{S}_2$ are *independent* on the same probability space $(\Omega, \mathcal{A}, \mathsf{P})$ if $\mathsf{P}(X \in A_1, Y \in A_2) = \mathsf{P}(X \in A_1)\mathsf{P}(Y \in A_2)$ for all $A_1 \in \mathcal{B}_1$, $A_2 \in \mathcal{B}_2$.

This definition can be applied to the independence of random functions X and Y with phase space $(\mathcal{S}_T, \mathcal{B}_T)$, since they can be considered as random elements with $\mathcal{S}_1 = \mathcal{S}_2 = \mathcal{S}_T$, $\mathcal{B}_1 = \mathcal{B}_2 = \mathcal{B}_T$ (cf. lemma 1.1.1). The same holds for the independence of a random element (or a random function) X and of a sub-σ-algebra $\mathcal{G} \in \mathcal{A}$: this is the case if $\mathsf{P}(\{X \in A\} \cap G) = \mathsf{P}(X \in A)\mathsf{P}(G)$, for all $A \in \mathcal{B}_1$, $G \in \mathcal{G}$ (or $A \in \mathcal{B}_T$, $G \in \mathcal{G}$).

1.7 Processes with independent increments

In this section we want to talk about the properties and existence of processes with independent increments.

Let $\{\varphi_{s,t},\ s, t \geq 0\}$ be a family of characteristic functions of probability measures $Q_{s,t},\ s, t \geq 0$ on $\mathcal{B}(\mathbb{R})$, i.e., for $z \in \mathbb{R}$, $s, t \geq 0$ it holds that $\varphi_{s,t}(z) = \int_{\mathbb{R}} e^{izx} Q_{s,t}(dx)$.

Theorem 1.7.1
It exists a stochastic process $X = \{X(t),\ t \geq 0\}$ with independent increments with the property that for all $s, t \geq 0$ the characteristic function of $X(t) - X(s)$ is equal to $\varphi_{s,t}$ if and only if

$$\varphi_{s,t} = \varphi_{s,u}\varphi_{u,t} \tag{1.7.1}$$

for all $0 \leq s < u < t < \infty$. Thereby the distribution of $X(0)$ can be chosen arbitrarily.

Proof The necessity of the condition 1.7.1 is clear since for all $s \in (0, \infty) : s < u < t$ it holds that: $X(t) - X(s) = \underbrace{X(t) - X(u)}_{Y_1} + \underbrace{X(u) - X(s)}_{Y_2}$ and $X(t) - X(u)$ and $X(u) - X(s)$ are pairwise independent. Then it holds $\varphi_{s,t} = \varphi_{Y_1+Y_2} = \varphi_{Y_1}\varphi_{Y_2} = \varphi_{s,u}\varphi_{u,t}$.
Now we proof the sufficiency.
If the existence of a process X with independent increments and property $\varphi_{X(t)-X(s)} = \varphi_{s,t}$ on a probability space $(\Omega, \mathcal{A}, \mathsf{P})$ had already been proven, one could declare the characteristic functions of all of its finite-dimensional distributions by $\{\varphi_{s,t}\}$.
Let $n \in \mathbb{N}$, $0 = t_0 < t_1 < \ldots < t_n < \infty$ and $Y = (X(t_0), X(t_1) - X(t_0), \ldots, X(t_n) - X(t_{n-1}))^\top$.
The independence of increments results in

$$\varphi_Y(\underbrace{z_0, z_1, \ldots, z_n}_{z}) = \mathsf{E}e^{i\langle z,Y\rangle} = \varphi_{X(t_0)}(z_0)\varphi_{t_0,t_1}(z_1)\ldots\varphi_{t_{n-1},t_n}(z_n),\ z \in \mathbb{R}^{n+1},$$

where the distribution of $X(t_0)$ is an arbitrary probability measure Q_0 on $\mathcal{B}(\mathbb{R})$. For $X_{t_0,\ldots,t_n} = (X(t_0), X(t_1), \ldots, X(t_n))^\top$ however it holds that $X_{t_0,\ldots,t_n} = AY$, where

$$A = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 1 & 1 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 1 & 1 & \ldots & 1 \end{pmatrix}.$$

Then $\varphi_{X_{t_0,\ldots,t_n}}(z) = \varphi_{AY}(z) = \mathsf{E}e^{i\langle z,AY\rangle} = \mathsf{E}e^{i\langle A^\top z,Y\rangle} = \varphi_Y(A^\top z)$ holds. Therefore the finite-dimensional distriution of $X_{t_0,\ldots,t_n}$ possesses the characteristic function $\varphi_{X_{t_0,\ldots,t_n}}(z) = \varphi_{Q_0}(l_0)\varphi_{t_0,t_1}(l_1)\ldots\varphi_{t_{n-1},t_n}(l_n)$, where $l = (l_1, l_1, \ldots, l_n)^\top = A^\top z$, thus

$$\begin{cases} l_0 &= z_0 + \ldots + z_n \\ l_1 &= z_1 + \ldots + z_n \\ & \vdots \\ l_n &= z_n \end{cases}$$

Thereby $\varphi_{X(t_0)} = \varphi_{Q_0}$ and $\varphi_{X_{t_1,\ldots,t_n}}(z_1, \ldots, z_n) = \varphi_{X_{t_0,\ldots,t_n}}(0, z_1, \ldots, z_n)$ holds for all $z_i \in \mathbb{R}$.
Now we proof the existence of such a process X.
For that we construct the family of characteristic functions

$$\{\varphi_{t_0}, \varphi_{t_0,t_1,\ldots,t_n}, \varphi_{t_1,\ldots,t_n},\quad 0 = t_0 < t_1 < \ldots < t_n < \infty,\ n \in \mathbb{N}\}$$

from φ_{Q_0} and $\{\varphi_{s,t},\ 0 \leq s < t\}$ as above, thus

$$\varphi_{t_0} = \varphi_{Q_0},\ \varphi_{t_1,\ldots,t_n}(0, z_1, \ldots, z_n) = \varphi_{t_0,t_1,\ldots,t_n}(0, z_1, \ldots, z_n),\ z_i \in \mathbb{R},$$

$$\varphi_{t_0,\ldots,t_n}(z) = \varphi_{t_0}(z_1 + \ldots + z_n)\varphi_{t_0,t_1}(z_1 + \ldots + z_n)\ldots\varphi_{t_{n-1},t_n}(z_n).$$

Now we have to check whether the corresponding probability measures of these characteristic functions fulfill the conditions of theorem 1.1.2. We will do that in equivalent form since according to exercise ... of exercise sheet ... the conditions of symmetry and consistency in theorem 1.1.2 are equivalent to:

a) $\varphi_{t_{i_0},\ldots,t_{i_n}}(z_{i_0},\ldots,z_{i_n}) = \varphi_{t_0,\ldots,t_n}(z_0,\ldots,z_n)$ for an arbitrary permutation $(0,1,\ldots,n) \mapsto (i_0,i_1,\ldots,i_n)$,

b) $\varphi_{t_0,\ldots,t_{m-1},t_{m+1},\ldots,t_n}(z_0,\ldots,z_{m-1},z_{m+1},\ldots,z_n) = \varphi_{t_0,\ldots,t_n}(z_0,\ldots,0,\ldots,z_n)$, for all $z_0,\ldots,z_n \in \mathbb{R}$, $m \in \{1,\ldots,n\}$.

The first condition a) is obvious. Conditon b) holds since

$$\varphi_{t_{m-1},t_m}(0 + z_{m+1} + \ldots + z_n)\varphi_{t_m,t_{m+1}}(z_{m+1} + \ldots + z_n) = \varphi_{t_{m-1},t_{m+1}}(z_{m+1},\ldots,z_n)$$

for all $m \in \{1,\ldots,n\}$. Thus, the existence of X is proven. □

Example 1.7.1
1. If $T = \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, then $X = \{X(t),\ t \in \mathbb{N}_0\}$ has independent increments if and only if $X(n) \stackrel{d}{=} \sum_{i=0}^{n} Y_i$, where $\{Y_i\}$ are independent random variables and $Y_n \stackrel{d}{=} X(n) - X(n-1)$, $n \in \mathbb{N}$. Such a process X is called *random walk*. It also may be defined for Y_i with values in $\mathbb{R}^m$.
2. The Poisson process with intensity λ has independent increments (we will show that later).
3. The Wiener process possesses independent increments.

Exercise 1.7.1
Proof that!

Exercise 1.7.2
Let $X = \{X(t),\ t \geq 0\}$ be a process with independent increments and $g : [0,\infty) \to \mathbb{R}$ an arbitrary (deterministic) function. Show that the process $Y = \{Y(t),\ t \geq 0\}$ with $Y(t) = X(t) + g(t)$, $t \geq 0$, also possesses independent increments.

1.8 Additional exercises

Exercise 1.8.1
Proof the following assertion: The family of probability measures $\mathsf{P}_{t_1,\ldots,t_n}$ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, $n \geq 1$, $t = (t_1,\ldots,t_n)^\top \in T^n$ fulfills the conditions of the theorem of Kolmogorov if and only if $n \geq 2$ and for all $s = (s_1,\ldots,s_n)^\top \in \mathbb{R}^n$ the following conditions are fulfilled:

a) $\varphi_{\mathsf{P}_{t_1,\ldots,t_n}}((s_1,\ldots,s_n)^\top) = \varphi_{\mathsf{P}_{t_{\pi(1)},\ldots,t_{\pi(n)}}}((s_{\pi(1)},\ldots,s_{\pi(n)})^\top)$ for all $\pi \in \mathcal{S}_n$.

b) $\varphi_{\mathsf{P}_{t_1,\ldots,t_{n-1}}}((s_1,\ldots,s_{n-1})^\top) = \varphi_{\mathsf{P}_{t_1,\ldots,t_n}}((s_1,\ldots,s_{n-1},0)^\top)$.

Remark: $\varphi(\cdot)$ denotes the characteristic function of the corresponding measure. $\mathcal{S}_n$ denotes the group of all permutations $\pi : \{1,\ldots,n\} \to \{1,\ldots,n\}$.

Exercise 1.8.2
Show the existence of a random function whose finite-dimensional distributions are multivariate-normally distributed and explicitly give the measurable spaces $(E_{t_1,\dots,t_n}, \mathcal{E}_{t_1,\dots,t_n})$.

Exercise 1.8.3
Give an example of a family of probability measures $\mathsf{P}_{t_1,\dots,t_n}$, which do not fulfill the conditions of the theorem of Kolmogorov.

Exercise 1.8.4
Let $X = \{X(t),\, t \in T\}$ and $Y = \{Y(t),\, t \in T\}$ be two stochastic processes which are defined on the same complete probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and which take values in the measurable space $(\mathsf{S}, \mathcal{B})$.

a) Proof that: X and Y are stochastically equivalent $\Longrightarrow \mathsf{P}_X = \mathsf{P}_Y$.

b) Give an example of two processes X and Y for which holds: $\mathsf{P}_X = \mathsf{P}_Y$, but X and Y are not stochastically equivalent.

c) Proof that: X and Y are stochastically indistinguishable $\Longrightarrow X$ and Y are stochastically equivalent.

d) Proof in the case of countability of T: X and Y are stochastically equivalent $\Longrightarrow X$ and Y are stochastically indistinguishable.

e) Give in the case of uncountability of T an example of two processes X and Y for which holds: X and Y are stochastically equivalent but not stochastically indistinguishable.

Exercise 1.8.5
Let $W = \{W(t),\, t \in \mathbb{R}\}$ be a Wiener Process. Which of the following processes are Wiener processes as well?

a) $W_1 = \{W_1(t) := -W(t),\, t \in \mathbb{R}\}$,

b) $W_2 = \{W_2(t) := \sqrt{t}W(1),\, t \in \mathbb{R}\}$,

c) $W_3 = \{W_3(t) := W(2t) - W(t),\, t \in \mathbb{R}\}$.

Exercise 1.8.6
Given a stochastic process $X = \{X(t),\, t \in [0,1]\}$ which consists of idependent and identically distributed random variables with density $f(x)$, $x \in \mathbb{R}$. Show that such a process can not be continuous in $t \in [0,1]$.

Exercise 1.8.7
Give an example of a stochastic process $X = \{X(t),\, t \in T\}$ which is stochastically continuous on T, and proof why this is the case.

Exercise 1.8.8
In connection with the continuity of stochastic processes the so-called *criterion of Kolmogorov* plays a central role. (see also theorem 1.3.1 in the lecture notes): Let $X = \{X(t),\, t \in [a,b]\}$ be a real-valued stochastic process. If constants $\alpha, \varepsilon > 0$ and $C := C(\alpha, \varepsilon) > 0$ exist such that

$$\mathsf{E}|X(t+h) - X(t)|^\alpha \leq C|h|^{1+\varepsilon} \tag{1.8.1}$$

for sufficient small h, then the process X possesses a continuous modification. Show that:

a) If you fix the variable $\varepsilon = 0$ in condition (1.8.1), then in general the condition is not sufficient for the existence of a continuous modification. *Hint: Consider the Poisson process.*

b) The Wiener process $W = \{W(t),\, t \in [0, \infty)\}$ possesses a continuous modification. *Hint: Consider the case* $\alpha = 4$.

Exercise 1.8.9
Show that the Wiener process W is not stochastically differentiable at any point $t \in [0, \infty)$.

Exercise 1.8.10
Show that the covariance function $C(s, t)$ of a complex-valued stochastic process $X = \{X(t),\, t \in T\}$

a) is symmetric, i.e. $C(s, t) = \overline{C(t, s)}$, $s, t \in T$,

b) fulfills the identity $C(t, t) = \mathsf{var}\, X(t)$, $t \in T$,

c) is *positive semidefinite*, i.e. for all $n \in \mathbb{N}$, $t_1, \dots, t_n \in T$, $z_1, \dots, z_n \in \mathbb{C}$ it holds that:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} C(t_i, t_j) z_i \bar{z}_j \geq 0.$$

Exercise 1.8.11
Show that it exists a random function $X = \{X(t),\, t \in T\}$ which simultaneously fulfills the conditions:

- The second moment $\mathsf{E}X^2$ does not exist.
- The variogram $\gamma(s, t)$ is finite for all $s, t \in T$.

Exercise 1.8.12
Give an example of a stochastic process $X = \{X(t),\, t \in T\}$ whose paths are simultaneously L^2-differentiable but not almost surely differentiable, and proof why this is the case.

Exercise 1.8.13
Give an example of a stochastic process $X = \{X(t),\, t \in T\}$ whose paths are simultaneously almost surely differentiable but not L^1-differentiable, and proof why this is the case.

Exercise 1.8.14
Proof that the Wiener process possesses independent increments.

Exercise 1.8.15
Proof: A (real-valued) stochastic process $X = \{X(t),\, t \in [0, \infty)\}$ with independent increments already has stationary increments if the distibution of the random variable $X(t + h) - X(h)$ is independent of h.

2 Counting processes

In this chapter we consider several examples of stochastic processes which model the counting of events and thus possess piecewise constant paths.

Let $(\Omega, \mathcal{A}, \mathsf{P})$ be a probability space and $\{S_n\}_{n\in\mathbb{N}}$ a non-decreasing sequence of a.s. non-negative random variables, i.e. $0 \le S_1 \le S_2 \le \ldots \le S_n \le \ldots$.

Definition 2.0.1
The stochastic process $N = \{N(t),\ t \ge 0\}$ is called *counting process* if

$$N(t) = \sum_{n=1}^{\infty} \mathbb{1}(S_n \le t),$$

where $\mathbb{1}(A)$ is the indicator function of the event $A \in \mathcal{A}$.

$N(t)$ counts the events which occur at S_n until time t. S_n e.g. may be the time of occurence of

1. the n-th elementary particle in the Geiger counter, or
2. a damage in the insurance of material damage, or
3. a data paket at a server within a computer network, etc.

A special case of the counting processes are the so-called *renewal processes*.

2.1 Renewal processes

Definition 2.1.1
Let $\{T_n\}_{n\in\mathbb{N}}$ be a sequence of i.i.d. non-negative random variables with $\mathsf{P}(T_1 > 0) > 0$. A counting process $N = \{N(t),\ t \ge 0\}$ with $N(0) = 0$ a.s., $S_n = \sum_{k=1}^{n} T_k$, $n \in \mathbb{N}$, is called *renewal process*. Thereby S_n is called the *time of the n-th renewal*, $n \in \mathbb{N}$.

The name „renewal process“ is given by the following interpretation. The „interarrival times“ T_n are interpreted as the lifetime of a technical spare part or mechanism within a system, thus S_n is the time of the n-th break down of the system. The defective part is immediately replaced by a new part (comparable with the exchange of a lightbulb). Thus, $N(t)$ is the number of repairs (the so-called „renewals“) of the system until time t.

Remark 2.1.1 1. It is $N(t) = \infty$ if $S_n \le t$ for all $n \in \mathbb{N}$.

2. Often it is assumed that only $T_2, T_3, \ldots$ are identically distributed with $\mathsf{E}T_n < \infty$. The distribution of T_1 is freely selectable. Such a process $N = \{N(t),\ t \ge 0\}$ is called *delayed renewal process* (with delay T_1).
3. Sometimes the requirement $T_n \ge 0$ is omitted.

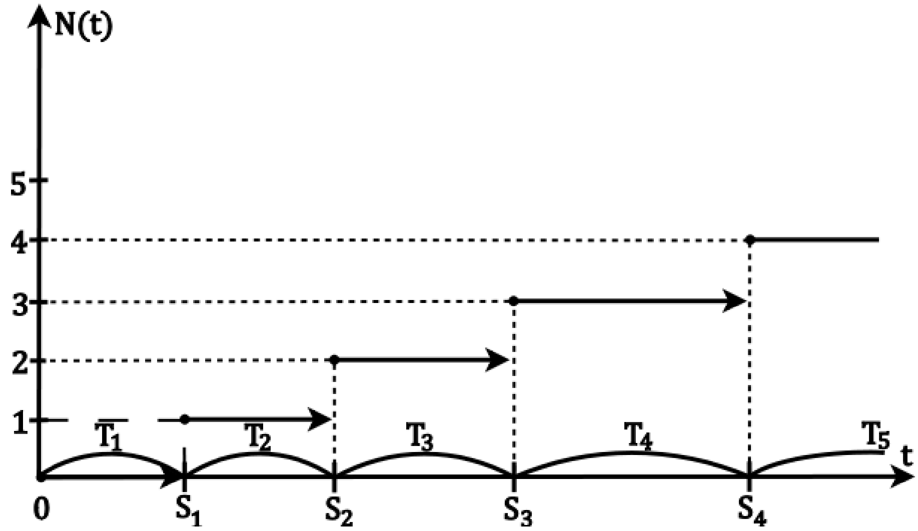

Abb. 2.1: Konstruktion und Trajektorien eines Erneuerungsprozesses

4. It is clear that $\{S_n\}_{n\in\mathbb{N}_0}$ with $S_0 = 0$ a.s., $S_n = \sum_{k=1}^{n} T_k$, $n \in \mathbb{N}$ is a *random walk*.

5. If one requires that the n-th exchange of a defective part in the system takes a time T'_n, then by $\tilde{T}_n = T_n + T'_n$, $n \in \mathbb{N}$ a different renewal process is given. Its stochastic property does not differ from the process which is given in definition 2.1.1.

In the following we assume that $\mu = \mathsf{E}T_n \in (0, \infty)$, $n \in \mathbb{N}$.

Theorem 2.1.1 (Individual ergodic theorem):
Let $N = \{N(t),\ t \geq 0\}$ be a renewal process. Then it holds that:

$$\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1}{\mu} \quad \text{a.s.}.$$

Proof For all $t \geq 0$ and $n \in \mathbb{N}$ it holds that $\{N(t) = n\} = \{S_n \leq t < S_{n+1}\}$, therefore $S_{N(t)} \leq t < S_{N(t)+1}$ and

$$\frac{S_{N(t)}}{N(t)} \leq \frac{t}{N(t)} \leq \frac{S_{N(t)+1}}{N(t)+1} \cdot \frac{N(t)+1}{N(t)}.$$

If we can show that $\frac{S_{N(t)}}{N(t)} \xrightarrow[t\to\infty]{a.s} \mu$ and $N(t) \xrightarrow[t\to\infty]{a.s.} \infty$, then $\frac{t}{N(t)} \xrightarrow[t\to\infty]{a.s} \mu$ holds and therefore the assertion of the theorem.
According to the strong law of large numbers of Kolmogorov (cf. lecture notes „Wahrscheinlichkeitsrechnung“ (WR), theorem 7.4) it holds that $\frac{S_n}{n} \xrightarrow[n\to\infty]{a.s.} \mu$, thus $S_n \xrightarrow[n\to\infty]{a.s.} \infty$ and therefore $\mathsf{P}(N(t) < \infty) = 1$ since $\mathsf{P}(N(t) = \infty) = \mathsf{P}(\ S_n \leq t, \forall n) = 1 - \underbrace{\mathsf{P}(\exists n : \forall m \in \mathbb{N}_0\ S_{n+m} > t)}_{=1, \text{ if } S_n \xrightarrow[n\to\infty]{a.s} \infty} = 1 - 1 = 0$. Then $N(t)$, $t \geq 0$, is a real random variable.
We show that $N(t) \xrightarrow[t\to\infty]{a.s.} \infty$. All trajectories of $N(t)$ are monotonously non-decreasing in

$t \geq 0$, thus $\exists \lim_{t\to\infty} N(\omega, t)$ for all $\omega \in \Omega$. Moreover it holds that

$$\begin{aligned}
\mathsf{P}(\lim_{t\to\infty} N(t) < \infty) &= \lim_{n\to\infty} \mathsf{P}(\lim_{t\to\infty} N(t) < n) \overset{(*)}{=} \lim_{n\to\infty} \lim_{t\to\infty} \mathsf{P}(N(t) < n) \\
&= \lim_{n\to\infty} \lim_{t\to\infty} \mathsf{P}(S_n > t) = \lim_{n\to\infty} \lim_{t\to\infty} \mathsf{P}(\sum_{k=1}^{n} T_k > t) \\
&\leq \lim_{n\to\infty} \lim_{t\to\infty} \sum_{k=1}^{n} \underbrace{\mathsf{P}(T_k > \frac{t}{n})}_{\xrightarrow[t\to\infty]{} 0} = 0.
\end{aligned}$$

The equality $(*)$ holds since $\{\lim_{t\to\infty} N(t) < n\} = \{\exists t_0 \in \mathbb{Q}_+ : \forall t \geq t_0 \; N(t) < n\} = \cup_{t_0 \in \mathbb{Q}_+} \cap_{\substack{t\in\mathbb{Q}_+ \\ t \geq t_0}} \{N(t) < n\} = \liminf_{\substack{t\in\mathbb{Q}_+ \\ t\to\infty}} \{N(t) < n\}$, then the continuity of the probability measure is used, where $\mathbb{Q}_+ = \mathbb{Q} \cap \mathbb{R}_+ = \{q \in \mathbb{Q} : q \geq 0\}$. Since for every $\omega \in \Omega$ it holds that $\lim_{n\to\infty} \frac{S_n}{n} = \lim_{t\to\infty} \frac{S_{N(t)}}{N(t)}$ (the codomain of a realization of $N(\cdot)$ is a subsequence of $\mathbb{N}$), it holds that $\lim_{t\to\infty} \frac{S_{N(t)}}{N(t)} \overset{a.s}{=} \mu$. □

Remark 2.1.2
One can generalize the ergodic theorem to the case of non-identically distributed T_n. Thereby we require that $\mu_n = \mathsf{E}T_n$, $\{T_n - \mu_n\}_{n\in\mathbb{N}}$ are uniformly integrable and $\frac{1}{n}\sum_{k=1}^{n} \mu_k \xrightarrow[n\to\infty]{} \mu > 0$. Then we can proof that $\frac{N(t)}{t} \xrightarrow[t\to\infty]{\mathsf{P}} \frac{1}{\mu}$ (cf. [2], page 276).

Theorem 2.1.2 (Central limit theorem):
If $\mu \in (0, \infty)$, $\sigma^2 = \mathsf{var}\, T_1 \in (0, \infty)$, it holds that

$$\mu^{\frac{3}{2}} \cdot \frac{N(t) - \frac{t}{\mu}}{\sigma\sqrt{t}} \xrightarrow[t\to\infty]{d} Y,$$

where $Y \sim \mathcal{N}(0, 1)$.

Proof According to the central limit theorem for sums of i.i.d. random variables (cf. theorem 7.5, WR) it holds that

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow[n\to\infty]{d} Y. \tag{2.1.1}$$

Let $[x]$ be the whole part of $x \in \mathbb{R}$. It holds for $a = \frac{\sigma^2}{\mu^3}$ that

$$\mathsf{P}\left(\frac{N(t) - \frac{t}{\mu}}{\sqrt{at}} \leq x\right) = \mathsf{P}\left(N(t) \leq x\sqrt{at} + \frac{t}{\mu}\right) = \mathsf{P}\left(S_{m(t)} > t\right),$$

where $m(t) = \left[x\sqrt{at} + \frac{t}{\mu}\right] + 1$, $t \geq 0$, and $\lim_{t\to\infty} m(t) = \infty$. Therefore we get that

$$\begin{aligned}
\left|\mathsf{P}\left(\frac{N(t) - \frac{t}{\mu}}{\sqrt{at}} \leq x\right) - \varphi(x)\right| &= \left|\mathsf{P}\left(S_{m(t)} > t\right) - \varphi(x)\right| \\
&= \left|\mathsf{P}\left(\frac{S_{m(t)} - \mu m(t)}{\sigma\sqrt{m(t)}} > \frac{t - \mu m(t)}{\sigma\sqrt{m(t)}}\right) - \varphi(x)\right| := I_t(x)
\end{aligned}$$

for arbitrary $t \geq 0$ and $x \in \mathbb{R}$, where φ is the distribution function of the $\mathcal{N}(0,1)$-distribution. For fixed $x \in \mathbb{R}$ we introduce $Z_t = -\frac{t-\mu m(t)}{\sigma\sqrt{m(t)}} - x$, $t \geq 0$. Then it holds that

$$I_t(x) = \left| \mathsf{P}\left(\frac{S_{m(t)} - \mu m(t)}{\sigma\sqrt{m(t)}} + Z_t > -x \right) - \varphi(x) \right|.$$

If we can proof that $Z_t \xrightarrow[t\to\infty]{} 0$, then applying (2.1.1) and the theorem of Slutsky (theorem 6.4.1, WR) would result in $\frac{S_{m(t)}-\mu m(t)}{\sigma\sqrt{m(t)}} + Z_t \xrightarrow[t\to\infty]{d} Y \sim \mathcal{N}(0,1)$ since $Z_t \xrightarrow[t\to\infty]{} 0$ a.s. results in $Z_t \xrightarrow[t\to\infty]{d} 0$. Therefore we could write $I_t(x) \xrightarrow[t\to\infty]{} |\bar{\varphi}(-x) - \varphi(x)| = |\varphi(x) - \varphi(x)| = 0$, where $\bar{\varphi}(x) = 1 - \varphi(x)$ is the tail function of the $\mathcal{N}(0,1)$-distribution, and the property of symmetry of $\mathcal{N}(0,1) : \bar{\varphi}(-x) = \varphi(x)$, $x \in \mathbb{R}$ was used.

Now we show that $Z_t \xrightarrow[t\to\infty]{} 0$, thus $\frac{t-\mu m(t)}{\sigma\sqrt{m(t)}} \xrightarrow[t\to\infty]{} -x$. It holds that $m(t) = x\sqrt{at} + \frac{t}{\mu} + \varepsilon(t)$, where $\varepsilon(t) \in [0,1)$. Then it holds that

$$\begin{aligned}
\frac{t-\mu m(t)}{\sigma\sqrt{m(t)}} &= \frac{t - \mu x\sqrt{at} - t - \mu\varepsilon(t)}{\sigma\sqrt{m(t)}} = -x\frac{\sqrt{at} - \mu}{\sigma\sqrt{x\sqrt{at} + \frac{t}{\mu} + \varepsilon(t)}} - \frac{\mu\varepsilon(t)}{\sigma\sqrt{m(t)}} \\
&= -\frac{x\mu}{\sigma\sqrt{\frac{x}{\sqrt{at}} + \frac{1}{\mu a} + \frac{\varepsilon(t)}{at}}} - \frac{\mu - \varepsilon(t)}{\sigma\sqrt{m(t)}} \\
&= \underbrace{-\frac{x\frac{\mu}{\sigma}}{\sqrt{\frac{\mu^2}{\sigma^2} + \frac{x}{\sqrt{at}} + \frac{\varepsilon(t)}{at}}}}_{\xrightarrow[t\to\infty]{} -x} - \underbrace{\frac{\mu\varepsilon(t)}{\sigma\sqrt{m(t)}}}_{\xrightarrow[t\to\infty]{} 0} \xrightarrow[t\to\infty]{} -x.
\end{aligned}$$

□

Remark 2.1.3
In Lineberg form, the central limit theorem can also be proven for non-identically distributed T_n, cf. [2], pages 276 - 277.

Definition 2.1.2
The function $H(t) = \mathsf{E}N(t)$, $t \geq 0$ is called *renewal function* of the process N (or of the sequence $\{S_n\}_{n\in\mathbb{N}}$).

Let $F_T(x) = \mathsf{P}(T_1 \leq x)$, $x \in \mathbb{R}$ be the distribution function of T_1. For arbitrary distribution functions $F, G : \mathbb{R} \to [0,1]$ the *convolution* $F * G$ is defined as $F * G(x) = \int_{-\infty}^{x} F(x-y)dG(y)$. The *k-fold convolution* F^{*k} of the distribution F with itself, $k \in \mathbb{N}_0$, is defined inductive:

$$\begin{aligned}
F^{*0}(x) &= \mathbb{1}(x \in [0,\infty)),\ x \in \mathbb{R}, \\
F^{*1}(x) &= F(x),\ x \in \mathbb{R}, \\
F^{*(k+1)}(x) &= F^{*k} * F(x),\ x \in \mathbb{R}.
\end{aligned}$$

Lemma 2.1.1
The renewal function H of a renewal process N is monotonously non-decreasing and right-sided continuous on $\mathbb{R}_+$. Moreover it holds that

$$H(t) = \sum_{n=1}^{\infty} \mathsf{P}(S_n \leq t) = \sum_{n=1}^{\infty} F_T^{*n}(t),\ t \geq 0. \tag{2.1.2}$$

Proof The monotony and right-sided continuity of H are consequences from the almost surely monotony and right-sided continuity of the trajectories of N. Now we proof (2.1.2):

$$H(t) = \mathsf{E}N(t) = \mathsf{E}\sum_{n=1}^{\infty} \mathbb{1}(S_n \le t) \overset{(*)}{=} \sum_{n=1}^{\infty} \mathsf{E}\mathbb{1}(S_n \le t) = \sum_{n=1}^{\infty} \mathsf{P}(S_n \le t) = \sum_{n=1}^{\infty} F_T^{*n}(t),$$

since $\mathsf{P}(S_n \le t) = \mathsf{P}(T_1 + \ldots + T_n \le t) = F_T^{*n}(t)$, $t \ge 0$. The equality $(*)$ holds for all partial sums on both sides, therefore in the limit as well. □

Except for exceptional cases it is impossible to calculate the renewal function H by the formula (2.1.2) analytically. Therefore the *Laplace transform* of H is often used in calulations. For a monotonously (e.g. monotonously non-decreasing) right-sided continuous function $G : [0,\infty) \to \mathbb{R}$ the *Laplace transform* is defined as $\hat{l}_G(s) = \int_0^\infty e^{-sx} dG(x)$, $s \ge 0$. Here the integral is to be understood as the Lebesgue-Stieltjes integral, thus as a Lebesgue integral with respect to the measure μ_G on $\mathcal{B}_{\mathbb{R}_+}$ defined by $\mu_G((x,y]) = G(y) - G(x)$, $0 \le x < y < \infty$, if G is monotonously non-decreasing.
Just to remind you: the Laplace transform $\hat{l}_X$ of a random variable $X \ge 0$ is defined by $\hat{l}_X(s) = \int_0^\infty e^{-sx} dF_X(x)$, $s \ge 0$.

Lemma 2.1.2
For $s > 0$ it holds that:

$$\hat{l}_H(s) = \frac{\hat{l}_{T_1}(s)}{1 - \hat{l}_{T_1}(s)}.$$

Proof It holds that:

$$\begin{aligned} \hat{l}_H(s) &= \int_0^\infty e^{-sx} dH(x) \overset{(2.1.2)}{=} \int_0^\infty e^{-sx} d\left(\sum_{n=1}^\infty F_T^{*n}(x)\right) = \sum_{n=1}^\infty \int_0^\infty e^{-sx} dF^{*n}(x) \\ &= \sum_{n=1}^\infty \hat{l}_{T_1+\ldots+T_n}(s) = \sum_{n=1}^\infty \left(\hat{l}_{T_1}(s)\right)^n = \frac{\hat{l}_{T_1}(s)}{1 - \hat{l}_{T_1}(s)}, \end{aligned}$$

where for $s > 0$ it holds that $\hat{l}_{T_1}(s) < 1$ and thus the geometric series $\sum_{n=1}^\infty \left(\hat{l}_{T_1}(s)\right)^n$ converges.
□

Remark 2.1.4
If $N = \{N(t),\ t \ge 0\}$ is a delayed renewal process (with delay T_1), the statements of lemmas 2.1.1 - 2.1.2 hold in the following form:

1.
$$H(t) = \sum_{n=0}^\infty (F_{T_1} * F_{T_2}^{*n})(t),\ t \ge 0,$$

   where F_{T_1} and F_{T_2}, respectively are the distribution functions of T_1 and T_n, $n \ge 2$, respectively.

2.
$$\hat{l}_H(s) = \frac{\hat{l}_{T_1}(s)}{1 - \hat{l}_{T_2}(s)},\ s \ge 0, \tag{2.1.3}$$

   where $\hat{l}_{T_1}$ and $\hat{l}_{T_2}$ are the Laplace transforms of the distribution of T_1 and T_n, $n \ge 2$.

For further observations we need a theorem (from Wald) about the expected value of a sum (with random number) of independent random variables.

Definition 2.1.3
Let ν be a $\mathbb{N}$-valued random variable and be $\{X_n\}_{n\in\mathbb{N}}$ a sequence of random variables defined on the same probability space. ν is called *independent of the future*, if for all $n \in \mathbb{N}$ the event $\{\nu \leq n\}$ does not depend on the σ-algebra $\sigma(\{X_k,\ k > n\})$.

Theorem 2.1.3 (Wald's identity):
Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of random variables with $\sup \mathsf{E}|X_n| < \infty$, $\mathsf{E}X_n = a$, $n \in \mathbb{N}$ and be ν a $\mathbb{N}$-valued random variable which is independent of the future, with $\mathsf{E}\nu < \infty$. Then it holds that

$$\mathsf{E}(\sum_{n=1}^{\nu} X_n) = a \cdot \mathsf{E}\nu.$$

Proof Calculate $S_n = \sum_{k=1}^{n} X_k$, $n \in \mathbb{N}$. Since $\mathsf{E}\nu = \sum_{n=1}^{\infty} \mathsf{P}(\nu \geq n)$, the theorem follows from lemma 2.1.3. □

Lemma 2.1.3 (Kolmogorov-Prokhorov):
Let ν be a $\mathbb{N}$-valued random variable which is independent of the future and it holds that

$$\sum_{n=1}^{\infty} \mathsf{P}(\nu \geq n)\mathsf{E}|X_n| < \infty. \tag{2.1.4}$$

Then $\mathsf{E}S_\nu = \sum_{n=1}^{\infty} \mathsf{P}(\nu \geq n)\mathsf{E}X_n$ holds. If $X_n \geq 0$ a.s., then condition (2.1.4) is not required.

Proof It holds that $S_\nu = \sum_{n=1}^{\nu} X_n = \sum_{n=1}^{\infty} X_n \mathbf{1}(\nu \geq n)$. We introduce the notation $S_{\nu,n} = \sum_{k=1}^{n} X_k \mathbf{1}(\nu \geq k)$, $n \in \mathbb{N}$. First, we proof the lemma for $X_n \geq 0$ f.s., $n \in \mathbb{N}$. It holds $S_{\nu,n} \uparrow S_\nu$, $n \to \infty$ for every $\omega \in \Omega$, and thus according to the monotone convergence theorem it holds that: $\mathsf{E}S_\nu = \lim_{n\to\infty} \mathsf{E}S_{\nu,n} = \lim \sum_{k=1}^{n} \mathsf{E}(X_k \mathbf{1}(\nu \geq k))$. Since $\{\nu \geq k\} = \{\nu \leq k-1\}^c$ does not depend on $\sigma(X_k) \subset \sigma(\{X_n,\ n \geq k\})$ it holds that $\mathsf{E}(X_k \mathbf{1}(\nu \geq k)) = \mathsf{E}X_k \mathsf{P}(\nu \geq k)$, $k \in \mathbb{N}$, and thus $\mathsf{E}S_\nu = \sum_{n=1}^{\infty} \mathsf{P}(\nu \geq n)\mathsf{E}X_n$.
Now, let X_n be arbitrary. Take $Y_n = |X_n|$, $Z_n = \sum_{n=1}^{n} Y_n$, $Z_{\nu,n} = \sum_{k=1}^{n} Y_k \mathbf{1}(\nu \geq k)$, $n \in \mathbb{N}$. Since $Y_n \geq 0$, $n \in \mathbb{N}$, it holds that $\mathsf{E}Z_\nu = \sum_{n=1}^{\infty} \mathsf{E}(X_n \mid \mathsf{P}(\nu \geq k)) < \infty$ from (2.1.4). Since $|S_{\nu,n}| \leq Z_{\nu,n} \leq Z_\nu$, $n \in \mathbb{N}$, according to the dominated convergence theorem of Lebesgue it holds that $\mathsf{E}S_\nu = \lim_{n\to\infty} \mathsf{E}S_{\nu,n} = \sum_{n=1}^{\infty} \mathsf{E}X_n \mathsf{P}(\nu \geq n)$, where this series converges absolutely.
□

Conclusion 2.1.1
1. For an arbitrary Borel measurable function $g : \mathbb{R}_+ \to \mathbb{R}_+$ and the renewal process $N = \{N(t),\ t \geq 0\}$ with interarrival times $\{T_n\}$, T_n i.i.d., $\mu = \mathsf{E}T_n \in (0, \infty)$ it holds that

$$\mathsf{E}\left(\sum_{k=1}^{N(t)+1} g(T_n)\right) = (1 + H(t))\mathsf{E}g(T_1),\ t \geq 0.$$

2. $H(t) < \infty$, $t \geq 0$.

Proof
1. For every $t \geq 0$ it is obvious that $\nu = 1 + H(t)$ does not depend on the future of $\{T_n\}_{n\in\mathbb{N}}$, the rest follows from theorem 2.1.3 with $X_n = g(T_n)$, $n \in \mathbb{N}$.

2. For $s > 0$ consider $T_n^{(s)} = \min\{T_n, s\}$, $n \in \mathbb{N}$. Choose $s > 0$ such that for freely selected (but fixed) $\varepsilon > 0 : \mu^{(s)} = \mathsf{E}T_1^{(s)} \geq \mu - \varepsilon > 0$. Let $N^{(s)}$ be the renewal process which is based on the sequence $\{T_n^{(s)}\}_{n\in\mathbb{N}}$ of interarrival times: $N^{(s)}(t) = \sum_{n=1}^{\infty} \mathbf{1}(T_n^{(s)} \leq t)$, $t \geq 0$. It holds $N(t) \leq N^{(s)}(t)$, $t \geq 0$, a.s., according to conclusion 2.1.1:

$$(\mu-\varepsilon)(\mathsf{E}N^{(s)}(t)+1) \leq \mu^{(s)}(\mathsf{E}N^{(s)}(t)+1) = \mathsf{E}S^{(s)}_{N^{(s)}(t)+1} = \mathsf{E}(\underbrace{S^{(s)}_{N^{(s)}(t)}}_{\leq t} + \underbrace{T^{(s)}_{N^{(s)}(t)+1}}_{\leq s}) \leq t+s,$$

$t \geq 0$, where $S_n^{(s)} = T_1^{(s)} + \ldots + T_n^{(s)}$, $n \in \mathbb{N}$. Thus $H(t) = \mathsf{E}N(t) \leq \mathsf{E}N^{(s)}(t) \leq \frac{t+s}{\mu-\varepsilon}$, $t \geq 0$. Since $\varepsilon > 0$ is arbitrary, it holds that $\limsup_{t\to\infty} \frac{H(t)}{t} \leq \frac{1}{\mu}$, and also our assertion $H(t) < \infty$, $t \geq 0$.

□

Conclusion 2.1.2 (Elementary renewal theorem):
For a renewal process N as defined in conclusion 2.1.1, 1) it holds:

$$\lim_{t\to\infty} \frac{H(t)}{t} = \frac{1}{\mu}.$$

Proof In conclusion 2.1.1, part 2) we already proved that $\limsup_{t\to\infty} \frac{H(t)}{t} \leq \frac{1}{\mu}$. If we show $\liminf_{t\to\infty} \frac{H(t)}{t} \geq \frac{1}{\mu}$, our assertion would be proven. According to theorem 2.1.1 it holds that $\frac{N(t)}{t} \xrightarrow[t\to\infty]{} \frac{1}{\mu}$ a.s., therefore according to Fatou's lemma

$$\frac{1}{\mu} = \mathsf{E}\liminf_{t\to\infty} \frac{N(t)}{t} \leq \liminf_{t\to\infty} \frac{\mathsf{E}N(t)}{t} = \liminf_{t\to\infty} \frac{H(t)}{t}.$$

□

Remark 2.1.5 1. We can prove that in the case of the finite second moment of T_n ($\mu_2 = \mathsf{E}T_1^2 < \infty$) we can derive a more exact asymptotics for $H(t)$, $t \to \infty$:

$$H(t) = \frac{t}{\mu} + \frac{\mu_2}{2\mu^2} + o(1),\ \ t \to \infty.$$

2. The elementary renewal theorem also holds for delayed renewal processes, where $\mu = \mathsf{E}T_2$. We define the *renewal measure* H on $\mathcal{B}(\mathbb{R}_+)$ by $H(B) = \sum_{n=1}^{\infty} \int_B dF_T^{*n}(x)$, $B \in \mathcal{B}(\mathbb{R}_+)$. It holds $H((-\infty, t]) = H(t)$, $H((s,t]) = H(t) - H(s)$, $s, t \geq 0$, if H is the renewal function as well as the renewal measure.

