# Contents


**1 Conditional Expectation** **2**

1.1 Weak convergence of measures . . . . . . . . . . . . . . . . . . 3

**2 Discrete time Martingales** **4**

**3 Stochastic processes** **9**

3.1 Relevant Exercises . . . . . . . . . . . . . . . . . . . . . . . . 11

**4 Markov Processes** **12**

4.1 Stationary measures . . . . . . . . . . . . . . . . . . . . . . . 15

**5 Ergodicity** **18**

**6 Continuous time Martingales** **20**

**7 Brownian Motion** **22**

7.1 Construction of Brownian Motion . . . . . . . . . . . . . . . . . 24

7.2 Sample path properties of a Brownian Motion . . . . . . . . . 28

7.3 Martingales derived from Brownian Motion . . . . . . . . . . 32

# 1 Conditional Expectation

We start with some preliminaries, things known or less known.

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

**Definition 1.1.** *We define the Lebesgue spaces*
$L^p(\Omega, \mathcal{F}, \mathbb{P}) = L^p = \{X : \Omega \to \mathbb{R} \mid X$ *is* $\mathcal{F}$ *measurable and* $\mathbb{E}(|X|^p) < \infty\}$ *where* $1 \leq p < \infty$.
$L^\infty = \{X : \Omega \to \mathbb{R} \mid X$ *is* $\mathcal{F}$ *measurable and* $\exists C > 0 : \mathbb{P}(|X| \leq C) = 1\}$ *is the space of almost surely bounded random variables.*

**Theorem 1.1** (Hölder's inequality)**.** *Let* $X \in L^P, Y \in L^q$ *for conjugate* $p, q \geq 1$ *(i.e.* $1/p + 1/q = 1$*) then we have* $XY \in L^1$ *moreover Hölder's inequality holds true*

$$\mathbb{E}(|XY|) \leq \mathbb{E}(|X|^p)^{\frac{1}{p}} \mathbb{E}(|X|^q)^{\frac{1}{q}}$$

**Remark 1.1.**

1. Hölder's inequality is the reason why probability theory behaves "nicer" in consideration of Lebesgue spaces, i.e. let $0 < r < s$ and set $p = \frac{s}{r}$ then for the conjugate $q = \frac{p}{p-1}$ we can apply Hölder's inequality to $|X|^r$ and the constant 1 function.

$$\mathbb{E}(|X|^r) \leq \mathbb{E}(|X|^{rp})^{1/p} = \mathbb{E}(|X|^s)^{r/s} \implies \mathbb{E}(|X|^r)^{\frac{1}{r}} \leq \mathbb{E}(|X|^s)^{\frac{1}{s}}$$

   in particular if $X$ is a random variable such that $X \in L^s$ then it always follows that $X \in L^r$ for any $r < s$. In general measure theory, where the spaces don't need to be of finite measure, this is not the case at all.

2. It is an important Theorem of Riesz that states that for all $1 \leq p \leq \infty$ the $L^p$ spaces are complete, in particular they are Banach spaces. Especially for $p = 2 = q$ the space $L^2$ is even a Hilbert space.

**Proposition 1.1** (Jensen's inequality)**.** *Let* $X \in L^1$ *and* $\varphi : \mathbb{R} \to \mathbb{R}$ *a convex function, then we have*

$$\varphi(\mathbb{E}(X)) \leq \mathbb{E}(\varphi(X))$$

**Proposition 1.2** (Existence of the conditional expection)**.** *Let* $\mathcal{G} \subset \mathcal{F}$ *be a subsigma-algebra. We define* $Z = \mathbb{E}(Z \mid \mathcal{G})$ *to be the unique* $\mathcal{G}$*-measurable variable such that*

$$\mathbb{E}(Z1_A) = \mathbb{E}(X1_A) \text{ for all } A \in \mathcal{G}$$

## 1.1 Weak convergence of measures

Weak convergence is one of many types of convergence relating to the convergence of measures. It depends on a topology on the underlying space and thus is not a purely measure theoretic notion.

There are several equivalent definitions of weak convergence of a sequence of measures, some of which are more general than others. The equivalence of these conditions is sometimes known as the Portmanteau theorem. It is an Exercise of Exercise Sheet 4 to prove this theorem.

In the following let $\mathcal{X}$ be a metric space and let it be endowed with its Borel $\sigma$-Algebra $\Sigma$. Let $\mathcal{M}(\mathcal{X})$ denote the space of all measures on $\mathcal{X}$.

**Definition 1.2.** *A sequence of probability measures $(\mu_n)_{n\in\mathbb{N}}$ converges weakly to $\mu$ in $\mathcal{M}(\mathcal{X})$ if for all $f \in C_b(\mathcal{X} \to \mathbb{R})$ we have*

$$\int f d\mu_n \to \int f d\mu, \text{ as } n \to \infty. \tag{*}$$

*We then simply denote this convergence by $\mu_n \implies \mu$.*

**Remark 1.2.** With this mode of convergence, we increasingly expect to see the next outcome in a sequence of random experiments becoming better and better modeled by a given probability distribution. Moreover, convergence in distribution or weak convergence is, as the name suggests, the weakest form of convergence.

**Exercise 4.5. (Portmanteau's theorem)** The following assertions are equivalent:

1. $\mu_n \implies \mu$.
2. (*) holds for all $f \in UC_b(\mathcal{X} \to \mathbb{R})$.
3. $\liminf \mu_n(O) \geq \mu(O)$ for all open sets $O \subset \mathcal{X}$.
4. $\limsup \mu_n(A) \leq \mu(A)$ for all closed sets $A \subset \mathcal{X}$.
5. $\lim \mu_n(B) = \mu(B)$ for all measurable sets $B \subset \mathcal{X}$ such that $\mu(B^\circ) = \mu(\overline{B})$.

**Theorem 1.2** (Lévy's continuity theorem)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables and let $\varphi_n(t) := \mathbb{E}(e^{itX_n})$ for $t \in \mathbb{R}$ denote the characteristic function. We then have $X_n \implies X$ if and only if $\varphi_n(t) \to \varphi(t)$.*

**Proposition 1.3.** *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables, then $X_n \implies X$ if and only if $\mathbb{P}(X_n \leq x) \to \mathbb{P}(X_n \leq x)$ for all cotninuity points of $x \mapsto \mathbb{P}(X \leq x)$.*

## 2 Discrete time Martingales

**Definition 2.1.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, an increasing sequence of $\sigma$-algebras $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}$ is called a filtration on $\Omega$. We call the space $(\Omega, \mathcal{F}, (\mathcal{F})_{n\in\mathbb{N}}, \mathbb{P})$ a filtered probability space.*

**Definition 2.2.** *A stochastic process $X = (X_n)_{n\in\mathbb{N}}$ is called adapted if $X_n$ is $\mathcal{F}_n$-measurable for all $n \in \mathbb{N}$. Moreoever, an adapted process is called a martingale if $X_n \in L^1$ for all $n \in \mathbb{N}$ and*

$$\mathbb{E}(X_n \mid \mathcal{F}_m) = X_{n\wedge m} \text{ for all } n, m \geq 0$$

**Remark 2.1.** Analogeously we define sub-martingales if instead of an equality we have $\geq$ above, or a super-martingale for $\leq$ respectively.

**Definition 2.3.** *A stopping time $T$ is a random variable $T : \Omega \to \mathbb{N}_0^\infty$ such that $\{T \leq n\} \in \mathcal{F}_n$ for all $n \in \mathbb{N}$. The sigma-algebra at the stopping time $T$ is then defined as*

$$\mathcal{F}_T := \{A \in \mathcal{F} : A \cap \{T \leq n\} \in \mathcal{F}_n \text{ for all } n \in \mathbb{N}\}$$

**Remark 2.2.** The sigma algebra at the stopping time $T$ $\mathcal{F}_T$ encodes the information up to the random time $T$. In other words, if we interprete our filtered probability space as an random experiment, then the maximum information that can be found until the random time $T$ is in $\mathcal{F}_T$.

**Theorem 2.1** (Optional stopping theorem - weak version)**.** *Let $X$ be a martingale. Let $X^T := (X_n^T)_{n\in\mathbb{N}}$ be given by $X_n^T := X_{n\wedge T}$ then $X^T$ is also a martingale. Moreover if $S, T$ are almost surely finite stopping times with $S \leq T$ almost surely, then*

$$\mathbb{E}(X_T \mid \mathcal{F}_S) = X_S$$

**Theorem 2.2** (Martingale convergence Theorem)**.** *Let $X$ be a super-martingale that is bounded in $L^1$ (i.e. $\sup_{n\in\mathbb{N}} \mathbb{E}(|X_n|) < \infty$). Then there exists a random variable $X_\infty \in L^1$ such that $X_n \to X_\infty$ almost surely, moreover*

$$\mathbb{E}(|X_\infty|) \leq \sup_{n\in\mathbb{N}} \mathbb{E}(|X_n|)$$

**Corollary 2.1.** *Let $X$ be a super-martingale that is bounded from below (i.e. $X_n \geq c$ a.s. for some $c \in \mathbb{R}$ for all $n \in \mathbb{N}$). Then $X_n$ converges almost surely to $X_\infty \in L^1$.*

**Theorem 2.3** (Behaviour of Martingales with bounded increments)**.** *Let $X = (X_n)_{n\in\mathbb{N}}$ be a martingale with bounded increments (i.e. $|X_{n+1} - X_n| \leq C$). Let $\mathcal{C} := \{\limsup X_n = \liminf X_n \in \mathbb{R}\}$ , $\mathcal{O} := \{\limsup X_n = +\infty\} \cap \{\liminf X_n = -\infty\}$, then $\mathbb{P}(\mathcal{O} \cup \mathcal{C}) = 1$.*

**Lemma 2.1** (Doob's inequalities)**.** *Let $X$ be a non-negative submartingale and $X_n^* := \max_{k\leq n} X_k$ for $n \in \mathbb{N}$, then we have for any $\lambda > 0$*

$$\lambda\mathbb{P}(X_n^* > \lambda) \leq \mathbb{E}(X_n 1_{X_n^*>\lambda}) \leq \mathbb{E}(X_n)$$

*In addition, for $p > 1$, we have*

$$\|X_n^*\|_p \leq \frac{p}{p-1}\|X_n\|_p$$

**Theorem 2.4** (Closed martingale convergence theorem)**.** *Let $X$ be a martingale and $p > 1$. Then the following statements are equivalent:*

1. $\sup_{n\in\mathbb{N}} \|X_n\|_p < \infty$
2. *$X_n$ converges almost surely and in $L^p$ to $X_\infty \in L^p$*
3. *There exists a random varaible $X_\infty \in L^p$ such that*

$$\mathbb{E}(X_\infty \mid \mathcal{F}_n) = X_n \textit{ for all } n \in \mathbb{N}$$

*If any (and consequently all) of these conditions hold true, we say that $X$ is a closed martingale in $L^p$.*

**Definition 2.4.** *A family $\mathcal{H}$ of random variables is called uniformly integrable (UI) if*

$$\lim_{\lambda\to\infty} \sup_{X\in\mathcal{H}} \mathbb{E}(|X|1_{|X|>\lambda}) = 0$$

**Exercise:** Prove that $\mathcal{H}$ is UI if there exists $G : [0,\infty) \to [0,\infty)$ non-decreasing such that

1. $\lim_{t\to\infty} \frac{G(t)}{t} = \infty$
2. $\sup_{X\in\mathcal{H}} \mathbb{E}(G(X))\infty$

**Theorem 2.5** (UI convergence Theorem)**.** *Let $(X_n)_{n\in\mathbb{N}}$ a stochastic process, then the following statements are equivalent:*

1. *$(X_n)_{n\in\mathbb{N}}$ is UI and $X_n \to X_\infty$ in probability.*
2. *$X_n \to X_\infty$ in $L^1$.*

**Theorem 2.6** (Characterisation of UI martingales through closedness)**.** *Let $X$ be a martingale. The following assertions are equivalent:*

1. *$X$ is UI.*
2. *$X$ converges almost surely and in $L^1$ to $X_\infty \in L^1$.*
3. *There exists $X_\infty \in L^1$ such that $\mathbb{E}(X_\infty \mid \mathcal{F}_n) = X_n$, i.e. the martingale is closed.*

**Theorem 2.7** (Optional stopping - strong version)**.** *Let $X$ be a closed (and thus UI) martingale. Then for any stopping times $S$ and $T$, one has*

$$\mathbb{E}(X_T \mid \mathcal{F}_S) = X_{T\wedge S}$$

**Definition 2.5.** *Let $\mathcal{F}$ be a sigma-algebra and consider a sequence $\cdots \subset \mathcal{H}_2 \subset \mathcal{H}_1 \subset \mathcal{F}$ of sigma-algebras (i.e. we lose information over time). Let $Y_1 \in L^p(\Omega, \mathcal{H}_1, \mathbb{P})$ for some $p \in [1, \infty)$. Then $Y = (Y_n)_{n\in\mathbb{N}}$ is called a backward-martingale if $Y_n = \mathbb{E}(Y_1 \mid \mathcal{H}_n)$ for all $n \in \mathbb{N}$.*

**Theorem 2.8** (Convergence theorem for backward-martingales)**.**
*Let $Y = (Y_n)_{n\in\mathbb{N}}$ be a backward-martingale. Then $Y$ converges almost surely and in $L^p$ to $Y_\infty = \mathbb{E}(Y_1 \mid \mathcal{H}_\infty)$ where $\mathcal{H}_\infty = \cap_{n\in\mathbb{N}}\mathcal{H}_n$.*

**Theorem 2.9** (Kolmogorov's Law of Large Numbers)**.** *Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables with $\xi_1 \in L^1$. Let $S_n = \xi_1 + \cdots + \xi_n$ be the simple random walk, then we have*

$$\frac{S_n}{n} \to \mathbb{E}(X_1) \textit{ almost surely and in } L^1$$

*Proof.* **Exercise 2.4.** Kolmogorov's LLN can be shown in it's full generality (i.e. for $L^1$ RV) by using backwards martingales and Kolomogorov's 0-1 Law. □