Theorem 2.1.4 (Fundamental theorem of the renewal theory):
Let $N = \{N(t),\ t \geq 0\}$ be a (delayed) renewal process associated with the sequence $\{T_n\}_{n\in\mathbb{N}}$, where T_n, $n \in \mathbb{N}$ are independent, $\{T_n,\ n \geq 2\}$ identically distributed, and the distribution of T_2 is not arithmetic, thus not concentrated on a regular lattice with probability 1. The distribution of T_1 is arbitrary. Let $\mathsf{E}T_2 = \mu \in (0, \infty)$. Then it holds that

$$\int_0^t g(t-x)dH(x) \xrightarrow[t\to\infty]{} \frac{1}{\mu}\int_0^{\infty} g(x)dx,$$

where $g : \mathbb{R}_+ \to \mathbb{R}$ is Riemann integrable $[0, n]$, for all $n \in \mathbb{N}$, and $\sum_{n=0}^{\infty} \max_{n\leq x\leq n+1} |g(x)| < \infty$.

Without proof.

In particular $H((t-u,t]) \xrightarrow[t\to\infty]{} \frac{u}{\mu}$ holds for an arbitrary $u \in \mathbb{R}_+$, thus H asymptotically (for $t \to \infty$) behaves as the Lebesgue measure.

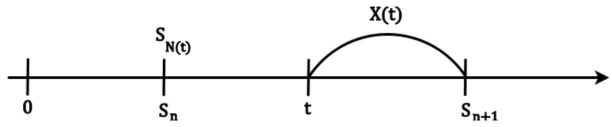

Abb. 2.2:

Definition 2.1.4
The random variable $\chi(t) = S_{N(t)+1} - t$ is called *excess* of N at time $t \geq 0$.

Obviously $\chi(0) = T_1$ holds. We now give an example of a renewal process with stationary increments.
Let $N = \{N(t),\ t \geq 0\}$ be a delayed renewal process associated with the sequence of interarrival times $\{T_n\}_{n\in\mathbb{N}}$. Let F_{T_1} and F_{T_2} be the distribution functions of the delays T_1 and T_n, $n \geq 2$. We assume that $\mu = \mathsf{E}T_2 \in (0,\infty)$, $F_{T_2}(0) = 0$, thus $T_2 > 0$ a.s. and

$$F_{T_1}(x) = \frac{1}{\mu}\int_0^x \bar{F}_{T_2}(y)dy,\ x \geq 0. \tag{2.1.5}$$

In this case F_{T_1} is called the *integrated tail distribution function* of T_2.

Theorem 2.1.5
Under the conditions we mentioned above, N is a process with stationary increments.

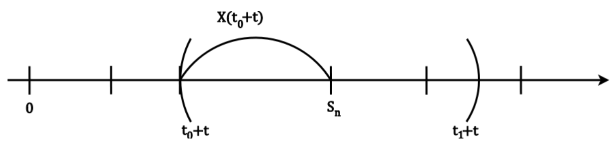

Abb. 2.3:

Proof Let $n \in \mathbb{N}$, $0 \leq t_0 < t_1 < \ldots < t_n < \infty$. Because N does not depend on T_n, $n \in \mathbb{N}$ the common distribution of $(N(t_1+t) - N(t_0+t), \ldots, N(t_n+t) - N(t_{n-1}+t))^\top$ does not depend on t, if the distribution of $\chi(t)$ does not depend on t, thus $\chi(t) \stackrel{d}{=} \chi(0) = T_1$, $t \geq 0$, see Figure ....

We show that $F_{T_1} = F_{X(t)}$, $t \geq 0$.

$$
\begin{aligned}
F_{\chi(t)}(x) &= \mathsf{P}(\chi(t) \leq x) = \sum_{n=0}^{\infty} \mathsf{P}(S_n \leq t,\ t < S_{n+1} \leq t+x) \\
&= \mathsf{P}(S_0 = 0 \leq t,\ t < S_1 = T_1 \leq t+x) \\
&\quad + \sum_{n=1}^{\infty} \mathsf{E}(\mathsf{E}(1(S_n \leq t,\ t < S_n + T_{n+1} \leq t+x) \mid S_n)) \\
&= F_{T_1}(t+x) - F_{T_1}(t) + \sum_{n=1}^{\infty} \int_0^t \mathsf{P}(t-y < T_{n+1} \leq t+x-y) dF_{S_n}(y) \\
&= F_{T_1}(t+x) - F_{T_1}(t) + \int_0^t \mathsf{P}(t-y < T_2 \leq t+x-y) d(\underbrace{\sum_{n=1}^{\infty} F_{S_n}(y)}_{H(y)}).
\end{aligned}
$$

If we can proof that $H(y) = \frac{y}{\mu}$, $y \geq 0$, then we would get

$$
\begin{aligned}
F_{\chi(t)}(x) &= F_{T_1}(t+x) - F_{T_1}(t) + \frac{1}{\mu} \int_t^0 (F_{T_2}(z+x) - 1 + 1 - F_{T_2}(z)) d(-z) \\
&= F_{T_1}(t+x) - F_{T_1}(t) + \frac{1}{\mu} \int_0^t (\bar{F}_{T_2}(z) - \bar{F}_{T_2}(z+x)) dz \\
&= F_{T_1}(t+x) - F_{T_1}(t) + F_{T_1}(t) - \frac{1}{\mu} \int_x^{t+x} \bar{F}_{T_2}(y) dy \\
&= F_{T_1}(t+x) - F_{T_1}(t+x) + F_{T_1}(x) = F_{T_1}(x),\ x \geq 0,
\end{aligned}
$$

according to the form (2.1.5) of the distribution of T_1.
Now we like to show that $H(t) = \frac{t}{\mu}$, $t \geq 0$. For that we use the formula (2.1.4): it holds that

$$
\begin{aligned}
\hat{l}_{T_1}(s) &= \frac{1}{\mu} \int_0^{\infty} e^{-st}(1 - F_{T_2}(t)) dt = \frac{1}{\mu} \underbrace{\int_0^{\infty} e^{-st} dt}_{\frac{1}{s}} - \frac{1}{\mu} \int_0^{\infty} e^{-st} F_{T_2}(t) dt \\
&= \frac{1}{\mu s} \left(1 + \int_0^{\infty} F_{T_2}(t) de^{-st} \right) = \frac{1}{\mu s} (1 + \underbrace{e^{-st} F_{T_2}(t) |_0^{\infty}}_{-F_{T_2}(0)=0} - \underbrace{\int_0^{\infty} e^{-st} dF_{T_2}(t))}_{\hat{l}_{T_2}(s)} \\
&= \frac{1}{\mu s} (1 - \hat{l}_{T_2}(s)),\ s \geq 0.
\end{aligned}
$$

Using the formula (2.1.4) we get

$$
\hat{l}_H(s) = \frac{\hat{l}_{T_1}(s)}{1 - \hat{l}_{T_2}(s)} = \frac{1}{\mu s} = \frac{1}{\mu} \int_0^{\infty} e^{-st} dt = \hat{l}_{\frac{t}{\mu}}(s),\ s \geq 0.
$$

Since the Laplace transform of a function uniquely determines this function, it holds that $H(t) = \frac{t}{\mu}$, $t \geq 0$. □

Remark 2.1.6
In the proof of theorem 2.1.5 we showed that for the renewal process with delay which possesses

the distribution (2.1.5), $H(t) \sim \frac{t}{\mu}$ not only asymptotical for $t \to \infty$ (as in the elementary renewal theorem) but it holds $H(t) = \frac{t}{\mu}$, for all $t \geq 0$. This means, per unit of the time interval we get an average of $\frac{1}{\mu}$ renewals. For that reason such a process N is called *homogeneous renewal process*.

We can proof the following theorem:

Theorem 2.1.6
If $N = \{N(t),\ t \geq 0\}$ is a delayed renewal process with arbitrary delay T_1 and non-arithmetic distribution of T_n, $n \geq 2$, $\mu = \mathsf{E}T_2 \in (0, \infty)$, then it holds that

$$\lim_{t\to\infty} F_{\chi(t)}(x) = \frac{1}{\mu}\int_0^x \bar{F}_{T_2}(y)dy,\ x \geq 0.$$

This means, the limit distribution of excess $\chi(t),\ t \to \infty$ is taken as the distribution of T_1 when defining a homogeneous renewal process.

2.2 Poisson-ish processes

2.2.1 Poisson processes

In this section we generalize the definition of a homogeneous Poisson process (see section 1.2, example 5)

Definition 2.2.1
The counting process $N = \{N(t),\ t \geq 0\}$ is called Poisson process with intensity measure Λ if

1. $N(0) = 0$ a.s.
2. Λ is a locally finite measure $\mathbb{R}_+$,i.e., $\Lambda : \mathcal{B}(\mathbb{R}_+) \to \mathbb{R}_+$ possesses the property $\Lambda(B) < \infty$ for every bounded set $B \in \mathcal{B}(\mathbb{R}_+)$.
3. N possesses independent increments.
4. $N(t) - N(s) \sim \text{Pois}(\Lambda((s,t]))$ for all $0 \leq s < t < \infty$.

Sometimes the Poisson process $N = \{N(t),\ t \geq 0\}$ is defined by the corresponding random Poisson counting measure $N = \{N(B),\ B \in \mathcal{B}(\mathbb{R}_+)\}$, i.e., $N = ([0,t])$, $t \geq 0$, where a counting measure is a locally finite measure with values in $\mathbb{N}_0$.

Definition 2.2.2
A random counting measure $N = \{N(B),\ B \in \mathcal{B}(\mathbb{R}_+)\}$ is called Poissonsh with locally finite intensity measure Λ if

1. For arbitrary $n \in \mathbb{N}$ and for arbitrary pairwise disjoint bounded sets $B_1, B_2, \ldots, B_n \in \mathcal{B}(\mathbb{R}_+)$ the random variables $N(B_1), N(B_2), \ldots, N(B_n)$ are independent.
2. $N(B) \sim \text{Pois}(\Lambda(B))$, $B \in \mathcal{B}(\mathbb{R}_+)$, B-bounded.

It is obvious that properties 3 and 4 of definition 2.2.1 follow from properties 1 and 2 of definition 2.2.2. Property 1 of definition 2.2.1 however is an autonomous assumption. $N(B)$, $B \in \mathcal{B}(\mathbb{R}_+)$ is interpreted as the number of points of N within the set B.

Remark 2.2.1
As stated in definition 2.2.2, a Poisson counting measure can also be defined on an arbitrary topological space E equipped with the Borel-σ-algebra $\mathcal{B}(E)$. Very often $E = \mathbb{R}^d$, $d \geq 1$ is chosen in applications.

Lemma 2.2.1
For every locally finite measure Λ on $\mathbb{R}_+$ there exists a Poisson process with intensity measure Λ.

Proof If such a Poisson process had existed, the characteristic function $\varphi_{N(t)-N(s)}(\cdot)$ of the increment $N(t) - N(s)$, $0 \leq s < t < \infty$ would have been equal to $\varphi_{s,t}(z) = \varphi_{\text{Pois}(\Lambda((s,t]))}(z) = e^{\Lambda((s,t])(e^{iz}-1)}$, $z \in \mathbb{R}$ according to property 4 of definition 2.2.1. We show that the family of characteristic functions $\{\varphi_{s,t},\ 0 \leq s < t < \infty\}$ possesses property 1.7.1: for all $n : 0 \leq s < u < t$, $\varphi_{s,u}(z)\varphi_{u,t}(z) = e^{\Lambda((s,u])(e^{iz}-1)}e^{\Lambda((u,t])(e^{iz}-1)} = e^{(\Lambda((s,u])+\Lambda((u,t]))(e^{iz}-1)} = e^{\Lambda((s,t])(e^{iz}-1)} = \varphi_{s,t}(z)$, $z \in \mathbb{R}$ since the measure Λ is additive. Thus, the existence of the Poisson process N follows from theorem 1.7.1. □

Remark 2.2.2
The existence of a Poisson counting measure can be proven with the help of the theorem of Kolmogorov, yet in a more general form than in theorem 1.1.2.

From the properties of the Poisson distribution it follows that $\mathsf{E}N(B) = \mathsf{var}\,N(B) = \Lambda(B)$, $B \in \mathcal{B}(\mathbb{R}_+)$. Thus $\Lambda(B)$ is interpreted as the mean number of points of N within the set B, $B \in \mathcal{B}(\mathbb{R}_+)$.
We get an important special case if $\Lambda(dx) = \lambda dx$ for $\lambda \in (0, \infty)$, i.e., Λ is proportional to the Lebesgue measure ν_1 on $\mathbb{R}_+$. Then we call $\lambda = \mathsf{E}N(1)$ the intensity of N.
Soon we will proof that in this case N is a homogeneous Poisson process with intensity λ. To remind you: In section 1.2 the homogeneous Poisson process was defined as a renewal process with interarrival times $T_N \sim \text{Exp}(\lambda)$: $N(t) = \sup\{n \in \mathbb{N}\ S_n \leq t\}$, $S_n = T_1 + \ldots + T_n$, $n \in \mathbb{N}$, $t \geq 0$.

Exercise 2.2.1
Show that the homogeneous Poisson process is a homogeneous renewal process with $T_1 \stackrel{d}{=} T_2 \sim \text{Exp}(\lambda)$. Hint: you have to show that for an arbitrary exponential distributed random variable X the integrated tail distribution function of X is equal to F_X.

Theorem 2.2.1
Let $N = \{N(t),\ t \geq 0\}$ be a counting process. The following statements are equivalent.

1. N is a homogeneous Poisson process with intensity $\lambda > 0$.

2. a) $N(t) \sim \text{Pois}(\lambda t)$, $t \geq 0$

   b) for an arbitrary $n \in \mathbb{N}$, $t \geq 0$, it holds that the random vector $(S_1, \ldots, S_n)$ under condition $\{N(t) = n\}$ possesses the same distribution as the order statistics of i.i.d. random variables $U_i \in \mathcal{U}([0,t])$, $i = 1, \ldots, n$.

3. a) N has independent increments,

   b) $\mathsf{E}N(1) = \lambda$, and

   c) property 2b) holds.

4. a) N has stationary and independent increments, and
   b) $\mathsf{P}(N(t) = 0) = 1 - \lambda t + o(t)$, $\mathsf{P}(N(t) = 1) = \lambda t + o(t)$, $t \downarrow 0$ holds.

5. a) N hast stationary and independent increments,
   b) property 2a) holds.

Remark 2.2.3 1. It is obvious that definition 2.2.1 with $\Lambda(dx) = \lambda dx$, $\lambda \in (0, \infty)$ is an equivalent definition of the homogeneous Poisson process according to theorem 2.2.1.

2. The homogeneous Poisson process N was introduced in the beginning of the 20th century from the physicists A. Einstein and M. Smoluchovsky to be able to model the counting process of elementary particle in the Geiger counter.

3. From 4b) it follows $\mathsf{P}(N(t) > 1) = o(t)$, $t \downarrow 0$.

4. The intensity of N has the following interpretation: $\lambda = \mathsf{E}N(1) = \frac{1}{\mathsf{E}T_n}$, thus the mean number of renewals of N within a time interval with length 1.

5. The renewal function of the homogeneous Poisson process is $H(t) = \lambda t$, $t \geq 0$. Thereby $H(t) = \Lambda([0, t])$, $t > 0$ holds.

Proof Structure of the proof: 1) $\Rightarrow$ 2) $\Rightarrow$ 3) $\Rightarrow$ 4) $\Rightarrow$ 5) $\Rightarrow$ 1)
1) $\Rightarrow$ 2):
From 1) follows $S_n = \sum_{k=1}^n T_k \sim Erl(n, \lambda)$ since $T_k \sim \text{Pois}(\lambda)$, $n \in \mathbb{N}$, thus $\mathsf{P}(N(t) = 0) = \mathsf{P}(T_1 > t) = e^{-\lambda t}$, $t \geq 0$, and for $n \in \mathbb{N}$

$$\begin{aligned}
\mathsf{P}(N(t) = n) &= \mathsf{P}(\{N(t) \geq n\} \setminus \{N(t) \geq n+1\}) = \mathsf{P}(N(t) \geq n) - \mathsf{P}(N(t) \geq n+1) \\
&= \mathsf{P}(S_n \leq t) - \mathsf{P}(S_{n+1} \leq t) = \int_0^t \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x} dx - \int_0^t \frac{\lambda^{n+1} x^n}{n!} e^{-\lambda x} dx \\
&= \int_0^t \frac{d}{dx} \left(\frac{(\lambda x)^n}{n!} e^{-\lambda x} \right) dx = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \ t \geq 0.
\end{aligned}$$

Thus 2a) is proven.
Now let's proof 2b). According to the transformation theorem of random variables (cf. theorem 3.6.1, WR), it follows from

$$\begin{cases} S_1 &= T_1 \\ S_2 &= T_1 + T_2 \\ &\vdots \\ S_{n+1} &= T_1 + \ldots + T_{n+1} \end{cases}$$

that the density $f_{(S_1, \ldots, S_n)}$ of $(S_1, \ldots, S_{n+1})^\top$ can be expressed by the density of $(T_1, \ldots, T_{n+1})^\top$, $T_i \sim \text{Exp}(\lambda)$, i.i.d.:

$$f_{(S_1, \ldots, S_{n+1})}(t_1, \ldots, t_{n+1}) = \prod_{k=1}^{n+1} f_{T_k}(t_k - t_{k-1}) = \prod_{k=1}^{n+1} \lambda e^{-\lambda(t_k - t_{k-1})} = \lambda^{n+1} e^{-\lambda t_{n+1}}$$

for arbitrary $0 \leq t_1 \leq \ldots \leq t_{n+1}$, $t_0 = 0$.
For all other $t_1, \ldots, t_{n+1}$ it holds $f_{(S_1, \ldots, S_{n+1})}(t_1, \ldots, t_{n+1}) = 0$.

Therefore

$$\begin{aligned}
f_{(S_1,\dots,S_n)}(t_1,\dots,t_n|N(t)=n) &= f_{(S_1,\dots,S_n)}(t_1,\dots,t_n|S_k\le t,\ k\le n,\ S_{n+1}>t)\\
&= \frac{\int_t^\infty f_{(S_1,\dots,S_{n+1})}(t_1,\dots,t_{n+1})dt_{n+1}}{\int_0^t\int_{t_1}^t\dots\int_{t_{n-1}}^t\int_t^\infty f_{(S_1,\dots,S_{n+1})}(t_1,\dots,t_{n+1})dt_{n+1}dt_n\dots dt_1}\\
&= \frac{\int_t^\infty \lambda^{n+1}e^{-\lambda t_{n+1}}dt_{n+1}}{\int_0^t\int_{t_1}^t\dots\int_{t_{n-1}}^t\int_t^\infty \lambda^{n+1}e^{-\lambda t_{n+1}}dt_{n+1}dt_n\dots dt_1}\times\\
&\quad \times \mathrm{I}(0\le \mathrm{t}_1\le \mathrm{t}_2\le\dots\le \mathrm{t}_\mathrm{n}\le \mathrm{t})\\
&= \frac{n!}{t^n}\mathrm{I}(0\le \mathrm{t}_1\le \mathrm{t}_2\le\dots\le \mathrm{t}_\mathrm{n}\le \mathrm{t}).
\end{aligned}$$

This is exactly the density of n i.i.d. $\mathcal{U}([0,t])$-random variables.

Exercise 2.2.2
Proof this.

2) $\Rightarrow$ 3)
From 2a) obviously follows 3b). Now we just have to proof the independence of the increments of N. For an arbitrary $n\in\mathbb{N}$, $x_1,\dots,x_n\in\mathbb{N}$, $t_0=0<t_1<\dots<t_n$ for $x=x_1+\dots+x_n$ it holds that

$$\begin{aligned}
\mathsf{P}(\cap_{k=1}^n\{N(t_k)-N(t_{k-1})=x_k\}) &= \underbrace{\mathsf{P}(\cap_{k=1}^n\{N(t_k)-N(t_{k-1})=x_k\}|N(t_k)=x)}_{\frac{x!}{x_1!\dots x_n!}\prod_{k=1}^n\left(\frac{t_k-t_{k-1}}{t_n}\right)^{x_k}\text{ according to 2c)}}\times\\
&\quad\times \underbrace{\mathsf{P}(N(t_k)=x)}_{e^{-\lambda t_n}\frac{(\lambda t_n)^x}{x!}\text{ according to 2a)}}\\
&= \prod_{k=1}^n\frac{(\lambda(t_k-t_{k-1}))^{x_k}}{x_k!}e^{-\lambda(t_k-t_{k-1})},
\end{aligned}$$

since the probability of $(*)$ belongs to the polynomial distribution with parameters n, $\left\{\frac{t_k-t_{k-1}}{t_n}\right\}_{k=1}^n$. Because the event $(*)$is that at the independent uniformly distributed toss of x points on $[0,t]$, exactly x_k points occur within the basket of length t_k-t_{k-1}, $k=1,\dots,n$:

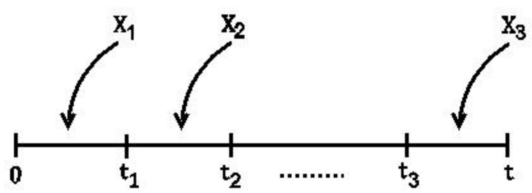

Abb. 2.4:

Thus 3a) is proven since $\mathsf{P}(\cap_{k=1}^n\{N(t_k)-N(t_{k-1})=x_k\})=\prod_{k=1}^n\mathsf{P}(\{N(t_k)-N(t_{k-1})=x_k\})$.

3) $\Rightarrow$ 4)
We proof that N possesses stationary increments. For an arbitrary $n \in \mathbb{N}_0$, $x_1, \ldots, x_n \in \mathbb{N}$, $t_0 = 0 < t_1 < \ldots < t_n$ and $h > 0$ we consider $I(h) = \mathsf{P}(\cap_{k=1}^n \{N(t_k + h) - N(t_{k-1} + h) = x_k\})$ and show that $I(h)$ does not depend on $h \in \mathbb{R}$. According to the formula of the total probability it holds that

$$\begin{aligned}
I(h) &= \sum_{m=0}^{\infty} \mathsf{P}(\cap_{k=1}^n \{N(t_k + h) - N(t_{k-1} + h) = x_k\} \mid N(t_n + h) = m) \cdot \mathsf{P}(N(t_n + h) = m) \\
&= \sum_{m=0}^{\infty} \frac{m!}{x_1! \ldots x_n!} \prod_{k=1}^{n} \left(\frac{t_k + h - t_{n-1} - h}{t_n + h - h} \right)^{x_k} e^{-\lambda(t_n + h)} \frac{(\lambda(t_n + h))}{m!} \\
&= \sum_{m=0}^{\infty} \mathsf{P}(\cap_{k=1}^n \{N(t_k) - N(t_{k-1}) = x_k \mid N(t_n + h) = m) \times \mathsf{P}(N(t_n + h) = m) = I(0).
\end{aligned}$$

We now show property 4b) for $h \in (0, 1)$:

$$\begin{aligned}
\mathsf{P}(N(h) = 0) &= \sum_{k=0}^{\infty} \mathsf{P}(N(h) = 0, N(1) = k) = \sum_{k=0}^{\infty} \mathsf{P}(N(h) = 0, N(1) - N(h) = k) \\
&= \sum_{k=0}^{\infty} \mathsf{P}(N(1) - N(h) = k, N(1) = k) \\
&= \sum_{k=0}^{\infty} \mathsf{P}(N(1) = k) \mathsf{P}(N(1) - N(h) = k \mid N(1) = k) \\
&= \sum_{k=0}^{\infty} \mathsf{P}(N(1) = k)(1 - h)^k.
\end{aligned}$$

We have to show that $\mathsf{P}(N(h) = 0) = 1 - \lambda h + o(h)$, i.e., $lim_{h \to \infty} \frac{1}{h}(1 - \mathsf{P}(N(h) = 0)) = \lambda$. Indeed it holds that

$$\begin{aligned}
\frac{1}{h}(1 - \mathsf{P}(N(h) = 0)) &= \frac{1}{h}\left(1 - \sum_{k=0}^{\infty} \mathsf{P}(N(1) = k)(1 - h)^k\right) = \sum_{k=1}^{\infty} \mathsf{P}(N(1) = k) \cdot \frac{1 - (1 - h)^k}{h} \\
&\xrightarrow[h \to 0]{} \sum_{k=1}^{\infty} \mathsf{P}(N(1) = k) \underbrace{\lim_{h \to 0} \frac{1 - (1 - h)^k}{h}}_{k} \\
&= \sum_{k=0}^{\infty} \mathsf{P}(N(1) = k) k = \mathsf{E} N(1) = \lambda,
\end{aligned}$$

since the series uniformly converges in h because it is dominated by $\sum_{k=0}^{\infty} \mathsf{P}(N(1) = k)k = \lambda < \infty$ because of the inequality $(1 - h)^k \geq 1 - kh$, $h \in (0, 1)$, $k \in \mathbb{N}$.
Similarly one can show that $\lim_{h \to 0} \frac{\mathsf{P}(N(h)=1)}{h} = \lim_{h \to 0} \sum_{k=1}^{\infty} \mathsf{P}(N(1) = k)k(1 - h)^{k-1} = \lambda$.
4) $\Rightarrow$ 5)
We have to show that for an arbitrary $n \in \mathbb{N}$ and $t \geq 0$

$$p_n(t) = \mathsf{P}(N(t) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \tag{2.2.1}$$

holds. We will proof that by induction with respect to n. First we show that $p_0(t) = e^{-\lambda t}$, $h = 0$. For that we consider $p_0(t + h) = \mathsf{P}(N(t + h) = 0) = \mathsf{P}(N(t) = 0, N(t + h) - N(t) =$

$0) = p_0(t)p_0(h) = p_0(t)(1 - \lambda h + o(h)),\ h \to 0$. Similarly one can show that $p_0(t) = p_0(t - h)(1 - \lambda h + o(h)),\ h \to 0$. Thus $p_0'(t) = \lim_{h\to 0} \frac{p_0(t+h)-p_0(t)}{h} = -\lambda p_0(t),\ t > 0$ holds. Since $p_0(0) = \mathsf{P}(N(0) = 0) = 1$, it follows from

$$\begin{cases} p_0'(t) &= -\lambda p_0(t) \\ p_0(0) &= 0, \end{cases}$$

that it exists an unique solution $p_0(t) = e^{-\lambda t}$, $t \geq 0$. Now for n the formular (2.2.1) be approved. Let's proof it for $n+1$.

$$\begin{aligned} p_{n+1}(t+h) &= \mathsf{P}(N(t+h) = n+1) \\ &= \mathsf{P}(N(t) = n, N(t+h) - N(t) = 1) + \mathsf{P}(N(t) = n+1, N(t+h) - N(t) = 0) \\ &= p_n(t) - p_1(h) + p_{n+1}(t) - p_0(h) \\ &= p_n(t)(\lambda h + o(h)) + p_{n+1}(t)(1 - \lambda h + o(h)),\ h \to 0, h > 0. \end{aligned}$$

Thus

$$\begin{cases} p_{n+1}'(t) &= -\lambda p_{n+1}(t) + \lambda p_n(t),\ t > 0 \\ p_{n+1}(0) &= 0 \end{cases} \tag{2.2.2}$$

Since $p_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$, we obtain $p_{n+1}(t) = e^{-\lambda t}\frac{(\lambda t)^{n+1}}{(n+1)!}$ as solution of (2.2.2). (Indeed
$p_{n+1}(t) = C(t)e^{-\lambda t} \Rightarrow C'(t)e^{-\lambda t} = \lambda C(t)e^{-\lambda t} \ldots\ldots\ldots\ldots + \lambda p_n(t)$
$C'(t) = \frac{\lambda^{n+1}t^n}{n!} \Rightarrow C(t) = \frac{\lambda^{n+1}t^{n+1}}{(n+1)!}$, $C(0) = 0)$
$5) \Rightarrow 1)$
Let N be a counting process $N(t) = \max\{n : S_n \leq t\}$, $t \geq 0$, which fulfills conditions 5a) and 5b). We show that $S_n = \sum_{k=1}^n T_k$, where T_k i.i.d. with $T_k \sim \text{Exp}(\lambda)$, $k \in \mathbb{N}$. Since $T_k = S_k - S_{k-1}$, $k \in \mathbb{N}$, $S_0 = 0$, we consider for $b_0 = 0 \leq a_1 < b_1 \leq \ldots \leq a_n < b_n$

$$\begin{aligned} \mathsf{P} & \left(\cap_{k=1}^n \{a_k < S_k \leq b_k\}\right) \\ =\ & \mathsf{P}(\cap_{k=1}^{n-1}\{N(a_k) - N(b_{k-1}) = 0, N(b_k) - N(a_k) = 1\} \\ & \cap\{N(a_n) - N(b_{n-1}) = 0, N(b_n) - N(a_n) \geq 1\}) \\ =\ & \prod_{k=1}^{n-1}(\underbrace{\mathsf{P}(N(a_k - b_{k-1}) = 0)}_{e^{-\lambda(a_k - b_{k-1})}}\underbrace{\mathsf{P}(N(b_k - a_k) = 1)}_{\lambda(b_k - a_k)e^{-\lambda(b_k - a_k)}}) \times \\ & \underbrace{\mathsf{P}(N(a_n - b_{n-1}) = 0)}_{e^{-\lambda(a_n - b_{n-1})}}\underbrace{\mathsf{P}(N(b_n - a_n) \geq 1)}_{(1 - e^{-\lambda(b_n - a_n)})} \\ =\ & e^{-\lambda(a_n - b_{n-1})}(1 - e^{-\lambda(b_n - a_n)})\prod_{k=1}^{n-1}\lambda(b_k - a_k)e^{-\lambda(b_k - b_{k-1})} \\ =\ & \lambda^{n-1}(e^{-\lambda a_n} - e^{-\lambda b_n})\prod_{k=1}^{n-1}(b_k - a_k) = \int_{a_1}^{b_1}\cdots\int_{a_n}^{b_n}\lambda^n e^{-\lambda y_n}dy_n \ldots y_1. \end{aligned}$$

The common density of $(S_1, \ldots, S_n)^\top$ therefore is given by $\lambda^n e^{-\lambda y_n}\mathbf{1}(y_1 \leq y_2 \leq \ldots \leq y_n)$. $\square$

2.2.2 Compound Poisson process

Definition 2.2.3
Let $N = \{N(t),\ t \geq 0\}$ be a homogeneous Poisson process with intensity $\lambda > 0$, build by

means of the sequence $\{T_n\}_{n\in\mathbb{N}}$ of interarrival times. Let $\{U_n\}_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables, independent of $\{T_n\}_{n\in\mathbb{N}}$. Let F_U be the distribution function of U_1. For an arbitrary $t \geq 0$ let $X(t) = \sum_{k=1}^{N(t)} U_k$. The stochastic process $X = \{X(t),\ t \geq 0\}$ is called *compound Poisson process with parameters* λ, F_U. The distribution of $X(t)$ thereby is called *compound Poisson distribution with parameters* λt, F_U.

The compound Poisson process $X(t)$, $t \geq 0$ can be interpreted as the sum of „marks“ U_n of a homogeneous marked Poisson process (N, U) until time t.
In queueing theory $X(t)$ is interpreted as the overall workload of a server until time t if the requests to the service occur at times $S_n = \sum_{k=1}^n T_k$, $n \in \mathbb{N}$ and represent the amount of work U_n, $n \in \mathbb{N}$.
In actuarial mathematics $X(t)$, $t \geq 0$ is the total damage in a portfolio until time $t \geq 0$ with number of damages $N(t)$ and amount of loss U_n, $n \in \mathbb{N}$.

Theorem 2.2.2
Let $X = \{X(t),\ t \geq 0\}$ be a compound Poisson process with parameters λ, F_U. The following properties hold:

1. X has independent increments.
2. If $\hat{m}_U(s) = \mathsf{E}e^{sU_1}$, $s \in \mathbb{R}$, is the moment generating function of U_1, such that $\hat{m}_U(s) < \infty$, $s \in \mathbb{R}$, then it holds that
$$\hat{m}_{X(t)}(s) = e^{\lambda t(\hat{m}_U(s)-1)},\ s \in \mathbb{R},\ t \geq 0, \quad \mathsf{E}X(t) = \lambda t \mathsf{E}U_1,\ \operatorname{var} X(t) = \lambda t \mathsf{E}U_1^2,\ t \geq 0.$$

Proof 1. We have to show that for arbitrary $n \in \mathbb{N}$, $0 \leq t_0 < t_1 < \ldots < t_n$ and h
$$\mathsf{P}\left(\sum_{i_1=N(t_0+h)+1}^{N(t_1+h)} U_{i_1} \leq x_1, \ldots, \sum_{i_n=N(t_{n-1}+h)+1}^{N(t_n+h)} U_{i_n} \leq x_n\right) = \prod_{k=1}^n \mathsf{P}\left(\sum_{i_k=N(t_{k-1})+1}^{N(t_k)} U_{i_k} \leq x_k\right)$$
for arbitrary $x_1, \ldots, x_n \in \mathbb{R}$. Indeed it holds that
$$\begin{aligned}
&\mathsf{P}\left(\sum_{i_1=N(t_0+h)+1}^{N(t_1+h)} U_{i_1} \leq x_1, \ldots, \sum_{i_n=N(t_{n-1}+h)+1}^{N(t_n+h)} U_{i_n} \leq x_n\right)\\
&= \sum_{k_1,\ldots,k_n=0}^{\infty} \left(\prod_{j=1}^n F_n^{*k_j}(x_j)\right) \mathsf{P}\left(\cap_{m=1}^n \{N(t_m+h) - N(t_{m-1}+h) = k_m\}\right)\\
&= \sum_{k_1,\ldots,k_n=0}^{\infty} \left(\prod_{j=1}^n F_n^{*k_j}(x_j)\right) \left(\prod_{m=1}^n \mathsf{P}(N(t_m) - N(t_{m-1}) = k_m)\right)\\
&= \prod_{m=1}^n \sum_{k_m=0}^{\infty} F_n^{*k_m}(x_m) \mathsf{P}(N(t_m) - N(t_{m-1}) = k_m)\\
&= \prod_{m=1}^n \mathsf{P}\left(\sum_{k_m=N(t_{m-1})+1}^{N(t_m)} \leq x_m\right)
\end{aligned}$$

2.

Exercise 2.2.3

□

2.2.3 Cox process

A Cox process is a (in general inhomogeneous) Poisson process with intensity measure Λ which as such is a random measure. The induitive idea is stated in the following definition.

Definition 2.2.4
Let $\Lambda = \{\Lambda(B),\ B \in \mathcal{B}(\mathbb{R}_+)\}$ be a random a.s. locally finite measure. The random counting measure $N = \{N(B),\ B \in \mathcal{B}(\mathbb{R}_+)\}$ is called *Cox counting measure (or doubly stochastic Poisson measure) with random intensity measure* Λ if for arbitrary $n \in \mathbb{N}$, $k_1, \ldots, k_n \in \mathbb{N}_0$ and $0 \le a_1 < b_1 \le a_2 < b_2 \le \ldots \le a_n < b_n$ it holds that $\mathsf{P}(\cap_{i=1}^n \{N((a_i,b_i]) = k_i\}) = \mathsf{E}\left(\prod_{i=1}^n e^{-\lambda((a_i,b_i])} \frac{\Lambda^{k_i}((a_i,b_i])}{k_i!}\right)$. The process $\{N(t),\ t \ge 0\}$ with $N(t) = N((0,t])$ is called *Cox process* (or *doubly stochastic Poisson process*) with random intensity measure Λ.

Example 2.2.1 1. If the random measure Λ is a.s. absolutly continuous with respect to the Lebesgue measure, i.e., $\Lambda(B) = \int_B \lambda(t)dt$, B - bounded, $B \in \mathcal{B}(\mathbb{R}_+)$, where $\{\lambda(t), t \ge 0\}$ is a stochastic process with a.s. Borel-measurable Borel-integrable trajectories, then $\lambda(t) \ge 0$ a.s. for all $t \ge 0$ is called the intensity process of N.

2. In particular, it can be that $\lambda(t) \equiv Y$ where Y is a non negative random variable. Then it holds that $\Lambda(B) = Y\nu_1(B)$, thus N has a random intensity Y. Such Cox processes are called *mixed Poisson processes.*

A Cox process $N = \{N(t),\ t \ge 0\}$ with intensity process $\{\lambda(t),\ t \ge 0\}$ can be build explicitly as the following. Let $\tilde{N} = \{\tilde{N}(t),\ t \ge 0\}$ be a homogeneous Poisson process with intensity 1, which is independent of $\{\lambda(t),\ t \ge 0\}$. Then $N \stackrel{d}{=} N_1$, where the process $N_1 = \{N_1(t),\ t \ge 0\}$ is given by $N_1(t) = \tilde{N}(\int_0^t \lambda(y)dy)$, $t \ge 0$. The assumption $N \stackrel{d}{=} N_1$ of course has to be proven. However, we will assume it without proof. It also is the basis for the simulation of the Cox process N.

2.3 Additional exercises

Exercise 2.3.1
Let $\{N_t\}_{t \ge 0}$ be a renewal process with interarrival times T_i, which are exponentially distributed, i.e. $T_i \sim \text{Exp}(\lambda)$.

a) Proof that: N_t is Poisson distributed for every $t > 0$.

b) Determine the parameter of this Poisson distribution.

c) Determine the renewal function $H(t) = \mathsf{E}\, N_t$.

Exercise 2.3.2
Proof that: A (real-valued) stochastic process $X = \{X(t),\, t \in [0,\infty)\}$ with independent increments already has stationary increments if the distribution of the random variable $X(t+h) - X(h)$ does not depend on h.

Exercise 2.3.3
Let $N = \{N(t),\, t \in [0,\infty)\}$ be a Poisson process with intensity λ. Calculate the probabilities that within the interval $[0,s]$ exactly i events occur under the condition that within the interval $[0,t]$ exactly n events occur, i.e. $\mathsf{P}(N(s) = i \mid N(t) = n)$ for $s < t$, $i = 0, 1, \ldots, n$.

Exercise 2.3.4
Let $N^{(1)} = \{N^{(1)}(t),\, t \in [0,\infty)\}$ and $N^{(2)} = \{N^{(2)}(t),\, t \in [0,\infty)\}$ be independent Poisson processes with intensities λ_1 and λ_2. In this case the independence indicates that the sequences $T_1^{(1)}, T_2^{(1)}, \dots$ and $T_1^{(2)}, T_2^{(2)}, \dots$ are independent. Show that $N = \{N(t) := N^{(1)}(t)+N^{(2)}(t),\, t \in [0,\infty)\}$ is a Poisson process with intensity $\lambda_1 + \lambda_2$.

Exercise 2.3.5 (Queuing paradox):
Let $N = \{N(t),\, t \in [0,\infty)\}$ be a renewal process. Then $T(t) = S_{N(t)+1} - t$ is called the *time of excess*, $C(t) = t - S_{N(t)}$ the *current lifetime* and $D(t) = T(t)+C(t) = T_{N(t)+1}$ the *lifetime* at time $t > 0$. Now let $N = \{N(t),\, t \in [0,\infty)\}$ be a Poisson process with intensity λ.

a) Calculate the distribution of the time of excess $T(t)$.

b) Show that the distribution of the current lifetime is given by $\mathsf{P}(C(t) = t) = \mathrm{e}^{-\lambda t}$ and the density is given by $f_{C(t)|N(t)>0}(s) = \lambda \mathrm{e}^{-\lambda s} \mathbb{1}\{s \le t\}$.

c) Show that $\mathsf{P}(D(t) \le x) = (1 - (1 + \lambda \min\{t, x\})\mathrm{e}^{-\lambda x})\mathbb{1}\{x \ge 0\}$.

d) To determine $\mathsf{E}T(t)$, one could argue like this: On average t lies in the middle of the surrounding interval of interarriving time $(S_{N(t)}, S_{N(t)+1})$, i.e. $\mathsf{E}T(t) = \frac{1}{2}\mathsf{E}(S_{N(t)+1} - S_{N(t)}) = \frac{1}{2}\mathsf{E}T_{N(t)+1} = \frac{1}{2\lambda}$. Considering the result from part (a) this reason is false. Where is the mistake in the reasoning?

Exercise 2.3.6
Gegeben sei ein zusammengesetzter Poisson-Prozess $X = \{X(t) := \sum_{i=1}^{N(t)} U_i,\, t \ge 0\}$. Sei $M_{N(t)}(s) = \mathsf{E}s^{N(t)}$, $s \in (0,1)$, die erzeugende Funktion des Poisson-Prozesses $N(t)$, $\mathcal{L}\{U\}(s) = \mathsf{E}\exp\{-sU\}$ die Laplace-Transformierte von U_i, $i \in \mathbb{N}$, und $\mathcal{L}\{X(t)\}(s)$ die Laplace-Transformierte von $X(t)$. Beweisen Sie, dass

$$\mathcal{L}\{X(t)\}(s) = M_{N(t)}(\mathcal{L}\{U\}(s)), \quad s \ge 0$$

gilt.

Exercise 2.3.7
Given is a compound Poisson process $X = \{X(t),\, t \in [0,\infty)\}$ with U_i i.i.d., $U_1 \sim \mathrm{Exp}(\gamma)$, where the intensity of $N(t)$ is given by λ. Show that for the Laplace transform $\mathcal{L}\{X(t)\}(s)$ of $X(t)$ it holds:

$$\mathcal{L}\{X(t)\}(s) = \exp\left\{-\frac{\lambda t s}{\gamma + s}\right\}.$$

Exercise 2.3.8
Write a function in **R** (alternatively: Java) to which we pass the parameters time t, intensity λ and a value γ. The return of the function is a random value of a compound Poisson process with characteristics $(\lambda, \mathrm{Exp}(\gamma))$ at time t. *Note: the results have to be printed in commented, structured and readable form.*

Exercise 2.3.9
The stochastic process $N = \{N(t),\, t \in [0,\infty)\}$ be a Cox process with intensity function $\lambda(t) = Z$, where Z is a discrete random variable which takes values λ_1 and λ_2 with probability $1/2$. Determine the moment generating function as well as the expected value and the variance of $N(t)$.