**Theorem 2.10** (Kolmogorov's 0-1 Law)**.** *Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables and let $\mathcal{H}_n = \sigma(\xi_j : j \geq n)$. Then, the sigma algebra $\mathcal{H}_\infty := \cup_{n\in\mathbb{N}}\mathcal{H}_n$ is trivial, i.e. for every event $A \in \mathcal{H}_\infty$ we have $\mathbb{P}(A) \in \{0, 1\}$.*

Next we will remind the very important Borel-Cantelli Lemmas. We recall the following definiton. Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of events in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ then we define

$$A = \limsup_{n\to\infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{i\geq n}^{\infty} A_i = \{A_n \text{ infinitely often}\}$$
$$A^c := \liminf_{n\to\infty} A_n := \bigcup_{n=1}^{\infty} \bigcap_{i\geq n}^{\infty} A_i = \{A_n \text{ eventually}\}$$

We remark that the following is true

$$\omega \in \bigcap_{n=1}^{\infty} \bigcup_{i\geq n}^{\infty} A_i \iff \forall n \in \mathbb{N}, \exists i \geq n : \omega \in A_i$$
$$\omega \in \bigcup_{n=1}^{\infty} \bigcap_{i\geq n}^{\infty} A_i^c \iff \exists n \in \mathbb{N}, \forall i \geq n : \omega \in A_i$$

**Theorem 2.11** (Borel-Cantelli Lemma 1)**.** *If $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < \infty$, then we have $\mathbb{P}(A_n \text{ i.o.}\} = 0$.*

**Theorem 2.12** (Borel-Cantelli Lemma 2)**.** *If the events $A_n$ are independent and $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = \infty$, then we have $\mathbb{P}(A_n \text{ i.o.}\} = 1$*

**Theorem 2.13** (Extension of Borel Cantelli Lemma). *Let $\mathcal{F} = (\mathcal{F}_n)_{n=1}^{\infty}$ be a Filtration and suppose $A_n \in \mathcal{F}_n$, then we have*

$$\{A_n \ i.o.\} = \left\{ \sum_{n \geq k} \mathbb{P}(A_n \mid \mathcal{F}_{n-1}) = \infty \right\}$$

*for all $k \in \mathbb{N}$ up to a set of measure zero.*

We now give some results that are related to the Central Limit Theorem (CLT). We will first recall the classical result.

**Theorem 2.14** (CLT). *Let $X_1, X_2, \ldots$ be a sequence of i.i.d. $L^2$ random variables with $\mathbb{E}(X) = \mu$ and $Var(X) = \sigma > 0$. Let $S_n = X_1 + \cdots + X_n$, then we have*

$$\frac{S_n - n\mu}{n\sigma} \implies \mathcal{N}(0,1)$$

It is our goal to generalize this result, i.e. weaken the conditions and then give a central limit theorem for martingales.

The outline is the following: Let $\{X_{n,k} : 1 \leq k \leq J(n), n \in \mathbb{N}_{\geq 1}\}$ be an array of centered random variables. Here $K$ can be random, however it is almost surely finite with $J(n) \to \infty$ as $n \to \infty$.

Let $(\xi_j)_{j \in \mathbb{N}}$ be a sequence of (centered?) i.i.d. random variables, we then set

$$X_{n,k} = \frac{\xi_k}{\sqrt{n}}$$

Suppose that $X_{n,k} \in L^2(\mathcal{F})$ and let

$$S_n := \sum_{k=1}^{J(n)} X_{n,k}$$

**Theorem 2.15** (Mc Leish). *In the situation as above, let us define $T_n = \prod_{k=1}^{J(n)}(1 + itX_{n,k})$ and suppose that*

1. *$T_n$ is UI and $\mathbb{E}(T_n) \to 1$ as $n \to \infty$.*
2. *$\sum_{k=1}^{J(n)} X_{n,k}^2 \to 1$ in Probability.*
3. *$\max_{k \leq n} |X_{n,k}| \to 0$ in Probability.*

*Then $S_n \implies \mathcal{N}(0,1)$ as $n \to \infty$.*

**Theorem 2.16** (Martingale Central Limit Theorem)**.** *Let $X_{n,k}$ where $1 \leq k \leq m_n$ ($m_n$ is deterministic with $m_n \to \infty$ as $n \to \infty$) be a $\mathcal{F}_{n,k}$-measurable martingale difference array. Suppose that*

1. $\mathbb{E}(X_{n,k} \mid \mathcal{F}_{n,k-1}) = 0$
2. $V_n := \sum_{j=1}^{m_n} X_{n,j}^2 \to 1$ *in Probability.*
3. *(Lindenberg Condition) For any $\epsilon > 0$ we have that*

$$\lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{E}(X_{n,k}^2 1_{|X_{n,k}|>\epsilon}) = 0$$

*Then we have as $n \to \infty$*

$$S_n = \sum_{k=1}^{m_n} X_{n,k} \implies \mathcal{N}(0,1)$$

# 3 Stochastic processes

Let $(H_i)_{i \in I}$ be a sequence of complete, separable metric spaces, where $I$ is an arbitrary index set.

**Definition 3.1** (Product space)**.** *We define the product topology for* $H = \prod_{i \in I} H_i$ *to be the coarsest (that is the weakest/smallest) topology such that all the maps* $\pi_i : H \to H_i$ *are continuous.*

**Remark 3.1.**

1. This construction is equivalent to the construction of the weak topology.
2. This guarantees that the Borel $\sigma$-Algebra $\mathcal{B}(H)$ coincides with the product $\sigma$-algebra

$$\sigma\left(\left\{\prod_{i \in I} A_i : A_i \in \mathcal{B}(H_i) \text{ for } i \in K, A_i = H_i \text{ for } i \notin K, K \subset F \text{ finite}\right\}\right)$$

that is, the product $\sigma$-algebra on $\mathbb{R}^{\mathbb{N}}$ is generated by

$$\bigcup_{n=1}^{\infty} \{A_1 \times A_2 \times \cdots \times A_n \times \mathbb{R} \times \mathbb{R} \times \cdots \mid A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R})\}$$

**Definition 3.2** (Product measure)**.** *Let* $\mu_i$ *be a probability measure on* $H_i$*, we define the product measure* $\mu = \bigotimes_{i \in I} \mu_i$ *on* $H$ *to be the unique measure such that*

$$\mu\left(\prod_{i \in I} A_i\right) = \int 1_{\prod_{i \in I} A_i} d\mu = \prod_{i \in I} \mu_i(A_i)$$

*This implies that*

$$\int \prod_{j \in K} f_j(\pi_j(\omega))\mu(d\omega) = \prod_{j \in K} \int f_j d\mu_j, \text{ for all } f_j \in L^1(H_j, \mathcal{B}(H_j), \mu_j)$$

*under the assumption that* $A_i \in \mathcal{B}(H_i)$ *for* $i \in K$ *and* $A_i = H_i$ *for* $i \notin K$*.*

It is by no means clear that such a unique product measure exists at all, this is however part of the celebrated Kolmogorov Extension theorem which we will state next.