Exercise 2.3.10
Given are two independent homogeneous Poisson processes $N^{(1)} = \{N^{(1)}(t),\, t \in [0,\infty)\}$ and $N^{(2)} = \{N^{(2)}(t),\, t \geq 0\}$ with intensities λ_1 and λ_2. Moreover, $X \geq 0$ be an arbitrary non negative random variable which is independent of $N^{(1)}$ and $N^{(2)}$. Show that the process $N = \{N(t),\, t \geq 0\}$ with

$$N(t) = \begin{cases} N^{(1)}(t), & t \leq X, \\ N^{(1)}(X) + N^{(2)}(t - X), & t > X \end{cases}$$

is a Cox process whose intensity processe $\lambda = \{\lambda(t),\, t \geq 0\}$ is given by

$$\lambda(t) = \begin{cases} \lambda_1, & t \leq X, \\ \lambda_2, & t > X. \end{cases}$$

3 Wiener process

3.1 Elementary properties

In example 2) of section 1.2 we defined the *Brownian motion* (or *Wiener process*) $W = \{W(t),\ t \geq 0\}$ as an Gaussian process with $\mathsf{E}W(t) = 0$ and $\mathsf{cov}(W(s), W(t)) = \min\{s, t\}$, $s, t \geq 0$. We now give a new (equivalent) definition.

Definition 3.1.1
A stochastic process $W = \{W(t),\ t \geq 0\}$ is called Wiener process (or Brownian motion) if

1. $W(0) = 0$ a.s.
2. W possesses independent increments
3. $W(t) - W(s) \sim \mathcal{N}(0, t - s)$, $0 \leq s < t$

The existence of W according to definition 3.1.1 follows from theorem 1.7.1 since $\varphi_{s,t}(z) = \mathsf{E}e^{iz(W(t)-W(s))} = e^{-\frac{(t-s)z^2}{2}}$, $z \in \mathbb{R}$, and $e^{-\frac{(t-u)z^2}{2}} e^{-\frac{(u-s)z^2}{2}} = e^{-\frac{(t-s)z^2}{2}}$ for $0 \leq s < u < t$, thus $\varphi_{s,u}(z)\varphi_{u,t}(z) = \varphi_{s,t}(z)$, $z \in \mathbb{R}$. From theorem 1.3.1 the existence of a version with continuous trajectories follows.

Exercise 3.1.1
Show that the theorem holds for $\alpha = 3$, $\sigma = \frac{1}{2}$.

The Wiener process is called after the mathematician Norbert Wiener (1894 - 1964). Why does the Brownian motion exist? According to theorem of Kolmogorov (theorem 1.1.2) it exists a real-valued Gaussian process $X = \{X(t),\ t \geq 0\}$ with mean value $\mathsf{E}X(t) = \mu(t)$, $t \geq 0$, and covariance function $\mathsf{cov}(X(s), X(t)) = C(s, t)$, $s, t \geq 0$ for every function $\mu : \mathbb{R}_+ \to \mathbb{R}$ and every positive semidefinite function $C : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$. We just have to show that $C(s, t) = \min\{s, t\}$, $s, t \geq 0$ is positive semidefinite.

Exercise 3.1.2
Proof this!

Therefore, often it is assumed that the Wiener process possesses continuous paths (just take its corresponding version)

Theorem 3.1.1
Both definitions of the Wiener process are equivalent.

Proof 1. From definition in section 1.2 follows definition 3.1.1.
$W(0) = 0$ a.s. follows from $\mathsf{var}(W(0)) = \min\{0, 0\} = 0$. Now we proof that the increments of W are independent. If $Y \sim \mathcal{N}(\mu, K)$ is a n-dimensional Gaussian random vector and A a $(n \times n)$-matrix, then $AY \sim \mathcal{N}(A\mu, AKA^\top)$ holds, this follows from the explicit form of the characteristic function of Y. Now let $n \in \mathbb{N}$, $0 = t_0 \leq t_1 < \ldots < t_n$, $Y =$

$(W(t_0), W(t_1), \dots, W(t_n))^\top$. For $Z = (W(t_0), W(t_1) - W(t_0), \dots, W(t_n) - W(t_{n-1}))^\top$ it holds that $Z = AY$, where

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & \dots & 0 \\ -1 & 1 & 0 & \dots & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix}.$$

Thus Z is also Gaussian with a covariance matrix which is diagonal. Indeed it holds $\mathsf{cov}(W(t_{i+1}) - W(t_i), W(t_{j+1}) - W(t_j)) = \min\{t_{i+1}, t_{j+1}\} - \min\{t_{i+1}, t_j\} - \min\{t_i, t_{j+1}\} + \min\{t_i, t_j\} = 0$ for $i \neq j$. Thus the coordinates of Z are uncorrelated, which means independence in case of a multivariate Gaussian distribution. Thus the increments of W are independent. Moreover, for arbitrary $0 \leq s < t$ it holds that $W(t) - W(s) \sim \mathcal{N}(0, t-s)$. The normal distribution follows since $Z = AY$ is Gaussian, obviously it holds that $\mathsf{E}W(t) - \mathsf{E}W(s) = 0$ and $\mathsf{var}(W(t) - W(s)) = \mathsf{var}(W(t)) - 2\,\mathsf{cov}(W(s), W(t)) + \mathsf{var}(W(s)) = t - 2\min\{s,t\} + s = t - s$.

2. From definition 3.1.1 the definition in section 1.2 follows.
   Since $W(t) - W(s) \sim \mathcal{N}(0, t-s)$ for $0 \leq s < t$, it holds $\mathsf{cov}(W(s), W(t)) = \mathsf{E}[W(s)(W(t) - W(s) + W(s))] = \mathsf{E}W(s)\mathsf{E}(W(t) - W(s)) + \mathsf{var}\,W(s) = s$, thus it holds $\mathsf{cov}(W(s), W(t)) = \min\{s,t\}$. From $W(t) - W(s) \sim \mathcal{N}(0, t-s)$ and $W(0) = 0$ it also follows that $\mathsf{E}W(t) = 0$, $t \geq 0$. Since W is a Gaussian process, point 1) of the proof follows from the relation $Y = A^{-1}Z$.

□

Definition 3.1.2
The process $\{W(t),\ t \geq 0\}$, $W(t) = (W_1(t), \dots, W_d(t))^\top$, $t \geq 0$, is called *d-dimensional Brownian motion* if $W_i = \{W_i(t),\ t \geq 0\}$ are independent Wiener processes, $i = 1, \dots, d$.

The definitions above and exercise 3.1.1 ensure the existence of a Wiener process with continuous paths. How do we find an explicit way of building these paths? We will talk about that in the proximate section.

3.2 Explicit construction of a Wiener process

First we construct the Wiener process on the interval $[0,1]$. The main idea of the construction is to introduce a stochastic process $X = \{X(t),\ t \in [0,1]\}$ which is defined on a subprobability space of $(\Omega, \mathcal{A}, \mathsf{P})$ with $X \stackrel{d}{=} W$, where $X(t) = \sum_{n=1}^\infty c_n(t) Y_n$, $t \in [0,1]$, $\{Y_n\}_{n\in\mathbb{N}}$ is a sequence of i.i.d. $\mathcal{N}(0,1)$-random variables and $c_n(t) = \int_0^t H_n(s)ds$, $t \in [0,1]$, $n \in \mathbb{N}$. Here, $\{H_n\}_{n\in\mathbb{N}}$ is the orthonormed hair basis in $L_2([0,1])$ which is introduced shortly now.

3.2.1 Hair- and Schauder-functions

Definition 3.2.1
The functions $H_n : [0,1] \to \mathbb{R}$, $n \in \mathbb{N}$, are called *hair functions* if $H_1(t) = 1$, $t \in [0,1]$, $H_2(t) = \mathbf{1}_{[0,\frac{1}{2}]}(t) - \mathbf{1}_{(\frac{1}{2},1]}(t)$, $H_k(t) = 2^{\frac{n}{2}}(\mathbf{1}_{I_{n,k}}(t) - \mathbf{1}_{J_{n,k}}(t))$, $t \in [0,1]$, $2^n < k \leq 2^{n+1}$, where $I_{n,k} = [a_{n,k}, a_{n,k} + 2^{-n-1}]$, $J_{n,k} = (a_{n,k} + 2^{-n-1}, a_{n,k} + 2^{-n}]$, $a_{n,k} = 2^{-n}(k - 2^n - 1)$, $n \in \mathbb{N}$.

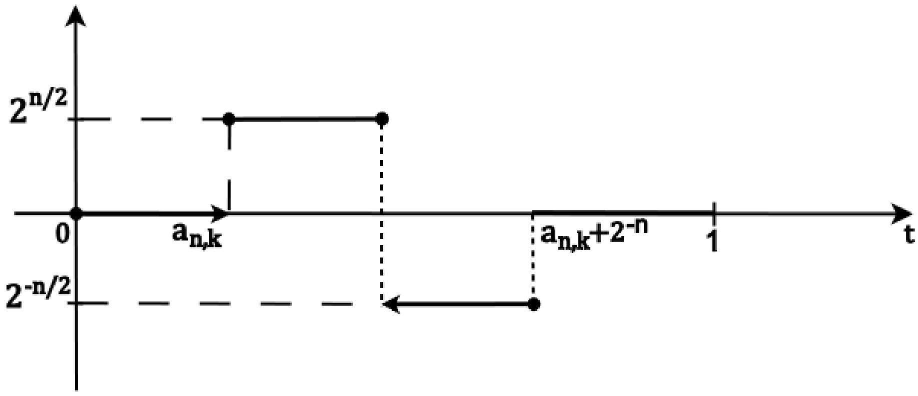

Abb. 3.1: Hair functions

Lemma 3.2.1
The function system $\{H_n\}_{n\in\mathbb{N}}$ is an orthonormed basis in $L^2([0,1])$ with scalar product $< f,g >= \int_0^1 f(t)g(t)dt,\ f,g \in L^2([0,1])$.

Proof The orthonormality of the system $< H_k, H_n >= \delta_{kn},\ k,n \in \mathbb{N}$ directly follows from definition 3.2.1. Now we proof the completeness of $\{H_n\}_{n\in\mathbb{N}}$. It is sufficient to show that for arbitrary function $g \in L^2([0,1])$ with $< g, H_n >= 0,\ n \in \mathbb{N}$, it holds $g = 0$ almost everywhere on $[0,1]$. In fact, we always can write the indicator function of an interval $1_{[a_{n,k}, a_{n,k}+2^{-n-1}]}$ as a linear combination of $H_n,\ n \in \mathbb{N}$.

$$\begin{aligned}
1_{[0,\frac{1}{2}]} &= \frac{(H_1 + H_2)}{2}, \\
1_{(\frac{1}{2},1]} &= \frac{(H_1 - H_2)}{2}, \\
1_{[0,\frac{1}{4}]} &= \frac{(1_{[0,\frac{1}{2}]} + \frac{1}{\sqrt{2}}H_2)}{2}, \\
1_{(\frac{1}{4},\frac{1}{2}]} &= \frac{(1_{[0,\frac{1}{2}]} - \frac{1}{\sqrt{2}}H_2)}{2}, \\
&\ \vdots \\
1_{[a_{n,k}, a_{n,k}+2^{-n-1}]} &= \frac{(1_{a_{n,k}, a_{n,k}+2^{-n}} + 2^{-\frac{n}{2}H_k})}{2},\ \ 2^n < k \le 2^{n+1}.
\end{aligned}$$

Therefore it holds $\int_{\frac{k}{2^n}}^{\frac{(k+1)}{2^n}} g(t)dt = 0,\ n \in \mathbb{N}_0,\ k = 1, \dots, 2^n - 1$, and thus $G(t) = \int_0^t g(s)ds = 0$ for $t = \frac{k}{2^n},\ n \in \mathbb{N}_0,\ k = 1, \dots, 2^n - 1$. Since G is continuous on $[0,1]$, it follows $G(t) = 0$, $t \in [0,1]$, and thus $g(s) = 0$ for almost every $s \in [0,1]$. □

From lemma 3.2.1 it follows that two arbitrary functions $f, g \in L^2([0,1])$ have notations $f = \sum_{n=1}^{\infty} < f, H_n > H_n$ and $g = \sum_{n=1}^{\infty} < g, H_n > H_n$ (these series converge in $L^2([0,1])$) and $< f,g >= \sum_{n=1}^{\infty} < f, H_n >< g, H_n >$ (Parseval identity).

Definition 3.2.2
The functions $S_n(t) = \int_0^t H_n(s)ds =< 1_{[0,t]}, H_n >,\ t \in [0,1],\ n \in \mathbb{N}$ are called *Schauder functions*.

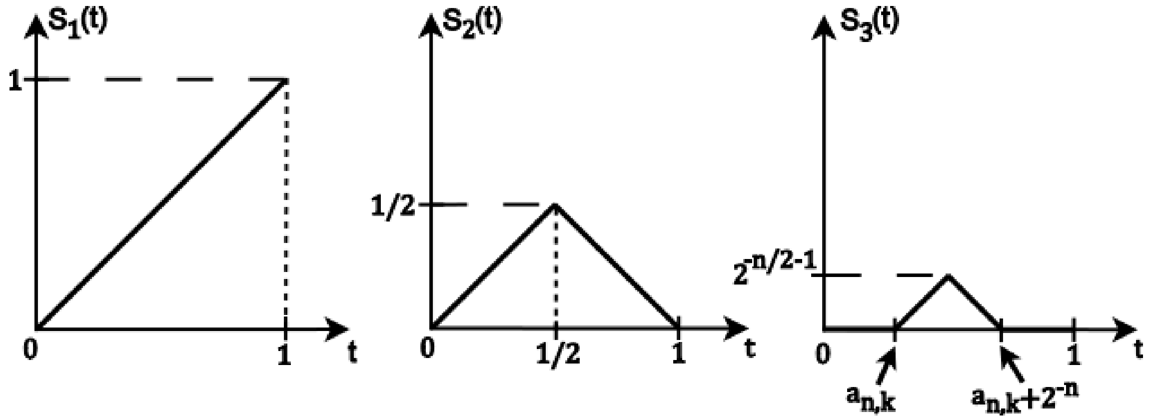

Abb. 3.2: Schauder functions

Lemma 3.2.2
It holds:

1. $S_n(t) \geq 0$, $t \in [0,1]$, $n \in \mathbb{N}$,
2. $\sum_{k=1}^{2^n} S_{2^n+k}(t) \leq \frac{1}{2} 2^{-\frac{n}{2}}$, $t \in [0,1]$, $n \in \mathbb{N}$,
3. Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence of real numbers with $a_n = O(n^\varepsilon)$, $\varepsilon < \frac{1}{2}$, $n \to \infty$. Then the series $\sum_{n=1}^{\infty} a_n S_n(t)$ converges absolutly and uniformly in $t \in [0,1]$ and therefore is a continuous function on $[0,1]$.

Proof 1. follows directly from definition 3.2.2.

2. follows since functions S_{2^n+k} for $k = 1, \ldots, 2^n$ have disjoint supports and $S_{2^n+k}(t) \leq S_{2^n+k}(\frac{2k-1}{2^{n-1}}) = 2^{-\frac{n}{2}-1}$, $t \in [0,1]$.
3. It is sufficien to show that $R_m = \sup_{t\in[0,1]} \sum_{k>2^n} |a_k| S_k(t) \xrightarrow[n\to\infty]{} 0$. For every $k \in \mathbb{N}$ and $c > 0$ it holds $|a_k| \leq ck^\varepsilon$. Therefore it holds for all $t \in [0,1]$, $n \in \mathbb{N}$

$$\sum_{2^n<k\leq 2^{n+1}} |a_k| S_k(t) \leq c \cdot 2^{(n+1)\varepsilon} \cdot \sum_{2^n<k\leq 2^{n+1}} S_k(t) \leq c \cdot 2^{(n+1)\varepsilon} \cdot 2^{-\frac{n}{2}-1} \leq c \cdot 2^{\varepsilon - n(\frac{1}{2}-\varepsilon)}.$$

Since $\varepsilon < \frac{1}{2}$, it holds $R_m \leq c \cdot 2^\varepsilon \sum_{n\geq m} 2^{-n(\frac{1}{2}-\varepsilon)} \xrightarrow[m\to\infty]{} 0$.

□

Lemma 3.2.3
Let $\{Y_n\}_{n\in\mathbb{N}}$ be a sequence of (not necessarily independent) random variables defined on $(\Omega, \mathcal{A}, \mathsf{P})$, $Y_n \sim \mathcal{N}(0,1)$, $n \in \mathbb{N}$. Then it holds $|Y_n| = O((\log n)^{\frac{1}{2}})$, $n \to \infty$.

Proof We have to show that for $c > \sqrt{2}$ and almost all $\omega \in \Omega$ it exists a $n_0 = n_0(\omega, c) \in \mathbb{N}$ such that $|Y_n| \leq c(\log n)^{\frac{1}{2}}$ for $n \geq n_0$. If $Y \sim \mathcal{N}(0,1)$, $x > 0$, it holds

$$\begin{aligned}
\mathsf{P}(Y > x) &= \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{y^2}{2}} dy = \frac{1}{\sqrt{2\pi}} \int_x^\infty \left(-\frac{1}{y}\right) d\left(e^{-\frac{y^2}{2}}\right) \\
&= \frac{1}{\sqrt{2\pi}} \left(\frac{1}{x} e^{-\frac{y^2}{2}} - \int_x^\infty e^{-\frac{y^2}{2}} \frac{1}{y^2} dy\right) \leq \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-\frac{x^2}{2}}.
\end{aligned}$$

(We also can show that $\bar{\Phi}(x) \sim \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-\frac{x^2}{2}}$, $x \to \infty$.) Thus for $c > \sqrt{2}$ it holds

$$\sum_{n \geq 2} \mathsf{P}(|Y_n| > c(\log n)^{\frac{1}{2}}) \leq c^{-1} \frac{2}{\sqrt{2\pi}} \sum_{n \geq 2} (\log n)^{-\frac{1}{2}} e^{-\frac{c^2}{2} \log n} = \frac{c^{-1}\sqrt{2}}{\sqrt{\pi}} \sum_{n \geq 2} (\log n)^{-\frac{1}{2}} n^{-\frac{c^2}{2}} < \infty.$$

According to the lemma of Borel-Cantelli (cf. WR, lemma 2.2.1) it holds $\mathsf{P}(\cap_n \cup_{k \geq n} A_k) = 0$ if $\sum_k \mathsf{P}(A_k) < \infty$ with $A_k = \{|Y_k| > e \cdot (\log k)^{\frac{1}{2}}\}$, $k \in \mathbb{N}$. Thus A_k occurs in infinite number only with probability 0, with $|Y_n| \leq c(\log n)^{\frac{1}{2}}$ for $n \geq n_0$. □

3.2.2 Wiener process with a.s. continuous paths

Lemma 3.2.4
Let $\{Y_n\}_{n \in \mathbb{N}}$ be a sequence of independent $\mathcal{N}(0,1)$-distributed random variables. Let $\{a_n\}_{n \in \mathbb{N}}$ and $\{b_n\}_{n \in \mathbb{N}}$ be sequences of numbers with $\sum_{k=1}^{2^m} |a_{2^m+k}| \leq 2^{-\frac{m}{2}}$, $\sum_{k=1}^{2^m} |b_{2^m+k}| \leq 2^{-\frac{m}{2}}$, $m \in \mathbb{N}$. Then the limits $U = \sum_{n=1}^{\infty} a_n Y_n$ and $V = \sum_{n=1}^{\infty} b_n Y_n$, $U \sim \mathcal{N}(0, \sum_{n=1}^{\infty} a_n^2)$, $V \sim \mathcal{N}(0, \sum_{n=1}^{\infty} b_n^2)$ exist a.s., where $\mathsf{cov}(U,V) = \sum_{n=1}^{\infty} a_n b_n$. U and V are independent if and only if $\mathsf{cov}(U,V) = 0$.

Proof Lemma 3.2.2 and 3.2.3 reveal the a.s. existence of the limits U and V (replace a_n by Y_n and S_n by e.g. b_n in lemma 3.2.2). From the stability under convolution of the normal distribution it follows for $U^{(m)} = \sum_{n=1}^{m} a_n Y_n$, $V^{(m)} = \sum_{n=1}^{m} b_n Y_n$, that $U^{(m)} \sim \mathcal{N}(0, \sum_{n=1}^{m} a_n^2)$, $V^{(m)} \sim \mathcal{N}(0, \sum_{n=1}^{m} b_n^2)$. Since $U^{(m)} \xrightarrow{d} U$, $V^{(m)} \xrightarrow{d} V$ it follows $U \sim \mathcal{N}(0, \sum_{n=1}^{\infty} a_n^2)$, $V \sim \mathcal{N}(0, \sum_{n=1}^{\infty} b_n^2)$. Moreover, it holds

$$\begin{aligned}
\mathsf{cov}(U,V) &= \lim_{m \to \infty} \mathsf{cov}(U^{(m)}, V^{(m)}) \\
&= \lim_{m \to \infty} \sum_{i,j=1}^{m} a_i b_j \, \mathsf{cov}(Y_i, Y_j) \\
&= \lim_{m \to \infty} \sum_{i=1}^{m} a_i b_i = \sum_{i=1}^{\infty} a_i b_i,
\end{aligned}$$

according to the dominated convergence theorem of Lebesgue, since according to lemma 3.2.3 it holds $|Y_n| \leq c \underbrace{(\log n)^{\frac{1}{2}}}_{\leq cn^{\varepsilon},\ \varepsilon < \frac{1}{2}}$, for $n \geq \mathbb{N}_0$, and the dominated series converges according to lemma 3.2.2:

$$\sum_{n,k=2^m}^{2^{m+1}} a_n b_k Y_n Y_k \overset{f.s.}{\leq} \sum_{n,k=2^m}^{2^{m+1}} a_n b_k c^2 n^{\varepsilon} k^{\varepsilon} \leq 2^{2\varepsilon(m+1)} \cdot 2^{-\frac{m}{2}} \cdot 2^{-\frac{m}{2}} = 2^{-(1-2\varepsilon)m}, \quad 1 - 2\varepsilon > 0.$$

For sufficient large m it holds $\sum_{n,k=m}^{\infty} a_n b_k Y_n Y_k \leq \sum_{j=m}^{\infty} 2^{-(1-2\varepsilon)j} < \infty$, and this series converges a.s.
Now we show

$$\mathsf{cov}(U,V) = 0 \iff U \text{ and } V \text{ are independent}$$

Independence always results in the uncorrelation of random variables. We proof the other direction. From $(U^{(m)}, V^{(m)}) \xrightarrow[m\to\infty]{d} (U,V)$ it follows $\varphi_{(U^{(m)},V^{(m)})} \xrightarrow[m\to\infty]{} \varphi_{(U,V)}$, thus

$$\begin{aligned}
\varphi_{(U^{(m)},V^{(m)})}(s,t) &= \lim_{m\to\infty} \mathsf{E}\exp\{i(t\sum_{k=1}^{m} a_k Y_k + s\sum_{n=1}^{m} b_n Y_n)\} \\
&= \lim_{m\to\infty} \mathsf{E}\exp\{i\sum_{k=1}^{m}(ta_k + sb_k)Y_k\} = \lim_{m\to\infty}\prod_{k=1}^{m}\mathsf{E}\exp\{i(ta_k+sb_k)Y_k\} \\
&= \lim_{m\to\infty}\prod_{k=1}^{m}\exp\{-\frac{(ta_k+sb_k)^2}{2}\} = \exp\{-\sum_{k=1}^{\infty}\frac{(ta_k+sb_k)^2}{2}\} \\
&= \exp\{-\frac{t^2}{2}\sum_{k=1}^{\infty}a_k^2\}\exp\{ts\underbrace{\sum_{k=1}^{\infty}a_k b_k}_{\mathsf{cov}(U,V)=0}\}\exp\{-\frac{s^2}{2}\sum_{k=1}^{\infty}b_k^2\} = \varphi_U(t)\varphi_V(s),
\end{aligned}$$

$s,t\in\mathbb{R}$. Thus, U and V are independent if $\mathsf{cov}(U,V)=0$. □

Theorem 3.2.1
Let $\{Y_n,\ n\in\mathbb{N}\}$ be a sequence of i.i.d. random variables that are $\mathcal{N}(0,1)$-distributed, defined on a probability space $(\Omega,\mathcal{A},\mathsf{P})$. Then it exists a subprobability space $(\Omega_0,\mathcal{A}_0,\mathsf{P})$ of $(\Omega,\mathcal{A},\mathsf{P})$ and a stochastic process $X=\{X(t),\ t\in[0,1]\}$ on it such that $X(t)=\sum_{n=1}^{\infty}Y_nS_n(t)$, $t\in[0,1]$, and $X\stackrel{d}{=}W$. Here, $\{S_n\}_{n\in\mathbb{N}}$ is the family of Schauder functions.

Proof According to lemma 3.2.2, 2) the coefficients $S_n(t)$ fulfill the conditions of lemma 3.2.4 for every $t\in[0,1]$. In addition to that it exists according to lemma 3.2.3 a subset $\Omega_0\subset\Omega$, $\Omega_0\in\mathcal{A}$ with $\mathsf{P}(\Omega_0)=1$, such that for every $\omega\in\Omega_0$ the relation $|Y_n(\omega)|=O(\sqrt{\log n})$, $n\to\infty$, holds. Let $\mathcal{A}_0=\mathcal{A}\cap\Omega_0$. We restrict the probability space to $(\Omega_0,\mathcal{A}_0,\mathsf{P})$. Then condition $a_n=Y_n(\omega)=O(n^{\varepsilon})$, $\varepsilon<\frac{1}{2}$, is fulfilled since $\sqrt{\log n}<n^{\varepsilon}$ for sufficient large n, and according to lemma 3.2.2, 3) the series $\sum_{n=1}^{\infty}Y_n(\omega)S_n(t)$ converges absolutly and uniformly in $t\in[0,1]$ to the function $X(\omega,t)$, $\omega\in\Omega_0$, which is a continuous function in t for every $\omega\in\Omega_0$. $X(\cdot,t)$ is a random variable since in lemma 3.2.4 the convergence of this series holds almost surely. Moreover it holds $X(t)\sim\mathcal{N}(0,\sum_{n=1}^{\infty}S_n^2(t))$, $t\in[0,1]$.
We show that this stochastic process, defined on $(\Omega_0,\mathcal{A}_0,\mathsf{P})$, is a Wiener process. For that we check the conditions of definition 3.1.1. We consider arbitrary times $0\le t_1<t_2,t_3<t_4\le 1$

and evaluate

$$\begin{aligned}
\mathsf{cov}(X(t_2)-X(t_1),X(t_4)-X(t_3)) &= \mathsf{cov}(\sum_{n=1}^{\infty} Y_n(S_n(t_2)-S_n(t_1)),\sum_{n=1}^{\infty} Y_n(S_n(t_4)-S_n(t_3))) \\
&= \sum_{n=1}^{\infty}(S_n(t_2)-S_n(t_1))(S_n(t_4)-S_n(t_3)) \\
&= \sum_{n=1}^{\infty}(< H_n, 1_{[0,t_2]} > - < H_n, 1_{[0,t_1]} >) \times \\
&\quad (< H_n, 1_{[0,t_4]} > - < H_n, 1_{[0,t_3]} >) \\
&= \sum_{n=1}^{\infty} < H_n, 1_{[0,t_2]} - 1_{[0,t_1]} >< H_n, 1_{[0,t_4]} - 1_{[0,t_3]} > \\
&= < 1_{[0,t_2]} - 1_{[0,t_1]}, 1_{[0,t_4]} - 1_{[0,t_3]} > \\
&= < 1_{[0,t_2]}, 1_{[0,t_4]} > - < 1_{[0,t_1]}, 1_{[0,t_4]} > \\
&\quad - < 1_{[0,t_2]}, 1_{[0,t_3]} > + < 1_{[0,t_1]}, 1_{[0,t_3]} > \\
&= \min\{t_2,t_4\} - \min\{t_1,t_4\} - \min\{t_2,t_3\} + \min\{t_1,t_3\},
\end{aligned}$$

since $< 1_{[0,s]}, 1_{[0,t]} >= \int_0^{\min\{s,t\}} du = \min\{s,t\}$, $s,t \in [0,1]$. If $0 \le t_1 < t_2 \le t_3 < t_4 < 1$, it holds $\mathsf{cov}(X(t_2)-X(t_1),X(t_4)-X(t_3)) = t_2-t_1-t_2+t_1 = 0$, thus the increments of X (according to lemma 3.2.4) are uncorrelated. Moreover it holds $X(0) \sim \mathcal{N}(0,\sum_{n=1}^{\infty} S_n^2(0)) = \mathcal{N}(0,0)$, therefore $X(0) \overset{f.s.}{=} 0$. From that for $t_1=0$, $t_2=t$, $t_3=0$, $t_4=t$ it follows that $\mathsf{var}(X(t)) = t$, $t \in [0,1]$, and for $t_1=t_3=s$, $t_2=t_4=t$, that $\mathsf{var}(X(t)-X(s)) = t-s-s+s = t-s$, $0 \le s < t \le 1$. Thus it holds $X(t)-X(s) \sim \mathcal{N}(0,t-s)$, and according to definition 3.1.1 it holds $X \overset{d}{=} W$. □

Remark 3.2.1

1. Theorem 3.2.1 is the basis for an approximative simulation of the paths of a Brownian motion through the partial sums $X^{(n)}(t) = \sum_{k=1}^{n} Y_k S_k(t)$, $t \in [0,1]$, for sufficient large $n \in \mathbb{N}$.
2. The construction in theorem 3.2.1 can be used to induce the Wiener process with continuous paths on the interval $[0,t_0]$ for arbitrary $t_0 > 0$. If $W = \{W(t),\ t \in [0,1]\}$ is a Wiener process on $[0,1]$ then $Y = \{Y(t),\ t \in [0,t_0]\}$ with $Y(t) = \sqrt{t_0}W(\frac{t}{t_0})$, $t \in [0,t_0]$, is a Wiener process on $[0,t_0]$.

   Exercise 3.2.1
   Proof that.
3. The Wiener process W with continuous paths on $\mathbb{R}_+$ can be constructed as the following. Let $W^{(n)} = \{W^{(n)}(t),\ t \in [0,1]\}$ be independent copies of the Wiener process as in theorem 3.2.1. Define $W(t) = \sum_{n=1}^{\infty} 1(t \in [n-1,n])[\sum_{k=1}^{n-1} W^{(k)}(1) - W^{(n)}(t-(n-1))]$, $t \ge 0$, thus,

$$W(t) = \begin{cases} W^{(1)}(t),\ t \in [0,1], \\ W^{(1)}(1) + W^{(2)}(t-1),\ t \in [1,2], \\ W^{(1)}(1) + W^{(2)}(1) + W^{(3)}(t-2),\ t \in [2,3], \\ \text{etc.} \end{cases}$$

Exercise 3.2.2
Show that the introduced stochastic process $W = \{W(t),\ t \ge 0\}$ is a Wiener process on $\mathbb{R}_+$.

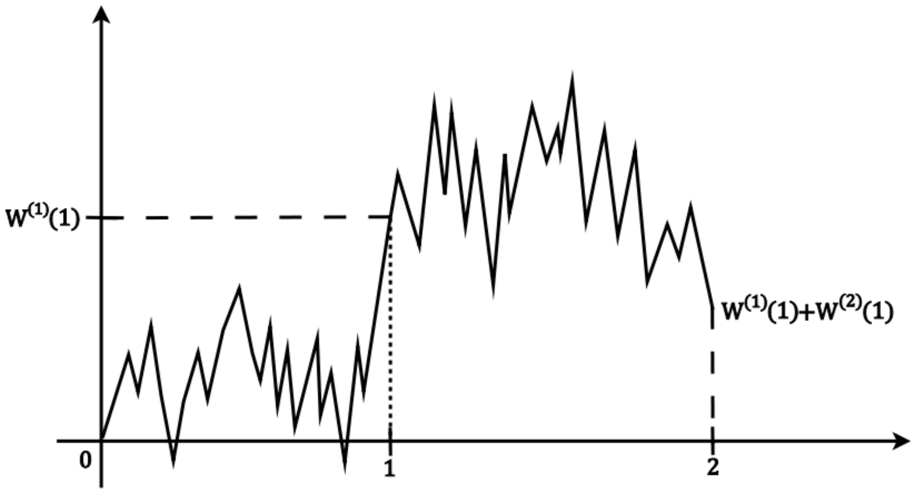

Abb. 3.3:

3.3 Distribution- and path properties of Wiener processes

3.3.1 Distribution of the maximum

Theorem 3.3.1
Let $W = \{W(t),\ t \in [0,1]\}$ be a Wiener process defined on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Then it holds:

$$\mathsf{P}\left(\max_{t\in[0,1]} W(t) > x\right) \leq \sqrt{\frac{2}{\pi}} \int_x^\infty e^{-\frac{y^2}{2}}\, dy, \tag{3.3.1}$$

for all $x \geq 0$.

The mapping $\max_{t\in[0,1]} W(t) : \Omega \to [0,\infty)$ given in 3.3.1 is a well-defined random variable since it holds: $\max_{t\in[0,1]} W(t,\omega) = \lim_{n\to\infty} \max_{i=1,\ldots,k} W(\frac{i}{k},\omega)$ for all $\omega \in \Omega$ since the trajectories of $\{W(t),\ t \in [0,1]\}$ are continuous. From 3.3.1 it follows that $\max_{t\in[0,1]} W(t)$ has an exponential bounded tail: thus $\max_{t\in[0,1]} W(t)$ has finite k-th moments.

Useful ideas for the proof of theorem 3.3.1
Let $\{W(t),\ t \in [0,1]\}$ be a Wiener process and $Z_1, Z_2, \ldots$ a sequence of independent random variables with $\mathsf{P}(Z_i = 1) = \mathsf{P}(Z_i = -1) = \frac{1}{2}$ for all $i \geq 1$. For every $n \in \mathbb{N}$ we define $\{\tilde{W}^n(t),\ t \in [0,1]\}$ by $\tilde{W}^n(t) = \frac{S_{\lfloor nt\rfloor}}{\sqrt{n}} + (nt - \lfloor nt\rfloor)\frac{Z_{\lfloor nt\rfloor+1}}{\sqrt{n}}$, where $S_i = Z_1 + \ldots + Z_i$, $i \geq 1$, $S_0 = 0$.

Lemma 3.3.1
For every $k \geq 1$ and arbitrary $t_1, \ldots, t_k \in [0,1]$ it holds:

$$\left(\tilde{W}^{(n)}(t_1), \ldots, \tilde{W}^{(n)}(t_k)\right)^\top \xrightarrow{d} (W(t_1), \ldots, W(t_k))^\top.$$

Proof Special case $k = 2$ (for $k > 2$ the proof is analogous). Let $t_1 < t_2$. For all $s_1, s_2 \in \mathbb{R}$ it holds:

$$\begin{aligned} s_1\tilde{W}^{(n)}(t_1) + s_2\tilde{W}^{(n)}(t_2) &= (s_1+s_2)\frac{S_{\lfloor nt_1 \rfloor}}{\sqrt{n}} + s_2\frac{(S_{\lfloor nt_2 \rfloor} - S_{\lfloor nt_1 \rfloor + 1})}{\sqrt{n}} \\ &\quad + Z_{\lfloor nt_1 \rfloor + 1}((nt_1 - \lfloor nt_1 \rfloor)\frac{s_1}{\sqrt{n}} + \frac{s_2}{\sqrt{n}}) \\ &\quad + Z_{\lfloor nt_2 \rfloor + 1}(nt_2 - \lfloor nt_2 \rfloor)\frac{s_2}{\sqrt{n}}. \end{aligned}$$

$$\begin{aligned} \lim_{n\to\infty} \mathsf{E} e^{i(s_1\tilde{W}^{(n)}(t_1)+s_2\tilde{W}^{(n)}(t_2))} &= \lim_{n\to\infty} \mathsf{E} e^{i\frac{s_1+s_2}{\sqrt{n}}S_{\lfloor nt_1\rfloor}} \mathsf{E} e^{i\frac{s_2}{\sqrt{n}}(S_{\lfloor nt_2\rfloor} - S_{\lfloor nt_1\rfloor+1})} \\ &= \left| \mathsf{E} e^{i\frac{s_1+s_2}{\sqrt{n}}s_{\lfloor nt_1\rfloor}} = \varphi_{S_{\lfloor nt_1\rfloor}}\left(\frac{s_1+s_2}{\sqrt{n}}\right) = \left(\varphi_{Z_1}\left(\frac{s_1+s_2}{\sqrt{n}}\right)\right)^{\lfloor nt_1\rfloor} \right| \\ &= \lim_{n\to\infty}\left(\varphi_{Z_1}\left(\frac{s_1+s_2}{\sqrt{n}}\right)\right)^{\lfloor nt_1\rfloor}\left(\varphi_{Z_1}\left(\frac{s_2}{\sqrt{n}}\right)\right)^{\lfloor nt_2\rfloor - \lfloor nt_1\rfloor - 1} \\ &= \left|\lim_{n\to\infty}\varphi^n\left(\frac{s}{\sqrt{n}}\right) = e^{-\frac{s^2}{2}}\right| \\ &= e^{-\frac{\left(s_1^2t_1+2s_1s_2\min\{t_1,t_2\}+s_2^2t_2\right)}{2}} \\ &= \varphi_{(W(t_1),W(t_2))}(s_1,s_2), \end{aligned}$$

where $\varphi_{(W(t_1),W(t_2))}$ is the characteristic function of $(W(t_1), W(t_2))$. □

Lemma 3.3.2
Let $\tilde{W}^{(n)} = \max_{t\in[0,1]} \tilde{W}^{(n)}(t)$. Then it holds:

$$\tilde{W}^{(n)} = \frac{1}{\sqrt{n}}\max_{k=1,\ldots,n} S_k, \quad \text{for all } n = 1, 2, \ldots$$

and

$$\lim_{n\to\infty} \mathsf{P}(\tilde{W}^{(n)} \le x) = \sqrt{\frac{2}{\pi}}\int_0^x e^{-\frac{y^2}{2}}\,dy, \quad \text{for all } x \ge 0.$$

Without proof

Proof of theorem 3.3.1
From lemma 3.3.1 it follows for $x \ge 0$, $k \ge 1$ and $t_1, \ldots, t_n \in [0,1]$

$$\begin{aligned} &\lim_{n\to\infty} \mathsf{P}\left(\max_{t\in\{t_1,\ldots,t_k\}} \tilde{W}^{(n)}(t) > x\right) = \mathsf{P}\left(\max_{t\in\{t_1,\ldots,t_k\}} W(t) > x\right) \\ &\Rightarrow \mathsf{P}\left(\max_{t\in[0,1]} \tilde{W}^{(n)}(t) > x\right) \ge \mathsf{P}\left(\max_{t\in\{t_1,\ldots,t_k\}} W(t) > x\right). \end{aligned}$$

With $(t_1,\ldots,t_k)^\top = \left(\frac{1}{k},\ldots,\frac{k}{k}\right)^\top$ and $\max_{t\in[0,1]} W(t,\omega) = \lim_{k\to\infty}\max_{i=1,\ldots,k} W\left(\frac{i}{k},\omega\right)$ it holds

$$\lim_{n\to\infty} \mathsf{P}\left(\max_{t\in[0,1]} \tilde{W}^{(n)}(t) > x\right) \ge \mathsf{P}\left(\max_{t\in[0,1]} W(t) > x\right).$$

The assertion follows from lemma 3.3.2. □

Corollary 3.3.1
Let $\{W(t),\ t \in [0,1]\}$ be a Wiener process. Then it holds:

$$\mathsf{P}\left(\lim_{t\to\infty} \frac{W(t)}{t} = 0\right) = 1.$$

Proof

$$\begin{aligned}
\left|\frac{W(t)}{t} - \frac{W(n)}{n}\right| &\leq \left|\frac{W(t)}{t} - \frac{W(n)}{t}\right| + \left|\frac{W(n)}{t} - \frac{W(n)}{n}\right| \\
&\leq |W(n)|\left|\frac{1}{t} - \frac{1}{n}\right| + \frac{1}{n}\sup_{t\in[n,n+1]}|W(t) - W(n)| \\
&\leq \frac{2}{n}|W(n)| - \frac{Z(n)}{n},
\end{aligned}$$

where $Z(n) = \sup_{t\in[0,1]}|W(n+t) - W(n)|$, $t \in [n, n+1)$. It holds

$$\frac{2}{n}|W(n)| = \frac{2}{n}\left|\sum_{i=1}^{\infty}(W(i) - W(i-1))\right| \xrightarrow{a.s.} 2\,|\mathsf{E}W(1)| = 0.$$

We show that $\mathsf{E}Z(1) < \infty$.

$$\mathsf{P}\left(Z(1) > x\right) \leq \mathsf{P}\left(\max_{t\in[0,1]} W(t) > x\right) + \mathsf{P}\left(\max_{t\in[0,1]}(-W(t)) > x\right) = 2\mathsf{P}\left(\max_{t\in[0,1]} W(t) > x\right),$$

since $\{-W(t),\ t \in [0,1]\}$ is also a Wiener process. It holds

$$\mathsf{P}\left(Z(1) > x\right) \leq 2\sqrt{\frac{2}{\pi}}\int_x^{\infty} e^{-\frac{y^2}{2}}\,dy \quad \text{and thus} \quad \frac{Z(n)}{n} \xrightarrow{a.s.} 0 \quad \text{for } n \to \infty.$$

Hence $\frac{W(t)}{t} \xrightarrow{a.s.} 0$ for $t \to \infty$. □

3.3.2 Invariance properties

Specific transformations of the Wiener process again reveal a Wiener process.

Theorem 3.3.2
Let $\{W(t),\ t \geq 0\}$ be a Wiener process. Then the stochastic processes $\{Y^{(i)}(t),\ t \geq 0\}$, $i = 1, \ldots, 4$, with

$$\begin{aligned}
Y^{(1)}(t) &= -W(t), && \text{(Symmetry)} \\
Y^{(2)}(t) &= W(t+t_0) - W(t_0) \quad \text{for a } t_0 > 0, && \text{Translation of the origin)} \\
Y^{(3)}(t) &= \sqrt{c}W(\tfrac{t}{c}) \quad \text{for a } c > 0, && \text{(Scaling)} \\
Y^{(4)}(t) &= \begin{cases} tW(\frac{1}{t}), & t > 0, \\ 0, & t = 0. \end{cases} && \text{(Reflection at } t\text{=0)}
\end{aligned}$$

are Wiener processes as well.