**Theorem 3.1** (Kolmogorov's Extension Theorem)**.** *Let $K \subset I$ be a finite subset and let $\mu_K$ be a probability measure on $\prod_{i\in K} H_i$. Suppose that for any $K \subset J$ where $J$ is another finite subset of $I$ the following consistency condition holds:*

$$\mu_J\left(\prod_{i\in J} A_i\right) = \mu_K\left(\prod_{i\in K} B_i\right), \text{ with } A_i = \begin{cases} B_i \in \mathcal{B}(H_i), & \text{if } i \in K \\ H_i, & \text{if } i \notin K \end{cases}$$

*Then there exists a unique measure on $H = \prod_{i\in I} H_i$ such that for all finite subsets $K \subset I$ one has*

$$\mu\left(\prod_{i\in I} A_i\right) = \mu_K\left(\prod_{i\in K} B_i\right), \text{ with } A_i = \begin{cases} B_i \in \mathcal{B}(H_i), & \text{if } i \in K \\ H_i, & \text{if } i \notin K \end{cases}$$

**Remark 3.2.**

1. Here $(\mu_K)_{K\subset I, K \text{ finite}}$ is called the collection of finite dimensional marginal distributions of $\mu$.

2. Kolmogorov extension theorem says that, given the finite dimensional marginales $(\mu_K)$ as above, if they satisfy the consistency condition as requested in the theorem, then there exists a unique extension $\mu$ to the whole product space $H$ with marginale $(\mu_K)$

3. The Theorem is also known as Kolmogorov's existence theorem or Kolmogorov's consistency theorem. It guarantees taht a suitably "consistent" (here consistency is meant as in the theorem) collection of finite-dimensional distributions will define a stochastic process.

**Definition 3.3.** *A stochastic process $X = (X_i)_{i\in I}$, where $I$ is an arbitraty index sex, is a collection of random variables defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, i.e. $X_i : \Omega \to H_i$ is $\mathcal{F}$-measurable.*
*The law of $X$ is a probability measure $\mu$ on $H = \prod_{i\in I} H_i$ such that for all finite $K \subset I$, the marginal law $\mu_K$ of $\mu$ is given by*

$$\mu_K\left(\prod_{j\in K} A_j\right) = \mathbb{P}(X_j \in A_j, j \in K) \text{ for all } A_j \in \mathcal{B}(H_j)$$

**Remark 3.3.** We can always choose $(\Omega, \mathcal{F}, \mathbb{P}) = (H = \prod_{i\in I} H_i, \mathcal{B}(H), \mu)$, then we have $X_i = \pi_i$ where $\pi_i$ are the coordinate maps, i.e.

$$\pi_i : \begin{cases} H & \longrightarrow H_i \\ \omega & \longmapsto \omega_i \end{cases}$$

This is sometimes called the *canonical process*.

There is a section about the Galton-Watson process, it is an important exercise and will be left open on this summary.

## 3.1 Relevant Exercises

These Exercises are related to stochastic processes and are from Exercise Sheet 3. In the following let $\mathcal{B}$ denote the Borel $\sigma$-algebra on the unit interval $[0,1]$ and let $\lambda$ be the Lebesgue measure on $[0,1]$.

**Exercise 3.1.** Construct on the probability space $([0,1],\mathcal{B},\lambda)$ a sequence of independent Bernoulli random variables, that is a measurable map $X : [0,1] \to \{0,1\}^{\mathbb{N}}$ whose law is the appropriate product measure.

**Exercise 3.2.** Construct on the probability space $([0,1],\mathcal{B},\lambda)$ a sequence of independent random variables uniformly distributed on $[0,1]$. Then, deduce that for any sequence of probability measures $(\mu_n)_{n\in\mathbb{N}}$ on $\mathbb{R}$, there exists a stochastic process $X$ with law $\otimes_{n=1}^{\infty}\mu_n$.

**Exercise 3.4.** Show that on the probability space $([0,1],\mathcal{B},\lambda)$ there can be no process $(X_t)_{t\geq 0}$ such that the $X_t$ are i.i.d. Bernoulli 1/2 random variables. Why isn't this a contradiction to Kolmogorov's Extension Theorem?

The next Exercise is from Exercise Sheet 5 and rather tough.

We define the standad $n$-dimensional Gaussian measure $\gamma_n$ to be the probability with density with respect to the Lebesgue measure $\lambda_n$ on $\mathbb{R}^n$ given by

$$\frac{d\gamma_n}{d\lambda_n}(x) = \frac{e^{-|x|^2/2}}{(2\pi)^{n/2}}, \text{ for all } x \in \mathbb{R}^n$$

Exercise 5.6. * Show that the sequence of probability measures $(\gamma_n)_{n=1}^{\infty}$ satisfies the Kolmogorov consistentcy condition and deduce that there exists a probability measure $\Gamma$ on $\mathbb{R}^{\mathbb{N}}$ such that if $X \sim \Gamma$, then $X = (X_1, X_2, \dots)$ is a sequence of independent, identically distributed stnadard Gaussian random variables.

# 4 Markov Processes

**Definition 4.1.** *A stochastic process $(X_t)_{t\in I}$ is a Markov if for all $f : \mathcal{H} \to \mathbb{R}_+$ measurable ($\mathcal{H}$ is a metric space) and for all $s < t$ we have*

$$\mathbb{E}(f(X_t) \mid \mathcal{F}_s^X) = \mathbb{E}(f(X_t) \mid \sigma(X_s)), \text{ where } \mathcal{F}_s^X := \sigma(X_t : t \leq s)$$

**Definition 4.2.** *A process $(X_n)_{n\in\mathbb{N}}$ is a time-homogenous Markov chain if there exists a measurable map $Q : \mathcal{H} \to \mathcal{M}(\mathcal{H}) := \{\text{probability measures on } \mathcal{H}\}$ such that $\mathbb{P}(X_n \in A \mid X_{n-1} = x) = Q(x, A)$ for all Borel sets $A \subset \mathcal{H}$. We call $Q$ the transition probability of the Markov chain. $\mathcal{H}$ is called the state space of the Chain, moreover $x \in \mathcal{H}$ is called a state of the chain.*

**Exercise 5.1.** Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. real-valued random variables with common law $\mu$. Let $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a measurable map. Show that the process given by

$$X_n := F(X_{n-1}, \xi_n)$$

is a time-homogenous Markov chain on $\mathbb{R}$ and find its transition probability. Is the converse of this statement true? I.e. does every time-homogenous Markov chain admit such a representation?

**Proposition 4.1.** *$\mathcal{M}(\mathcal{H})$ equipped with the topology of weak convergence is a completely metrizable, separable space.*

Recall that a sequence $\mu_n \in \mathcal{M}(\mathcal{H})$ converges weakly to $\mu \in \mathcal{M}(\mathcal{H})$ if and only if

$$\int f d\mu_n \to \int f d\mu, \text{ for all } f \in C_b(\mathcal{H} \to \mathbb{R})$$

Given $Q$ as above, then there exists a Markov chain with $\mathbb{P}_\nu(X_0 \in A) = \nu(A)$ and transition probability $Q$. We define for all $n \in \mathbb{N}$

$$\mathbb{P}_\nu(x_0 \in A_0, x_1 \in A_1, \dots, x_n \in A_n) = \int\limits_{A_0\times\cdots\times A_n} \nu(dx_0)Q(x_0, dx_1)\dots Q(dx_{n-1}, dx_n)$$

By Kolmogorov's extension theorem, there exists a unique probability measure $\mathbb{P}_\nu$ on $(\Omega, \mathcal{F}_\infty)$, where $\mathcal{F}_\infty = \bigvee_{n=1}^\infty \mathcal{F}_n$ is the product $\sigma$-algebra on $\Omega$, with marginal distribution as above.

**Definition 4.3.** *We define the shift operator acting on $\Omega$ by $\Theta : \Omega \to \Omega$, $\Theta(\omega) = (\omega_1, \omega_2, \omega_3, \dots)$ where $\omega = (\omega_0, \omega_1, \dots)$.*
*We then have $\Theta^n(\omega) = (\omega_{n+j})_{j=0}^\infty$ for all $n \in \mathbb{N}$.*

**Theorem 4.1** (Markov Property)**.** *Under $\mathbb{P}_\nu$, conditionally on the canonical filtration $\mathcal{F}_n = \sigma(\omega_0, \omega_1, \dots, \omega_n)$, $\Theta^n(\omega)$ has law given by $\mathbb{P}_{\omega_n}$ and is independent of $\mathcal{F}_{n-1}$ for all $n \in \mathbb{N}_{\geq 1}$.*

**Theorem 4.2** (Strong Markov Property)**.** *Let $T$ be a stopping time such that $\mathbb{P}_\nu(T < \infty) > 0$. Conditionally on the event $\{T < \infty\}$ and $\mathcal{F}_T$ we have that $\Theta^n$ is a Markov Chain with law given by $\mathbb{P}_{X_T}$ and independent of $\mathcal{F}_{T-1}$*

We now work (hopefully) with $\mathcal{H}$ a countable (or finite) state space, the transition probability $Q$ is then represented as a matrix, i.e. we have

$$Q(x, A) = \sum_{y \in A} Q_{xy}, \text{ for all } x \in \mathcal{H}$$

A probability measure on $\mathcal{H}$ is a vector such that we have

$$\nu(A) = \sum_{y \in A} \nu_y$$

We have

$$\mathcal{M}(\mathcal{H}) = \{\mu \in \mathbb{R}^{\mathcal{H}} : \mu_x \geq 0 \text{ for all } x \in \mathcal{H} \text{ and } \sum_{x \in \mathcal{H}} \mu_x = 1\}$$

**Definition 4.4.** *We define for all $x \in \mathcal{H}$*

$$R_x := \inf\{n \geq 1 : \omega_n = x\}, \textit{ i.e. the first return to the state } x.$$
$$N_x := \#\{n \geq 0 : \omega_n = x\}, \textit{ i.e. the number of visits of the state } x.$$

*A state $x \in \mathcal{H}$ is called recurrent if $\mathbb{P}_x(R_x < \infty) = 1$, else $x$ is called transient.*

**Proposition 4.2.** *If $x$ is recurrent, then $\mathbb{P}(N_x = \infty) = 1$.*
*If $x$ is transient, then we have*

$$\mathbb{E}_x(N_x) = \frac{1}{\mathbb{P}_x(R_x = \infty)} < \infty$$

**Definition 4.5.** *The Green Kernel of the Markov Chain is defined as*

$$G_{xy} := \mathbb{E}_x(N_y) = \sum_{n \in \mathbb{N}} Q^n_{xy} = \sum_{n \in \mathbb{N}} \mathbb{P}_x(\omega_n = y)$$

*i.e. the expected number of visits of $y$ starting from the state $x$.*
*(Recall that $Q^n_{xy} = \mathbb{P}_x(\omega_n = y)$.)*

**Proposition 4.3.** *Let $x, y \in \mathcal{H}$, $x \neq y$, then the following holds*

1. *$G_{xx} = \infty$ if and only if $x$ is recurrent.*
2. *$G(x, y) = 0$ if and only if $\mathbb{P}_x(R_y < \infty) = 0$.*
   *(i.e. impossible to go from $x$ to $y$.)*
3. *If $\mathbb{P}_x(R_y < \infty) > 0$, then $G(x, y) = G(y, y)\mathbb{P}_x(R_y < \infty)$.*

*Proof.* See Exercises! □

**Corollary 4.1.** *Let* $R := \{x \in \mathcal{H} : \mathbb{P}_x(R_x < \infty) = 1\}$ *the set of recurrent states. We have that if* $x \in R$ *and* $G(x,y) > 0$*, then* $y \in R$ *and also* $G(y,x) > 0,\ \mathbb{P}_y(R_x < \infty) = 1$.

**Remark 4.1.** The above corollary yields that the being recurrent gives equivalent classes on $R$. We have $R = \cup_{j\in I} R_j$ and the sets $R_j$ are disjoint, the $R_j$'s are called the recurrence classes. We define the equivalence relation $x \sim y$ if $G(x,y) > 0$. It follows that $x \in R_j$, then $G(x,y) > 0$ if and only if $y \in R_j$.

**Exercise 5.3.** ** We define the simple random walk on $\mathbb{Z}^d$ to be the Markov process with transition probability given by

$$Q_{x,y} = 2^{-d} 1_{|x-y|=1}, \text{ for all } x, y \in \mathbb{Z}^d$$

Compute the Green function $G_{00}$ for the simple random walk on $\mathbb{Z}^d$. Deduce that the walk is recurrent if and only if $d \leq 2$.

**Theorem 4.3** (Classification of states). *The following holds true*

1. *If* $x \in R_j$ *for some* $j \in I$*, then*

$$\begin{cases} \mathbb{P}_x(N_y = \infty) = 1, & \textit{for all } y \in R_j \\ \mathbb{P}_x(N_y = 0) = 1, & \textit{for all } y \notin R_j \end{cases}$$

2. *Let* $T = \inf\{n \geq 0 : \omega_n \in R\}$ *be the first time, then some weird statement that should be reformulated.*

**Definition 4.6.** *The chain is called irreducible if* $G_{xy} > 0$ *for all* $x, y \in \mathcal{H}$*. Irreducibility means that either we have* $R = \mathcal{H}$ *or* $R = \emptyset$.

**Corollary 4.2.** *If the chain is irreducible, one of the following applies*

1. *All states are recurrent, in particular* $\mathbb{P}_x(N_y = \infty,\ \textit{for all } y \in \mathcal{H}\} = 1$ *for all* $x \in \mathcal{H}$.

2. *All states are transient, in particular* $\mathbb{P}_x(N_y < \infty,\ \textit{for all } y \in \mathcal{H}\} = 1$ *for all* $x \in \mathcal{H}$.