Proof 1. $Y^{(i)}$, $i = 1, \ldots, 4$, have independent increments with $Y^{(i)}(t_2) - Y^{(i)}(t_1) \sim \mathcal{N}(0, t_2 - t_1)$.

2. $Y^{(i)}(0) = 0$, $i = 1, \dots, 4$.
3. $Y^{(i)}$, $i = 1, \dots, 3$, have continuous trajectories. $\{Y^{(i)}(t),\ t \geq 0\}$ has continuous trajectories for $t > 0$.
4. We have to proof that $\lim_{t\to 0} tW(\frac{1}{t}) = 0$.
   $\lim_{t\to 0} tW(\frac{1}{t}) = \lim_{t\to\infty} \frac{W(t)}{t} \overset{a.s.}{=} 0$ because of corollary 3.3.1.

□

Corollary 3.3.2
Let $\{W(t),\ t \geq 0\}$ be a Wiener process. Then it holds:

$$\mathsf{P}\left(\sup_{t\geq 0} W(t) = \infty\right) = \mathsf{P}\left(\inf_{t\geq 0} W(t) = -\infty\right) = 1.$$

Proof For $x, c > 0$ it holds:

$$\mathsf{P}\left(\sup_{t\geq 0} W(t) > x\right) = \mathsf{P}\left(\sup_{t\geq 0} W\left(\frac{t}{c}\right) > \frac{x}{\sqrt{c}}\right) = \mathsf{P}\left(\sup_{t\geq 0} W(t) > \frac{x}{\sqrt{c}}\right)$$

$$\Rightarrow \mathsf{P}\left(\{\sup_{t\geq 0} W(t) = 0\} \cup \{\sup_{t\geq 0} W(t) = \infty\}\right) = \mathsf{P}(\sup_{t\geq 0} W(t) = 0) + \mathsf{P}(\sup_{t\geq 0} W(t) = \infty) = 1.$$

Moreover it holds

$$\begin{aligned}
\mathsf{P}\left(\sup_{t\geq 0} W(t) = 0\right) &= \mathsf{P}\left(\sup_{t\geq 0} W(t) \leq 0\right) \leq \mathsf{P}\left(W(t) \leq 0, \sup_{t\geq 1} W(t) \leq 0\right) \\
&= \mathsf{P}\left(W(1) \leq 0, \sup_{t\geq 1}(W(t) - W(1)) \leq -W(1)\right) \\
&= \int_{-\infty}^{0} \mathsf{P}\left(\sup_{t\geq 1} W(t) - W(1) \leq -W(t) \mid W(1) = x\right) \mathsf{P}\left(W(1) \in dx\right) \\
&= \int_{-\infty}^{0} \mathsf{P}\left(\sup_{t\geq 0}(W(t) - W(1)) \leq -x \mid W(1) = x\right) \mathsf{P}\left(W(1) \in dx\right) \\
&= \int_{-\infty}^{0} \mathsf{P}\left(\sup_{t\geq 0} W(t) = 0\right) \mathsf{P}\left(W(1) \in dx\right) \\
&= \mathsf{P}\left(\sup_{t\geq 0} W(t) = 0\right) \frac{1}{2},
\end{aligned}$$

thus $\mathsf{P}\left(\sup_{t\geq 0} W(t) = 0\right) = 0$ and thus $\mathsf{P}\left(\sup_{t\geq 0} W(t) = \infty\right) = 1$.
Analogous one can show that $\mathsf{P}\left(\inf_{t\geq 0} W(t) = -\infty\right) = 1$. □

Remark 3.3.1
$\mathsf{P}\left(\sup_{t\geq 0} X(t) = \infty, \inf_{t\geq 0} X(t) = -\infty\right) = 1$ implies that the trajectories of W oscillate between positive and negative values on $[0, \infty)$ an infinite number of times.

Corollary 3.3.3
Let $\{W(t),\ t \geq 0\}$ be a Wiener process. Then it holds

$$\mathsf{P}\left(\omega \in \Omega : W(\omega) \text{ is nowhere differentiable in } [0, \infty)\right) = 1.$$

Proof

$$\begin{aligned}&\{\omega \in \Omega : W(\omega) \text{ is nowhere differentiable in } [0,\infty)\}\\&= \cap_{n=0}^{\infty}\{\omega \in \Omega : W(\omega) \text{ is nowhere differentiable in } [n,n+1)\}.\end{aligned}$$

It is sufficient to show that $\mathsf{P}(\omega \in \Omega : W(\omega)$ is differentiable for a $t_0 = t_0(\omega) \in [0,1]) = 0$. Define the set

$$A_{nm} = \left\{\omega \in \Omega : \text{ it exists a } t_0 = t_0(\omega) \in [0,1] \text{ with } |X(t_0\omega + h,\omega) - W(t_0(\omega,\omega))| \leq mh,\ \forall h \in \left[0, \frac{4}{k}\right]\right\}.$$

Then it holds

$$\{\omega \in \Omega : W(\omega) \text{ differentiable for a } t_0 = t_0(\omega)\} = \cup_{m\geq 1} \cup_{n\geq 1} A_n m.$$

We still have to show $\mathsf{P}(\cup_{m\geq 1} \cup_{n\geq 1} A_{nm}) = 0$.
Let $k_0(\omega) = \min_{k=1,2,\ldots}\{\frac{k}{n} \geq t_0(\omega)\}$. Then it holds for $\omega \in A_{nm}$ and $j = 0,1,2$

$$\begin{aligned}\left|W\left(\frac{k_0(\omega)+j+1}{n},\omega\right) - W\left(\frac{k_0(\omega)+j}{n},\omega\right)\right| &\leq \left|W\left(\frac{k_0(\omega)+j+1}{n},\omega\right) - W\left(t_0(\omega),\omega\right)\right|\\&\quad + \left|W\left(\frac{k_0(\omega)+j}{n},\omega\right) - W\left(t_1(\omega),\omega\right)\right|\\&\leq \frac{8m}{n}.\end{aligned}$$

Let $\Delta_0(k) = W(\frac{k+1}{n}) - W(\frac{k}{n})$. Then it holds

$$\begin{aligned}\mathsf{P}(A_{nm}) &\leq \mathsf{P}\left(\cup_{k=0}^{n} \cup_{j=0}^{2} |\Delta_n(k+j)| \leq \frac{8m}{n}\right)\\&\leq \sum_{k=0}^{n} \mathsf{P}\left(\cap_{j=0}^{2} |\Delta_n(k+j)| \leq \frac{8m}{n}\right) = \mathsf{P}\left(|\Delta_n(0)| \leq \frac{8m}{n}\right)\\&\leq (n+1)\left(\frac{16m}{\sqrt{2\pi n}}\right)^3 \to 0, \quad n \to \infty,\end{aligned}$$

and since $A_{nm} \subset A_{n+1,m}$ holds, it follows $\mathsf{P}\left(A_{nm}\right) = 0$. □

Corollary 3.3.4
With probability 1 it holds:

$$\sup_{n\geq 1} \sup_{0\leq t_0 < \ldots < t_n \leq 1} \sum_{i=1}^{n} |W(t_i) - W(t_{i-1})| = \infty,$$

i.e. $\{W(t),\ t \in [0,1]\}$ possesses a.s. trajectories with unbounded variation.

Proof Since every continuous function $g : [0,1] \to \mathbb{R}$ with bounded variation is differentiable almost everywhere, the assertion follows from corollary 3.3.3.

Alternative proof
It is sufficient to show that $\lim_{n\to\infty} \sum_{i=1}^{2^n} \left|W\left(\frac{it}{2^n}\right) - W\left(\frac{(i-1)t}{2^n}\right)\right| = \infty$.

Let $Z_n = \sum_{i=1}^{2^n} \left(W\left(\frac{it}{2^n}\right) - W\left(\frac{(i-1)t}{2^n}\right)\right)^2 - t$. Hence $\mathsf{E}Z_n = 0$ and $\mathsf{E}Z_n^2 = t^2 2^{-n+1}$ and with Tchebysheff's inequality

$$\mathsf{P}\left(|Z_n| < \varepsilon\right) \leq \frac{\mathsf{E}Z_n^2}{\varepsilon^2} = \left(\frac{t}{\varepsilon}\right)^2 2^{-n+1}, \quad \text{i.e.} \quad \sum_{i=1}^{\infty} \mathsf{P}\left(|Z_n| > \varepsilon\right) \stackrel{a.s.}{=} 0.$$

From lemma of Borel-Cantelli it follows that $\lim_{n\to\infty} Z_n = 0$ almost surely and thus

$$\begin{aligned} 0 \leq t &\leq \sum_{i=1}^{2^n} \left(W\left(\frac{it}{2^n}\right) - W\left(\frac{(i-1)t}{2^n}\right)\right)^2 \\ &\leq \liminf_{n\to\infty} \max_{1\leq k\leq 2^n} \left|W\left(\frac{kt}{2^n}\right) - W\left(\frac{(k-1)t}{2^n}\right)\right| \sum_{i=1}^{2^n} \left|W\left(\frac{it}{2^n}\right) - W\left(\frac{(i-1)t}{2^n}\right)\right|. \end{aligned}$$

Hence the assertion follows since W has continuous trajectories and therefore

$$\lim_{n\to\infty} \max_{1\leq k\leq 2^n} \left|W\left(\frac{kt}{2^n}\right) - W\left(\frac{(k-1)t}{2^n}\right)\right| = 0.$$

□

3.4 Additional exercises

Exercise 3.4.1
Give an intuitive (exact!) method to realize trajectories of a Wiener process $W = \{W(t),\, t \in [0,1]\}$. Thereby use the independence and the distribution of the increments of W. Additionally, write a program in **R** for the simulation of paths of W. Draw three paths $t \mapsto W(t,\omega)$ for $t \in [0,1]$ in a common diagram.

Exercise 3.4.2
Given are the Wiener process $W = \{W(t),\, t \in [0,1]\}$ and $L := \text{argmax}_{t\in[0,1]} W(t)$. Show that it holds:

$$\mathsf{P}(L \leq x) = \frac{2}{\pi} \arcsin \sqrt{x}, \quad x \in [0,1].$$

Hint: Use relation $\max_{r\in[0,t]} W(r) \stackrel{d}{=} |W(t)|$.

Exercise 3.4.3
For the simulation of a Wiener process $W = \{W(t),\, t \in [0,1]\}$ we also can use the approximation

$$W_n(t) = \sum_{k=1}^{n} S_k(t) z_k$$

where $S_k(t)$, $t \in [0,1]$, $k \geq 1$ are the *Schauder functions*, and $z_k \sim \mathcal{N}(0,1)$ i.i.d. random variables and the series converges almost surely for all $t \in [0,1]$ $(n \to \infty)$.

a) Show that for all $t \in [0,1]$ the approximation $W_n(t)$ also converges in the L^2-sense to $W(t)$.

b) Write a program in **R** (alternative: C) for the simulation of a Wiener process $W = \{W(t),\, t \in [0,1]\}$.

c) Simulate three paths $t \mapsto W(t,\omega)$ for $t \in [0,1]$ and draw these paths into a common diagram. Hereby consider the sampling points $t_k = \frac{k}{n}$, $k = 0,\dots,n$ with $n = 2^8 - 1$.

Exercise 3.4.4
For the Wiener process $W = \{W(t),\, t \geq 0\}$ we define the process of the maximum that is given by $M = \{M(t) := \max_{s\in[0,t]} W(s),\, t \geq 0\}$. Show that it holds:

a) The density $f_{M(t)}$ of the maximum $M(t)$ is given by

$$f_{M(t)}(x) = \sqrt{\frac{2}{\pi t}} \exp\left\{-\frac{x^2}{2t}\right\} 1\{x \geq 0\}.$$

Hint: Use property $\mathsf{P}(M(t) > x) = 2\mathsf{P}(W(t) > x)$.

b) Expected value and variance of $M(t)$ are given by

$$\mathsf{E}M(t) = \sqrt{\frac{2t}{\pi}}, \quad \mathsf{var}\, M(t) = t(1 - 2/\pi).$$

Now we define $\tau(x) := \operatorname{argmin}_{s\in\mathbb{R}}\{W(s) = x\}$ as the first point in time for which the Wiener process takes value x.

c) Determine the density of $\tau(x)$ and show that: $\mathsf{E}\tau(x) = \infty$.

Exercise 3.4.5
Let $W = \{W(t),\, t \geq 0\}$ be a Wiener process. Show that the following processes are Wiener processes as well:

$$W_1(t) = \begin{cases} 0, & t = 0, \\ tW(1/t), & t > 0, \end{cases} \qquad W_2(t) = \sqrt{c}W(t/c), \quad c > 0.$$

Exercise 3.4.6
The Wiener process $W = \{W(t),\, t \geq 0\}$ is given. Size $Q(a,b)$ denotes the probability that the process exceeds the half line $y = at + b$, $t \geq 0$, $a, b > 0$. Proof that:

a) $Q(a,b) = Q(b,a)$ and $Q(a, b_1 + b_2) = Q(a,b_1)Q(a,b_2)$,

b) $Q(a,b)$ is given by $Q(a,b) = \exp\{-2ab\}$.

4 Lèvy Prozesse

4.1 Lèvy processe

Definition 4.1.1
A stochastic process $\{X(t),\ t \geq 0\}$ is called Lèvy process, if

1. $X(0) = 0$,
2. $\{X(t)\}$ has stationary and independent increments,
3. $\{X(t)\}$ is stochastically continuous, i.e for an arbitrary $\varepsilon > 0$, $t_0 \geq 0$:

$$\lim_{t \to t_0} \mathsf{P}(|X(t) - X(t_0)| > \varepsilon) = 0.$$

<u>Note</u>

- One can easily consider, that compound Poisson processes fulfil the 3 conditions, since for arb. $\varepsilon > 0$ it holds

$$\mathsf{P}\left(|X(t) - X(t_0)| < \varepsilon\right) \geq \mathsf{P}\left(|X(t) - X(t_0)| > 0\right) \leq 1 - e^{-\lambda|t-t_0|} \xrightarrow[t \to t_0]{} 0.$$

- Further holds for the Wiener process for arb. $\varepsilon > 0$

$$\begin{aligned} \mathsf{P}\left(|X(t) - X(t_0)| > \varepsilon\right) &= \sqrt{\frac{2}{\pi(t-t_0)}} \int_t^\infty \exp\left(-\frac{y^2}{2(t-t_0)}\right) dy \\ &\overset{x = \frac{y}{\sqrt{t-t_0}}}{=} \frac{2}{\pi} \int_{\frac{t}{\sqrt{t-t_0}}}^\infty e^{-\frac{x^2}{2}} dx \xrightarrow[t \to t_0]{} 0. \end{aligned}$$

4.1.1 Infinitely Divisibility

Definition 4.1.2
Let $X : \Omega \to \mathbb{R}$ be an arbitrary random variable. Then X is called *infinitely divisible*, if for arbitrary $n \in \mathbb{N}$ there exist random variables $Y_1, Y_2, \ldots, Y_n$ with $X \overset{d}{=} Y_1^{(n)} + \ldots + Y_n^{(n)}$.

Theorem 4.1.1
Let $\{X(t),\ t \geq 0\}$ be a Lèvy process. Then the random variable $X(t)$ is infinitely divisible for every $t \geq 0$.

Proof For arbitrary $t \geq 0$ and $n \in \mathbb{N}$ it obviously holds that

$$X(t) = X\left(\frac{t}{n}\right) + \left(X\left(\frac{2t}{n}\right) - X\left(\frac{t}{n}\right)\right) + \ldots + \left(X\left(\frac{nt}{n}\right) - X\left(\frac{(n-1)t}{n}\right)\right).$$

Since $\{X(t)\}$ has independent and stationary increments, summands are obviously independent and identically distributed random variables. $\square$

Lemma 4.1.1
The random variable $X : \Omega \to \mathbb{R}$ is infinitely divisible if and only if the characteristic function φ_X of X can be expressed for every $n \geq 1$ in the form

$$\varphi_X(s) = (\varphi_n(s))^n \quad \text{for all } s \in \mathbb{R},$$

where φ_n are characteristic functions of random variables.

Proof „ $\Rightarrow$ "
$Y_1^{(n)}, \dots, Y_n^{(n)}$ i.i.d., $X \stackrel{d}{=} Y_1^{(n)} + \ldots + Y_n^{(n)}$. Hence, it follows that $\varphi_X(s) = \prod_{i=1}^n \varphi_{Y_i^{(n)}}(s) = (\varphi_n(s))^n$.
" $\Leftarrow$ "
$\varphi_X(s) = (\varphi_n(s))^n \Rightarrow$ there exist $Y_1^{(n)}, \dots, Y_n^{(n)}$ i.i.d. with characteristic function φ_n and $\varphi_{Y_1,\dots,Y_n}(s) = (\varphi_n(s))^n = \varphi_X(s)$. With the uniqueness theorem for characteristic functions it follows that $X \stackrel{d}{=} Y_1^{(n)} + \ldots + Y_n^{(n)}$. □

Lemma 4.1.2
Let $X_1, X_2, \ldots : \Omega \to \mathbb{R}$ be a sequence of random variables. If there exists a function $\varphi : \mathbb{R} \to \mathbb{C}$, such that $\varphi(s)$ is continuous in $s = 0$ and $\lim_{n\to\infty} \varphi_{X_n}(s) = \varphi(s)$ for all $s \in \mathbb{R}$, then φ is the characteristic function of a random variable X and it holds that $X_n \xrightarrow{d} X$.

Definition 4.1.3
Let ν be a measure on the measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Then ν is called a Lèvy measure, if $\nu(\{0\}) = 0$ and

$$\int_{\mathbb{R}} \min\left\{y^2, 1\right\} \nu(dy) < \infty.$$

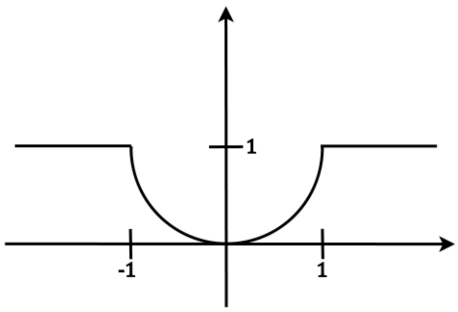

Abb. 4.1:

<u>Note</u>

- Apparently every Lèvy measure is σ-finite and

$$\nu\left((-\varepsilon, \varepsilon)^c\right) < \varepsilon, \quad \text{for all } \varepsilon > 0, \tag{4.1.1}$$

  where $(-\varepsilon, \varepsilon)^c = \mathbb{R} \mid (-\varepsilon, \varepsilon)$.

- In particular every finite measure ν is a Lèvy measure, if $\nu(\{0\}) = 0$.

- An equivalent condition to (4.1.1) is
$$\int_{\mathbb{R}} \frac{y^2}{1+y^2}\nu(dy) < \infty, \quad \text{since} \quad \frac{y^2}{1+y^2} \le \min\left\{y^2, 1\right\} \le 2\frac{y^2}{1+y^2}. \tag{4.1.2}$$

Theorem 4.1.2
Let $a \in \mathbb{R}$, $b \ge 0$ be arbitrary and let ν be an arbitrary Lèvy measure. Then the characteristic function of a infinitely divisible random variable is given through the function $\varphi : \mathbb{R} \to \mathbb{C}$ with
$$\varphi(s) = \exp\left\{ias - \frac{bs^2}{2} + \int_{\mathbb{R}} \left(e^{isy} - 1 - isy\mathbb{1}(y \in (-1,1))\right)\nu(dy)\right\} \quad \text{for all } s \in \mathbb{R}. \tag{4.1.3}$$

Remark 4.1.1 • The formula (4.1.3) is also called *Lèvy-Chintschin formula.*

- The inversion of theorem 4.1.2 also holds, hence every infinitely divisible random variable has such a representation. Therefore the characteristic triplet (a, b, ν) is also called *Lèvy characteristic* of an infinitely divisible random variable.

Proof des Theorems 4.1.2 1st step
Show that φ is a characteristic function.

- $$\left|e^{isy} - 1 - isy\right| = \left|\sum_{k=0}^{\infty} \frac{(isy)^k}{k!} - 1 - isy\right| = \left|\sum_{k=2}^{\infty} \frac{(isy)^k}{k!}\right| \le y^2 \underbrace{\left|\sum_{k=2}^{\infty} \frac{s^k}{k!}\right|}_{:=c} \le y^2 c$$
  Hence follows with (4.1.1) and (4.1.2) that the integral in (4.1.3) exists and therefore it is well-defined.

- Let now $\{c_n\}$ be an arbitrary sequence of numbers with $c_n > c_{n+1} > \ldots > 0$ and $\lim_{n\to\infty} c_n = 0$. Then the function $\varphi_n : \mathbb{R} \to \mathbb{C}$ with
  $$\varphi_n(s) := \exp\left\{is\left(a - \int_{[-c_n,c_n]^c \cap (-1,1)} y\nu(dy)\right) - \frac{bs^2}{2}\right\} \exp\left\{\int_{[-c_n,c_n]^c} \left(e^{isy} - 1\right)\nu(dy)\right\}$$
  is the characteristic function of the sum from $Z_1^{(n)}$ and $Z_2^{(n)}$, 2 independent random variables, since
  - the first factor is the characteristic function of the normal distribution with expectation $a - \int_{[-c_n,c_n]^c \cap (-1,1)} y\nu(dy)$ and variance b.
  - the second factor is the characteristic function of a compound Poisson process with characteristics
    $$\lambda = \nu([-c_n, c_n]^c) \quad \text{and} \quad \mathsf{P}_U(\cdot) = \nu(\cdot \cap [-c_n, c_n]^c / \nu([-c_n, c_n]^c))$$
    .

- Furthermore $\lim_{n\to\infty} \varphi_n(s) = \varphi(s)$ for all $s \in \mathbb{R}$, where φ is obviously continuous in 0, since it holds for the function $\varphi : \mathbb{R} \to \mathbb{C}$ which is the exponent of (4.1.3), thus
  $$\psi(s) = \int_{\mathbb{R}} \left(e^{isy} - 1 - isy\mathbb{1}\left(y \in (-1,1)\right)\right)\nu(dy) \quad \text{for all } s \in \mathbb{R}$$
  that $|\psi(s)| = cs^2 \int_{(-1,1)} y^2\nu(dy) + \int_{(-1,1)^c} \left|e^{isy} - 1\right|\nu(dy)$. Out of this and from (4.1.2) follows with the theorem of Lebesgue that $\lim_{s\to\infty} \psi(s) = 0$.

- Lemma 4.1.2 gives that the function φ given in (4.1.3) is the characteristic function of a random variable.

2nd step
The infinitely divisibility of this random variable follows from lemma 4.1.1 and out of the fact, that for arbitrary $n \in \mathbb{N}$ $\frac{\nu}{n}$ is also a Lèvy measure and that

$$\varphi(s) = \exp\left\{ i\frac{a}{n}s - \frac{\frac{b}{n}s^2}{2} + \int_{\mathbb{R}} \left(e^{isy} - 1 - isy1(y \in (-1,1))\right) \left(\frac{\nu}{n}\right)(dy) \right\} \quad \text{for all } s \in \mathbb{R}.$$

□

Remark 4.1.2
The map $\eta : \mathbb{R} \to \mathbb{C}$ with

$$\eta(s) = ias - \frac{bs^2}{2} + \int_{\mathbb{R}} \left(e^{isy} - 1 - isy1(y \in (-1,1))\right) \nu(dy)$$

from (4.1.3) is called *Lèvy exponent* of this infinitely divisible distribution.

4.1.2 Lèvy-Chintschin Representation

$\{X(t),\ t \geq 0\}$ – Lèvy process. We want to represent the characteristic function of $X(t)$, $t \geq 0$, through the Lèvy-Chintschin formula.

Lemma 4.1.3
Let $\{X(t),\ t \geq 0\}$ be a stochastic continuous process, i.e. for all $\varepsilon > 0$ and $t_0 \geq 0$ it holds that $\lim_{t \to t_0} \mathsf{P}(|X(t) - X(t_0)| > \varepsilon) = 0$. Then for every $s \in \mathbb{R}$, $t \longmapsto \varphi_{X(t)}(s)$ is a continuous map from $[0, \infty)$ to $\mathbb{C}$.

Proof

- $y \longmapsto e^{isy}$ continuous in 0, i.e. for all $\varepsilon > 0$ there exists a $\delta_1 > 0$, such that

$$\sup_{y \in (-\delta_1, \delta_1)} \left| e^{isy} - 1 \right| < \frac{\varepsilon}{2}.$$

- $\{X(t),\ t \geq 0\}$ is stochastic continuous, i.e. for all $t_0 \geq 0$ there exists a $\delta_2 > 0$, such that

$$\sup_{t \geq 0,\ |t - t_0| < \delta_2} \mathsf{P}\left(|X(t) - X(t_0)| > \delta_1\right) < \frac{\varepsilon}{4}.$$

Hence, it follows that for $s \in \mathbb{R}$, $t \geq 0$ and $|t - t_0| < \delta_2$ it holds

$$\begin{aligned}
\left|\varphi_{X(t)}(s) - \varphi_{X(t_0)}(s)\right| &= \left|\mathsf{E}\left(e^{isX(t)} - e^{isX(t_0)}\right)\right| \leq \mathsf{E}\left|e^{isX(t_0)}\left(e^{is(X(t)-X(t_0))} - 1\right)\right| \\
&\leq \mathsf{E}\left|e^{is(X(t)-X(t_0))} - 1\right| = \int_{\mathbb{R}} \left|e^{isy} - 1\right| \mathsf{P}_{X(t)-X(t_0)}(dy) \\
&\leq \int_{(-\delta_1,\delta_1)} \left|e^{isy} - 1\right| \mathsf{P}_{X(t)-X(t_0)}(dy) \\
&\quad + \int_{(-\delta_1,\delta_1)^c} \underbrace{\left|e^{isy} - 1\right|}_{=2} \mathsf{P}_{X(t)-X(t_0)}(dy) \\
&\leq \sup_{y \in (-\delta_1,\delta_1)} \left|e^{isy} - 1\right| + 2\mathsf{P}\left(|X(t) - X(t_0)| > \delta_1\right) \leq \varepsilon.
\end{aligned}$$

□

Theorem 4.1.3
Let $\{X(t),\ t \geq 0\}$ be a Lèvy process. For all $t \geq 0$

$$\varphi_{X(t)}(s) = e^{t\eta(s)}, \quad s \in \mathbb{R},$$

holds, where $\eta : \mathbb{R} \to \mathbb{C}$ is a continuous function. In particular it holds that

$$\varphi_{X(t)}(s) = e^{t\eta(s)} = \left(e^{\eta(s)}\right)^t = \left(\varphi_{X(1)}(s)\right)^t, \quad \text{for all } s \in \mathbb{R},\ t \geq 0.$$

Proof

$$\varphi_{X(t+t')}(s) = \mathsf{E}e^{isX(t+t')} = \mathsf{E}\left(e^{isX(t)}e^{is(X(t+t')-X(t))}\right) = \varphi_{X(t)}(s)\varphi_{X(t')}(s)$$

Let $g_s : [0,\infty) \to \mathbb{C}$ be defined by $g_s(t) = \varphi_{X(t)}(s)$, $s \in \mathbb{R}$, $g_s(t+t') = g_s(t)g_s(t')$, $t, t' \geq 0$. $X(0) = 0$.

$$\begin{cases} g_s(t+t') = g_s(t)g_s(t'), \quad t, t' \geq 0, \\ g_s(0) = 1, \\ g_s : [0,\infty) \to \mathbb{C} \text{ continuous.} \end{cases}$$

Hence follows: $\eta : \mathbb{R} \to \mathbb{C}$ exists, such that $g_s(t) = e^{\eta(s)t}$ for all $s \in \mathbb{R}$, $t \geq 0$. $\varphi_{X(1)}(s) = e^{\eta(s)}$ and it follows that η is continuous. □

Lemma 4.1.4
Let $\mu_1, \mu_2, \ldots$ be a sequence of finite measures (on $\mathcal{B}(\mathbb{R})$) with

1. $\sup_{n\geq 1} \mu_n(\mathbb{R}) < c$, $c = \mathit{const} < \infty$ (uniformly bounded)
2. for all $\varepsilon > 0$ there exists $B_\varepsilon \in \mathcal{B}(\mathbb{R})$ compact, such that $\sup_{n\geq 1} \mu_n(B_\varepsilon^c) \leq \varepsilon$. Hence follows that there exists a subsequence $\mu_{n_1}, \mu_{n_2}, \ldots$ and a finite measure over $\mathcal{B}(\mathbb{R})$, such that for all $f : \mathbb{R} \to \mathbb{C}$, bounded, continous, it holds that

$$\lim_{k\to\infty} \int_{\mathbb{R}} f(y)\mu_{n_k}(dy) = \int_{\mathbb{R}} \lim_{k\to\infty} f(y)\mu_{n_k}(dy) = \int_{\mathbb{R}} f(y)\mu(dy)$$

Proof See [14], page 122 - 123. □

Theorem 4.1.4
Let $\{X(t),\ t \geq 0\}$ be a Lèvy process. Then there exist $a \in \mathbb{R}$, $b \geq 0$ and a Lèvy measure ν, such that

$$\varphi_{X(1)}(s) = e^{ias - \frac{bs^2}{2}} + \int_{\mathbb{R}} \left(e^{isy} - 1 - iy\mathbb{1}(y \in (-1,1))\right)\nu(dy), \quad \text{for all } s \in \mathbb{R}.$$

Proof For all null sequences $t_1, t_2, \ldots$ it holds

$$\eta(s) = \left(e^{t\eta(s)}\right)'\Big|_{t=0} = \lim_{n\to\infty} \frac{e^{t_n\eta(s)} - 1}{t_n} = \lim_{n\to\infty} \frac{\varphi_{X(t_n)}(s) - 1}{t_n}. \tag{4.1.4}$$

$\eta : \mathbb{R} \to \mathbb{C}$ continuous $\Rightarrow$ The convergence in (4.1.4) is uniformly in $s \in [-t_0, t_0]$ for an $s_0 > 0$ (Taylor-expansion of $e^{t_n\eta(s)}$). Let $t_n = \frac{1}{n}$ and P_n be the distribution of $X(\frac{1}{n})$. Hence it follows that

$$\begin{aligned}
&\lim_{n\to\infty} n\int_{\mathbb{R}}(e^{isy}-1)\mathsf{P}_n(ds) = \lim_{n\to\infty} n\frac{\varphi_{X(\frac{1}{n})}(s)-1}{\frac{1}{n}} = \eta(s)\\
&\lim_{n\to\infty}\int_{\mathbb{R}} n\int_{-s_0}^{s_0}\left(e^{isy}-1\right)\mathsf{P}_n(dy)ds = \int_{-s_0}^{s_0}\eta(s)ds\\
&\Rightarrow \lim_{n\to\infty} n\int_{\mathbb{R}}\left(1-\frac{\sin(s_0 y)}{s_0 y}\right)\mathsf{P}_n(dy) = -\frac{1}{2s_0}\int_{-s_0}^{s_0}\eta(s)ds
\end{aligned}$$

$\eta:\mathbb{R}\to\mathbb{C}$ is continuous with $\eta(0)=0$ and it follows from that, that it exists for all $\varepsilon>0$ $\delta_0>0$, such that $\left|-\frac{1}{2s_0}\int_{-s_0}^{s_0}\eta(s)ds\right|<\varepsilon$. Since $1-\frac{\sin(s_0y)}{s_0y}\geq\frac{1}{2}$, $|s_0y|\geq 2$, it holds: for all $\varepsilon>0$ there exist $s_0>0$, $n_0>0$, such that

$$\limsup_{n\to\infty}\frac{n}{2}\int_{\left\{y:|y|\geq\frac{2}{s_0}\right\}}\mathsf{P}_n(dy)\leq\limsup_{n\to\infty} n\int_{\mathbb{R}}\left(1-\frac{sin(s_0y)}{s_0y}\right)\mathsf{P}_n(dy)<\varepsilon.$$

For all $\varepsilon>0$ there exist $s_0>0$, $n_0>0$, such that

$$n\int_{\left\{y:|y|\geq\frac{2}{s_0}\right\}}\mathsf{P}_n(dy)\leq 4\varepsilon,\quad\text{for all } n\geq n_0.$$

Decreasing s_0 gives

$$n\int_{\left\{y:|y|\geq\frac{2}{s_0}\right\}}\mathsf{P}_n(dy)\leq 4\varepsilon,\quad\text{for all } n\geq 1.$$

$$\frac{y^2}{1+y^2}\leq c\left(1-\frac{\sin y}{y}\right),\quad\text{for all } y\neq 0\quad\text{and a } c>0.$$

Hence, it follows that

$$\sup_{n\geq1} n\int_{\mathbb{R}}\frac{y^2}{1+y^2}\mathsf{P}_n(dy)\leq c'\quad\text{for a } c'<\infty.$$

Let now $\mu_n:\mathcal{B}(\mathbb{R})\to[0,\infty)$ be defined as

$$\mu_n(B)=n\int_B\frac{y^2}{1+y^2}\mathsf{P}_n(dy)\quad\text{for all } B\in\mathcal{B}(\mathbb{R}).$$

It follows that $\{\mu_n\}_{n\in\mathbb{N}}$ is uniformly bounded, $\sup_{n\geq1}\mu_n(\mathbb{R})<c'$. Furthermore holds $\frac{y^2}{1+y^2}\leq 1$, $\sup_{n\geq1}\mu_n\left(\left\{y:|y|>\frac{2}{s_0}\right\}\right)\leq 4\varepsilon$ and $\{\mu_n\}_{n\in\mathbb{N}}$ relatively compact. After lemma 4.1.3 it holds: there exists $\{\mu_{n_k}\}_{k\in\mathbb{N}}$, such that

$$\lim_{k\to\infty}\int_{\mathbb{R}}f(y)\mu_{n_k}(dy)=\int_{\mathbb{R}}f(y)\mu(dy)$$

for a measure μ and f continuous and bounded. Let for $s\in\mathbb{R}$ the function $f_s:\mathbb{R}\to\mathbb{C}$ be defined as

$$f_s(y)=\begin{cases}\left(e^{isy}-1-isy\right)\frac{1+y^2}{y^2}, & y\neq 0,\\ -\frac{s^2}{2} &, \text{otherwise.}\end{cases}$$

Hence follows that f_s is bounded and continuous and

$$\begin{aligned}
\eta(s) &= \lim_{n\to\infty}\int_{\mathbb{R}}\left(e^{isy}-1\right)\mathsf{P}_n(dy)\\
&= \lim_{n\to\infty}\left(\int_{\mathbb{R}} f_s(y)\mu_n(dy)+isn\int_{\mathbb{R}}\sin y\mathsf{P}_n(dy)\right)\\
&= \lim_{n\to\infty}\left(\int_{\mathbb{R}} f_s(y)\mu_{n_k}(dy)+isn_k\int_{\mathbb{R}}\sin y\mathsf{P}_{n_k}(dy)\right)\\
&= \int_{\mathbb{R}} f_s(y)\mu(dy)+\lim_{k\to\infty} isn_k\int_{\mathbb{R}}\sin y\mathsf{P}_{n_k}(dy)\\
\eta(s) &= ia's-\frac{bs^2}{2}+\int_{\mathbb{R}}\left(e^{isy}-1-is\sin y\right)\nu(dy),
\end{aligned}$$

for all $s\in\mathbb{R}$ with $a'=\lim_{k\to\infty} isn_k\int_{\mathbb{R}}\sin y\mathsf{P}_{n_k}(dy)$, $b=\mu(\{0\})$, $\nu:\mathcal{B}(\mathbb{R})\to[0,\infty)$,

$$\nu(dy)=\begin{cases}\frac{1+y^2}{y^2}\mu(dy), & y\neq 0,\\ 0\qquad\qquad , & y=0.\end{cases}$$

$$\int_{\mathbb{R}}|y\mathbf{1}(y\in(-1,1))-\sin y|\,\nu(dy)<\infty.$$

$$|y\mathbf{1}(y\in(-1,1))-\sin y|\,\frac{1+y^2}{y^2}<c'',\quad \text{for all } y\neq 0 \quad \text{and a } c''>0.$$

Hence follows that

$$\eta(s)=ias-\frac{bs^2}{2}+\int_{\mathbb{R}}\left(e^{isy}-1-isy\mathbf{1}\left(y\in(-1,1)\right)\right)\nu(dy),\quad \text{for all } s\in\mathbb{R}.$$

$$a=a'+\int_{\mathbb{R}}\left(y\mathbf{1}(y\in(-1,1))-\sin y\right)\nu(dy).$$

□

4.1.3 Examples

1. Wiener process (it is enough to look at $X(1)$)
   $X(1)\sim\mathcal{N}(0,1)$, $\varphi_{X(1)}(s)=e^{-\frac{s^2}{2}}$ and hence follows

   $$(a,b,\nu)=(0,1,0).$$

   Let $X=\{X(t),\ t\geq 0\}$ be a Wiener proess with drift μ, i.e. $X(t)=\mu t+\sigma W(t)$, $W=\{W(t),\ t\geq 0\}$ – Brownian motion. It follows

   $$(a,b,\nu)=(\mu,\sigma^2,0).$$

   $$\varphi_{X(1)}(s)=\mathsf{E}e^{isX(1)}=\mathsf{E}e^{(\mu+\sigma W(1))is}=e^{\mu is}\varphi_{W(1)}(\sigma s)=e^{is\mu-\sigma^2\frac{s^2}{2}},\quad s\in\mathbb{R}.$$

2. Compound Poisson process with parameters (λ, P_n)
   $X(t) = \sum_{i=1}^{N(t)} U_i$, $N(t) \sim \text{Pois}(\lambda t)$, U_i i.i.d. $\sim \mathsf{P}_U$.

$$\begin{aligned}\varphi_{X(1)}(s) &= \exp\left\{\lambda \int_{\mathbb{R}} \left(e^{isX} - 1\right) \mathsf{P}_U(dx)\right\} \\ &= \exp\left\{\lambda is \int_{\mathbb{R}} x\mathbb{1}(x \in [-1,1])\mathsf{P}_U(dx) + \lambda \int_{\mathbb{R}} \left(e^{isx} - 1 - isx\mathbb{1}(x \in [-1,1])\right) \mathsf{P}_U(dx)\right\} \\ &= \exp\left\{\lambda is \int_{-1}^{1} x\mathsf{P}_U(dx) + \lambda \int_{\mathbb{R}} \left(e^{isx} - 1 - isx\mathbb{1}(x \in [-1,1])\right) \mathsf{P}_U(dx)\right\}, \quad s \in \mathbb{R}.\end{aligned}$$

   Hence follows

$$(a, b, \nu) = \left(\lambda \int_{-1}^{1} x\mathsf{P}_U(dx), 0, \lambda\mathsf{P}_U\right), \quad \mathsf{P}_U - \text{finite on } \mathbb{R}.$$

3. Processes of Gauss-Poisson type
   $X = \{X(t),\ t \geq 0\}$, $X(t) = X_1(t) + X_2(t)$, $t \geq 0$.
   $X_1 = \{X_1(t),\ t \geq 0\}$ und $X_2 = \{X_2(t),\ t \geq 0\}$ independent.
   X_1 – Wiener process with drift μ and variance σ^2,
   X_2 – Compound Poisson process with parameters λ, P_U.

$$\begin{aligned}\varphi_{X(t)}(s) &= \varphi_{X_1(t)}(s)\varphi_{X_2(t)}(s) \\ &= \exp\left\{is\left(\mu + \lambda \int_{-1}^{1} x\mathsf{P}_U(dx)\right) - \frac{\sigma^2 s^2}{2}\right. \\ &\quad \left. + \int_{\mathbb{R}} \lambda \left(e^{isx} - 1 - isx\mathbb{1}(x \in [-1,1])\right) \mathsf{P}_U(dx)\right\}, \quad s \in \mathbb{R}.\end{aligned}$$

   Hence follows

$$(a, b, \nu) = \left(\mu + \lambda \int_{-1}^{1} x\mathsf{P}_U(dx), \sigma^2, \lambda\mathsf{P}_U\right).$$

4. Stable Lèvy processes
   $X = \{X(t),\ t \geq 0\}$ – Lèvy process with $X(t) \sim \alpha$ stable distribution, $\alpha \in (0, 2]$. If $X = W$ (Wiener process), then $X(1) \sim \mathcal{N}(0, 1)$. Let $Y, Y_1, \ldots, Y_n$ be i.i.d. $\mathcal{N}(\mu, \sigma^2)$-variables. Since the convolution of the normal distribution is stable it holds

$$\begin{aligned}Y_1 + \ldots + Y_n \sim \mathcal{N}(n\mu, n\sigma^2) &\stackrel{d}{=} \sqrt{n}Y + n\mu - \sqrt{n}\mu \\ &= \sqrt{n}Y + \mu\left(n - \sqrt{n}\right) \\ &= n^{\frac{1}{2}}Y + \mu\left(n^{\frac{2}{2}} - n^{\frac{1}{2}}\right), \quad \alpha = 2.\end{aligned}$$

Definition 4.1.4
The distribution of a random variable Y is called α-stable, if for all $n \in \mathbb{N}$ only copies $Y_1, \ldots, Y_n$ exist (of Y)

$$Y_1 + \ldots + Y_n \stackrel{d}{=} n^{\frac{1}{\alpha}}Y + d_n,$$

where d_n is deterministic (thus a constant w.r.t. W, i.e. not random). The constant $\alpha \in (0, 2]$ is called *index of stability.*

$$d_n = \begin{cases} \mu\left(n - n^{\frac{1}{\alpha}}\right), & \alpha \neq 1, \\ \mu n \log n \quad, & \alpha = 1. \end{cases}$$

Without proof

Example 4.1.1 • $\alpha = 2$: Normal distribution

- $\alpha = 1$: Cauchy distribution with parameters (μ, σ^2). The density:

$$f_Y(x) = \frac{\sigma}{\pi\left((x-\mu)^2 + \sigma^2\right)}, \quad x \in \mathbb{R}.$$

It holds $\mathsf{E}Y^2 = \infty$, $\mathsf{E}Y$ does not exist.