## 4.1 Stationary measures

We consider a Markov Chain with transition probability $Q$.

**Definition 4.7.** *Let $\mu$ be a non-trivial measure on $\mathcal{H}$ such that $\mu_x < \infty$ for all $x \in \mathcal{H}$. We say that $\mu$ is stationary if $\mu = \mu Q$. That is*

$$\mu(x) = \sum_{y \in \mathcal{H}} \mu(y) Q(y, x), \text{ for all } x \in \mathcal{H}$$

**Definition 4.8.** *We say that $\mu$ is reversible if it satisfies the detailed balance, i.e.*

$$\mu_x Q_{xy} = \mu_y Q_{yx}, \text{ for all } x, y \in \mathcal{H}$$

**Exercise 6.6.**: Check that if $\mu$ is reversible, then it is also stationary.

**Lemma 4.1.** *Let $z \in R$ and define for all $x \in \mathcal{H}$*

$$\hat{\pi}(x) := \sum_{n=0}^{\infty} \mathbb{P}_z(\omega_n = x, R_z > n)$$

*Then $\hat{\pi}$ is a stationary measure and we have $\hat{\pi}(x) > 0$ if and only if $x \sim z$ i.e. $G_{zx} > 0$.*

**Proposition 4.4.** *If the chain is irreducible the solution $\mu = \mu Q$ is unique up to a multiplicative constant.*

*Proof.* Exercise! □

**Theorem 4.4.** *Assume that the chain is irreducible and recurrent. Then we have the following dichotomy.*

1. *The chain is positive recurrent, i.e. there exists a stationary measure with finite mass. Moreover we then have, the probability measure $\pi_x = 1/\mathbb{E}_x(R_x)$ is staionary, $\pi$ is called the equilibrium measure.*
2. *The chain is null recurrent, i.e. all stationary measures have infinite mass. Moreover we then have, $\mathbb{E}_x(R_x) = \infty$ for all $x \in \mathcal{H}$.*

**Proposition 4.5.** *Assume the chain is irreducible ($G_{xy} > 0$) and there exists a equilibrium measure $\pi$. Then the chain is positive recurrent, in particular $\pi$ is unique.*

**Theorem 4.5** (Ergodic Theorem)**.** *Let $f, g : \mathcal{H} \to \mathbb{R}_+$. and $z \in R$. Assume that $\int f d\hat{\pi} < \infty$ and $g(x) > 0$ for some $x \sim z$. Then we have $\mathbb{P}_z$-almost surely*

$$\frac{\sum_{j=0}^{n} f(\omega_j)}{\sum_{j=0}^{n} g(\omega_j)} \to \frac{\int f d\hat{\pi}}{\int g d\hat{\pi}} \text{ as } n \to \infty$$

**Corollary 4.3.** *The following holds*

1. *If the chain is positive recurrent, then for $\mu \in \mathcal{M}(\mathcal{H})$ one has $\mathbb{P}_\mu$ almost surely that*

   $$\frac{1}{n}\sum_{j=0}^{n} f(\omega_j) \to \int f d\pi, \text{ as } n \to \infty$$

   *where $\pi$ is the equilibrium measure.*

2. *If the chain is null-recurrent, then for any $\mu \in \mathcal{M}(\mathcal{H})$ one has $\mathbb{P}_\mu$ almost surely that*

   $$\frac{1}{n}\sum_{j=0}^{n} 1_{\omega_j = x} \to 0, \text{ as } n \to \infty \text{ for all } x \in \mathcal{H}.$$

**Definition 4.9.** *Let $x \in R$ be a recurrent state and consider the set $P_x := \{n \in \mathbb{N}_{\geq 1} : Q^n_{x,x} > 0\}$. This set is then infinite (because $x$ is recurrent) and the period of $x$ is defined to be the greatest common divisor of the set $P_x$. That is, the period of $x$ is the number $d = \gcd P_x$. A state $x$ is called aperiodic if $d = 1$, else it is called periodic with period $d$. A Markov chain is called aperiodic if every state is aperiodic.*

**Proposition 4.6.** *A chain is aperiodic if and only if there exists a positive integer $k > 0$ such that $Q^k_{xy} > 0$ for all $x, y \in \mathcal{H}$.*

**Proposition 4.7.** *Let $X$ be an irreducible markov chain, then all states $x$ have the same period $d$. In particular for an irreducible chain it is enough to find one recurrent state to conclude that the chain is aperiodic.*

**Theorem 4.6.** *Suppose the chain is positive recurrent and aperiodic, then*

$$\sum_{x \in \mathcal{H}} |\mathbb{P}_z(\omega_n = x) - \pi(x)| \to 0, \text{ as } n \to \infty$$

*i.e. the law of $\omega_n$ under $\mathbb{P}_z$ converges in total variation to $\pi$.*

**Exercise:** Given a transition probability $Q$. Show that $\tilde{Q} = (I + Q)/2$ defines an irreducible aperiodic chain, it is called the lazy chain.

**Exercise 6.1.** * Let $Q$ be an irreducible transition matrix and assume that the chain has a stationary measure which is finite. Show that the chain is positive recurrent.

**Exercise 6.2.** Let $Q$ be a transition matrix of an irreducible recurrent chain. Show that $\widehat{Q} = \epsilon I + (1-\epsilon)Q$ is also a transition matrix for any $0 < \epsilon < 1$ and that the corresponding chain is irreducible and aperiodic.

**Theorem 4.7.** *Let $Z = (Z_k)_{k\in\mathbb{N}}$ identically distributed random variables such that*

1. $\mathbb{E}(Z_k^2) = 1$
2. $\mathbb{E}(Zk \mid \mathcal{F}_{k-1}) = 0$ *where* $\mathcal{F}_k = \sigma(Z_1, \ldots, Z_k)$
3. *$Z$ is erogidic, in this regime this means*

$$\frac{1}{n}\sum_{j=1}^{n} f(Z_j) \to \mathbb{E}(f(Z)), \text{ almost surely, for all integrable } f.$$

*Then we have that*

$$\frac{Z_1 + \cdots + Z_n}{\sqrt{n}} \implies \mathcal{N}(0,1), \text{ as } n \to \infty$$

*Proof.* Follows by McLeish CLT. □

**Corollary 4.4.** *Let $\mathcal{H}$ be a finite state space and $Q$ be an irreducible transition Matrix with equilibrium measure $\pi$. If $\int f d\pi = 0$ and $\sigma = \int f^2 d\pi < \infty$ then we have that*

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{n} f(\omega_j) \implies \mathcal{N}(0,\sigma^2), \text{ as } n \to \infty$$

# 5 Ergodicity

In this section we consider $H$ to be complete, separable metric space with Borel $\sigma$-algebra $\mathcal{F}$. For all $t \in \mathbb{R}$ let $T_t : H \to H$ be a family of bijective measurable transformations such that they satisfy the semigroup property, i.e.

$$\begin{cases} T_t \circ T_s = T_{t+s} \\ T_0(x) = x \end{cases}$$

**Definition 5.1.** *Let $\mu$ be a measure on $H$. We say that $\mu$ is an invariant measure if $\mu(T_t^{-1}(A)) = \mu(A)$ for all $A \in \mathcal{F}$ and for all $t \in \mathbb{R}$.*
*$(H, \mathcal{F}, \mu, T_t)$ is called a dynamical system.*

**Definition 5.2.** *$\mu$ is called ergodic if for every $A \in \mathcal{F}$ with $T_t^{-1}(A) = A$ for all $t \in \mathbb{R}$ we have $\mu(A) \in \{0, 1\}$.*

We consider the functional expression

$$R_t(f) = \frac{1}{t} \int_0^t f(T_s x) ds, \text{ for } f : H \to \mathbb{R} \text{ measurable}$$

and we wonder under what conditions we have that $R_t(f) \to \int f d\mu$ as $t \to \infty$ and for which $\mu$? We start by discussing some functional analytic properties.