- $\alpha = \frac{1}{2}$: Lèvy distribution with parameters (μ, σ^2). The density:

$$f_Y(x) = \begin{cases} \left(\frac{\sigma}{2\pi}\right)^{\frac{1}{2}} \frac{1}{(x-\mu)^{\frac{3}{2}}} \exp\left\{-\frac{\sigma}{2(x-\mu)}\right\}, & x > \mu, \\ 0 \qquad\qquad\qquad\qquad\qquad\quad , & \text{otherwise.} \end{cases}$$

These examples are the few examples of α-stable distribution, which have an explicit form of the density. For other $\alpha \in (0,2)$, $\alpha \neq \frac{1}{2}$, 1, the α-stable distribution will be introduced through its characteristic function. In general holds: If Y α-stable, $\alpha \in (0,2]$, then $\mathsf{E}|Y|^p < \infty$, $0 < p < \alpha$.

Definition 4.1.5
The distribution of a random variable is called symmetric, if $Y \stackrel{d}{=} -Y$. If Y has a symmetric α-stable distribution, $\alpha \in (0,2]$,

$$\varphi_Y(s) = \exp\left\{-c\,|s|^\alpha\right\}.$$

In fact, it follows from the stability of Y that

$$(\varphi_Y(s))^n = e^{id_n s}\varphi_Y\left(n^{\frac{1}{\alpha}}s\right), \quad s \in \mathbb{R}.$$

It follows that $d_n = 0$, since $\varphi_{-Y}(s) = \varphi_Y(s)$. It holds: $e^{id_n s} = e^{-id_n s}$, $s \in \mathbb{R}$ and $d_n = 0$. The rest is left as an exercise.

Lemma 4.1.5
Lèvy-Chintschin representation of the characteristic function is a stable distribution. A Lèvy characteristic (a, b, ν), $a \in \mathbb{R}$ arbitrary.

$$b = \begin{cases} \sigma^2, & \alpha = 2, \\ 0\;, & \alpha < 2. \end{cases}$$

$$\nu(dx) = \begin{cases} 0 & , \quad \alpha = 2, \\ \frac{c_1}{x^{1+\alpha}} 1(x \geq 0)dx + \frac{c_2}{|x|^{1+\alpha}} 1(x < 0)dx, & \alpha < 2, \\ c_1, c_2 \geq 0 : c_1 + c_2 > 0 & , \end{cases}$$

Without proof
You can show that

$$\mathsf{P}\left(|Y| \geq x\right) \underset{x\to\infty}{\sim} \begin{cases} e^{-\frac{x^2}{2\sigma^2}}, & \alpha = 2, \\ \frac{c}{x^\alpha} \quad , & \alpha < 2. \end{cases}$$

Definition 4.1.6
The Lèvy process $X = \{X(t),\ t \geq 0\}$ is called stable, if $X(1)$ has an α-stable distribution, $\alpha \in (0,2]$ ($\alpha = 2$: Brownian motion (with drift)).

4.1.4 Subordinators

Definition 4.1.7
A Lèvy process $X = \{X(t),\ t \geq 0\}$ is called *subordinator*, if for all $0 < t_1 < t_2$, $X(t_1) \leq X(t_2)$ a.s. holds

$$X(0) = 0 \quad \text{a.s.} \quad \Rightarrow \quad X(t) \geq 0, \quad t \geq 0.$$

This class of Lèvy processes is important, since you can easily introduce $\int_a^b g(t)dX(t)$ as a Lebesgue-Stieltjes-integral.

Theorem 4.1.5
The Lèvy process $X = X(t),\ t \geq 0$ is a subordinator if and only if the Lèvy-Chintschin representation can be expressed in the form

$$\varphi_{X(t)}(s) = \exp\left\{ias + \int_{\mathbb{R}} \left(e^{isx} - 1\right)\nu(dx)\right\}, \quad s \in \mathbb{R},$$

where ν is the Lèvy measure, with

$$\nu\left((-\infty, 0)\right) = 0, \quad \int_0^\infty \min\left\{1, y^2\right\}\nu(dy) < \infty.$$

Proof Sufficiency
It has to be shown that $X(t_2) \geq X(t_1)$ a.s., if $t_2 \geq t_1 \geq 0$.
First of all we show that $X(1) \geq 0$ a.s.. If $\nu \equiv 0$, then $X(1) = a$ a.s., hence

$$\varphi_{X(t)}(s) = \begin{pmatrix} \varphi(s) \\ X(t) \end{pmatrix}^t = e^{iats}, \quad s \in \mathbb{R}.$$

$X(t) = at$ a.s. and therefore it follows that $X(t) \uparrow$ and X is a subordinator. If $\nu([0,\infty)) > 0$, then there exists $N > 0$, such that $n \geq N$, $0 < \nu\left(\left[\frac{1}{n}, \infty\right)\right) < \infty$. It follows

$$\varphi_{X(t)}(s) = \exp\left\{ias + \lim_{n\to\infty} \int_{\frac{1}{n}}^\infty \left(e^{isx} - 1\right)\nu(dx)\right\} = e^{ias} \lim_{n\to\infty} \varphi_n(s), \quad s \in \mathbb{R},$$

where $\varphi_n(s) = \int_{\frac{1}{n}}^\infty \left(e^{isx} - 1\right)\nu(dx)$ is the characteristic function of a compound Poisson process distribution with parameters $\left(\nu\left(\left[\frac{1}{n}, \infty\right)\right), \frac{\nu(\cap[\frac{1}{n},\infty))}{\nu([\frac{1}{n},\infty))}\right)$ for all $n \in \mathbb{N}$. Let Z_n be the random variable with characteristic function φ_n. It holds: $Z_n = \sum_{i=1}^{N_n} U_i$, $N_n \sim \text{Pois}\left(\nu\left(\left[\frac{1}{n}, \infty\right)\right)\right)$, $U_i \sim \frac{\nu(\cap[\frac{1}{n},\infty))}{\nu([\frac{1}{n},\infty))}$ and hence follows $Z_n \geq 0$ a.s. and $X(1) = \underbrace{a}_{=0} + \underbrace{\lim Z_n}_{\geq 0} \geq 0$ a.s.. Since X is a Lèvy process, it holds

$$X(1) = X\left(\frac{1}{n}\right) + \left(X\left(\frac{2}{n}\right) - X\left(\frac{1}{n}\right)\right) + \ldots + \left(X\left(\frac{n}{n}\right) - X\left(\frac{n-1}{n}\right)\right),$$

where, because of stationarity and independence of the increments $X\left(\frac{k}{n}\right) - X\left(\frac{k-1}{n}\right) \overset{a.s.}{\geq} 0$ for $1 \leq k \leq n$ for all n. $X(q_2) - X(q_1) \geq 0$ a.s. for all $q_1, q_2 \in \mathbb{Q}$, $q_2 \geq q_1 \geq 0$. Now let $t_1, t_2 \in \mathbb{Q}$, such that $0 \leq t_1 \leq t_2$. Let $\left\{q_1^{(n)}, q_2^{(n)}\right\}$ be sequences of numbers from $\mathbb{Q}$ with $q_1^{(n)} \leq q_2^{(n)}$. $q_1^{(n)} \downarrow t_1$, $q_2^{(n)} \uparrow t_2$, $n \to \infty$. For $\varepsilon > 0$

$$\begin{aligned}
\mathsf{P}\left(X(t_2) - X(t_1) < -\varepsilon\right) &= \mathsf{P}\left(X(t_2) - X\left(q_2^{(n)}\right) + X\left(q_2^{(n)}\right) - X\left(q_1^{(n)}\right) + X\left(q_1^{(n)}\right) - X\left(t_1\right) < -\varepsilon\right) \\
&\leq \mathsf{P}\left(X(t_2) - X\left(q_2^{(n)}\right) + X\left(q_1^{(n)}\right) - X\left(t_1\right) < -\varepsilon\right) \\
&\leq \mathsf{P}\left(X(t_2) - X\left(q_2^{(n)}\right) < -\varepsilon\right) + \mathsf{P}\left(X\left(q_1^{(n)}\right) - X(t_1) \leq -\frac{\varepsilon}{2}\right) \xrightarrow[n\to\infty]{} 0.
\end{aligned}$$

$$\begin{aligned}
&\Rightarrow \mathsf{P}\left(X(t_2) - X(t_1) < \varepsilon\right) = 0 \quad \text{for all } \varepsilon > 0 \\
&\Rightarrow \mathsf{P}\left(X(t_2) - X(t_1) < 0\right) = \lim_{\varepsilon \to +0} \mathsf{P}\left(X(t_2) - X(t_1) < \varepsilon\right) = 0 \\
&\Rightarrow X(t_2) \geq X(t_1) \quad \text{a.s.}
\end{aligned}$$

Necessity

Let X be a Lèvy process, which is a subordinator. It has to be shown that $\varphi_{X_1(t)}(\cdot)$ has the above form.
After the Lèvy-Chintschin representation for $X_1(t)$ it holds that

$$\varphi_{X(1)}(s) = \exp\left\{ias - \frac{b^2 s^2}{2} + \int_0^\infty \left(e^{isx} - 1 - isx\mathbb{1}(x \in [-1,1])\right)\nu(dx)\right\}, \quad s \in \mathbb{R}.$$

The measure ν is concentrated on $[0, \infty)$, since $X(1) \overset{a.s.}{\geq} 0$ and from the proof of theorem 4.1.4 $\nu\left((-\infty, 0)\right) = 0$ can be chosen.

$$\varphi_{X(1)}(s) \leq \underbrace{\exp\left\{ias - \frac{b^2 s^2}{2}\right\}}_{:=\varphi_{Y_1(s)}} \underbrace{\exp\left\{\int_0^\infty \left(e^{isx} - 1 - isx\mathbb{1}\left(x \in [-1,1]\right)\right)\nu(dx)\right\}}_{:=\varphi_{Y_2(s)}}$$

Hence, it follows that $X(1) = Y_1 + Y_2$, Y_1 and Y_2 are independent, $Y_1 \sim \mathcal{N}(a, b^2)$ and therefore $b = 0$. For all $\varepsilon \in (0,1)$

$$\varphi_{X_1}(s) = \exp\left\{is\left(a - \int_\varepsilon^1 x\nu(dx)\right) + \int_0^\varepsilon \left(e^{isx} - 1 - isx\right)\nu(dx) + \int_0^\infty \left(e^{isx} - 1\right)\nu(dx)\right\}$$

It has to be shown that for $\varepsilon \to 0$ it holds that $\int_\varepsilon^\infty \left(e^{isx} - 1\right)\nu(dx) \to \int_0^\infty \left(e^{isx} - 1\right)\nu(dx) < \infty$ with $\int_0^1 \min\{x, 1\}\nu(dx) < \infty$. $\varphi_{X(1)}(s) = \exp\left\{is\left(a - \int_\varepsilon^1 x\nu(dx)\right)\right\}\varphi_{Z_1}(s)\varphi_{Z_2}(s)$, where Z_1 and Z_2 are independent, $\varphi_{Z_1}(s) = \exp\left\{\left(e^{isx} - 1 - isx\right)\nu(dx)\right\}$, $\varphi_{Z_2}(s) = \exp\left\{\int_\varepsilon^\infty \left(e^{isx} - 1\right)\nu(dx)\right\}$, $s \in \mathbb{R}$. $X(1) \overset{d}{=} a - \int_\varepsilon^1 x\nu(dx) + Z_1 + Z_2$. There exist $\varphi_{Z_1}^{(2)}(0) = \frac{-\mathsf{E}Z_1^2}{2} < \infty$, $\varphi_{Z_1}^{(1)}(0) = 0 = i\mathsf{E}Z_1$ and it therefore follows that $\mathsf{E}Z_1 = 0$ and $\mathsf{P}(Z_1 \leq 0) > 0$. On the other hand, Z_2 has a compound Poisson distribution with parameters $\left(\nu\left([\varepsilon,\infty)\right),\frac{\nu(\cap[\varepsilon,+\infty])}{\nu([\varepsilon,+\infty))}\right)$, $\varepsilon\in(0,1)$.

$$\Rightarrow \mathsf{P}\left(Z_2\le 0\right)>0$$
$$\Rightarrow \mathsf{P}\left(Z_1+Z_2\le 0\right)\ge \mathsf{P}\left(Z_1\le 0, Z_2\le 0\right)=\mathsf{P}\left(Z_1\le 0\right)\mathsf{P}\left(Z_2\le 0\right)>0$$
$$\Rightarrow a-\int_\varepsilon^1 x\nu(dx)\ge 0 \quad \text{for all } \varepsilon\in(0,1)$$
$$\Rightarrow \int_0^a \min\{x,1\}\,dx<\infty$$
$$\Rightarrow \text{for } \varepsilon\to\infty \quad Z_1\xrightarrow{d}0$$
$$\varphi_{X(1)}(s)=\exp\left\{is\left(a-\int_0^1 x\nu(dx)\right)+\int_0^\infty\left(e^{isx}-1\right)\nu(dx)\right\},\quad s\in\mathbb{R}.$$

□

Example 4.1.2 (α-stable subordinator):
$X=\{X(t),\ t\ge 0\}$ a Lèvy process, subordinator, with $a=0$ – Lévy measure.

$$\nu(dx)=\begin{cases}\frac{\alpha}{\Gamma(1-\alpha)}\frac{1}{x^{1+\alpha}}dx\,, & x>0,\\ 0\cdot\frac{1}{x^{1+\alpha}}dx=0, & x\le 0.\end{cases}$$

Therefore, it follows that X is a α-stable Lèvy process.
We show that $\hat{l}_{X(\cdot)}(s)=\mathsf{E}e^{-sX(t)}=e^{-ts^\alpha}$ for all $s,t\ge 0$.

$$\varphi_{X(t)}(s)=\left(\varphi_{X(1)}(s)\right)^t=\exp\left\{t\int_0^\infty\left(e^{isx}-1\right)\frac{\alpha}{\Gamma(1-\alpha)}\frac{1}{x^{1+\alpha}}dx\right\},\quad s\in\mathbb{R}.$$

It has to be shown that

$$U^d=\frac{\alpha}{\Gamma(1-\alpha)}\int_0^\infty\left(1-e^{-ux}\right)\frac{dx}{x^{1+\alpha}},\quad u\ge 0.$$

This is enough since $\varphi_{X(t)}(\cdot)$ can be continued analytically on $\{Z\in\mathbb{C}:\Im Z\ge 0\}$, i.e. $\varphi_{X(t)}(iu)=\hat{l}_{X(t)}$, $u\ge 0$. In fact, it holds that

$$\begin{aligned}\int_0^\infty\left(1-e^{-ux}\right)\frac{dx}{x^{1+d}} &= \int_0^\infty u\int_0^x e^{-uy}dyx^{-1-\alpha}dx\\ &= \int_0^\infty\int_y^\infty ue^{-uy}x^{-1-\alpha}dxdy\\ &= \int_0^\infty\int_y^\infty x^{-1-\alpha}dxue^{-uy}dy\\ &= \frac{u}{\alpha}\int_0^\infty e^{-uy}y^{-\alpha}dy\\ &= \frac{u}{\alpha}\int_0^\infty e^{-z}z^{-\alpha}\frac{1}{u^{-\alpha}}d\left(\frac{z}{u}\right)\\ &= \frac{u^\alpha}{\alpha}\int_0^\infty e^{-z}z^{(1-\alpha)-1}dz\\ &= \frac{u^\alpha}{\alpha}\Gamma(1-\alpha)\end{aligned}$$

and hence follows $\hat{l}_{X(t)}(s)=e^{-ts^\alpha}$, $t,s\ge 0$.

4.2 Additional Exercises

Exercise 4.2.1
Given a real-valued random variable X with distribution function F and characteristic function φ. Show that the following statements hold:

a) If X is infinitely divisible, then it holds $\varphi(t) \neq 0$ for all $t \in \mathbb{R}$. *Hint: Show that* $\lim_{n\to\infty} |\varphi_n(s)|^2 = 1$ *for all* $s \in \mathbb{R}$, *if* $\varphi(s) = (\varphi_n(s))^n$. *Note further, that* $|\varphi_n(s)|^2$ *is again a characteristic function and* $\lim_{n\to\infty} x^{\frac{1}{n}} = 1$ *holds for* $x > 0$.

b) Give an example (with explanation) for a distribution, which is not infinitely divisible.

Exercise 4.2.2
Let $X = \{X(t),\, t \geq 0\}$ be a Lévy process. Show that the random variable $X(t)$ is then infinitely divisible for every $t \geq 0$.

Exercise 4.2.3
Show that the sum of two indepenent Lévy processes is again a Lévy process, and state the corresponding Lévy characteristic.

Exercise 4.2.4
Look at the following function $\varphi : \mathbb{R} \to \mathbb{C}$ with

$$\varphi(t) = \mathrm{e}^{\psi(t)}, \quad \text{where } \psi(t) = 2\sum_{k=-\infty}^{\infty} 2^{-k}(\cos(2^k t) - 1).$$

Show that $\varphi(t)$ is the characteristic function of an infinitely divisible distribution. *Hint: Look at the Lévy-Chintschin representation with measure* $\nu(\{\pm 2^k\}) = 2^{-k}$, $k \in \mathbb{Z}$.

Exercise 4.2.5
The Lévy process $\{X(t), t \geq 0\}$ be a Gamma process with parameters $b, p > 0$, that is, for every $t \geq 0$ holds $X(t) \sim \Gamma(b, pt)$. Show that $\{X(t), t \geq 0\}$ is a subordinator with the Laplace exponent $\xi(u) = \int_0^\infty (1 - e^{-uy})\nu(dy)$ für $\nu(dy) = py^{-1}e^{-by}dy$, $y > 0$. (The Laplace exponent of $\{X(t), t \geq 0\}$ is the function $\xi : [0, \infty) \to [0, \infty)$, for which holds that $\mathsf{E}e^{-uX(t)} = e^{-t\xi(u)}$ for arbitrary $t, u \geq 0$)

Exercise 4.2.6
Let $\{X(t), t \geq 0\}$ be a Lévy process with charactersistic Lévy exponent η and $\{\tau(s), s \geq 0\}$ a independent subordinator witch characteristic Lévy exponent γ. The stochastic process Y be defined as $Y = \{X(\tau(s)), s \geq 0\}$.

(a) Show that

$$\mathsf{E}\left(e^{i\theta Y(\tau(s))}\right) = e^{\gamma(-i\eta(\theta))s}, \quad \theta \in \mathbb{R},$$

where $\Im z$ describes the imaginary part of z.

Hint: Since τ is a process with non-negative values, it holds $\mathsf{E}e^{i\theta\tau(s)} = e^{\gamma(\theta)s}$ for all $\theta \in \{z \in \mathbb{C} : \Im z \geq 0\}$ through analytical continuation of theorem 4.1.3.

(b) Show that Y is a Lèvy process with characteristic Lèvy exponent $\gamma(-i\eta(\cdot))$.

Exercise 4.2.7
Let $\{X(t),\ t \geq 0\}$ be a compound Poisson process with Lèvy measure

$$\nu(dx) = \frac{\lambda\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{x^2}{2\sigma^2}}\, dx, \quad x \in \mathbb{R},$$

where $\lambda, \sigma > 0$. Show that $\{\sigma W(N(t)),\ t \geq 0\}$ has the same finite-dimensionale distribution as X, where $\{N(s),\ s \geq 0\}$ is a Poisson process with intensity 2λ and W is a standard Wiener process independent from N.

Hint to exercise 4.2.6 a) and exercise 4.2.7

- In order to calculate the expectation for the characteristic function, the identity $\mathsf{E}(X) = \mathsf{E}(\mathsf{E}(X|Y)) = \int_{\mathbb{R}} \mathsf{E}(X|Y = y) F_Y(dy)$ for two random variables X and Y can be used. In doing so, it should be conditioned on $\tau(s)$.
- $\int_{-\infty}^{\infty} \cos(sy) e^{-\frac{y^2}{2a}}\, dy = \sqrt{2\pi a} \cdot e^{-\frac{as^2}{2}}$ for $a > 0$ and $s \in \mathbb{R}$.

Exercise 4.2.8
Let W be a standard Wiener process and τ an independent $\frac{\alpha}{2}$-stable subordinator, where $\alpha \in (0, 2)$. Show that $\{W(\tau(s)), s \geq 0\}$ is a α-stable Lévy process.

Exercise 4.2.9
Show that the subordinator T with marginal density

$$f_{T(t)}(s) = \frac{t}{2\sqrt{\pi}} s^{-\frac{3}{2}} e^{-\frac{t^2}{4s}} \mathbb{1}\{s > 0\}$$

is a $\frac{1}{2}$-stable subordinator. (Hint: Differentiate the Laplace transform of $T(t)$ and solve the differential equation)

5 Martingales

5.1 Basic Ideas

Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a complete probability space.

Definition 5.1.1
Let $\{\mathcal{F}_t,\ t \geq 0\}$ be a family of σ-algebras $\mathcal{F}_t \subset \mathcal{F}$. It is called

1. *a filtration*, if $\mathcal{F}_s \subseteq \mathcal{F}_t$, $0 \leq s < t$.

2. *a complete filtration*, if it is a filtration, such that $\mathcal{F}_0$ (and therefore all $\mathcal{F}_s$, $s > 0$) contains all the probability measure null sets.
   Later on we will always assume, that we have a complete filtration.

3. *a right-continuous filtration*, if for all $t \geq 0$ $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$.

4. *a natural filtration* for a stochastic process $\{X(t),\ t \geq 0\}$, if it is generated by the past of the process untill time $t \geq 0$, i.e. for all $t \geq 0$ $\mathcal{F}_t$ is the smallest σ-algebra ($\subset \mathcal{F}_t$), which contains the sets $\{\omega \in \Omega : (X(t_1), \ldots, X(t_n))^\top \subset B\}$, for all $n \in \mathbb{N}$, $0 \leq t_1, \ldots, t_n \leq t$, $B \in \mathcal{B}(\mathbb{R}^n)$.

A random variable $\tau : \Omega \to \mathbb{R}_+$ is called *stopping time (w.r.t. the filtration* $\{\mathcal{F}_t,\ t \geq 0\}$*)*, if for all $t \geq 0$ $\{\omega \in \Omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$, i.e. by looking at the process X (up to the natural filtration $\{\mathcal{F}_t,\ t \geq 0\}$) you can tell, if the moment τ occured.

Lemma 5.1.1
Let $\{\mathcal{F}_t,\ t \geq 0\}$ be a right-continuous filtration. τ is a stopping time w.r.t. $\{\mathcal{F}_t,\ t \geq 0\}$ if and only if $\underbrace{\{\tau < t\} \in \mathcal{F}_t}_{\{\omega\in\Omega:\tau(\omega)\leq t\}\in\mathcal{F}_t}$, for all $t \geq 0$.

Proof „ $\Leftarrow$ "
Let $\{\tau < t\} \in \mathcal{F}_t$, $t \geq 0$. To show: $\{\tau \leq t\} \in \mathcal{F}_t$.
$\{\tau \leq t\} = \cap_{s\in(t,t+\varepsilon)}\{\tau < s\}$ for all $\varepsilon > 0 \Rightarrow \{\tau \leq t\} \in \cap_{s>t}\mathcal{F}_s = \mathcal{F}_t$

„ $\Rightarrow$ "
To show: $\{\tau \leq t\} \in \mathcal{F}_t$, $t \geq 0 \Rightarrow \{\tau < t\} \in \mathcal{F}_t$, $t \geq 0$.
$\{\tau < t\} = \cup_{s\in(0,t)}\{\tau \leq t - s\} \in \cup_{s\in(0,t)}\mathcal{F}_{t-s} \subset \mathcal{F}_t$ □

Definition 5.1.2
Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space, $\{\mathcal{F}_t,\ t \geq 0\}$ a filtration ($\mathcal{F}_t \subset \mathcal{F},\ t \geq 0$) and $X = \{X(t),\ t \geq 0\}$ a stochastic process on $(\Omega, \mathcal{F}, \mathsf{P})$. X is adapted w.r.t. the filtration $\{\mathcal{F}_t, t \geq 0\}$, if $X(t)$ is $\mathcal{F}_t$-measurable, for all $t \geq 0$, i.e., for all $B \subset \mathcal{B}(\mathbb{R})$ $\{X(t) \in B\} \in \mathcal{F}_t$.

Definition 5.1.3
The time $\tau_B(\omega) = \inf\{t \geq 0 : X(t) \in B\}$, $\omega \in \Omega$, is called *first hitting time* to the set $B \in \mathcal{B}(\mathbb{R})$ by the stochastic process $X = \{X(t),\ t \geq 0\}$ (also called: first passage time, first entrance time).

Theorem 5.1.1
Let $\{\mathcal{F}_t,\ t \geq 0\}$ be a right-continuous filtration and $X = \{X(t),\ t \geq 0\}$ an adapted (w.r.t. $\{\mathcal{F}_t,\ t \geq 0\}$) càdlàg process. For open $B \subset \mathbb{R}$, τ_B is a stopping time. If B is closed, then $\tilde{\tau}_B(\omega) = \inf\{t \geq 0 : X(t) \in B \text{ or } X(t-) \in B\}$ is a stopping time, where $X(t-) = \lim_{s \uparrow t} X(s)$.

Proof 1. Let $B \in \mathcal{B}(\mathbb{R})$ be open.
Because of lemma 5.1.1 it is enough to show that $\{\tau_B < t\} \in \mathcal{F}_t,\ t \geq 0$. Because of right-continuity of the trajectories of X it holds:

$$\{\tau_B < t\} = \cup_{s \in \mathbb{Q} \cap (0,t)} \{X(s) \in B\} \in \cup_{s \in \mathbb{Q} \cap (0,t)} \mathcal{F}_s \subseteq \mathcal{F}_t, \quad \text{since } \mathcal{F}_s \subseteq \mathcal{F}_t,\ s < t.$$

2. Let $B \in \mathcal{B}(\mathbb{R})$ be closed.
For all $\varepsilon > 0$. Let $B_\varepsilon = \{x \in \mathbb{R} : d(x, B) < \varepsilon\}$ be a parallel set of B, where $d(x, B) = \inf_{y \in B} |x - y|$. B_ε is open, for all $t \geq 0$.
$\{\tilde{\tau}_B \leq t\} = \{X(t) \in B\} \cup \cap_{n \geq 1, s \in \mathbb{Q} \cap (0,t)} \cup \{X(s) \in B_{\frac{1}{n}}\} \in \mathcal{F}_t$, since X is adapted w.r.t. $\{\mathcal{F}_t,\ t \geq 0\}$.

□

Lemma 5.1.2
Let τ_1, τ_2 be stopping times w.r.t. the filtration $\{\mathcal{F}_t,\ t \geq 0\}$. Then $\min\{\tau_1, \tau_2\}$, $\tau_1 + \tau_2$ and $\alpha\tau_1$, $\alpha \geq 1$, are stopping times (w.r.t. $\{\mathcal{F}_t,\ t \geq 0\}$).

Proof For all $t \geq 0$ holds:
$\{\min\{\tau_1, \tau_2\} \leq t\} = \underbrace{\{\tau_1 \leq t\}}_{\in \mathcal{F}_t} \cup \underbrace{\{\tau_2 \leq t\}}_{\in \mathcal{F}_t} \in \mathcal{F}_t$,
$\{\max\{\tau_1, \tau_2\} \leq t\} = \{\tau_1 \leq t\} \cap \{\tau_2 \leq t\} \in \mathcal{F}_t$,
$\{\alpha\tau_1 \leq t\} = \{\tau_1 \leq \frac{t}{\alpha}\} \in \mathcal{F}_{\frac{t}{\alpha}} \subset \mathcal{F}_t$, since $\frac{t}{\alpha} \leq t$,
$\{\tau_1 + \tau_2 \leq t\} = \underbrace{\{\tau_1 > t\}}_{\in \mathcal{F}_t} \cup \underbrace{\{\tau_2 > t\}}_{\in \mathcal{F}_t} \cup \underbrace{\tau_1 \geq t, \tau_2 > 0}_{\mathcal{F}_t} \cup \{0 < \tau_2 < t, \tau_1 - \tau_2 > t\}$,
To show: $\{0 < \tau_2 < t, \tau_1 - \tau_2 > t\} \in \mathcal{F}_t$.
$\{0 < \tau_2 < t, \tau_1 + \tau_2 > t\} = \cup_{s \in \mathbb{Q} \cap (0,t)} \{s < \tau_1 < t, \tau_2 > t - s\} \in \mathcal{F}_t$ □

Theorem 5.1.2
Let τ be an a.s. finite stopping time w.r.t. the filtration $\{\mathcal{F}_t, t \geq 0\}$ on the probability space $(\Omega, \mathcal{F}, \mathsf{P})$, i.e. $\mathsf{P}(\tau = \infty) = 1$. Then there exists a sequence of discrete stopping times $\{\tau_n\}_{n \in \mathbb{N}}$, $\tau_1 \geq \tau_2 \geq \tau_3 \geq \ldots$, such that $\tau_n \downarrow \tau$, $n \to \infty$ a.s.

Proof For all $n \in \mathbb{N}$ let

$$\tau_n = \begin{cases} 0, & \text{if } \tau(\omega) = 0 \\ \frac{k+1}{2^n}, & \text{if } \frac{k}{2^n} < \tau(\omega) \leq \frac{k+1}{2^n}, \quad \text{for a } k \in \mathbb{N}_0 \end{cases}$$

For all $t \geq 0$ and for all $n \in \mathbb{N}$ $\exists k \in \mathbb{N}_0 : \frac{k}{2^n} \leq t \leq \frac{k+1}{2^n}$ holds $\{\tau_n \leq t\} = \{\tau_n \leq \frac{k}{2^n}\} = \{\tau \leq \frac{k}{2^n}\} \in \mathcal{F}_{\frac{k}{2^n}} \subset \mathcal{F}_t \Rightarrow \tau_n$ is a stopping time. Therefore it is obvious, that $\tau_n \downarrow \tau$, $n \to \infty$ a.s. □

Conclusion 5.1.1
Let τ be an a.s. finite stopping time w.r.t. the filtration $\{\mathcal{F}_t,\ t \geq 0\}$ and $X = \{X(t),\ t \geq 0\}$ a càdlàg process on $(\Omega, \mathcal{F}, \mathsf{P})$, $\mathcal{F}_t \subset \mathcal{F}$ for all $t \geq 0$. Then $X(\omega, \tau(\omega))$, $\omega \in \Omega$ is a random variable on $(\Omega, \mathcal{F}, \mathsf{P})$.

Proof To show: $X(\tau) : \Omega \to \mathbb{R}$ measurable, i.e. for all $B \in \mathcal{B}(\mathbb{R})$ $\{X(\tau) \in B\} \in \mathcal{F}$. Let $\tau_n \downarrow \tau$, $n \to \infty$ be as in theorem 5.1.2. Since X is càdlàg, it holds that $X(\tau_n) \xrightarrow[n\to\infty]{} X(\tau)$ a.s.. Then $X(\tau)$ is $\mathcal{F}$-measurable as the limit of $X(\tau_n)$, which are themselves $\mathcal{F}$-measurable. For all $B \in \mathcal{B}(\mathbb{R})$ holds

$$\{X(\tau_n) \in B\} = \cup_{k=0}^{\infty}(\underbrace{\{\tau_n = \frac{k}{2^n}\}}_{\in\mathcal{F}} \cap \underbrace{\{X(\frac{k}{2^n}) \in B\}}_{\in\mathcal{F}}) \in \mathcal{F}$$

□

5.2 (Sub-, Super-)Martingales

Definition 5.2.1
Let $X = \{X(t),\ t \geq 0\}$ be a stochastic process adapted w.r.t. to a filtration $\{\mathcal{F}_t,\ t \geq 0\}$, $\mathcal{F}_t \subset \mathcal{F}$, $t \geq 0$, on the probability space $(\Omega, \mathcal{F}, \mathsf{P})$, $\mathsf{E}\,|X(t)| < \infty$, $t \geq 0$. X is called martingale (resp. sub- or supermartingale), if $\mathsf{E}(X(t) \mid \mathcal{F}_s) \overset{\geq}{\underset{\leq}{=}} X(s)$ for all $s, t \geq 0$ with $t \geq s$: $\Rightarrow$ $\mathsf{E}(X(t)) = \mathsf{E}(X(s)) = \text{const}$ for all s, t.

Examples

Very often martingales are constructed on the basis of a stochastic process $Y = \{Y(t),\ t \geq 0\}$ as follows: $X(t) = Y(t) - \mathsf{E}Y(t)$.

1. *Poisson process*
   Let $Y = \{Y(t),\ t \geq 0\}$ be the homogeneous Poisson process with intensity $\lambda > 0$. $\mathsf{E}Y(t) = \operatorname{var} Y(t) = \lambda t$, weil $Y(t) \sim \text{Pois}(\lambda t)$, $t \geq 0$.
   a) $X(t) = Y(t) - \lambda t$, $t \geq 0 \Rightarrow X(t)$ is a martingale w.r.t. the natural filtration $\{\mathcal{F}_s,\ s \geq 0\}$.
   $$\begin{aligned} \mathsf{E}(X(t) \mid \mathcal{F}_s)_{s\leq t} &= \mathsf{E}(Y(t) - \lambda t - (Y(s) - \lambda s + (Y(s) - \lambda s)) \mid \mathcal{F}_s) \\ &= Y(s) - \lambda s + \mathsf{E}(Y(t) - Y(s) - \lambda(t-s) \mid \mathcal{F}_s) \\ &= Y(s) - \lambda s + \mathsf{E}(Y(t) - Y(s)) + Y(s) - \lambda s \\ &= Y(s) - \lambda s + \underbrace{\mathsf{E}(Y(t-s))}_{=\lambda(t-s)} - \lambda(t-s) \\ &= Y(s) - \lambda \overset{a.s.}{=} X(s) \end{aligned}$$
   b) $X'(t) = X^2(t) - \lambda(t)$, $t \geq 0 \Rightarrow X'(t)$ is a martingale w.r.t. $\{\mathcal{F}_s,\ s \geq 0\}$.
   $$\begin{aligned} \mathsf{E}(X'(t) \mid \mathcal{F}_s) &= \mathsf{E}(X^2(t) - \lambda t \mid \mathcal{F}_s) = \mathsf{E}((X(t) - X(s) + X(s))^2 - \lambda t \mid \mathcal{F}_s) \\ &= \mathsf{E}((X(t) - X(s))^2 + 2((X(t) - X(s))X(s)) + X^2(s) - \lambda s - \lambda(t-s) \mid \mathcal{F}_s) \\ &= X'(s) + \underbrace{\mathsf{E}((X(t) - X(s)))^2}_{=\operatorname{var}(Y(t)-Y(s))=\lambda(t-s)} + 2X(s)\underbrace{\mathsf{E}(X(t) - X(s))}_{=0} - \lambda(t-s) \\ &\overset{a.s.}{=} X'(s), \quad s \leq t. \end{aligned}$$

2. *Compound Poisson process*
   $Y(t)=\sum_{i=1}^{N(t)} U_i$, $t\geq 0$, N – homogeneous Poisson process with intensity $\lambda>0$, U_i – independent identically distributed random variables, $\mathsf{E}|U_i|<\infty$, $\{U_i\}$ independent of N.
   $X(t)=Y(t)-\mathsf{E}Y(t)=Y(t)-\lambda t\mathsf{E}U_1$, $t\geq 0$.

   Exercise 5.2.1
   Show that $X=\{X(t),\ t\geq 0\}$ is a martingale w.r.t. the natural filtration.

3. *Wiener process*
   Let $W=\{W(t),\ t\geq 0\}$ be a Wiener process, $\{\mathcal{F}_s,\ s\geq 0\}$ be the natural filtration.
   a) $Y=\{Y(t),\ t\geq 0\}$
   $Y(t)=W^2(t)-\mathsf{E}W^2(t)=W^2(t)-t$, $t\geq 0$, is a martingale w.r.t. $\{\mathcal{F}_s,\ s\geq 0\}$.

$$\begin{aligned}\mathsf{E}(Y(t)\mid\mathcal{F}_s) &= \mathsf{E}((W(t)-W(s)+W(s))^2-s-(t-s)\mid\mathcal{F}_s)\\ &= \text{see example 1b, use the independence and stationarity of the increments of } W\\ &= W^2(s)-s\overset{a.s.}{=}Y(s),\quad s\leq t.\end{aligned}$$

   b) $Y'(t)=e^{uW(t)-u^2\frac{t}{2}}$, $t\geq 0$ and a fixed $u\in\mathbb{R}$.
   $\mathsf{E}|Y'(t)|=e^{-u^2\frac{t}{2}}\mathsf{E}e^{uW(t)}=e^{u^2\frac{t}{2}}e^{u^2\frac{t}{2}}=1<\infty$. We show that $Y'=\{Y'(t),\ t\geq 0\}$ is a martingale w.r.t. $\{\mathcal{F}_s,\ s\geq 0\}$.

$$\begin{aligned}\mathsf{E}(Y'(t)\mid\mathcal{F}_s) &= \mathsf{E}(e^{u(W(t)-W(s)+W(s))-u^2\frac{s}{2}-u^2\frac{(t-s)}{2}}\mid\mathcal{F}_s)\\ &= \underbrace{e^{-u^2\frac{s}{2}}e^{uW(s)}}_{=Y'(s)}e^{-u^2\frac{(t-s)}{2}}\underbrace{\mathsf{E}(e^{u(W(t)-W(s))}\mid\mathcal{F}_s)}_{=\mathsf{E}(e^{uW(t-s)})=e^{u^2\frac{(t-s)}{2}}}\\ &= Y'(s)e^{-u^2\frac{(t-s)}{2}}e^{u^2\frac{(t-s)}{2}}=Y'(s),\quad s\leq t.\end{aligned}$$

4. *Closed martingale*
   Let X be a random variable (on $(\Omega,\mathcal{F},\mathsf{P})$) with $\mathsf{E}|X|<\infty$. Let $\{\mathcal{F}_s,\ s\geq 0\}$ be a filtration on $(\Omega,\mathcal{F},\mathsf{P})$.
   $Y(t)=\mathsf{E}(X\mid\mathcal{F}_t)$, $t\geq 0$. $Y=\{Y(t),\ t\geq 0\}$ is a martingale.
   $\mathsf{E}|Y(t)|=\mathsf{E}|\mathsf{E}(X\mid\mathcal{F}_t)|\leq\mathsf{E}(\mathsf{E}(X\mid\mathcal{F}_t))=\mathsf{E}|X|<\infty$, $t\geq 0$.
   $\mathsf{E}(Y(t)\mid\mathcal{F}_s)=\mathsf{E}((X\mid\mathcal{F}_t)\mid\mathcal{F}_s)=\mathsf{E}(X\mid\mathcal{F}_s)\overset{a.s.}{=}Y(s)$, $s\leq t\Rightarrow\mathcal{F}_s\subseteq\mathcal{F}_t$.

5. *Lèvy processes*
   Let $X=\{X(t),\ t\geq 0\}$ be a Lèvy process with Lèvy exponent η and natural filtration $\{\mathcal{F}_s,\ s\geq 0\}$.
   a) If $\mathsf{E}|X(1)|<\infty$, define $Y(t)=X(t)-\underbrace{t\mathsf{E}X(1)}_{=\mathsf{E}X(t)}$, $t\geq 0$. As in the previous cases it can be shown that $Y=\{Y(t),\ t\geq 0\}$ is martingale w.r.t. the filtration $\{\mathcal{F}_s,\ s\geq 0\}$.
   b) Use the combination from example 3b – normalize the characteristic function of $X(t)$ without expectation, through the value $Y(t)=\frac{e^{iuX(t)}}{\varphi_{X(t)}^{(u)}}=\frac{e^{iuX(t)}}{t\eta(u)}=e^{iuX(t)-t\eta(u)}$, $t\geq 0$, $u\in\mathbb{R}$.

To show: $Y = \{Y(t),\ t \geq 0\}$ is a complexvalued martingale.
$\mathsf{E}|Y(t)| = |e^{-t\eta(u)}| < \infty$, since $\eta : \mathbb{R}_+ \to \mathbb{C}$. $\mathsf{E}Y(t) = 1$, $t \geq 0$.

$$\begin{aligned} \mathsf{E}(Y(t) \mid \mathcal{F}_s) &= \mathsf{E}(e^{iu(X(t)-X(s))(t-s)\eta(u)} e^{iuX(s)-s\eta(u)} \mid \mathcal{F}_s) \\ &= e^{iuX(s)-s\eta(u)} e^{-(t-s)\eta(u)} \mathsf{E}(e^{iu(X(t)-X(s))}) \\ &= Y(s) e^{-(t-s)\eta(u)} e^{(t-s)\eta(u)} \stackrel{a.s.}{=} Y(s) \end{aligned}$$

6. *Submartingale/Supermartingale*
   Every integrable stochastich process $X = \{X(t),\ t \geq 0\}$, which is adapted w.r.t. to a filtration $\{\mathcal{F}_s,\ s \geq 0\}$ and has a.s. monotone nondecreasing (resp. nonincreasing) trajectories, is a sub- (resp. a super-)martingale.
   In fact, it hols $X(t) \stackrel{a.s.}{\geq} X(s)$, $t \geq s \Rightarrow \mathsf{E}(X(t) \mid \mathcal{F}_s) \stackrel{a.s.}{\geq} \mathsf{E}(X(s) \mid \mathcal{F}_s) \stackrel{a.s.}{=} X(s)$. In particular, every subordinator is a submartingale.

Lemma 5.2.1
Let $X = \{X(t),\ t \geq 0\}$ be a stochastic process, which is adapted w.r.t. a filtration $\{\mathcal{F}_t,\ t \geq 0\}$ and let $f : \mathbb{R} \to \mathbb{R}$ be convex, such that $\mathsf{E}|f(X(t))| < \infty$, $t \geq 0$. Then $Y = \{f(X(t),\ t \geq 0)\}$ is a submartingale, if

a) X is a submartingale, or

b) X is a submartingale and f is monotone nondecreasing.