**Definition 5.3.** *A unitary operator is a bounded linear operator $U : H \to H$ on a Hilbert space $H$ that satisfies the following:*

1. *$U$ is surjective, and*
2. *$U$ preserves the inner product of the Hilbert space $H$. That is we have*

$$\langle Ux, Uy \rangle_H = \langle x, y \rangle_H, \text{ for all } x, y \in H.$$

**Lemma 5.1.** *Let $U_s : L^2(H, F, \mu) =: H \to H$ be defines as $H_s f(x) = f(T_s x)$ for all $s \in \mathbb{R}$. Then $U_s$ is an unitary operator on $L^2(\mu)$.*

**Lemma 5.2.** *If $U$ is a unitary operator on a Hilbert space $H$, then we have ran$(I - U)^\perp = \ker(I - U)$ where $I : H \to H$ is the identity operator.*

**Theorem 5.1** (Von Neumann Ergodic Theorem)**.** *Let $U$ be a unitary operator on a Hilbert space $H$. Let $P$ be the orthogonal projection on $\ker(I - U)$. Then we have for all $x \in H$*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} U^k f = Pf \text{ in } H$$

**Remark 5.1.** The convergence above is of course meant with respect to the norm in $H$ which is induced by the scalar product that is given on $H$.

**Lemma 5.3.** *Let us define* $\mathcal{I} := \{A \in \mathcal{F} : T_t^{-1}(A) = A \text{ for all } t \in \mathbb{R}\}$, *then* $\mathcal{I}$ *is a* $\sigma$*-algebra, it is called the* $\sigma$*-algebra of invariant sets. Moreover,* $f : H \to \mathbb{R}$ *is* $\mathcal{I}$*-measurable if and only if* $f \circ T_t = f$.

*Proof.* Exercise! (Exercise 7.6) □

**Corollary 5.1.** *Let* $\mu$ *be an invariant measure under* $(T_t)_{t\in\mathbb{R}}$, $\mathcal{I}$ *the* $\sigma$*-algebra of invariant sets as above, then we have*

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t f(T_s\omega)ds = \mathbb{E}(f \mid \mathcal{I}) \text{ in } L^2(\mu).$$

Next we discuss the Poincaré recurrence theorem, for dynamical systems it states that after a sufficiently long but finite time the system will return to a state very close to the initial state.

Let $\Omega$ be a separable complete metric space and $\theta : \Omega \to \Omega$ be a continuous map with a continuous inverse (i.e. a Homeomorphism). Let $\mu$ be an invariant probability measure on $\Omega$, that is $\mu$ satisfies $\mu(\theta^{-1}A) = \mu(A)$ for all $A \in \mathcal{B}(\Omega)$. We then call $(\Omega, \theta, \mu)$ a dynamical system.

**Theorem 5.2** (Poincaré recurrence Theorem)**.** *Given a dynamical system as above. For any* $A \in \mathcal{B}(\Omega)$ *we have for* $\mu$ *almost all* $x \in A$*, that* $\theta^n(x) \in A$ *infinitely often.*

*Proof.* Exercise 7.5. ! □

**Definition 5.4.** *A dynamical system* $(\Omega, \theta, \mu)$ *is called mixing if for any* $A, B \in \mathcal{B}(\Omega)$ *we have*

$$\lim_{n\to\infty} \mu(\theta^{-n}(A \cap B)) = \mu(A)\mu(B)$$

**Proposition 5.1.** *Show that a dynamical system* $(\Omega, \theta, \mu)$ *is ergodic if and only if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(\theta^{-k}(A \cap B)) = \mu(A)\mu(B)$$

*Deduce that mixing implies ergodicity.*

*Proof.* Exercise! (Exercise 8.1) □

**Exercise 8.4.** Let $(\Omega, \theta, \mu_1)$ and $(\Omega, \theta, \mu_2)$ be two ergodic dynamical systems. Show that if $\mu_1 \neq \mu_2$, then these two measures are mutually singular.

## 6 Continuous time Martingales

**Definition 6.1.** *Let $X = (X_t)_{t\in T}$ be a stochastic process on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in T}, \mathbb{P})$. Then we say that $X$ is a martingale if*

1. *$X$ is $\mathcal{F}_t$-adapted, i.e. $X_t$ is $\mathcal{F}_t$-measurable for all $t \in T$.*
2. *$X_t$ is $\mathbb{P}$-integrable for all $t \in T$.*
3. *$\mathbb{E}(X_t \mid \mathcal{F}_s) = X_s$ for all $s < t$.*

**Definition 6.2.** *Let $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ be stochastic processes. We say*

1. *$X_t$ and $Y_t$ are indistinguishable if $\mathbb{P}(X_t = Y_t,\ \text{for all } t \in \mathbb{R}_0^+) = 1$*
2. *$X_t$ is a modification of $Y_t$ (and vice versa) if $\mathbb{P}(X_t = Y_t) = 1$ for all $t \in \mathbb{R}_0^+$*
3. *$X_t$ and $Y_t$ have the same FDD (Finite Dimensional Distributions) if for all $n \in \mathbb{N}$ and for all $H \in \mathcal{B}(\mathbb{R}^n)$ we have*

$$\mathbb{P}((X_{t_1}, \dots, X_{t_n}) \in H) = \mathbb{P}((Y_{t_1}, \dots, Y_{t_n}) \in H)$$

**Remark 6.1.** We do indeed have that 1) $\implies$ 2) $\implies$ 3).

**Definition 6.3.** *Let $(\mathcal{F}_t)_{t\geq 0}$ be a Filtration. We define the left respectively right continuous Filtrations as*

$$\mathcal{F}_{t^-} := \bigvee_{s<t} \mathcal{F}_s \text{ and } \mathcal{F}_{t^+} := \bigcap_{s>t} \mathcal{F}_s$$

**Remark 6.2.** We have the inclusions $\mathcal{F}_{t^-} \subset \mathcal{F}_t \subset \mathcal{F}_{t^+}$ for all $t \in \mathbb{R}_0^+$ and it is very well possible that these inclusions are strict.