Proof Use the Jensen inequality for conditional expectations. $\mathsf{E}(f(X(t)) \mid \mathcal{F}_s) \geq f(\underbrace{\mathsf{E}(X(t)\mathcal{F}_s)}_{\geq X(s)}) \geq f(X(s))$, since f is monotone nondecreasing (case b)) or the equation holds (case a)). □

5.3 Uniform Integrability

Question: It is known, that in general $X_n \xrightarrow[n\to\infty]{a.s.} X$ does not give $X_n \xrightarrow[n\to\infty]{L_1} X$. Here $X, X_1, X_2, \dots$ are random variables, defined on the probability space $(\Omega, \mathcal{F}, \mathsf{P})$. When does „$X_n \xrightarrow[n\to\infty]{L_1} X$“ $\Rightarrow$ „$X_n \xrightarrow[n\to\infty]{L_1} X$“hold? The answer for this provides the term uniformly integrability of $\{X_n,\ n \in \mathbb{N}\}$.

Definition 5.3.1
The sequence $\{X_n,\ n \in \mathbb{N}\}$ of random variables is called uniformly integrable, if $\mathsf{E}|X_n| < \infty$, $n \in \mathbb{N}$, and $\sup_n \mathsf{E}(|X_n| 1(|X_n| > \varepsilon)) \xrightarrow[\varepsilon\to+\infty]{} 0$.

Lemma 5.3.1
The sequence $\{X_n,\ n \in \mathbb{N}\}$ of random variables is uniformly integrable, if and only if

1. $\sup_n \mathsf{E}|X_n| < \infty$ (uniformly bounded),

2. if for every $\varepsilon > 0$ there is a $\delta > 0$, such that $\mathsf{E}(|X_n| 1(A)) < \varepsilon$ for all $n \in \mathbb{N}$ and all $A \in \mathcal{F}$ with $\mathsf{P}(A) < \delta$.

Proof Let $\{X_n\}$ be a sequence of random variables.
It has to be shown that

$$\sup_n \mathsf{E}(|X_n| 1(|X_n| > x)) \xrightarrow[n\to\infty]{} 0 \Longleftrightarrow \begin{array}{ll} 1) & \sup_n \mathsf{E}|X_n| < \infty \\ 2) & \forall \varepsilon > 0\ \exists \delta > 0 : \mathsf{E}(|X_n| 1(A)) < \varepsilon\ \forall A \in \mathcal{F} : \mathsf{P}(A) < \delta \end{array}$$

„$\Leftarrow$"
$A_n = \{|X_n| > x\}$. From the Markov inequality: $\mathsf{P}(A_n) \le \frac{1}{x}\mathsf{E}|X_n|$ for all $n \Rightarrow \sup_n \mathsf{P}(A_n) \le \frac{1}{x}\sup_n \mathsf{E}|X_n| \le \frac{c}{x} \xrightarrow[x\to\infty]{} 0 \Rightarrow \exists N > 0 : \forall x > N\ \mathsf{P}(A_n) < \delta \overset{2)}{\Rightarrow} \sup_n \mathsf{E}(|X_n|1(A_n)) \le \varepsilon \Rightarrow$ since $\varepsilon > 0$ can be chosen arbitrarily small $\Rightarrow \sup_n \mathsf{E}(|X_n|1(|X_n| > x)) \xrightarrow[x\to\infty]{} 0$.

„$\Rightarrow$"

1.

$$\begin{aligned}
\sup_n \mathsf{E}|X_n| &\le \sup_n(\mathsf{E}(|X_n|1(|X_n| > x)) + \mathsf{E}(|X_n|1(|X_n| \le x)))\\
&\le \sup_n(\mathsf{E}(|X_n|1(|X_n| > x)) + x\underbrace{\mathsf{P}(|X_n| \le x)}_{\le 1})\\
&\le \varepsilon + x < \infty
\end{aligned}$$

2.

$$\begin{aligned}
\mathsf{E}(|X_n|1(A)) &= \mathsf{E}(\underbrace{|X_n|}_{\le x}\underbrace{1(|X_n| \le x)}_{\le 1}1(A)) + \mathsf{E}(|X_n|1(|X_n| > x)\underbrace{1(A)}_{\le 1})\\
&\le \underbrace{x\mathsf{P}(A)}_{\le\frac{\varepsilon}{2}} + \underbrace{\mathsf{E}(|X_n|1(|X_n| > x))}_{\le\frac{\varepsilon}{2}},
\end{aligned}$$

for all $\varepsilon > 0\ \exists x > 0$, such that $\mathsf{E}(|X_n|1(|X_n| > x)) < \frac{\varepsilon}{2}$ because of uniformly integrability. Choose $\delta > 0$, $x\delta < \frac{\varepsilon}{2}$.

□

Lemma 5.3.2
Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of random variables with $\mathsf{E}|X_n| < \infty$, $n \in \mathbb{N}$, $X_n \xrightarrow[n\to\infty]{a.s.} X$. $X_n \xrightarrow[n\to\infty]{L^1} X$ if and only if $\{X_n\}_{n\in\mathbb{N}}$ is uniformly integrable. In particular follows from $X_n \xrightarrow[n\to\infty]{L^1} X$ the convergence $\mathsf{E}X_n \xrightarrow[n\to\infty]{} \mathsf{E}X$.

Proof Let $\{X_n\}_{n\in\mathbb{N}}$ be uniformly integrable. It has to be shown, that $\mathsf{E}|X_n - X| \xrightarrow[n\to\infty]{} 0$.
$X_n \xrightarrow[n\to\infty]{a.s.} X \Rightarrow X_n \xrightarrow[n\to\infty]{\mathsf{P}} X \Rightarrow \mathsf{P}(|X_n - X| > \varepsilon) \xrightarrow[n\to\infty]{} 0$ for all ε.

$$\begin{aligned}
\mathsf{E}\,|X_n - X| &\le \mathsf{E}\left(|X_n - X|\,1\left(|X_n - X| \le \varepsilon\right)\right) + \mathsf{E}\left(|X_n - X|\,1\left(|X_n - X| > \varepsilon\right)\right)\\
&\le \varepsilon + \underbrace{\mathsf{E}\left(|X_n - X|\,1\left(|X_n - X| > \varepsilon\right)\right)}_{\xrightarrow[n\to\infty]{}0,\text{ because of lemma 5.3.1, 2) for } A_n=\{|X_n-X|>\varepsilon\}}\\
&\quad + \underbrace{\mathsf{E}\left(|X|\,1\left(|X_n - X| > \varepsilon\right)\right)}_{\xrightarrow[n\to\infty]{}0,\text{ weil } \mathsf{E}|X|<\infty,\text{ after the theorem of Lebesgue}} \xrightarrow[n\to\infty]{} 0
\end{aligned}$$

Why $\mathsf{E}|X| < \infty$? It holds $X_n \xrightarrow[n\to\infty]{a.s.} X$, from Lemma 5.3.1, 1): $\sup_n \mathsf{E}|X_n| < \infty$. After the lemma of Fatou it holds $\mathsf{E}|X| < \infty$, since for all $\varepsilon_0 > 0\ \exists N$: for all $n > N$ $|X_n - X| < \varepsilon_0 \Rightarrow$

$X_n \le \eta_1$, $\eta_1 = |X| + \varepsilon_0$, $X_n \ge \eta_2$, $\eta_2 = |X| - \varepsilon_0$, for all $n > N$. $\mathsf{E}|X| = \mathsf{E}|\lim_{n\to\infty} X_n| \le \lim_{n\to\infty} \mathsf{E}|X_n| < \infty$. Thus we have proven that $X_n \xrightarrow[n\to\infty]{L^1} X$.
Now let $\mathsf{E}|X_n - X| \xrightarrow[n\to\infty]{} 0$. The properties 1) and 2) of lemma 5.3.1 have to be shown.

1. $\sup_n \mathsf{E}|X_n| \le \sup_n \mathsf{E}|X_n - X| + \mathsf{E}|X| < \infty$, since $X_n \xrightarrow{L^1} X$.
2. For all $A \subset \mathcal{F}$, $\mathsf{P}(A) \le \delta$: $\mathsf{E}(|X_n|1(A)) \le \mathsf{E}(|X_n - X|\underbrace{1(A)}_{\le 1}) + \mathsf{E}(|X|1(A)) \le \underbrace{\mathsf{E}|X_n - X|}_{<\frac{\varepsilon}{2}} + \frac{\varepsilon}{2} = \varepsilon$ with an appropriate choice of δ, since $\mathsf{E}|X| < \infty$ and since for all $\varepsilon > 0$ $\exists N$, such that for all $n > N$ $\mathsf{E}|X_n - X| < \frac{\varepsilon}{2}$.

□

5.4 Stopped Martingales

Notation: $x_+ = (x)_+ = \max(x, 0)$, $x \in \mathbb{R}$.

Theorem 5.4.1 (Doob's inequality):
Let $X = \{X(t),\ t \ge 0\}$ be a càdlàg process, adapted w.r.t. the filtration $\{\mathcal{F},\ t \ge 0\}$. Let X be a submartingale. Then for arbitary $t > 0$ and arbitrary $x > 0$ it holds:

$$\mathsf{P}\left(\sup_{0\le s\le t} X(x) > x\right) \le \frac{\mathsf{E}(X(t))_+}{x}$$

Proof W.l.o.g. assume $X(t) \ge 0$, $t \ge 0$ a.s..
$\mathsf{P}(\sup_{0\le s\le t} X(s) > x) = \mathsf{P}(\sup_{0\le s\le t}((X(s))_+ > x))$, for all $t \ge 0$, $x > 0$. $A = \{\sup_{t_1,\dots,t_n} X(s) > x\}$, $0 \le t_1 < t_2 < \ldots < t_n \le t$ – arbitrary times. $A = \cup_{k=1}^n A_k$,

$$\begin{aligned}
A_1 &= \{X(t_1) > x\}\\
A_2 &= \{X(t_2) \le x, X(t_2) > x\}\\
&\ \vdots\\
A_k &= \{X(t_1) \le x, X(t_1) \le x, \ldots, X(t_{k-1}) \le x, X(t_k) > x\},
\end{aligned}$$

$k = 2, \ldots, n$, $A_i \cap A_j = \emptyset$, $i \ne j$.
It has to be shown that $\mathsf{P}(A) \le \frac{\mathsf{E}(X(t_n))}{x}$.
$\mathsf{E}(X(t_n)) \ge \mathsf{E}(X(t_n)1(A)) = \sum_{k=1}^n \mathsf{E}(X(t_n)1(A_k)) \ge x\sum_{k=1}^n \mathsf{P}(A_k) = x\mathsf{P}(A)$, $k = 1, \ldots, n-1$, since X is a martingale and thus follows that $\mathsf{E}(X(t_n)1(A_k)) \ge \mathsf{E}(X(t_k)1(A_k)) \ge \mathsf{E}(x1(A_k)) = x\mathsf{P}(A_k)$, $k = 1, \ldots, n-1$, $t_n > t_k$.
Let $B \subset [0, t]$ be a finite subset, $0 \in B$, $t \in B \Rightarrow$ it is proven similarly that $\mathsf{P}(\max_{s\in B} X(s) > x) \le \frac{\mathsf{E}X(t)}{x}$.
$\mathbb{Q}$ is dense in $\mathbb{R} \Rightarrow [0, t) \cap \mathbb{Q} \cup \{t\} = \cup_{k=1}^\infty B_k$, $B_k \subset [0, t) \cap \mathbb{Q} \cup \{t\}$ finite, $B_k \subset B_n$, $k < n$. By the monotonicity of the probability measure it holds:

$$\lim_{n\to\infty} \mathsf{P}\left(\max_{s\in B} X(s) \ge x\right) = \mathsf{P}\left(\cup_n\{\max_{s\in B_n} X(s) > x\}\right) = \mathsf{P}\left(\sup_{s\in\cup_n B_n} X(s) > x\right) \le \frac{\mathsf{E}X(t)}{x}$$

By the right-continuity of the paths of X it holds $\mathsf{P}(\sup_{0\le s\le t} X(s) > x) \le \frac{\mathsf{E}X(t)}{x}$. □

Conclusion 5.4.1
For the Wiener process $W = \{W(t),\ t \geq 0\}$ we are looking at the Wiener process with negative drift: $Y(t) = W(t) - \mu t$, $\mu > 0$, $t \geq 0$. From example nr.3 of section 5.3 $X(t) = \exp\{u(Y(t)+t\mu) - \frac{u^2 t}{2}\}$, $t \geq 0$ is a martingale w.r.t. the natural filtration of W. For $u = 2\mu$ it holds

$$X(t) = \exp\{2\mu Y(t)\}, \quad t \geq 0.$$

$\mathsf{P}\left\{\sup_{0\leq s\leq t} Y(s) > x\right\} = \mathsf{P}\left\{\sup_{0\leq s\leq t} e^{2\mu Y(s)} > e^{2\mu x}\right\} \leq \frac{\mathsf{E}e^{2\mu Y(t)}}{e^{2\mu x}} = e^{-2\mu x}$, $x > 0$
$\Rightarrow \lim_{t\to\infty} \mathsf{P}\{\sup_{0\leq s\leq t} Y(s) > x\}$. From example nr.3 $\mathsf{E}e^{2\mu Y(t)} = \mathsf{P}(\sup_{t\geq 0} Y(t) > x) \leq e^{-2\mu x}$ holds.

Theorem 5.4.2
Let $X = \{X(t),\ t \geq 0\}$ be a martingale w.r.t. the filtration $\{\mathcal{F}_t,\ t \geq 0\}$ with càdlàg paths. If $T : \Omega \to [0, \infty)$ is a finite stopping time w.r.t. the filtration $\{\mathcal{F}_t,\ t \geq 0\}$, then, the stochastic process $\{X_{T\wedge t}(t) \geq 0\}$ is also a martingale, which is also called a stopped martingale. Where $a \wedge b = \min\{a, b\}$.

Lemma 5.4.1
Let $X = \{X(t),\ t \geq 0\}$ be a martingale with càdlàg-trajectories w.r.t. the filtration $\{\mathcal{F}_t,\ t \geq 0\}$. Let T be a finite stopping time and let $\{T_n\}_{n\in\mathbb{N}}$ be the sequence of discrete stopping times out of theorem 5.1.2, for which $T_n \downarrow T$, $n \to \infty$, holds. Then $\{X(T_n \wedge t)\}_{n\in\mathbb{N}}$ is uniformly integrable for every $t \geq 0$.

Proof

$$T_n = \begin{cases} 0 \quad , & \text{if } T = 0 \\ \frac{k+1}{2^n}, & \text{if } \frac{k}{2^n} < T \leq \frac{k+1}{2^n}, \text{ for a } k \in \mathbb{N}_0 \end{cases}$$

1. It is to be shown: $\mathsf{E}|X(T_n \wedge t)| < \infty$ for all n.
   $\mathsf{E}|X(T_n \wedge t)| \leq \sum_{k:\frac{k}{2^n}<t} \mathsf{E}|X(\frac{k}{2^n})| + \mathsf{E}|X(t)| < \infty$, since X is a martingale, therefore integrable.

2. It is to be shown: $\sup_n \mathsf{E}(|X(T_n \wedge t)| 1(\underbrace{|X(T_n \wedge t)| > x}_{A_n})) \xrightarrow[x\to\infty]{} 0$.

$$\begin{aligned}
&\sup_n \mathsf{E}(|X(T_n \wedge t)| 1(A_n)) \\
= \ &\sup_n \left(\sum_{k:\frac{k}{2^n}<t} \mathsf{E}\left(\left|X\left(\frac{k}{2^n}\right)\right| 1\left(\left\{T_n = \frac{k}{2^n}\right\} \cap A_n\right)\right) + \mathsf{E}\left(|X(t)|\, 1\left(T_n > t\right) 1\left(A_n\right)\right)\right) \\
\leq \ &\sup_n \left(\sum_{k:\frac{k}{2^n}<t} \mathsf{E}\left(|X(t)|\, 1\left(\left\{T_n = \frac{k}{2^n}\right\} \cap A_n\right)\right) + \mathsf{E}\left(|X(t)|\, 1\left(\{T_n > t\} \cap A_n\right)\right)\right) \\
= \ &\sup_n \mathsf{E}\left(|X(t)|\, 1\left(A_n\right)\right) \leq \sup_n \mathsf{E}\left(|X(t)|\, 1\left(Y > x\right)\right) \\
= \ &\mathsf{E}\left(|X(t)|\, 1\left(Y > x\right)\right),
\end{aligned}$$

   where $1(A_n) \leq 1(\underbrace{\sup_n |X(T_n \wedge t)|}_{Y} > x)$. It is to be shown: $\mathsf{P}(Y > x) \xrightarrow[n\to\infty]{} 0$ with help of Doob's inequality.

$\mathsf{P}(Y > x) \leq \mathsf{P}(\sup_{0\leq s\leq t}|X(s)| > x) \leq \frac{\mathsf{E}|X(t)|}{x} \xrightarrow[x\to+\infty]{} 0$. Since $\mathsf{E}|X(t)| < \infty$ for all $t \geq 0$ and $\mathsf{P}(Y > x) \xrightarrow[x\to\infty]{} 0$, this gives $\mathsf{E}(|X(t)|1(Y > x)) \xrightarrow[n\to\infty]{} 0 \Rightarrow \sup_n \mathsf{E}|X(T_n \wedge t)1(A)| \xrightarrow[x\to\infty]{} 0$ $\{X(T_n \wedge t)\}_{n\in\mathbb{N}}$ is uniformly integrable.

□

Proof of theorem 5.4.2
It is to be shown that $\{X(T \wedge t),\ t \geq 0\}$ is a martingale.

1. $\mathsf{E}|X(T \wedge t)| < \infty$ for all $t \geq 0$. As in conclusion 5.1.1 $T_n \downarrow T$, $n \to \infty \Rightarrow X(T_n \wedge t) \xrightarrow[n\to\infty]{a.s.} X(T \wedge t)$ is approximated, but since $\mathsf{E}|X(T_n \wedge t)| < \infty$ for all n it follows $\mathsf{E}|X(T \wedge t)| < \infty$ because of lemma 5.4.1, since uniform integrability gives L^1-convergence.

2. *Martingale property*
It is to be shown:
$$\begin{array}{ccc} \mathsf{E}(X(T \wedge t) \mid \mathcal{F}_s) & \overset{a.s.}{=} & X(T \wedge s), \quad s \leq t \\ & \Updownarrow & \\ \mathsf{E}(X(T \wedge t)1(A)) & \overset{a.s.}{=} & \mathsf{E}(X(T \wedge s)1(A)),\ A \in \mathcal{F}_s \end{array}$$
First of all, we show that $\mathsf{E}(|X(T_n \wedge t)|1(A)) = \mathsf{E}(|X(T_n \wedge s)|1(A))$, $A \in \mathcal{F}_s$, $n \in \mathbb{N}$. Let $t_1, \ldots, t_k \in (s, t)$ be discrete values, which T_n takes with positive probability in (s, t).
$$\begin{aligned} \mathsf{E}(X(T_n \wedge t) \mid \mathcal{F}_s) &= \mathsf{E}(\mathsf{E}(X(T_n \wedge t) \mid \mathcal{F}_{t_k}) \mid \mathcal{F}_s) \\ &= \mathsf{E}(\mathsf{E}(\underbrace{X(T_n \wedge t)}_{X(t_k)} 1(T_n \leq t_k) \mid \mathcal{F}_{t_k}) \mid \mathcal{F}_s) \\ &\quad + \mathsf{E}(\mathsf{E}(\underbrace{X(T_n \wedge t)}_{X(t)} 1(T_n > t_k) \mid \mathcal{F}_{t_k}) \mid \mathcal{F}_s) \\ &= \mathsf{E}(X(t_k)1(T_n \leq t_k) \mid \mathcal{F}_s) + \mathsf{E}(1(T_n > t_k)\mathsf{E}(X(t) \mid \mathcal{F}_{t_k}) \mid \mathcal{F}_s) \\ &= \mathsf{E}(X(t_k \wedge T_n) \mid \mathcal{F}_s) = \ldots = \mathsf{E}(X(t_{k-1} \wedge T_n) \mid \mathcal{F}_s) = \ldots \\ &= \mathsf{E}(X(t_1 \wedge T_n) \mid \mathcal{F}_s) = \ldots = \mathsf{E}(X(T_n \wedge s) \mid \mathcal{F}_s) \\ &\overset{a.s.}{=} X(T_n \wedge s) \end{aligned}$$
Since X is càdlàg and $T_n \downarrow T$, $n \to \infty$, it holds $X(T_n \wedge t) \xrightarrow[n\to\infty]{a.s.} X(T_n \wedge t)$. Furthermore $\{X(T_n \wedge t)\}_{n\in\mathbb{N}}$ are uniformly integrable because of L^1-convergence. Therefore follows that
$$\begin{array}{ccc} \mathsf{E}(X(T_n \wedge t)1(A)) & = & \mathsf{E}(X(T_n \wedge s)1(A)) \quad \text{for all } A \in \mathcal{F}_s \\ \downarrow & & \downarrow \\ \mathsf{E}(X(T \wedge t)1(A)) & = & \mathsf{E}(X(T \wedge s)1(A)) \end{array}$$
$\Rightarrow \{X(T \wedge t),\ t \geq 0\}$ is a martingale.

□

Definition 5.4.1
Let $T : \Omega \to \mathbb{R}_+$ be a stopping time w.r.t. the filtration $\{\mathcal{F}_t,\ t \geq 0\}$, $\mathcal{F}_t \subset \mathcal{F}$, $t \geq 0$. The „stopped“ σ-algebra $\mathcal{F}_T$ is defined by $A \in \mathcal{F}_T \Rightarrow A \cap \{T \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$.

Lemma 5.4.2 1. Let S, T – stopping times w.r.t. the filtration $\{\mathcal{F}_t,\ t \geq 0\}$, $S \overset{a.s.}{\leq} T$. Then it holds $\mathcal{F}_S \subset \mathcal{F}_T$.

2. Let $X = \{X(t),\ t \geq 0\}$ be a martingale with càdlàg-trajectories w.r.t. the filtration $\{\mathcal{F}_t,\ t \geq 0\}$ and let T be a stopping time w.r.t. $\{\mathcal{F}_t,\ t \geq 0\}$. Then $X(T)$ is $\mathcal{F}_T$-measurable.

Proof 1. $A \in \mathcal{F}_s \Rightarrow A \cap \{S \leq t\} \in \mathcal{F}_t$, $t \geq 0$. $A \cap \{T \leq t\} = \underbrace{A \cap \{S \leq t\}}_{\in \mathcal{F}_t} \cap \underbrace{\{T \leq t\}}_{\in \mathcal{F}_t} \in \mathcal{F}_t$ for all $t \geq 0 \Rightarrow A \in \mathcal{F}_T$.

2. $X(T) = g \circ f$, $f : \Omega \to \Omega \times \mathbb{R}_+$, $f(\omega) = (\omega, T(\omega))$, $g : \Omega \times \mathbb{R}_+ \to \mathbb{R}$, $g(\omega, s) = X(s, \omega)$.
It has to be shown: f-$\mathcal{F} \mid \mathcal{F} \times B_{\mathbb{R}_+}$-measurable, g-$\mathcal{F} \times B_{\mathbb{R}_+} \mid \mathcal{F}_T$-measurable $\Rightarrow g \circ f$-$\mathcal{F} \mid \mathcal{F}_T$-measurable.
f-$\mathcal{F} \mid \mathcal{F} \times B_{\mathbb{R}_+}$-measurable is obvious, since T is a random variable. If we're looking at the restriction of $X = \{X(s),\ s \geq 0\}$ on $s \in [0, t]$, $t \geq 0$.
It has to be shown: $\{X(T) \in B\} \cap \{T \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$, $B \in \mathcal{B}(\mathbb{R})$.
X – càdlàg $\Rightarrow X(s, \omega) = X(0, \omega)1(s = 0) + \lim_{n\to\infty} \sum_{k=1}^{2^n} X(t\frac{k}{2^n}, \omega)1(\frac{k-1}{2^n}t < s \leq \frac{k}{2^n}t) \Rightarrow$ $X(s, \omega)$ is $B_{[0,t]} \times \mathcal{F}_t$-measurable $\Rightarrow X(T)$ is $\mathcal{F} \mid \mathcal{F}_T$-measurable.

□

Theorem 5.4.3 (Optional sampling-theorem):
Let $X = \{X(t),\ t \geq 0\}$ be a martingale with càdlàg-trajectories w.r.t. a filtration $\{\mathcal{F}_t,\ t \geq 0\}$ and let T be a finite stopping time w.r.t. $\{\mathcal{F}_t,\ t \geq 0\} \Rightarrow \mathsf{E}(X(t) \mid \mathcal{F}_T) \overset{a.s.}{=} X(T \wedge t)$, $t > 0$.

Proof First of all we show that $\mathsf{E}(X(t) \mid \mathcal{F}_{T_n}) \overset{a.s.}{=} X(T_n \wedge t)$, $t \geq 0$, $n \in \mathbb{N}$, where $T_n \downarrow T$, $n \to \infty$ is the discrete approximation of T. Let $t_1 \leq t_2 \leq \ldots \leq t_k = t$ be the values, which $T_n \wedge t$ takes with positive probability. It is to be shown, that for all $A \in \mathcal{F}_{T_n}$ it holds: $\mathsf{E}(X(t)1(A)) = \mathsf{E}(X(T_n \wedge t)1(A))$.

$$\begin{aligned}
(X(t) - X(T_n \wedge t))1(A) &= \sum_{i=1}^{k-1} X(t_k) - X(t_i)1(\{T_n \wedge t = t_i\} \cap A) \\
&= \sum_{i=2}^{k} (X(t_i) - X(t_{i-1}))1(A)1(\{T_n \wedge t < t_i\}) \\
\mathsf{E}((X(t) - X(T_n \wedge t))1(A)) &= \sum_{i=2}^{k} \mathsf{E}((X(t_i) - X(t_{i-1}))1(T_n \wedge t < t_i)1(A)) \\
&= \sum_{i=2}^{k} \mathsf{E}(\mathsf{E}(X(t_i) - X(t_{i-1}))1(T_n \wedge t < t_i)1(A) \mid \mathcal{F}_{t_{i-1}}) \\
&= \sum_{i=2}^{k} \mathsf{E}(1(T_n \wedge t < t_i)1(A))\mathsf{E}((X(t_i) - X(t_{i-1})) \mid \mathcal{F}_{t_{i-1}}) = 0
\end{aligned}$$

$\mathsf{E}(X(t) \mid \mathcal{F}_{T_n}) \overset{a.s.}{=} \mathsf{E}(X(T_n \wedge t) \mid \mathcal{F}_{T_n}) \overset{a.s.}{=} X(T_n \wedge t)$, since $X(T_n)$ is $\mathcal{F}_{T_n}$-measurable. $T \leq T_n \Rightarrow \mathcal{F}_T \subseteq \mathcal{F}_{T_n}$. Since $\{X(T_n \wedge t)\}_{n\in\mathbb{N}}$ is uniformly integrable for $t \in [0, \infty)$, it holds

$$\mathsf{E}(X(t) \mid \mathcal{F}_T) = \mathsf{E}(X(t) \mid \mathcal{F}_{T_n}) = \lim_{n\to\infty} \mathsf{E}(X(T_n \wedge t) \mid \mathcal{F}_{T_n}) = \lim_{n\to\infty} X(T_n \wedge t) = X(T \wedge t),$$

since X is càdlàg. □

Conclusion 5.4.2
Let $X = \{X(t),\ t \geq 0\}$ be a càdlàg-martingale and let S, T be finite stopping times, such that $\mathsf{P}(S \leq T) = 1$. Then it holds $\mathsf{E}(X(t \wedge T) \mid \mathcal{F}_s) \overset{a.s.}{=} \mathsf{E}(X(S \wedge t)),\ t \geq 0$. In particular $\mathsf{E}(X(T \wedge t))) = \mathsf{E}(X(0))$ holds.

Proof X – martingale. From theorem 5.4.2 $\{X(T \wedge t),\ t > 0\}$ is also a martingale. Use theorem 5.4.3 for this true martingale:

$$\mathsf{E}(X(T \wedge t) \mid \mathcal{F}_s) \overset{a.s.}{=} X(T \wedge S \wedge t) \overset{a.s.}{=} X(S \wedge t),$$

since $S \overset{a.s.}{\leq} T$. Set $S = 0$, then $\mathsf{E}(\mathsf{E}(X(T \wedge t) \mid \mathcal{F}_0)) = \mathsf{E}X(0 \wedge t) = \mathsf{E}X(0)$. □

5.5 Lèvy processes and Martingales

Theorem 5.5.1
Let $X = \{X(t),\ t \geq 0\}$ be a Lèvy process with characteristics (a, b, ν).

1. There exists a càdlàg-modification of $\tilde{X} = \{\tilde{X}(t),\ t \geq 0\}$ of X with the same characteristics (a, b, ν).

2. The natural filtration of a càdlàg-Lèvy processes ist right-continuous.

Without proof

Theorem 5.5.2 (Regeneration theorem for Lèvy processes):
Let $X = \{X(t),\ t > 0\}$ be a càdlàg-Lèvy process with natural filtration $\{\mathcal{F}_t^X,\ t \geq 0\}$ and let T be a finite stopping time w.r.t. $\{\mathcal{F}_t^X,\ t \geq 0\}$. The process $Y = \{Y(t),\ t \geq 0\}$, given by $Y(t) = X(T + t) - X(T),\ t \geq 0$, is also a Lèvy process, adapted w.r.t. the filtration $\{\mathcal{F}_{T+t}^X,\ t \geq 0\}$, which is independent from $\mathcal{F}_T^X$ and has the same characteristics as X. T is called regeneration time. Since $Y \overset{d}{=} X$, Y is independent of $\mathcal{F}_T^X$.

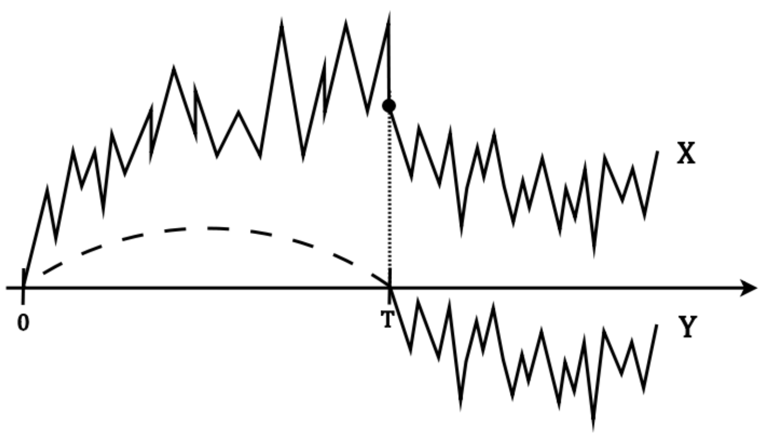

Abb. 5.1:

Proof 1. Assumption: There $\exists c > 0$, such that $\mathsf{P}(T \leq c) = 1$. Let $u_1, \ldots, u_n \in \mathbb{R}$. After example nr.5 in section 5.2 $\tilde{Y}_j = \{\tilde{Y}_j(t) = \exp\{iu_j X(t) - t\eta(u_j)\},\ t \geq 0\}$, $j = 1, \ldots, n$, is a complexvalued martingale, where $\eta(\cdot)$ is the Lèvy-exponente of $X(t)$. Let $0 \leq t_0 < t_1 < \ldots < t_n$ be arbitrary times. For all $A \in \mathcal{F}_T^X$ it holds

$$\mathsf{E}(1(A)\exp\{\sum_{j=1}^{n} iu_j(Y(t_j) - Y(t_{j-1}))\}) \overset{Z.z.}{=} \mathsf{P}(A)\mathsf{E}(\exp\{\sum_{j=1}^{n} iu_j(X(t_j) - X(t_{j-1}))\})$$

$$\begin{aligned}
& \mathsf{E}(1(A)\exp\{\sum_{j=1}^{n} iu_j(Y(t_j) - Y(t_{j-1}))\}) \\
= \;& \mathsf{E}(1(A)\exp\{\sum_{j=1}^{n} iu_j(X(T+t_j) - X(T) - X(T+t_{j-1}) - X(T)))\}) \\
= \;& \mathsf{E}\left(1(A)\prod_{j=1}^{n}\frac{\tilde{Y}_j(T+t_j)}{\tilde{Y}_j(T+t_{j-1})}\frac{\exp\{\eta(u_j)(T+t_j)\}}{\exp\{\eta(u_j)(T+t_{j-1})\}}\right) \\
= \;& \mathsf{E}\left(\mathsf{E}\left(1(A)\prod_{j=1}^{n}\frac{\tilde{Y}_j(T+t_j)}{\tilde{Y}_j(T+t_{j-1})}\exp\{(t_j - t_{j-1})\eta(u_j)\} \mid \mathcal{F}_{T+t_{j-1}}^X\right)\right) \\
= \;& \mathsf{E}\left(1(A)\prod_{j=1}^{n-1}\frac{\tilde{Y}_j(T+t_j)}{\tilde{Y}_j(T+t_{j-1})}e^{(t_j-t_{j-1})\eta(u_j)}\frac{e^{(t_n-t_{n-1})\eta(u_n)}}{\tilde{Y}_n(T+t_{n-1})}\mathsf{E}(\tilde{Y}_n(T+t_n) \mid \mathcal{F}_{T+t_{n-1}}^X)\right) \\
= \;& \mathsf{E}\left(1(A)\prod_{j=1}^{n-1}\frac{\tilde{Y}_j(T+t_j)}{\tilde{Y}_j(T+t_{j-1})}e^{(t_j-t_{j-1})\eta(u_j)} \cdot \ldots \cdot e^{(t_n-t_{n-1})\eta(u_n)}\right) \\
= \;& \ldots = \mathsf{E}(1(A)\prod_{j=1}^{n} e^{(t_j-t_{j-1})\eta(u_j)}) = \mathsf{P}(A)\prod_{j=1}^{n} e^{(t_j-t_{j-1})\eta(u_j)} \\
= \;& \mathsf{P}(A)\mathsf{E}(\exp\{i\sum_{j=1}^{n}(u_j(X(t_j) - X(t_{j-1})))\})
\end{aligned}$$

$\square$

Conclusion 5.5.1
$T_1 = T + t_n$, $S_1 = T + t_{n-1} \leq T_1$ a.s., $T_1, S_1 \leq t$, since $t > c + t_n$, $T \geq c$.

Exercise 5.5.1
Show that the statement of the theorem follows from $\mathsf{E}(1(A)\exp\{\sum_{j=1}^{n} iu_j(Y(t_j)-Y(t_{j-1}))\}) = \mathsf{P}(A)\mathsf{E}(\exp\{\sum_{j=1}^{n} iu_j(X(t_j) - X(t_{j-1}))\})$.

5.6 Martingales und Wiener Processes

Our goal: If $W = \{W(t),\ t \geq 0\}$ is a Wiener process, then it holds

$$\mathsf{P}(\max_{s\in[0,t]} W(s) > x) = \sqrt{\frac{2}{\pi t}}\int_x^{+\infty} e^{-\frac{y^2}{2t}}\,dy, \quad \text{for all } x \geq 0.$$

Theorem 5.6.1 (Reflection principle):
Let T be an arbitrary stopping time w.r.t. the natural filtration $\{\mathcal{F}_t^W,\ t \geq 0\}$. Let $X = \{X(t), t \geq 0\}$ be the reflected Wiener process at time T, i.e. $X(t) = W(T\wedge t) - (W(t) - W(T\wedge t))$, $t \geq 0$. Then $X \overset{d}{=} W$ holds.

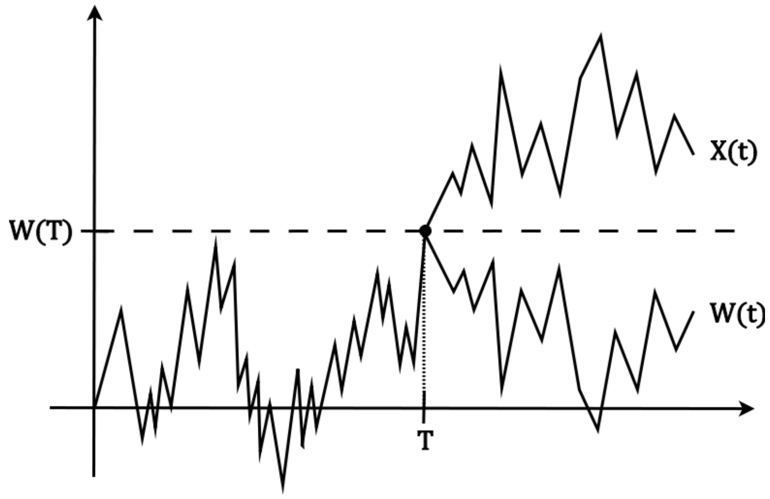

Abb. 5.2:

Proof Let $X_1(t) = W(T \wedge t)$, $X_2(t) = W(T+t) - W(T)$, $t \geq 0$. From theorem 5.5.2 follows that X_2 is independent from (T_1, X_1) (W – Lèvy process and T – regeneration time). It holds $W(t) \stackrel{g}{=} X_1(t) + X_2((t-T)_+)$, $X(t) \stackrel{g}{=} X_1(t) - X_2((t-T)_+)$, $t \geq 0$. From theorem **??** follows that

$$\begin{array}{ccc} (T_1, X_1, X_2) & \stackrel{d}{=} & (T, X_1, -X_2) \\ \downarrow & & \downarrow \\ W & \stackrel{d}{=} & X \end{array}$$

$\square$

Let $W = \{W(t),\ t \geq 0\}$ be a Wiener process on $(\Omega, \mathcal{F}, \mathsf{P})$, let $\{\mathcal{F}_t^W,\ t \geq 0\}$ the natural filtration w.r.t. W. For $z \in \mathbb{R}$ let $T_{\{z\}}^W = \inf\{t \geq 0 : W(t) = z\}$. $T_{\{z\}}^W := T_z^W$ is an a.s. finite stopping time w.r.t. $\{\mathcal{F}_t^W,\ t \geq 0\}$, $z > 0$. It obviously holds $\{\mathcal{F}_z^W \leq t\} \in \mathcal{F}_t^W$. Since W has continuous paths (a.s.), $\{\mathcal{F}_t^W,\ t \geq 0\}$ is right-continuous.

Conclusion 5.6.1
Let $M_t = \max_{s \in [0,t]} W(s)$, $t \geq 0$. Then it follows for all $z > 0$, $y \geq 0$, that $\mathsf{P}(M_t \geq z, W(t) \leq z - y) = \mathsf{P}(W(t) > y + z)$.

Proof M_t be a random variable, since W has continuous paths. $T := T_z^W$. After theorem 5.6.1 it holds: for $Y(t) = W(T \wedge t) - (W(t) - W(T \wedge t))$, $t \geq 0$, $Y \stackrel{d}{=} W$ resp. $\{T_z^W, W\} \stackrel{d}{=} \{T_z^Y, Y\}$, since $W(t) = z$, $T_z^W = T_z^Y$. Therefore

$$\mathsf{P}(T \leq t, W(t) < z - y) = \mathsf{P}(T_z^Y \leq t, Y(t) < z - y)$$

$\{T_z^Y \leq t\} \cap \{Y(t) < z - y\} = \{T_z^Y \leq t\} \cap \{2z - W(t) < z - y\}$. If $T = T_z^Y \leq t$, then $Y(t) = W(T) - W(t) + W(T) = 2z - W(t)$ and hence follows that

$$\mathsf{P}(T \leq t, W(t) < z-y) = \mathsf{P}(T \leq t, 2z-W(t) < z-y) = \mathsf{P}(T \leq t, W(t) > z+y) = \mathsf{P}(W(t) > z+y).$$

Per definition in $T = T_z^W$ it holds:

$$\mathsf{P}(T \leq t, W(t) < z - y) = \mathsf{P}(M_t \geq z, W(t) < z - y) = \mathsf{P}(W(t) > y + z)$$

$\Rightarrow T_z^W \le t \Longleftrightarrow \max_{s\in[0,t]} W(s) \ge z.$ □

Theorem 5.6.2 (Distribution of the maximum of W):
For $t > 0$ and $x \ge 0$ it holds

$$\mathsf{P}(M_t > x) = \sqrt{\frac{2}{\pi t}} \int_x^\infty e^{-\frac{y^2}{2t}} dy$$

Proof In conclusion 5.6.1 set $y = 0 \Rightarrow \mathsf{P}(M_t \ge z, W(t) < z) = \mathsf{P}(W(t) > z)$. It holds $\mathsf{P}(W(t) > z) = \mathsf{P}(W(t) \ge z)$ for all t and all z, since $W(t) \sim \mathcal{N}(0,t)$, thus continuously distributed
$\Rightarrow \mathsf{P}(M_t \ge z, W(t) < z) + \mathsf{P}(W(t) \ge z) = \mathsf{P}(W(t) > z) + \mathsf{P}(W(t) > z)$
$\Rightarrow \mathsf{P}(M_t \ge z, W(t) < z) + \mathsf{P}(M_t \ge z, W(t) \ge z) = \mathsf{P}(M_t \ge z) = 2\mathsf{P}(W(t) > z)$
$\Rightarrow \mathsf{P}(M_t > z) = 2\mathsf{P}(W(t) > z) = 2\frac{1}{\sqrt{2\pi t}} \int_z^\infty e^{-\frac{y^2}{2z}} dy = \sqrt{\frac{2}{\pi t}} \int_z^\infty e^{-\frac{y^2}{2t}} dy$ □

Let $X(t) = W(t) - t\mu$, $t \ge 0$, $\mu > 0$, be the Wiener process with negative drift. Consider $\mathsf{P}(\sup_{t\ge 0} X(t) > x) = e^{-2\mu x}$, $x \ge 0$.

Motivation Calculation of the ruin-probability in risk theory.

Assumptions Initial capital $x \ge 0$. Let μ be the volume of premiums per time unit.$\Rightarrow$ μt – earned premiums at time $t \ge 0$. Let $W(t)$ be the loss process (price development). $\Rightarrow$ $Y(t) = x + t\mu - W(t)$ – remaining capital at time t. The ruin probability is $\mathsf{P}(\inf_{t\ge 0} Y(t) < 0) = \mathsf{P}(x - \sup_{t\ge 0} X(t) < 0) = \mathsf{P}(\sup_{t\ge 0} X(t) > x)$

Theorem 5.6.3
It holds

$$\mathsf{P}(\sup_{t\ge 0} X(t) > x) = e^{-2\mu x}, \quad x \ge 0,\ \mu > 0.$$

Proof Let $T = T_z^X = \inf\{t \ge 0 : X(t) = z\}$. It is known that $Y(t) = \exp\{uX(t) - t(\frac{u^2}{2} - \mu u)\}$, $t \ge 0$, $u \ge 0$, is a martingale. Let $T' = T \wedge t$ – a finite stopping time w.r.t. $\{\mathcal{F}_t^X,\ t \ge 0\}$. From conclusion 5.4.1: $\mathsf{E}Y(T') = \mathsf{E}Y(0) = \mathsf{E}e^0 = 1$

$$\Rightarrow \mathsf{E}(Y(T')1(T < t)) + \mathsf{E}(Y(T')1(T \ge t)) = \mathsf{E}(Y(T)1(T < t)) + \mathsf{E}(Y(T')1(t \ge t))$$

It is to be shown that $\mathsf{E}(Y(T')1(T \ge t)) \xrightarrow[t\to\infty]{} 0$.
From conclusion **??** it is known that

$$\frac{W(t)}{t} \xrightarrow[t\to\infty]{a.s.} 0 \Rightarrow \lim_{t\to\infty} \frac{X(t)}{t} = \lim_{t\to\infty} \frac{W(t)}{t} - \mu = -\mu \Rightarrow X(t) \xrightarrow[t\to+\infty]{a.s.} -\infty$$

$Y(T')1(T \ge t) = \exp\{uX(t) - t(\frac{u^2}{2} - \mu u)\}1(T \ge t) \xrightarrow[t\to+\infty]{a.s.} 0$, if $\frac{u^2}{2} - \mu u > 0 \Rightarrow u \ge 2\mu$. Otherwise, $Y(T')1(T \le t) \le \exp\{uz\} \Rightarrow$ after Lebesgue's theorem it holds:

$$\mathsf{E}(Y(T')1(T \ge t)) \xrightarrow[t\to+\infty]{} 0$$

$$
\begin{aligned}
&\Rightarrow \quad \lim_{t\to+\infty} \mathsf{E}(Y(T)1(T<t)) = 1,\ Y(T) = \exp\{uz - T(\frac{u^2}{2} - \mu u)\}\\
&\Rightarrow \quad \lim_{t\to+\infty} \mathsf{E}(\exp\{-T(\frac{u^2}{2} - \mu u)\}1(T<t)) = e^{-uz}\\
&\overset{u=2\mu}{\Rightarrow} \quad \lim_{t\to+\infty} \mathsf{P}(T<t) = \mathsf{P}(T<\infty) = e^{-2\mu z}\\
&\Rightarrow \quad \mathsf{P}(\sup_{t\geq 0} X(t) > z) = \mathsf{P}(T_z^X < \infty) = e^{-2\mu z}
\end{aligned}
$$

□

Theorem 5.6.4
Let $\mu \in \mathbb{R}$, $\delta > 0$, $T(t) = \inf\{s \geq 0 : W(s) + \mu s = \delta t\}$, $t \geq 0$. Then $T = \{T(t),\ t \geq 0\}$ is a Lèvy process with $\hat{m}_{T(t)}(z) = \mathsf{E}e^{-zT(t)} = \exp\{-t\delta(\sqrt{2z+\mu^2} - \mu)\}$, $t \geq 0$, $z \geq 0$.

Special case: For $\mu = 0$, $\delta = \frac{1}{\sqrt{2}}$, $T = \{T(t),\ t \geq 0\}$ is a $\frac{1}{2}$-stable subordinator, which is sometimes also called Lèvy-subordinator. Here holds $\hat{m}_{X(t)}(z) = e^{-t\sqrt{z}}$. (For α-stable subordinators holds: $\hat{m}_{T(t)}(z) = e^{-tz^\alpha}$, $\alpha \in (0,1)$)

To remind you: The Lèvy-measure of a α-stable subordinator is

$$\nu(dx) = \frac{\alpha}{\Gamma(1-\alpha)} \frac{dx}{x^{1+\alpha}} 1(x>0), \quad \alpha \in (0,1).$$

Proof of theorem 5.6.4 for the special case (general case analog)
Let $T(t) = \inf\{s \geq 0 : W(s) = \frac{t}{\sqrt{2}}\}$, $t \geq 0$. It is to be shown, that $T = \{T(t),\ t \geq 0\}$ is a Lèvy process.
$T(0) \overset{a.s.}{=} 0$. It follows from theorem 5.5.2, that T has independent and stationary increments. T is stochastically continuous, since

$$\lim_{t\to 0} \mathsf{P}(T(t) > \varepsilon) = \lim_{t\to\infty} \mathsf{P}(\max_{s\in[0,\varepsilon]} W(s) < \frac{t}{\sqrt{2}}) = \lim_{t\to 0}(1 - \sqrt{\frac{2}{\pi\varepsilon}} \int_{\frac{t}{\sqrt{2}}}^{\infty} e^{-\frac{y^2}{2\varepsilon}} dy) = 1 - 1 = 0.$$

Thus we have proven that T is a Lèvy process.
It now is to be shown, that $T(t)$ is α-stable for $\alpha = \frac{1}{2}$, i.e. $\mathsf{E}e^{-zT(t)} = e^{-t\sqrt{z}}$, for all z and $t \geq 0$. Similar to the proof of theorem 5.6.3 we are considering the martingale $X = \{X(s),\ s \geq 0\}$, $X(s) = \exp\{zW(s) - s\frac{z^2}{2}\}$, $s \geq 0$.
Let $Y_{n,t} = T(t) \wedge n$, for all $n \in \mathbb{N}$, $t \geq 0$, a sequence of stopping times w.r.t. $\{\mathcal{F}_t,\ t \geq 0\}$. From conclusion 5.4.1

$\{X(Y_{n,t}),\ n \in \mathbb{N}\}$ for all t, $z > 0$, is also a martingale.

$$
\begin{aligned}
&\mathsf{E}X(Y_{n,t}) = \mathsf{E}X(Y_{0,t}) = \mathsf{E}X(0) = e^{z0} = 1\\
&\mathsf{E}(X(Y_{n,z})1(T(t)<n)) + \mathsf{E}(X(Y_{n,t}))1(T(t)\geq n)\\
&= \mathsf{E}(\exp\{z\underbrace{W(T(t))}_{=\frac{t}{\sqrt{2}}} - T(t)\tfrac{z^2}{2}\}1(T(t)<n)) + \mathsf{E}(\exp\{zW(n) - n\tfrac{z^2}{2}1(T(t)\geq n))
\end{aligned}
$$

It is to be shown, that $\mathsf{E}(\exp\{zW(n) - n\frac{z^2}{2}\}1(T(t)\geq n)) \xrightarrow[n\to\infty]{} 0$. It will follow from that, that $1 = \lim_{n\to\infty} \mathsf{E}(\exp\{z\frac{t}{\sqrt{2}} - T(t)\frac{z^2}{2}\}\underbrace{1(T(t)<n)}_{\xrightarrow[n\to\infty]{a.s.} 1}) = \mathsf{E}\exp\{z\frac{t}{\sqrt{2}} - T(t)\frac{z^2}{2}\}$, since $T(t)$ is a finite

stopping time, i.e. $\mathsf{P}(T(t) < \infty) = 1$ for all $t \geq 0$.
The convergence above holds after Lebesgue's theorem over majorised convergence

$$\Rightarrow \mathsf{E}\exp\{-T(t)\underbrace{\frac{z^2}{2}}_{=u}\} - e^{-t\frac{z}{\sqrt{2}}} \Rightarrow \mathsf{E}e^{-uT(t)} = e^{-t\sqrt{u}},\ u \geq 0.$$

It is yet to be shown, that $\mathsf{E}(\exp\{zW(n) - n\frac{z^2}{2}\}\underbrace{1(T(t) \geq n))}_{\xrightarrow[n\to\infty]{a.s.} 0} \xrightarrow[n\to\infty]{} 0$.

In addition holds: $T(t) \geq n \Rightarrow W(n) \leq \frac{t}{\sqrt{2}}$.
$\exp\{zW(n) - n\frac{z^2}{2}\}1(T(t) \geq n) \leq \exp\{t\frac{t}{\sqrt{2}}\}$ for all $n \in \mathbb{N}_0$.
$\Rightarrow$ Lebesgue's theorem gives the convergence. $\square$

Remark 5.6.1
If $T(t) = \min\{s \geq 0 : W(s) + \mu s = \delta t\}$, $\mu \in \mathbb{R}$, $\delta > 0$, $t \geq 0$, then the Laplace transform of $T(t)$, $\mathsf{E}e^{-zT(t)} = \exp\{-t\delta(\sqrt{2z + \mu^2} - \mu)\}$ can be explicitly inverted into (compare theorem 5.6.4): the density of $T(t)$ can be written as

$$f_{T(t)}(x) = \frac{\delta t}{\sqrt{2\pi}} e^{\delta t\mu} x^{-\frac{3}{2}} \exp\{-\frac{1}{2}(t^2\delta^2\frac{1}{x} + \mu^2 x)\}1(x \geq 0).$$

That is the density of the so called *inverse Gauss distribution*.

Theorem 5.6.5
Let $X = \{X(t),\ t \geq 0\}$ be a Lèvy process and let $T = \{T(t),\ t \geq 0\}$ be a subordinator, which are both defined on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Let X and T be independent. Then $Y = \{Y(t),\ t \geq 0\}$ is definde by $Y(t) = X(T(t))$, $t \geq 0$, which is also a Lèvy process.

Without proof

5.7 Additional Exercises

Exercise 5.7.1
Let $X, Y : \Omega \to \mathbb{R}$ be arbitrary random variables on $(\Omega, \mathcal{F}, \mathsf{P})$ with

$$\mathsf{E}|X| < \infty, \quad \mathsf{E}|Y| < \infty, \quad \mathsf{E}|XY| < \infty,$$

and let $\mathcal{G} \subset \mathcal{F}$ be an arbitrary sub-σ-Algebra of $\mathcal{F}$. Then it holds

(a) $\mathsf{E}(X|\{\emptyset, \Omega\}) = \mathsf{E}X, \mathsf{E}(X|\mathcal{F}) = X$,

(b) $\mathsf{E}(aX + bY|\mathcal{G}) = a\mathsf{E}(X|\mathcal{G}) + b\mathsf{E}(Y|\mathcal{G})$ for arbitrary $a, b \in \mathbb{R}$,

(c) $\mathsf{E}(X|\mathcal{G}) \leq \mathsf{E}(Y|\mathcal{G})$, if $X \leq Y$,

(d) $\mathsf{E}(XY|\mathcal{G}) = Y\mathsf{E}(X|\mathcal{G})$, if Y is a $(\mathcal{G}, \mathcal{B}(\mathbb{R}))$-measurable random variable,

(e) $\mathsf{E}(\mathsf{E}(X|\mathcal{G}_2)|\mathcal{G}_1) = \mathsf{E}(X|\mathcal{G}_1)$, if $\mathcal{G}_1$ and $\mathcal{G}_2$ are sub-σ-algebras of $\mathcal{F}$ with $\mathcal{G}_1 \subset \mathcal{G}_2$,

(f) $\mathsf{E}(X|\mathcal{G}) = \mathsf{E}X$, if the σ-algebra $\mathcal{G}$ and $\sigma(X) = X^{-1}(\mathcal{B}(\mathbb{R}))$ are independent, i.e., if $\mathsf{P}(A \cap A') = \mathsf{P}(A)\mathsf{P}(A')$ for arbitrary $A \in \mathcal{G}$ and $A' \in \sigma(X)$.

(g) $\mathsf{E}(f(X)|\mathcal{G}) \geq f(\mathsf{E}(X|\mathcal{G}))$, if $f : \mathbb{R} \to \mathbb{R}$ is a convex function, such that $\mathsf{E}|f(X)| < \infty$.

Exercise 5.7.2
Look at the two random variables X and Y on the probability space $([-1,1], \mathcal{B}([-1,1]), \frac{1}{2}\nu)$ with $\mathsf{E}|X| < \infty$, where ν is the Lebesgue measure of $[-1,1]$. Determine $\sigma(Y)$ and a version of the conditional expectation $\mathsf{E}(X|Y)$ for the following random variables.

(a) $Y(\omega) = \omega^5$ (Hint: Show first that $\sigma(Y) = \mathcal{B}([-1,1])$)

(b) $Y(\omega) = (-1)^k$ for $\omega \in \left[\frac{k-3}{2}, \frac{k-2}{2}\right)$, $k = 1, \dots, 4$ and $Y(1) = 1$
(Hint: It holds $\mathsf{E}(X|B) = \frac{\mathsf{E}(X 1_B)}{\mathsf{P}(B)}$ for $B \in \sigma(Y)$ with $\mathsf{P}(B) > 0$)

(c) Calculate the distribution of $\mathsf{E}(X|Y)$ in (a) and (b), if $X \sim U[-1,1]$.

Exercise 5.7.3
Let X and Y be random variables on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. The conditional variance $\mathsf{var}(Y|X)$ is defined by

$$\mathsf{var}(Y|X) = \mathsf{E}((Y - \mathsf{E}(Y|X))^2|X).$$

Show that

$$\mathsf{var}\, Y = \mathsf{E}(\mathsf{var}(Y|X)) + \mathsf{var}(\mathsf{E}(Y|X)).$$

Exercise 5.7.4
For a stopping time τ define the stopped σ-algebra $\mathcal{F}_\tau$ as follows:

$$\mathcal{F}_\tau = \{B \in \mathcal{F} : B \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for arbitrary } t \geq 0\}.$$

Let now S and T be stopping times w.r.t. the filtration $\{\mathcal{F}_t, t \geq 0\}$. Show:

(a) $A \cap \{S \leq T\} \in \mathcal{F}_T \ \forall A \in \mathcal{F}_S$

(b) $\mathcal{F}_{\min\{S,T\}} = \mathcal{F}_S \cap \mathcal{F}_T$

Exercise 5.7.5 (a) Let $\{X(t), t \geq 0\}$ be a martingale. Show that $\mathsf{E}X(t) = \mathsf{E}X(0)$ holds for all $t \geq 0$.

(b) Let $\{X(t), t \geq 0\}$ be a sub- resp. supermartingale. Show that $\mathsf{E}X(t) \geq \mathsf{E}X(0)$ resp. $\mathsf{E}X(t) \leq \mathsf{E}X(0)$ holds for all $t \geq 0$.

Exercise 5.7.6
The stochastisc process $X = \{X(t), t \geq 0\}$ be adapted and càdlàg. Show that

$$\mathsf{P}(\sup_{0 \leq v \leq t} X(v) > x) \leq \frac{\mathsf{E}X(t)^2}{x^2 + \mathsf{E}X(t)^2}$$

holds for arbitrary $x > 0$ and $t \geq 0$, if X is a submartingale with $\mathsf{E}X(t) = 0$ and $\mathsf{E}X(t)^2 < \infty$.

Exercise 5.7.7 (a) Let $g : [0,\infty) \to [0,\infty)$ be a monotone increasing function with

$$\frac{g(x)}{x} \to \infty, \quad x \to \infty.$$

Show that the sequence $X_1, X_2, \dots$ of random variables is uniformly integrable, if $\sup_{n \in \mathbb{N}} \mathsf{E}g(|X_n|) < \infty$.

(b) Let $X = \{X(n), n \in \mathbb{N}\}$ be a martingale. Show that the sequence of random variables $X(T \wedge 1), X(T \wedge 2), \ldots$ is uniformly integrable for every finite stopping time T, if $\mathsf{E}|X(T)| < \infty$ and $\mathsf{E}(|X(n)| 1_{\{T>n\}}) \to 0$ for $n \to \infty$.

Exercise 5.7.8
Let $S = \{S_n = a + \sum_{i=1}^n X_i, n \in \mathbb{N}\}$ be a symmetric random walk with $a > 0$ and $\mathsf{P}(X_i = 1) = \mathsf{P}(X_i = -1) = 1/2$ for $i \in \mathbb{N}$. The random walk is stopped at the time T, when it exceeds or falls below one of the two values 0 and $K > a$ for the first time, i.e.

$$T = \min_{k \geq 0}\{S_k \leq 0 \text{ or } S_k \geq K\}.$$

Show that $M_n = \sum_{i=0}^n S_i - \frac{1}{3} S_n^3$ is a martingale and $\mathsf{E}(\sum_{i=0}^T S_i) = \frac{1}{3}(K^2 - a^2)a + a$ holds.

Hint: To calculate $\mathsf{E}(M_n | \mathcal{F}_m^M)$, $n > m$, you can use $\mathsf{E}(\sum_{i=k}^l X_i)^3 = 0$, $1 \leq k \leq l$, $M_n = \sum_{r=0}^m S_r + \sum_{r=m+1}^n S_r - \frac{1}{3} S_n^3$ and $S_n = S_n - S_m + S_m$.

A discrete martingale w.r.t. a filtration $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ is a sequence of random variables $\{X_n\}_{n \in \mathbb{N}}$ on a probability space $\{\Omega, \mathcal{F}, \mathsf{P}\}$, such that X_n is measurable w.r.t. $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ and $\mathsf{E}(X_{n+1} | X_n) = X_n$ for all $n \in \mathbb{N}$. A discrete stopping time w.r.t. $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ is a random variable $T : \Omega \to \mathbb{N} \cup \{\infty\}$, such that $\{T \leq n\} \in \mathcal{F}_n$ for all $n \in \mathbb{N} \cup \{\infty\}$, where $\mathcal{F}_\infty = \sigma\{\bigcup_{n=1}^\infty \mathcal{F}_n\}$.

Exercise 5.7.9
Let $\{X_n\}_{n \in \mathbb{N}}$ be a discrete martingale and T a discrete stopping time w.r.t. $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$. Show that$\{X_{\min\{T,n\}}\}_{n \in \mathbb{N}}$ is also a martingale w.r.t. $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$.

Exercise 5.7.10
Let $\{S_n\}_{n \in \mathbb{N}}$ be a symmetric random walk with $S_n = \sum_{i=1}^n X_i$ for a sequence of independent and identically distributed random variables $X_1, X_2, \ldots$, such that $\mathsf{P}(X_1 = 1) = \mathsf{P}(X_1 = -1) = \frac{1}{2}$. Let $T = \inf\{n : |S_n| > \sqrt{n}\}$ and $\mathcal{F}_n = \sigma\{X_1, \ldots, X_n\}$, $n \in \mathbb{N}$.

(a) Show that T is a stopping time w.r.t. $\{F_n\}_{n \in \mathbb{N}}$.

(b) Show that $\{G_n\}_{n \in \mathbb{N}}$ with $G_n = S^2_{\min\{T,n\}} - \min\{T, n\}$ is a martingale w.r.t. $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$. (Hint: Use exercise 5.7.9)

(c) Show that $|G_n| \leq 4T$ holds for all $n \in \mathbb{N}$.
(Hint: It holds $|G_n| \leq |S^2_{\min\{T,n\}}| + |\min\{T, n\}| \leq S^2_{\min\{T,n\}} + T$)

Exercise 5.7.11
Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables with $\mathsf{E}|X_1| < \infty$. Let $\mathcal{F}_n = \sigma\{X_1, \ldots, X_n\}$, $n \in \mathbb{N}$, and let T be a stopping time w.r.t. $\{F_n\}_{n \in \mathbb{N}}$ with $\mathsf{E}T < \infty$.

(a) Let T be independent of $X_1, X_2, \ldots$. Derive a formula for the characteristic function of $S_T = \sum_{i=1}^T X_i$ her verify the Wald's identity with it, i.e. $\mathsf{E}S_T = \mathsf{E}T \mathsf{E}X_1$.

(b) Let additionally $\mathsf{E}X_1 = 0$ and $T = \inf\{n : S_n < 0\}$. Use theorem 2.1.3 from the lecture, to show that $\mathsf{E}T = \infty$. (Hint: Proof by contradiction)

6 Stationary Sequences of Random Variables

6.1 Sequences of Independent Random Variables

It is known, that the series

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{1}{n^{\alpha}}<\infty &\iff \alpha>1,\\ \sum_{n=1}^{\infty}\frac{(-1)^n}{n^{\alpha}}<\infty &\iff \alpha>0,\end{aligned}$$

since the drift of neighboring terms have oder $\frac{1}{n^{1+\alpha}}$.

When (for which $\alpha > 0$) converges the series $\sum_{n=1}^{\infty}\frac{\delta_n}{n^{\alpha}}$, where δ_n are i.i.d. random variables with $\mathsf{E}\delta_n = 0$, e.g. $\mathsf{P}(\delta_n = \pm 1) = \frac{1}{2}$?

More general question: Under which conditions converges (a.s.) the series $\sum_{n=1}^{\infty} X_n$, where X_n are independent?

It is known, that for a sequencce of random variables $\{Y_n\}$ with $Y_n \xrightarrow[n\to\infty]{a.s.} Y$ it holds that $Y_n \xrightarrow[n\to\infty]{\mathsf{P}} Y$. The opposite is in general not true.

Theorem 6.1.1
Let X_n, $n \in \mathbb{N}$, be independent random variables. If $S_n = \sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{\mathsf{P}} S$, then $S_n \xrightarrow[n\to\infty]{a.s.} S$.

Without proof

Conclusion 6.1.1
If the sequences X_n, $n \in \mathbb{N}$, are independent, $\operatorname{var} X_n < \infty$, $n \in \mathbb{N}$, $\mathsf{E}X_n = 0$, $\sum_{n=1}^{\infty}\operatorname{var} X_n < \infty$, then $\sum_{n=1}^{\infty} X_n$ converges a.s.

Proof $S_n = \sum_{i=1}^{n} X_i$, $S = \sum_{i=1}^{\infty} X_i$, $m < n$,

$$\mathsf{E}(S_n - S_m)^2 = \|S_n - S_m\|_{L^2}^2 = \sum_{i=m+1}^{n}\operatorname{var} X_i \xrightarrow[n,m\to\infty]{} 0,$$

since $\sum_{i=1}^{\infty}\operatorname{var} X_i < \infty \Rightarrow \{S_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega, \mathcal{F}, \mathsf{P})$

$$\Rightarrow \exists S = \lim_{n\to\infty} S_n = \sum_{i=1}^{\infty} X_i \Rightarrow S_n \xrightarrow[n\to\infty]{\mathsf{P}} S \overset{\text{Theorem 6.1.1}}{\Rightarrow} S_n \xrightarrow[n\to\infty]{a.s.} S.$$

□

Conclusion 6.1.2
If $\sum_{n=1}^{\infty} a_n^2 < \infty$, where $\{a_n\}_{n\in\mathbb{N}}$ is a deterministic sequence, and $\{\delta_n\}$ is a sequence of i.i.d. random variables with $\mathsf{E}\delta_n = 0$, $\operatorname{var}\delta_n = \sigma^2 < \infty$, $n \in \mathbb{N}$, then the sequence $\sum_{n=1}^{\infty} a_n\delta_n$ converges a.s.

Exercise 6.1.1
Derive conclusion 6.1.2 from theorem 6.1.1.

For us: δ_n i.i.d., $\mathsf{E}\delta_n = 0$, $\operatorname{var}\delta_n = \sigma^2 > 0$ (e.g. $\delta_n \sim Bernouli(\frac{1}{2})$), $a_n = \frac{1}{n^\alpha}$, $n \in \mathbb{N}$. $\sum_{n=1}^{\infty} \frac{1}{n^\alpha} < \infty$, if $\sum_{n=1}^{\infty} \frac{1}{n^{2\alpha}} < \infty$, i.e. for $\alpha < \frac{1}{2}$.

Conclusion 6.1.3
Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of independent random variables with $\sum_{n=1}^{\infty} \mathsf{E}X_n$, $\sum_{n=1}^{\infty} \operatorname{var} X_n < \infty$ $\Rightarrow \sum_{n=1}^{\infty} X_n \overset{a.s.}{<} \infty$.

Proof Let $Y_n = X_n - \mathsf{E}X_n$, thus $X_n = \underbrace{\mathsf{E}X_n}_{=a_n} + Y_n$, $n \in \mathbb{N}$, and $\mathsf{E}Y_n = 0$, $\sum_{n=1}^{\infty} a_n < \infty$ after the condition. $\sum_{n=1}^{\infty} Y_n \overset{a.s.}{<} \infty$ after conclusion 6.1.1, since $\operatorname{var} X_n = \operatorname{var} Y_n$, $n \in \mathbb{N}$, $\sum_{n=1}^{\infty} \operatorname{var} X_n < \infty \Rightarrow \sum_n X_n = \sum_n a_n + \sum_n Y_n \overset{a.s.}{<} \infty$. □

6.2 Stationarity in the Narrow Sense and Ergodic Theory

6.2.1 Basic Ideas

Let $\{X_n\}_{n\in\mathbb{N}}$ be stationary in the narrow sense sequence of random variables, i.e. for all $n, k \in \mathbb{N}$ the distribution of $(X_n, \ldots, X_{n+k})^\top$ is independent of $n \in \mathbb{N}$. In particular, this means that all X_n are identically distributed. In the language of Kolmogorov's theorem:

$$\mathsf{P}((X_n, X_{n+1}, \ldots) \in B) = \mathsf{P}((X_1, X_2, \ldots) \in B),$$

for all $n \in \mathbb{N}$, for all $B \in \mathcal{B}(\mathbb{R}^\infty)$, $R^\infty = \mathbb{R} \times \mathbb{R} \times \ldots \times \ldots$.

Example 6.2.1 (of stationary sequences of random variables):

1. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables, then $\{X_n\}_{n\in\mathbb{N}}$ is stationary.

2. Let $Y_n = a_0 X_n + \ldots + a_k X_{n+k}$, k – fixed number out of $\mathbb{N}$, $\{X_n\}_{n\in\mathbb{N}}$ from 1), $a_0, \ldots, a_k \in \mathbb{R}$ (fixed), $n \in \mathbb{N}$. Y_n are not stationary anymore, bud identically distributed. The sequence $\{Y_n\}_{n\in\mathbb{N}}$ is stationary.

3. Let $Y_n = \sum_{j=0}^{\infty} a_j X_{n+j}$ for arbitrary $n \in \mathbb{N}$. The sequence $\{a_j\}_{j\in\mathbb{N}}$ is a sequence of numbers from $\mathbb{R}$ with the property that $\sum_{j=1}^{\infty} |a_j| < \infty$ and $\mathsf{E}X_n = 0$, $\sum_{n=1}^{\infty} \operatorname{var} X_n < \infty$, $\sum_{j=1}^{\infty} a_j^2 < \infty$ (compare conclusion 6.1.2).
It is obvious that $\{Y_n\}_{n\in\mathbb{N}}$ is a stationary sequence. (This construction is important for autoregressive time series (AR processes), e.g. in econometrics).

4. Let $Y_n = g(X_n, X_{n+1}, \ldots)$, $n \in \mathbb{N}$, $g : \mathbb{R}^\infty \to \mathbb{R}$ measurable, $\{X_n\}_{n\in\mathbb{N}}$ from 1). Then $\{Y_n\}_{n\in\mathbb{N}}$ is stationary.

Remark 6.2.1

1. An arbitrary stationary sequence of random variables $X = \{X_n\}_{n\in\mathbb{N}}$ can be extended to a stationären sequence $\bar{X} = \{X_n\}_{n\in\mathbb{Z}}$. In fact, the finite dimensional distribution of $\bar{X}$ can be defined after the theorem of Kolmogorov as:

$$(X_n, \ldots, X_{n+k}) \overset{d}{=} (X_1, \ldots, X_{k+1}), \quad n \in \mathbb{Z},\ k \in \mathbb{N}.$$

Therefore (after Kolmogorov's theorem) there exists a probability space and a sequence $\{Y_n\}_{n\in\mathbb{Z}}$ with the above distribution. We set $\bar{X} = \{Y_n\}_{n\in\mathbb{Z}}$ and hence follows that $\{Y_n\}_{n\in\mathbb{N}} \overset{d}{=} \{X_n\}_{n\in\mathbb{N}}$.

2. We define a shift of coordinates. Let $x \in \mathbb{R}^{\infty}_{-\infty}$, $x = (x_k, k \in \mathbb{N})$, $x = (x_k, k \in \mathbb{Z})$. Define the mapping $\theta : \mathbb{R}^{\infty}_{-\infty} \to R^{\infty}_{-\infty}$, $(\theta x)_k = x_{k+1}$ (shift of the coordinates by 1), $k \in \mathbb{N}$, $k \in \mathbb{Z}$. If θ is considered on $\mathbb{R}^{\infty}_{-\infty}$, then it is bijective and the backwards mapping would be $(\theta^{-1}x)_k = x_{k-1}$, $k \in \mathbb{Z}$.
   Let now $X = \{X_n,\ n \in \mathbb{Z}\}$ be a stationary sequence of random variables. Let $\bar{X} = \theta X$. It is obvious that $\bar{X}$ is again stationary and $\bar{X} \overset{d}{=} X$. Hence follows that
   $$\mathsf{P}(\theta X \in B) = \mathsf{P}(X \in B), \quad B \in \mathcal{B}(\mathbb{R}^{\infty}_{-\infty}).$$
   θ is called a *measure preserving map*. There are also other maps which have a measure preserving effect.

Definition 6.2.1
Let $(\Omega, \mathcal{F}, \mathsf{P})$ be an arbitary probability space. A map $T : \Omega \to \Omega$ is called *measure preserving*, if

1. T is measurable, i.e. $T^{-1}A \in \mathcal{F}$ for all $A \in \mathcal{F}$,
2. $\mathsf{P}(T^{-1}A) = \mathsf{P}(A), \quad A \in \mathcal{F}$.

Lemma 6.2.1
Let T be a measure preserving mapping X_0 – a random variable. We define a sequence of random variables X_n. Let $UY(\omega) = Y(T(\omega))$, $\omega \in \Omega$, be the map for an arbitrary random variable Z to $(\Omega, \mathcal{F}, \mathsf{P})$. Define $X_n(\omega) = U^n X_0(\omega) = X_0(T^n(\omega))$, $\omega \in \Omega$, $n \in \mathbb{N}$. Then the sequence of random variables $X = \{X_0, X_1, X_2, \dots\}$ is stationary.

Proof Let $B \in \mathcal{B}(\mathbb{R}^{\infty})$, $A = \{\omega \in \Omega : X(\omega) \in B\}$, $A_1 = \{\omega \in \Omega : \theta X(\omega) \in B\}$.

$$X(\omega) = (X_0(\omega), X_0(T(\omega)), X_0(T^2(\omega)), \dots)$$
$$\theta X(\omega) = (X_0(T(\omega)), X_0(T^2(\omega)), \dots)$$

Therefore $\omega \in A_1 \Leftrightarrow T(\omega) \in A$. Since $\mathsf{P}(T^{-1}A) = \mathsf{P}(A)$, it holds $\mathsf{P}(A_1) = \mathsf{P}(A)$. For $A_n = \{\omega \in \Omega : \theta^n X(\omega) \in B\}$ the same holds, $\mathsf{P}(A_n) = \mathsf{P}(A)$, $n \in \mathbb{N}$ (Induction). And hence follows that the sequence X is stationary. □

The sequence X in lemma 6.2.1 is called the *sequence generated by T*.

Definition 6.2.2
A map $T : \Omega \to \Omega$ is called *measure preserving in both directions*, if

1. T is bijective and $T(\Omega) = \Omega$,
2. T and T^{-1} are measurable,
3. $\mathsf{P}(T^{-1}A) = \mathsf{P}(A)$, $A \in \mathcal{F}$, and therefore $\mathsf{P}(TA) = \mathsf{P}(A)$.

Thus, exactly as in lemma 6.2.1, we can construct stationary sequences of random variables with time parameter $n \in \mathbb{Z}$:

$$X(\omega) = \{X_0(T^n(\omega))\}_{n \in \mathbb{N}}, \quad \omega \in \Omega,$$

where T is a measure preserving map (in both directions), $X_0(T^0(\omega)) = X_0(\omega)$, $(T^0 = Id)$.

Lemma 6.2.2
For an arbitrary stationary sequence of random variables $X = (X_0, X_1, \ldots)$ there exists a measure preserving map T and a random variable Y_0, such that $Y(\omega) = \{Y_0(T^n(\omega))\}_{n\in\mathbb{N}}$ has the same distribution as X: $X \stackrel{d}{=} Y$. The same statement holds for sequences with time parameter $n \in \mathbb{Z}$.

Proof We are considering the canonical probability space $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty), \mathsf{P}_X)$, $Y(\omega) = \omega$, $\omega \in \mathbb{R}$, $T = \theta$. With that, Y is constructed, since $\mathsf{P}_X(A) = \mathsf{P}_Y(A) = \mathsf{P}_X(Y \in A)$, $A \in \mathcal{B}(\mathbb{R}^\infty)$. □

Example 6.2.2 (Measure preserving maps):
1. Let $\Omega = \{\omega_1, \ldots, \omega_k\}$, $k \geq 2$, $\mathcal{F} = 2^\Omega$, $\mathsf{P}(\omega_i) = \frac{1}{k}$, $i = 1, \ldots, k$, be a Laplace probability space. $T\omega_i = \omega_{i+1}$ for all $i = 1, \ldots, k-1$, $T\omega_k = \omega_1$.
2. Let $\Omega = [0, 1)$, $\mathcal{F} = \mathcal{B}([0, 1))$, $\mathsf{P} = \nu_1$ – Lebesgue-measure on $[0, 1)$. $T\omega = (\omega + s) \mod 1$, $s \geq 0$. T is measure preserving in both directions.

Sequences of random variables, which in these examples can be generated by the map T, are mostly determinisitc resp. zyclic. In example 1) we can consider a random variable $X_0 : \Omega \to \mathbb{R}$, such that $X(\omega_i) = x_i$ are all pairwise distinct. Therefore $X_n(\omega) = X_0(T^n(\omega))$ uniquely defines the value of $X_{n+1}(\omega) = X_0(T^{n+1}(\omega))$, for all $n \in \mathbb{N}$.

Remark 6.2.2
Measure preserving maps play a big role in physics. There, T is interpreted as the change of state of a physical system and the measure can e.g. be the volume. (Ex.: T – Change of temperature, measure P – volume of the gas.) Therefore the to be developed ergodic theory is transfered to some physical processes.

Theorem 6.2.1 (Poincarè):
If T is a measure preserving map on $(\Omega, \mathcal{F}, \mathsf{P})$, $A \in \mathcal{F}$, then for almost all $\omega \in A$ the relation $\{T^n(\omega) \in A\}$ holds for infinitely many $n \in \mathbb{N}$.

That means, the trajectory $\{T^n(\omega),\ n \in \mathbb{N}\}$ returns to A infinitely often, if $\omega \in \Omega$, $\mathsf{P}(A) > 0$.

Proof It is to be shown, that $A \in \mathcal{F}$, $T : \Omega \to \Omega$ is measure preserving. Show, that for almost all $\omega \in \Omega$, $T^(\omega) \in A$ for infinitely many $n \in \mathbb{N}$. Let $N = \{\omega \in A : T^n(\omega) \notin A \forall n \geq 1\}$. It is obvious that $N \in \mathcal{F}$, since $\{\omega \in \Omega : T^n(\omega) \notin A\} \in \mathcal{F}$ for all $n \geq 1$. $N \cap T^{-n}N = \emptyset$ for all $n \geq 1$. In fact, if $\omega \in N \cap T^{-n}N$, then $\omega \in A$, $T^n(\omega) \notin A$ for all $n \geq 1$, $\omega_1 = T^n(\omega)$, $\omega_1 \in N$. Hence follows, that $\omega_1 \in A$ and $T^n(\omega) \in A$. That a contradiction.
$T^{-n}N = \{\omega \in \Omega : T^n(\omega) \in N\}$. For arbitrary $m \in \mathbb{N}$ it holds

$$T^{-m}N \cap T^{-(n+m)}N = T^{-m}(N \cap T^{-n}N) = T^{-m}(\emptyset) = \emptyset.$$

Hence follows that the sets $T^{-n}N$, $n \in \mathbb{N}$, are pairwise disjoint, belong to $\mathcal{F}$ and $\mathsf{P}(T^{-n}N) = \mathsf{P}(A) = a \geq 0$ holds.

$$1 \geq \mathsf{P}(\cup_n T^{-n}N) = \sum \mathsf{P}(T^{-n}N) = \sum_{n=0}^{\infty} a \Rightarrow a = 0 \Rightarrow \mathsf{P}(N) = 0.$$

Hence follows that for almost all $\omega \in A$ $(\omega \in A\{\}N)$ there exists a $n_1 = n_1(\omega)$, such that $T^m(\omega) \in A$. Let now T^k be instead of T, $k \in \mathbb{N}$. It holds $\mathsf{P}(N_k) = 0$ and for all $\omega \in A\{\}N_k$ there exists $n_k = n_k(\omega)$, such that $(T^k)^{n_k}(\omega) \in A$. Since $kn_k \geq k$ it follows for almost all $\omega \in A$, that $T^n(\omega) \in A$ for infinitely many n. □

Conclusion 6.2.1
Let $X \geq 0$ be a random variable, $A = \{\omega \in \Omega : X(\omega) > 0\}$. Then it holds for almost all $\omega \in \Omega$ that $\sum_{n=0}^{\infty} X(T^(n)(\omega)) = +\infty$, where T is a measure preserving map.

Exercise 6.2.1
Proof it.

Remark 6.2.3
The proof of theorem 6.2.1 holds for the sets $A \in \mathcal{F} : \mathsf{P}(A) \geq 0$. If however $\mathsf{P}(A) = 0$, it is possible that $A\{\}N = \emptyset$ and thus the statement of the theoreme is trivial.
As an example we are considering $\Omega = [0, 1)$, $\mathcal{F} = \mathcal{B}_([0, 1))$, $\mathsf{P} = \nu_1$ – Lebesgue-measure, $T(\omega) = \omega + s \mod 1$, $s \in \mathbb{Q}$. As set a A we are considering $A = \omega_0$, $\omega_0 \in \Omega$. Then $T^n(\omega_0) \neq \omega_0$ holds for all n, because otherwise there exists $k, m \in \mathbb{N}$, such that $\omega_0 + ks - m = \omega_0$ and hence follows $s = \frac{m}{k} \in \mathbb{Q}$. Thus we get a contradiction.

6.2.2 Mixing Properties and Ergodicity

Here we study the dependency structure in a stationary sequence of random variables, which is generated by a measure preserving map T.

Let $X = \{X_n\}_{n\in\mathbb{N}}$ be a stationary sequence (in the narrow sense) of random variables. Then there exists a measure preserving map $T : \Omega \to \Omega$, such that $X_n(\omega) \stackrel{d}{=} X_0(T^(n)(\omega))$ and $X_n \stackrel{d}{=} X_0$, and thus X_s gives the marginal distribution of the sequence X. In return the map T responsiblbe for the dependency within X (it indicates the properties of multidimensional distributions). We will therefore now examine the depedency properties, which are generated by T.

Definition 6.2.3
1. Event $A \in \mathcal{F}$ is called invariant w.r.t. (a measure preserving map) $T : \Omega \to \Omega$, if $T^{-1}A = A$.
2. Event $A \in \mathcal{F}$ is called almost invariant w.r.t. T, if $\mathsf{P}(T^{-1}A \triangle A) = 0$. $\triangle$ is the symmetric difference.

Exercise 6.2.2
Show that the set of all (almost) invariant events T is a σ-algebra $J(J^*)$.

Lemma 6.2.3
Let $A \in J^*$. Then there exists $B \in J^*$, such that $\mathsf{P}(A \triangle B) = 0$

Proof Let $B = \limsup_{n\to\infty} T^n A = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} T^{-k}A$. It is to be shown, that $B \in J$, $\mathsf{P}(A \triangle B) = 0$. It is obvious, that $T^{-1}(B) = \limsup_{n\to\infty} T^{-(n+1)}A = B$ and hence follows that $B \in J$.
It is easy to see, that $A \triangle B \subset \cup_{k=0}^{\infty}(T^{-k}A \triangle T^{-(k+1)}A)$. Since $\mathsf{P}(T^{-k}A \triangle T^{-(k+1)}A) = 0$ for all $k \geq 1$ due to $A \in J^*$, it follows that $\mathsf{P}(A \triangle B) = 0$. □

Definition 6.2.4
1. The measure preserving map $T : \Omega \to \Omega$ is called ergodic, if for every $A \in J$
$$\mathsf{P}(A) = \begin{cases} 0 \\ 1 \end{cases}.$$
2. The stationary sequence of random variables $X = \{X_n\}_{n\in\mathbb{N}}$ is called ergodic, if the measure preserving map $T : \Omega \to \Omega$, which generates X, is ergodic.

Lemma 6.2.4
The measure preserving map T is ergodic if and only if the probability of arbitrary almost invariant sets

$$\mathsf{P}(A) = \left\{ \begin{array}{l} 0 \\ 1 \end{array} \right. \text{for all } A \in J^*.$$

Proof „ $\Leftarrow$ “
Obvious, since arbitrary invariant set are also alsmost invariant, i.e. $J \subset J^*$
„ $\Rightarrow$ “
T – ergodic. Let $A \in J^*$. It follows that there exists $B \in J$, such that $\mathsf{P}(A \triangle B) = 0$ nach Lemma 6.2.3. T – ergodic and hence follows

$$\mathsf{P}(B) = \left\{ \begin{array}{l} 0 \\ 1 \end{array} \right. \quad \text{and} \quad \mathsf{P}(A) = \left\{ \begin{array}{l} 0 \\ 1 \end{array} \right. .$$

□

Definition 6.2.5
A random variable $Y : \Omega \to \mathbb{R}$ is called (almost) invariant w.r.t. $T : \Omega \to \Omega$ (measure preserving map), if $Y(\omega) = Y(T(\omega))$ for (almost) all $\omega \in \Omega$.

Theorem 6.2.2
Let $T : \Omega \to \mathbb{R}$ be a measure preserving map. The following statements are equivalent:

1. T – ergodic
2. If Y is invariant w.r.t. T, then $Y = const$ a.s.
3. If Y is invarinat w.r.t. T, then $Y = const$ a.s.

Proof 1) $\Rightarrow$ 2) $\Rightarrow$ 3) $\Rightarrow$ 1)
1) $\Rightarrow$ 2)
T – ergodic, Y – almost invariant. It is to be shown, that $Y(\omega) = const$ for almost all $\omega \in \Omega$. $Y(T(\omega)) = Y(\omega)$ almost surely. Let $A_v = \{\omega \in \Omega : Y(\omega) \leq \omega\}$, $\omega \in \mathbb{R}$. Hence follows, that $A_v \in J^*$ for all $v \in \mathbb{R}$ and after lemma 6.2.4

$$\mathsf{P}(A_v) = \left\{ \begin{array}{l} 0 \\ 1 \end{array} \right. \quad \text{for all } v.$$

Let $c = \sup \{v : \mathsf{P}(A_v) = 0\}$. Show that $\mathsf{P}(Y = c) = 1$.
$A_v \uparrow \Omega$, $v \to \infty$, $A_v \downarrow \emptyset$, $v \to -\infty \Rightarrow |c| < \infty$.

$$\mathsf{P}(Y < c) = \mathsf{P}\left(\cup_{n=1}^{\infty} \left\{ Y \leq c - \frac{1}{n} \right\} \right) \leq \sum_{n=1}^{\infty} \mathsf{P}\left(A_{c - \frac{1}{n}} \right) = 0.$$

Just the same $\mathsf{P}(Y > c) = 0$ and $\mathsf{P}(Y = c) = 1$.
2) $\Rightarrow$ 3)
Obviously.
3) $\Rightarrow$ 1) It is to be shown, that T is ergodic, i.e. for all $A \in J$ $\mathsf{P}(A) = \left\{ \begin{array}{l} 0 \\ 1 \end{array} \right.$.

Let $Y = 1_A$ – invariant w.r.t. T, it hence follows that $1_A = const = \left\{ \begin{array}{l} 0 \\ 1 \end{array} \right.$ and $\mathsf{P}(A) = \left\{ \begin{array}{l} 0 \\ 1 \end{array} \right.$. □

Remark 6.2.4 1. The statements of theorem 6.2.2 stays true, if you demand 3) for a.s. bounded random variables Y.

2. If Y is invariant w.r.t. T, then $Y_n = \min\{Y, n\}$, $n \in \mathbb{N}$, is also invariant w.r.t. T.

Example 6.2.3 1. Let $\Omega = \{\omega_1, \ldots, \omega_d\}$, $\mathcal{F} = 2^\Omega$, $\mathsf{P}(\{\omega_i\}) = \frac{1}{d}$, $i = 1, \ldots, d$. Let $T(\omega_i) = \omega_{i+1} \mod d$, i.e. $\omega_d \xmapsto{T} \omega_1$. T is obviously ergodic and every other invariant random variable is constant.

2. Let $\Omega = [0,1)$, $\mathcal{F} = \mathcal{B}_{[0,1)}$, $\mathsf{P} = \nu_1$, $T(\omega) = (\omega + s) \mod 1$. Show that T is ergodic $\Longleftrightarrow$ $s \notin \mathbb{Q}$.

Proof „ $\Leftarrow$ “
Let $s \notin \mathbb{Q}$, Y – an arbitrary invariant random variable. Let $EY^2 < \infty$. We decompose the random variable Y into a Fourier-series. The Fourier series of Y is $Y(\omega) = \sum_{n=0}^{\infty} a_n e^{2\pi i n w}$. We want to show that $a_n = 0$, $n > 0$, and hence follows that $Y(\omega) = a_0$ a.s.. Then T is ergodic and after theorem 6.2.2.

$$a_n = < Y(\omega), e^{2\pi i n w} >_{L^2} = \mathsf{E}(Y(\omega)e^{-2\pi i n w}) = \mathsf{E}(Y(T(\omega))e^{-2\pi i n w})e^{-2\pi i n s} = e^{-2\pi i n s} a_n,$$

$s \notin \mathbb{Q} \Rightarrow a_n = 0$.
„ $\Rightarrow$ “
If $s = \frac{m}{n} \in \mathbb{Q}$, then T is not ergodic, i.e. there exists $A \in J$, such that $0 < P(A) < 1$. Let $A = \cup_{k=0}^{n-1} \left\{ \omega \in \Omega : \frac{2k}{2n} \le \omega < \frac{2k+1}{2n} \right\}$ and $\mathsf{P}(A) = \frac{1}{2}$. A is invariant, since $T(A) = \left(A + \frac{2m}{2n}\right) \mod 1 = A$. $\square$

Definition 6.2.6 1. The measure preserving map $T : \Omega \to \Omega$ is called *mixing*, if for all $A_1, A_2 \in \mathcal{F}$ it holds: $\mathsf{P}(A_1 \cap T^{-n} A_2) \xrightarrow[n\to\infty]{} \mathsf{P}(A_1)\mathsf{P}(A_2)$, i.e. by repeated application from T on A_2, A_1 and A_2 are getting asymptotically independent.

2. Let $X = \{X_n\}_{n \in \mathbb{N}_0}$ be a stationary sequence of random variables which are generated by a random variable X_0 and a measure preserving map T. X is called *weak dependent*, if the random variable X_k and X_{k+n} are getting asymptotically independent for $n \to \infty$, i.e. for all $B_1, B_2 \in \mathcal{B}_\mathbb{R}$

$$\mathsf{P}(X_k \in B_1, X_{k+n} \in B_2) \xrightarrow[n\to\infty]{} \mathsf{P}(X_0 \in B_1)\mathsf{P}(X_0 \in B_2).$$

Theorem 6.2.3
A stationary sequence of random variables $X = \{X_n\}_{n \in \mathbb{N}_0}$, generated by the measure preserving map T, is weak dependent on average, if and only if T is mixing on average.

Exercise 6.2.3
Proof the theorem.

Theorem 6.2.4
Let T be a measure preserving map. It is ergodic, if and only if it is mixing on average.

Proof „ $\Leftarrow$ “
It is to be shown, that if T is mixing on average, hence follows that T is ergodic, i.e. for all $A \in J$

it holds $\mathsf{P}(A) = \begin{cases} 0 \\ 1 \end{cases}$. $A_1 \in \mathcal{F}$, $A_2 = A = J$, $\frac{1}{n}\sum_{k=1}^{n} \mathsf{P}(A_1 \cap \underbrace{T^{-n}(A_2)}_{=A_2}) = \mathsf{P}(A_1 \cap A_2) \xrightarrow[n\to\infty]{}$ $\mathsf{P}(A_1)\mathsf{P}(A_2)$. $\mathsf{P}(A_1 \cap A_2) = \mathsf{P}(A_1)\mathsf{P}(A_2)$ for $A_1 = A$, $\mathsf{P}(A) = \mathsf{P}^2(A)$ and $\mathsf{P}(A) = \begin{cases} 0 \\ 1 \end{cases}$

„ $\Rightarrow$ “

Later. □

Now we give the motivation for the term „mixing mapping“.

Theorem 6.2.5

Let $A \in \mathcal{F}$, $\mathsf{P}(A) > 0$. The measure preserving map $T : \Omega \to \Omega$ is ergodic (i.e. mixing on average), if and only if

$$\mathsf{P}\left(\cup_{n=0}^{\infty} T^{-n}A\right) = 1.$$

I.e. the archetypes $T^{-n}A$, $n \in \mathbb{N}_0$, are covering almost the whole Ω.

Proof „ $\Leftarrow$ “

Let $B = \cup_{n=0}^{\infty} T^{-n}A$. Obviously, $T^{-1}B = \cup_{n=1}^{\infty} T^{-n}A \subset B$. Since T is measure preserving, i.e. $\mathsf{P}(T^{-1}B) = \mathsf{P}(B)$, it follows that $\mathsf{P}(T^{-1}B \triangle B) = 0$, $B \in J^*$ (B – almost invariant w.r.t. T) and $\mathsf{P}(B) = \begin{cases} 0 \\ 1 \end{cases}$. $\mathsf{P}(B) \geq \mathsf{P}(A) > 0 \Rightarrow \mathsf{P}(B) = 1$.

„ $\Rightarrow$ “

Let T be non-ergodic. It is to be shown, that $\mathsf{P}(B) < 1$.

If T is not ergodic, there exists $A \in J$, such that $0 < \mathsf{P}(A) < 1$. $B = \cup_{n=0}^{\infty} T^{-n}A = A$ and $\mathsf{P}(B) < 1$. □

Remark 6.2.5

So far, the fact that the random variables X are realvalued was never explicitly used. Therefore the above observations can be transfered without modifications to sequences of random elements with values in an arbitrary measurable space $\mathcal{M}$.

6.2.3 Ergodic Theorem

Let $X = \{X_n\}_{n=0}^{\infty}$ be a sequence of random variables on the probability space $(\Omega, \mathcal{F}, \mathsf{P})$. If X_n are i.i.d. , then

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \xrightarrow[n\to\infty]{a.s.} \mathsf{E}X_0, \quad \mathsf{E}|X_0| < \infty.$$

We want to prove a similar statement about stationary sequences.

Theorem 6.2.6 (Ergodic theorem, Birkkoff-Kchintchin):

Let $X = \{X_n\}_{n\in\mathbb{N}_0}$ be a stationary sequence of random variables, generated by the random variable X_0 and a measure preserving map $T : \Omega \to \Omega$. Let J be the σ-algebra of the invariant sets from T, i.e. $\mathsf{E}|X_0| < \infty$. Then

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \xrightarrow[n\to\infty]{a.s.} \mathsf{E}(X_0 \mid J).$$

If X is weak dependent on average (i.e. T – ergodic), then $\mathsf{E}(X_0 \mid J) = \mathsf{E}(X_0)$.

Lemma 6.2.5
Let $\{X_n\}$, T be as above. Let $S_n(\omega) = \sum_{k=0}^{n-1} X_0(T^k(\omega))$, $M_k(\omega) = \max\{0, S_1(\omega), \ldots, S_k(\omega)\}$. Under the condition of theorems 6.2.6 it holds

$$\mathsf{E}(X_0 1(M_n > 0)) \geq 0, \quad n \in \mathbb{N}.$$

Proof For all $k \leq n$ it holds $\underbrace{S_k(\omega)}_{S_k(T(\omega))} \leq \underbrace{M_n(\omega)}_{M_n(T(\omega))}$. We can add X_0 and get

$$X_0(\omega) + M_n(T(\omega)) \geq X_0(\omega) + S_k(T(\omega)) = S_{k+1}(\omega).$$

For $k = 0$ it holds $X_0(\omega) \geq S_1(\omega) - M_n(T(\omega))$. The same holds for $k = 0, \ldots, n-1$. Hence follows that $X_0(\omega) \geq \underbrace{\max\{S_1(\omega), \ldots, S_n(\omega)\}}_{=M_n(\omega)} - M_n(T(\omega))$. Since $M_n(\omega) > 0$, then $M_n = \max\{S_1, \ldots, S_n\}$. It follows that

$$\mathsf{E}(X_0 1(M_n > 0)) \geq \mathsf{E}((M_n - M_n(T))1(M_n > 0)) \geq \mathsf{E}(M_n - M_n(T\omega)) = 0.$$

□

Proof of the Ergordic theorem The statement $\mathsf{E}(X_0 \mid J) = \mathsf{E}(X_0)$ is trivial, since for ergodic T it holds $J = \{\emptyset, \Omega\}$. w.l.o.g. let $\mathsf{E}(X_0 \mid J) = 0$, otherwise consider $X_0 = \mathsf{E}(X \mid J)$. It has to be shown: $\lim_{n\to\infty} \frac{S_n}{n} \overset{a.s.}{=} 0$, $S_n = \sum_{k=0}^{n-1} X_k$. It is enough to show that

$$0 \leq \liminf_{n\to\infty} \frac{S_n}{n} \leq \limsup_{n\to\infty} \frac{S_n}{n} \leq 0.$$

We first of all show that $\overline{S} = \limsup_{n\to\infty} \frac{S_n}{n} \leq 0$. It is enough to show that $\mathsf{P}(\underbrace{\overline{S} > \varepsilon}_{A_\varepsilon}) = 0$ for all $\varepsilon > 0$. Let $X_0^* = (X_0 - \varepsilon)1_{A_\varepsilon}$, $S_k^* = \sum_{j=0}^{k-1} X_0^*(T^j(\omega))$, $M_k^* = \max\{0, S_1^*, \ldots, S_k^*\}$. From lemma 6.2.5 it follows $\mathsf{E}(X_0^* 1(M_n^* > 0)) \geq 0$ for all $n \geq 1$. But,

$$\{M_n^* > 0\} = \left\{\max_{1\leq k\leq n} S_k^* > 0\right\} \uparrow_{n\to\infty} \left\{\sup_{k\geq 1} S_k^* > 0\right\} = \left\{\sup_{k\ geq 1} \frac{S_k^*}{k} > 0\right\} = \left\{\sup_{k\geq 1} \frac{S_k}{k} > \varepsilon\right\} \cap A_\varepsilon = A_\varepsilon,$$

since $\left\{\sup_{k\geq 1} \frac{S_k}{k} > \varepsilon\right\} \supset \left\{\overline{S} > \varepsilon\right\} = A_\varepsilon$. After Lebesgue's theorem: $0 \leq \mathsf{E}(X_0^* 1(M_n^* > 0)) \xrightarrow[n\to\infty]{} \mathsf{E}(X_0^* 1_{A_\varepsilon})$, since $\mathsf{E}|X_0^*| \leq \mathsf{E}|X_0| + \varepsilon$. Hence $0 \leq \mathsf{E}(X_0^* 1_{A_\varepsilon}) = \mathsf{E}((X_0 - \varepsilon)1_{A_\varepsilon}) = \mathsf{E}(X_0 1_{A_\varepsilon}) - \varepsilon \mathsf{P}(A_\varepsilon) = \mathsf{E}(\mathsf{E}(X_0 1_{A_\varepsilon} \mid J)) - \varepsilon \mathsf{P}(A_\varepsilon) = \mathsf{E}(1_{A_\varepsilon} \underbrace{\mathsf{E}(X_0 \mid J)}_{=0}) - \varepsilon \mathsf{P}(A_\varepsilon) = -\varepsilon \mathsf{P}(A_\varepsilon)$ and hence follows $\mathsf{P}(A_\varepsilon) \leq 0$ and $\mathsf{P}(A_\varepsilon) = 0$ for all $\varepsilon > 0$.
In oder to show $0 \leq \liminf_{n\to\infty} \frac{S_n}{n} = \underline{S}$ it is enough to look at $-X_0$ instead of X_0, since $\limsup_{n\to\infty}(-\frac{S_n}{n}) = \liminf_{n\to\infty}(\frac{S_n}{n})$. Since $\mathsf{P}(-\underline{S} \leq 0) = 1$ it holds $\mathsf{P}(\underline{S} \geq 0) = 1$. □

Remark 6.2.6
The speciality about the Ergodic theorem 6.2.6, in comparision with the common law of large numbers lies in the fact that the limit $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k \xrightarrow{a.s.} \mathsf{E}(X_0 \mid J)$ is random.

Example 6.2.4
We are considering the probability space from example 6.2.3 a). $\Omega = \{\omega_1, \dots, \omega_d\}$, $d = 2l \in \mathbb{N}$. $T : \Omega \to \Omega$ be defined by

$$\begin{cases} T(\omega_i) &= \omega_{i+2} \quad , \quad i = 1, \dots, d-2, \\ T(\omega_{d-1}) &= \omega_1 \quad , \\ T(\omega_d) &= \omega_2 \quad . \end{cases}$$

Let $A_1 = \{\omega_1, \omega_3, \dots, \omega_{2l-1}\}$, $A_2 = \{\omega_2, \omega_4, \dots, \omega_{2l}\}$. Since $(\Omega, \mathcal{F}, \mathsf{P})$ is a Laplace probability space ($\mathsf{P}(\{\omega_i\}) = \frac{1}{d}$, for all i) it follows that $\mathsf{P}(A_i) = \frac{1}{2}$, $i = 1, 2$. On the other hand, $A_1, A_2 \in J$ w.r.t. T and therefore T is not ergodic. For an arbitrary random variable $X_0 : \Omega \to \mathbb{R}$ it holds

$$\frac{1}{n}\sum_{k=0}^{n-1}(T^n(\omega)) \xrightarrow[n\to\infty]{} \begin{cases} \frac{2}{d}\sum_{j=0}^{l-1} X_0(\omega_{2j+1}), & \text{with probability } \frac{1}{2}, \text{ if } \omega \in A_1, \\ \frac{2}{d}\sum_{j=1}^{l} X_0(\omega_{2j}) \quad , & \text{with probability } \frac{1}{2}, \text{ if } \omega \in A_2. \end{cases}$$

Proof of theorem 6.2.4 It has to be shown: If $T : \Omega \to \Omega$ is ergodic, then T is mixing on average, i.e. for all $A_1, A_2 \in \mathcal{F}$

$$\frac{1}{n}\sum_{k=0}^{n-1}\mathsf{P}(A_1 \cap T^{-k}A_2) \xrightarrow[n\to\infty]{} \mathsf{P}(A_1)\mathsf{P}(A_2).$$

Let $Y_n = \frac{1}{n}\sum_{k=0}^{n-1} 1(T^{-k}A_2) \xrightarrow[n\to\infty]{\text{Thoerem } 6.2.6} \mathsf{P}(A_2)$, since T is ergodic, thus also the sequence $\left\{1(T^{-k}A_2)\right\}_{k\in\mathbb{N}}$. After Lebesgue's theorem from $1(A_1)Y_n \xrightarrow[n\to\infty]{} 1(A_1)\mathsf{P}(A_2)$ it follows that

$$\mathsf{E}(1(A_1)Y_n) = \frac{1}{n}\sum_{k=0}^{n-1}\mathsf{P}(A_1 \cap T^{-k}A_2) \xrightarrow[n\to\infty]{} \mathsf{P}(A_1)\mathsf{P}(A_2).$$

□

Lemma 6.2.6
If $\{X_n\}_{n\in\mathbb{N}}$ is a uniformly integrable sequence of random variables and $p_{n,i} \geq 0$, such that $\sum_{i=1}^{n} p_{n,i} = 1$ for all $n \in \mathbb{N}$, then the sequence of random variables $Y_n = \sum_{i=1}^{n} p_{n,i}\,|X_i|$, $n \in \mathbb{N}$, uniformly integrable as well.

Without proof

Conclusion 6.2.2
Under the conditions of theorem 6.2.6 it holds

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \xrightarrow[n\to\infty]{L^2} \mathsf{E}(X_0 \mid J)$$

resp.

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \xrightarrow[n\to\infty]{L^2} \mathsf{E}(X_0)$$

in the ergodic case.

Proof If $\{X_n\}_{n\in\mathbb{N}_0}$ is stationary, it then holds $\sup_n \mathsf{E}(|X_n|1(|X_n| > \varepsilon)) = \mathsf{E}(|X_0|1(|X_0| > \varepsilon)) \xrightarrow[\varepsilon\to 0]{} 0$, since $\mathsf{E}|X_0| < \infty$. Let $S_n = \frac{1}{n}\sum_{k=0}^{n-1} X_k = \sum_{i=1}^{n} p_{n,i}X_{i-1}$, $p_{n,i} = \frac{1}{n}$, $\tilde{S}_n = \frac{1}{n}\sum_{k=0}^{n-1} X_k = \sum_{i=1}^{n} p_{n,i}|X_{i-1}|$. From lemma 6.2.6, $\left\{\tilde{S}_n\right\}_{n\in\mathbb{Z}}$ is also uniformly integrable and after Lemma 5.3.2 it follows from $S_k \xrightarrow[k\to\infty]{a.s.} 0$ that $\mathsf{E}|S_n| \leq \frac{1}{n}\sum_{k=0}^{n-1}\mathsf{E}|X_k| \xrightarrow[n\to\infty]{} 0$. □

6.3 Stationarity in the Wide Sense

Let $\{X_n\}_{n\in\mathbb{Z}}$ be a sequence of random variables, which is stationary in the wide sense: $\mathsf{E}|X_n|^2 < \infty$, $n \in \mathbb{N}$. $\mathsf{E}|X_n| = const$, $n \in \mathbb{N}$, $\mathsf{cov}(X_n, X_m) = C(n-m)$, $n, m \in \mathbb{Z}$.

6.3.1 Correlation Theory

Theorem 6.3.1 (Herglotz):
Let $C : \mathbb{Z} \to \mathbb{R}$ be a positive semi-definite function. Then there exists a finite measure μ on $(-\pi, \pi)$, such that

$$C(n) = \int_{-\pi}^{\pi} e^{inx} \mu(dx), \quad n \in \mathbb{Z}.$$

μ is called *spectral measure* of C.

Remark 6.3.1
Since covariance function of a stationary sequence is positive semi-definit, the the above representation holds for an arbitrary covariance function C.

Definition 6.3.1
A family $\{Q_\lambda,\ \lambda \in \Lambda\}$ of probability measures is called *weakly relatively compact*, if an arbitrary sequence of measures $\{Q_{\lambda n}\}_{n\in\mathbb{N}}$ has a subsequence $\{Q_{\lambda n_k}\}_{n\in\mathbb{N}}$, which converges weakly.

Definition 6.3.2
A family of probability measures $Q = \{Q_\lambda,\ \lambda \in \Lambda\}$ on $(\mathcal{S}, \mathcal{B})$, $\mathcal{B}$ – Borel σ-algebra on a metric space $\mathcal{S}$ is called *tight*, if for all $\varepsilon > 0$ there exists a compactum, such that $K_\varepsilon \in \mathcal{B}$ and $Q_\lambda(K_\varepsilon) > 1 - \varepsilon$ for all $\lambda \in \Lambda$.

Theorem 6.3.2 (Prokhorov):
If the family of probability measures $Q = \{Q_\lambda,\ \lambda \in \Lambda\}$ on the metric measurable space $(\mathcal{S}, \mathcal{B})$ is tight, then it is weakly relatively compact. If $\mathcal{S}$ is a Banach space, then every weakly relatively compact familiy $Q = \{Q_\lambda,\ \lambda \in \Lambda\}$ of measures is also tight.

Without proof

The theorem of Prokhorov is used to prove the weak convergence of a sequence of probability measures, by checking the tightness among other things. In particular, if $\mathcal{S}$ is compact, then every family of probability measures on $(\mathcal{S}, \mathcal{B})$ is tight, since $K_\varepsilon = \mathcal{S}$ for all $\varepsilon > 0$.

Proof of theorem 6.3.2 „ $\Leftarrow$ “
If $C(n) = \int_{-\pi}^{\pi} e^{inx}\mu(dx)$, $n \in \mathbb{Z}$, then for all $n \in \mathbb{N}$, for all $z_1, \ldots, z_n \in \mathbb{C}$ and $t_1, \ldots, t_n \in \mathbb{Z}$

$$\sum_{i,j=1}^{n} z_j \bar{z}_j C(t_i - t_j) = \int_{-\pi}^{\pi} \left| \sum_{i=1}^{n} z_i e^{i z_i x} \right|^2 \mu(dx) \geq 0.$$

Hence follows that C is positive semi-definit.
„ $\Rightarrow$ “
For all $N \geq 1$, $x \in [-\pi, \pi]$, define the function $g_N(x) = \frac{1}{2\pi N} \sum_{k,j=1}^{N} C(k-j) e^{-ikx} e^{ijx} \geq 0$, which is continuous in x, since C is positive semi-definit. It holds

$$g_N(x) = \frac{1}{2\pi} \sum_{|n|<N} \left(1 - \frac{|n|}{N}\right) C(n) e^{-inx},$$

sine there are $N - |n|$ pairs $(k, j) \in \{1, \dots, N\}^2$, such that $k - j = n$. Define the measure μ_N on $([-\pi, \pi], \mathcal{B}_{[-\pi,\pi]})$ by $\mu_N(B) = \int g_N(x)dx$, $B \in \mathcal{B}([-\pi, \pi])$.

$$\int_{-\pi}^{\pi} e^{inx} Q_N(dx) = \int_{-\pi}^{\pi} e^{inx} g_N(x)dx = \begin{cases} \left(1 - \frac{|n|}{N}\right) C(n), & |n| < N, \\ 0, & \text{otherwise}, \end{cases}$$

since $\{e^{inx}\}_{n\in\mathbb{Z}}$ is a orthogonale system in $L^2[-\pi, \pi]$. For $n = 0$ it holds $Q_N([-\pi, \pi]) = C(0) < \infty$, hence $\left\{\frac{Q_N}{C(0)}\right\}_{n\in\mathbb{N}}$ is a family of probability measures, which is tight. After theorem 6.3.2 there exists a subsequence $\{N_k\}_{k\in\mathbb{N}}$,

such that $Q_N \xrightarrow[k\to\infty]{\omega} \mu$. μ – finite measure on $[-\pi, \pi]$ and hence follows

$$\lim_{k\to\infty} \int_{-\pi}^{\pi} e^{inx} g_N(x)dx = \lim_{k\to\infty} \left(1 - \frac{|n|}{N_k}\right) C(n) = C(n), \quad \text{for all } n \in \mathbb{Z}.$$

□

Let $X = \{X_n\}_{n\in\mathbb{Z}}$ be a stationary in the wide sense sequence of random variables. Then the following spectral representation holds:

$$X_n \stackrel{d}{=} \int_{-\pi}^{\pi} e^{inx} Z(dx), \quad n \in \mathbb{Z},$$

where Z is an orthogonal random measure on $([-\pi, \pi], \mathcal{B}([-\pi, \pi]))$. Therefore both Z and $I(f) = \int_{-\pi}^{\pi} f(x)Z(dx)$ are to be introduced for deterministic functions $f : [-\pi, \pi] \to \mathbb{C}$.

6.3.2 Orthogonal Random Measures

Construction scheme of Z resp. $I(\cdot)$:

1. Z is defined on a semiring $\mathcal{K}$ (the sumbset of Λ).

2. Z is defined on the algebra A, which is generated by $\mathcal{K}$.

3. Define the integral$I$ w.r.t. Z for a simple function on $\sigma(A)$, if the measure $\mu(\Lambda) < \infty$, μ – given measure.

4. Define I as $\lim_{n\to\infty} I(f_n)$ for arbitrary random functions f, $f = \lim_{n\to\infty} f_n$, f_n simple, $\mu(\Lambda) < \infty$.

5. Define I on a σ-finite space $\Lambda = \cup_n \Lambda_n$, $\mu(\Lambda_n) < \infty$, $\Lambda_n \cap \Lambda_m = \emptyset$, $n \neq m$, as $I(f) = \sum_n I(f \mid \Lambda_n)$, I_n – integralw.r.t. Z on Λ_n. Hence Z is extended on $\{A \in \sigma(A) : \mu(A) < \infty\}$ as $Z(A) = I(1(A))$.

Step 1

Let $\mathcal{K}$ be a semiring of the subsets of Λ (Λ – arbitrary space), i.e. for all $A, B \in \mathcal{K}$ it holds $A \cap B \in \mathcal{K}$; if $A \subset B$, then there exist $A_1, \dots, A_n \in \mathcal{K}$, $A_i \cap A_j = \emptyset$, $i \neq j$, such that $B = A \cup \cup_{i=1}^n A_i$.

Definition 6.3.3 1. A complexvalued random measure $Z = \{Z(B),\ B \in \mathcal{K}\}$, given on the probability space $(\Omega, \mathcal{F}, \mathsf{P})$, is called orthogonal, if

a) all $Z(B) \in L^2(\Omega, \mathcal{F}, \mathsf{P})$, $B \in \mathcal{K}$,

b) $A, B \in \mathcal{K}$, $A \cap B = \emptyset \Rightarrow \langle Z(A), Z(B) \rangle_{L^2(\Omega,\mathcal{F},\mathsf{P})} = \mathsf{E}(Z(A), \overline{Z(B)}) = 0$,

c) as a random measure the σ-additivity of Z holds: If $B, B_1, \ldots, B_n, \ldots \in \mathcal{K}$, $B = \cup_n B_n$, $B_i \cap B_j = \emptyset$, $i \neq j$, $Z(B) \overset{a.s.}{=} \sum_n Z(B_n)$, where the convergence of this series is intepreted in $L^2(\Omega, \mathcal{F}, \mathsf{P})$ terms.

2. The term $\mu = \{\mu(B),\ B \in \mathcal{K}\}$ defined by $\mu(B) = \mathsf{E}|Z(B)|^2 = \langle Z(B), Z(B) \rangle_{L^2(\Omega,\mathcal{F},\mathsf{P})}$, $B \in \mathcal{K}$, is called stucture measure of Z. It is easy to see that μ is in fact a measure on $\mathcal{K}$. If $\Lambda \in \mathcal{K}$, then μ is finite, otherwise σ-finite, $\Lambda = \cup_n \Lambda_n$, $\Lambda_n \in \mathcal{K}$, $\Lambda_n \cap \Lambda_m = \emptyset$, such that $\mu(\Lambda_n) < \infty$.

3. The orthogonal random measure Z is called *centered*, if $\mathsf{E}Z(B) = 0$, $B \in \mathcal{K}$.

Example 6.3.1
Let $\Lambda = [0, \infty)$, $\mathcal{K} = \{[a, b),\ 0 \leq a < b < \infty\}$, $Z([a, b)) = W(b) - W(a)$, $0 \leq a < b < \infty$, where $W = \{W(t),\ t \geq 0\}$ is the Wiener process. Z is an orthogonal random measure on $\mathcal{K}$, since W has independent increments. Analog, this definition can be transfered to an arbitrary quadratic integrable stochastic process X with independent increments instead of W.

Step 2

Theorem 6.3.3
Let μ be a σ-finite measure on the algebra A, which is generated by $\mathcal{K}$ (after the theorem of Caratheodon μ is uniquely continued on $\sigma(A)$). Then there exists a probability space $(\Omega, \mathcal{F}, \mathsf{P})$ and a centered orthogonal random measure Z on $(\Omega, \mathcal{F}, \mathsf{P})$, defined on $\{B \in A : \mu(B) < \infty\}$, with structure measure (or control measure) μ.

Without proof
To the definition of Z on A: for $B \in A$, $B = \cup_{i=1}^n B_i$, $B_i \in \mathcal{K}$, $B_i \cap B_j = \emptyset$, $i \neq j$, we set $Z(B) = \sum_{i=1}^n Z(B_i)$.

6.3.3 Integral regarding an Orthogonal Random Measure

Step 3
Let $f : \Lambda \to \mathbb{C}$ be a simple function, i.e. $f(x) = \sum_{i=1}^n c_i 1(x \in B_i)$, for $c_i \in \mathbb{C}$ and $B_i \in \mathcal{E}$, $i = 1, \ldots, n$, such that $\cup_{i=1}^n B_i = \Lambda$, $B_i \cap B_j = \emptyset$, $i \neq j$, and $(\Lambda, \mathcal{E}, \mu)$ be a measurable space with $\mu(\Lambda) < \infty$.

Definition 6.3.4
The integral of f w.r.t. an orthogonal random measure Z defined on $(\Omega, \mathcal{F}, \mathsf{P})$ is given by $I(f) := \int_\Lambda f(x) Z(dx) = \sum_{i=1}^n c_i Z(B_i)$.

Exercise 6.3.1
Show that the definition is correct, i.e. $I(f)$ does not depend on the representation of f as a simple function.

Lemma 6.3.1 (Properties of I):
Let $I(\cdot)$ be the integral w.r.t. the orthogonal random measure, defined on a simple function $\Lambda \to \mathbb{C}$ as abovev. The following properties hold:

1. *Isometry*: $\langle I(t), I(g) \rangle_{L^2(\Omega)} = \langle f, g \rangle_{L^2(\Omega)}$, where f and g are simple functions $\Lambda \to \mathbb{C}$, $\langle f, g \rangle_{L^2(\Omega)} = \int_\Lambda f(x) \overline{g(x)} \Lambda(dx)$.

2. *Linearity*: For every simple function $f, g : \Lambda \to \mathbb{C}$ holds $I(f+g) \overset{a.s.}{=} I(f) + I(g)$.

Exercise 6.3.2
Proof it.

Step 4
Let now $f \in L^2(\Omega, \mathcal{E}, \mu)$. Then there exists a sequence of simple functions $f_n : \Lambda \to \mathbb{C}$, such that $f_n \xrightarrow[n\to\infty]{L^2(\Lambda)} f$ (simple functions are tight in $L^2(\Lambda)$). Then define $I(f) = \lim_{n\to\infty} I(f_n)$, whereas this limit is to be understood in the $L^2(\Omega, \mathcal{F}, \mathsf{P})$ sense. You can show, that the definition of $I(f)$ is independent of the choice of the sequence $\{f_n\}$.

Lemma 6.3.2
The statements of lemma 6.3.1 hold for the general case.

Proof Use the continuity $\langle \cdot, \cdot \rangle$. □

Remark 6.3.2
If Z is centered, then $\mathsf{E}I(f) = 0$ holds for arbitrary functions $f \in L^2(\Lambda, \mathcal{E}, \mu)$.

Step 5
Let now Λ be σ-finite, i.e. $\Lambda = \cup_n \Lambda_n$, $\mu(\Lambda_n) < \infty$, $\Lambda_n \cap \Lambda_m = \emptyset$, $n \neq m$. Then for all $f \in L^2(\Lambda, \mathcal{E}, \mu)$ holds $f = \sum_n f|_{\Lambda_n}$. On $L^2(\Lambda_n, \mathcal{E} \cap \Lambda_n, \mu)$ the integral I_n w.r.t. Z is defined as in 1)- 4). Now set $I(f) := \sum_n I_n(f|_{\Lambda_n})$.

Theorem 6.3.4
The map $g : L^2(\Lambda, \mathcal{E}, \mu) \to L^2(\Omega, \mathcal{F}, \mathsf{P})$ is an isometry. In particular, as a result, the random measure Z on $\{B \in \varepsilon : \mu(B) < \mathcal{E}\}$ can be continued as $Z(B) := I(1_B)$, $B \in \mathcal{E} : \mu(B) < \infty$.

6.3.4 Spectral Representation

Let $X = \{X(t),\ t \in T\}$ be an arbitrary complexvalued stochastic process on $(\Omega, \mathcal{F}, \mathsf{P})$, T – an arbitrary index set, $\mathsf{E}|X(t)|^2 < \infty$, $t \in T$, $\mathsf{E}X(t) = 0$, $t \in T$ (w.l.o.g., otherwise consider $\tilde{X}(t) = X(t) - \mathsf{E}X(t)$), $t \in T$, with $C(s,t) = \mathsf{E}(X(s), \overline{X(t)})$, $s, t \in T$).

Theorem 6.3.5 (Karhunen):
X has the spectral representation $X(t) = \int_\Lambda f(t,x) Z(dx)$, $t \in T$ (i.e., there exists a centered orthogonal random measure on $\{B \in \mathcal{E} : \mu(B) < \infty\}$, where $L^2(\Lambda, \mathcal{E}, \mu)$ is an as above defined space), if and only if there exists a system of the functions$f(t, \cdot) \in L^2(\Lambda, \mathcal{E}, \mu)$, $t \in T$, such that $C(s,t) = \int_\Lambda f(s,x) f(t,x) \mu(dx)$, $s, t \in T$, and this system F is completely in $L^2(\Lambda, \mathcal{E}, \mu)$ (i.e. $\langle f(t, \cdot), \psi \rangle_{L^2(\Omega)} = 0$, $\psi \in L^2(\Omega, \mathcal{E}, \mu)$, for all $t \in T$ and $\psi \equiv 0$, μ almost everywhere).

Without proof

Theorem 6.3.6
Let $\{X_n,\ n \in \mathbb{Z}\}$ be a centered complexvalued stationary in the wide sense sequence of random variables on $(\Omega, \mathcal{F}, \mathsf{P})$. Then there exists an orthogonal centered random measure on $([-\pi, \pi], \mathcal{B}([-\pi, \pi]))$ (defined on $(\Omega, \mathcal{F}, \mathsf{P})$), such that $X_n \overset{a.s.}{=} \int_{-\pi}^{\pi} e^{inx} Z(dx)$, $n \in \mathbb{Z}$.

Proof Let $F = \{e^{inx},\ x \in [-\pi, \pi],\ n \in \mathbb{Z}\}$. This system in complete on $L^2([-\pi, \pi])$ (comp. the theory of the Fourier-series). From the theorem of Herglotz follows that

$$C(n,m) = \mathsf{E}(X_n \overline{X}_m) = \int_{-\pi}^{\pi} e^{inx} e^{imx} \mu(dx),$$

where μ is the spectral measure of X, thus a finite measure on $([-\pi,\pi],\mathcal{B}([-\pi,\pi]))$. After theorem 6.3.5 there exists an orthogonal random measure on $(\Omega,\mathcal{F},\mathsf{P})$, such that $X_n \overset{a.s.}{=} \int_{-\pi}^{\pi} e^{inx} Z(dx)$, $n\in\mathbb{Z}$. □

Theorem 6.3.7 (Ergodic theorem for stationary (in the wide sense) sequences of random variables):
Unter the conditions of theorem 6.3.6 it holds

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \xrightarrow{L^2(\Omega)} Z(\{0\}).$$

In particular if X is not centered, i.e. $\mathsf{E}X_n = a$, $n\in\mathbb{Z}$, then $\frac{1}{n}\sum_{k=0}^{n-1} X_k \xrightarrow{L^2(\Omega)} a$ converges, if $\underbrace{\mathsf{E}|Z(\{0\})|^2}_{\mu(\{0\})} = 0$, thus Z and therefore μ has no atom in zero.

Proof $S_n = \frac{1}{n}\sum_{k=0}^{n-1} X_k = \int \underbrace{\frac{1}{n}\sum_{k=0}^{n-1} e^{ikx}}_{\psi_n(x)} Z(dx)$. $\psi_n(x) = \begin{cases} \frac{1}{n}\frac{1-e^{inx}}{1-e^{ix}}, & x\neq 0\\ 1, & x=0\end{cases}$, for all $n\in\mathbb{N}$. $S_n - Z(\{0\}) = \int_{-\pi}^{\pi} \underbrace{(\psi_n(x) - 1(x=0))}_{\varphi_n(x)} Z(dx) = \int_{-\pi}^{\pi}\varphi_n(x)Z(dx)$. $\|S_n - Z(\{0\})\|^2_{L^2(\Omega)} = \|\varphi_n(x)\|^2_{L^2([-\pi,\pi],\mu)} = \int_{-\pi}^{\pi}|\varphi_n(x)|^2\mu(dx) \xrightarrow[n\to\infty]{} 0$ after the theorem of Lebesgue, since $|\varphi_n(x)| \leq \frac{2}{n|1-e^{ix}|} \xrightarrow[n\to\infty]{} 0$ for all $x\in[-\pi,\pi]$. □

6.4 Additional Exercises

Exercise 6.4.1
Let $Z_1, Z_2, \ldots$ be a sequence of random variables, such that the series $\sum_{i=1}^{\infty} Z_i$ converges almost surely. Let $a_1, a_2, \ldots$ be a monotone increasing sequence of positive (deterministic) numbers with $a_n\to\infty$, $n\to\infty$. Show that

$$\frac{1}{a_n}\sum_{k=1}^{n} a_k Z_k \overset{a.s.}{\to} 0,\quad n\to\infty.$$

Exercise 6.4.2
Let X be a non-negative variable on a probability space $(\Omega,\mathcal{F},\mathsf{P})$ an $T:\Omega\to\Omega$ a measure preserving map. Show that

$$\sum_{k=0}^{\infty} X(T^k(\omega)) = \infty \quad a.s.$$

for almost all $\omega\in\Omega$ with $X(w) > 0$.

Exercise 6.4.3
Let X be a random variable on a probability space $(\Omega,\mathcal{F},\mathsf{P})$ and $T:\Omega\to\Omega$ a measure preserving map. Show that $\mathsf{E}X = \mathsf{E}(X\circ T)$, i.e.

$$\int_{\Omega} X(T(\omega))\mathsf{P}(d\omega) = \int_{\Omega} X(\omega)\mathsf{P}(d\omega).$$

(Hint: algebraic induction)

Exercise 6.4.4
Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space, where $\Omega = [0, 1)$, $\mathcal{F} = \mathcal{B}([0, 1))$ and P is the Lebesgue measure. Let $\lambda \in (0, 1)$.

(a) Show that $T(x) = (x + \lambda) \bmod 1$ is a measure preserving map, where $a \bmod m = a - \lfloor \frac{a}{b} \rfloor m$ for $a \in \mathbb{R}$ and $b \in \mathbb{Z}$ and $\lfloor\rfloor$ is the Gauss bracket.

(b) Show that $T(x) = \lambda x$ and $T(x) = x^2$ are no measure preserving maps.

Bibliography

[1] D. Applebaum. *Lévy processes and stochastic calculus.* Cambridge University Press, Cambridge, 2004.

[2] A. A. Borovkov. *Wahrscheinlichkeitstheorie. Eine Einfürung.* Birkhäuser, Basel, 1976.

[3] W. Feller. *An introduction to probability theory and its applications. Vol. I/II.* John Wiley & Sons Inc., New York, 1970/71.

[4] P. Gänssler and W. Stute. *Wahrscheinlichkeitstheorie.* Springer, Berlin, 1977.

[5] G. Grimmett and D. Stirzaker. *Probability and random processes.* Oxford University Press, New York, 2001.

[6] C. Hesse. *Angewandte Wahrscheinlichkeitstheorie.* Vieweg, Braunschweig, 2003.

[7] O. Kallenberg. *Foundations of modern probability.* Springer, New York, 2002.

[8] J. F. C. Kingman. *Poisson processes.* Oxford University Press, New York, 1993.

[9] N. Krylov. *Introduction to the theory of random processes.* American Mathematical Society, Providence, RI, 2002.

[10] S. Resnick. *Adventures in stochastic processes.* Birkhäuser, Boston, MA, 1992.

[11] G. Samorodnitsky and M. Taqqu. *Stable non-Gaussian random processes.* Chapman & Hall, New York, 1994.

[12] K.-I. Sato. *Lévy Processes and Infinite Divisibility.* Cambridge University Press, Cambridge, 1999.

[13] A. N. Shiryaev. *Probability.* Springer, New York, 1996.

[14] A. V. Skorokhod. *Basic principles and applications of probability theory.* Springer, Berlin, 2005.

[15] J. Stoyanov. *Counterexamples in probability.* John Wiley & Sons Ltd., Chichester, 1987.

[16] A. D. Wentzel. *Theorie zufälliger Prozesse.* Birkhäuser, Basel, 1979.