**Definition 6.4.** *Let $(\mathcal{F}_t)_{t\geq 0}$ be a Filtration. We define*

$$\tilde{\mathcal{F}}_t := \sigma(\mathcal{F}_{t^+}, \mathcal{N}), \text{ where } \mathcal{N} \text{ is the set of } \mathbb{P}\text{-negligible sets.}$$

*We say that out Filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfies the usual conditions if $\mathcal{F}_t = \tilde{\mathcal{F}}_t$ for all $t \in \mathbb{R}_0^+$.*

We come now to our main result in this section. It states that we can regulate martingales, i.e. represent them as modifications through cadlag (continue a droite, limites a gauche) martingales.

**Theorem 6.1** (Martingale regularization theorem)**.**
*Let $(X_t)_{t\geq 0}$ be a $\mathcal{F}_t$-martingale, then there exists $(\widetilde{X_t})_{t\geq 0}$ a $\widetilde{\mathcal{F}_t}$-martingale which is cadlag such that $X_t = \mathbb{E}(\widetilde{X_t} \mid \mathcal{F}_t)$ almost surely for all $t \geq 0$.*
*Moreover, if $(\mathcal{F}_t)_{t\geq 0}$ satisfies the usual conditions, then $\widetilde{X_t}$ is a cadlag modification of $X_t$.*

The Martingale regularization theorem is a powerful tool when it comes to extending statements we already know from discrete time martingales to continuous time martingales, because we can regularize them by cadlag representations and then approximate them (using for instance that $\mathbb{Q}$ is dense in $\mathbb{R}$).

**Theorem 6.2** (Convergence Theorem)**.** *Let $X$ be a cadlag martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$. If $X$ is bounded in $L^1$, i.e. if $\sup_{t\geq 0}\|X_t\|_{L^1} < \infty$, then $X_t \to X_\infty$ as $t \to \infty$ for some $X_\infty \in L^1(\Omega, \mathcal{F}_\infty, \mathbb{P})$.*
*Moreover if $X$ is UI, then $X_t \to X_\infty$ in $L^1(\Omega, \mathcal{F}, \mathbb{P})$.*

*Proof.* Exercise! (Exercise 9.3.) $\square$

**Theorem 6.3** (Stopping Theorem)**.** *Let $X$ be a cadlag martingale and let $\tau \leq \Lambda$ be stopping times with respect to $(\mathcal{F}_t)_{t\geq 0}$. If $\Lambda \leq C$ for some constant $C > 0$, then*

$$\mathbb{E}(X_\Lambda \mid \mathcal{F}_\tau) = X_\tau \tag{$\star$}$$

*Moreover, if $X$ is UI, then $(\star)$ above is true for all stopping times $\tau \leq \Lambda$.*

*Proof.* Exercise! $\square$

# 7 Brownian Motion

Before we introduce Brownian Motion we start with some preliminary reminders that might be useful later on.

**Definition 7.1.** *Let $X$ be a real valued random variable, we say that $X$ is standard Gaussian and denote this by $X \sim \mathcal{N}(0,1)$ if its probability density function is given by*

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right), \text{ for all } x \in \mathbb{R}.$$

The concept of a standard Gaussian (also known as normal) random variable can be extended to the multivariate i.e. the vector case.

**Definition 7.2.** *A random vector $X = (X_1, \ldots, X_n)$, $n \in \mathbb{N}$ is said to have the multivariate Gaussian distribution if every linear combination of its components $Y = a_1X_1 + \cdots + a_nX_n$ where $a_i \in \mathbb{R}$ for all $i \in \mathbb{N}$ is standard Gaussian distributed (i.e. Gaussian in dimension 1).*

We extend the definition of the mean and covariance of a random vector in the most natural way

**Definition 7.3.** *Let $X = (X_1, \ldots, X_n)$, $n \in \mathbb{N}$ be a random vector in $L^2$ (that is by definition that all its components are in $L^2$) then we define the mean of $X$, $\mathbb{E}(X) := (\mathbb{E}(X_1), \ldots, \mathbb{E}(X_n))$ and its Covariance/Dispersion Matrix $D = (D_{i,j})_{1 \leq i,j \leq n}$ where $D_{i,j} := Cov(X_i, X_j) := \mathbb{E}(X_iX_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)$. We notice that the Covariance Matrix is symmetric.*

**Theorem 7.1.** *Let $X$ be a Gaussian vector (i.e. $X$ has the multivariate Gaussian distribution), then the law of $X$ is fully characterized by its mean and its covariance matrix as defined above.*

We are now ready to start on the relevant definitions in order to introduce what a Brownian motion is.

**Definition 7.4.** *A stochastic process $X = (X_t)_{t \geq 0}$ is Gaussian if for every $n \in \mathbb{N}_{\geq 1}$ and times $t_1, \ldots, t_n \in [0, \infty)$ the vector $(X_{t_1}, \ldots, X_{t_n})$ has a multivariate Gaussian distribution.*

**Remark 7.1.**

1. A Gaussian process $X$ is fully characterized by its mean $\mathbb{E}(X_t)$ and by its covariance kernel $K(t,t') := \text{Cov}(X_t, X_{t'}) = \mathbb{E}(X_tX_{t'}) - \mathbb{E}(X_t)\mathbb{E}(X_{t'})$
2. If $\mathbb{E}(X_t) = 0$, then we have $K(t,t') = \text{Cov}(X_t, X_{t'}) = \mathbb{E}(X_tX_{t'})$

**Proposition 7.1** (Characterization)**.** *The following two statements are equivalent:*

1. $(X_t)_{t\geq 0}$ *has stationary, independent increments and* $X_t \sim \mathcal{N}(0,t)$ *for all* $t \geq 0$.

2. $(X_t)_{t\geq 0}$ *is a Gaussian process with mean zero and covariance kernel* $\mathbb{E}(X_t X_{t'}) = t \wedge t'$ *for all* $t, t' \in \mathbb{R}_0^+$.

**Remark 7.2.** We recall the following

1. Independence of the increments means that for all times $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n < \infty$ where $n \in \mathbb{N}$, $X_{t_n} - X_{t_{n-1}}, \ldots, X_{t_3} - X_{t_2}, X_{t_2} - X_{t_1}$ are independent.

2. Stationarity of the increments means that for all $s < t$ we have that $X_t - X_s$ is equal in distribution as $X_{t-s}$.

3. We can also summarize 1) and 2) above as, if for all $t \in \mathbb{R}_0^+$ we have

$$\begin{aligned}&(X_{t_n} - X_{t_{n-1}}, \ldots, X_{t_3} - X_{t_2}, X_{t_2} - X_{t_1})\\ &\sim (X_{t+t_n} - X_{t+t_{n-1}}, \ldots, X_{t+t_3} - X_{t+t_2}, X_{t+t_2} - X_{t+t_1})\end{aligned}$$

4. Since $X_t \sim \mathcal{N}(0,t)$ we have by the stationarity $X_t - X_s \sim X_{t-s} \sim \mathcal{N}(0, t-s)$.

**Definition 7.5.** *A Brownian motion is a stochastic process* $(B_t)_{t\geq 0}$ *with continuous sample paths (i.e. a continuous stochastic process) which satisfies either property 1) or 2) (and consequently both of them) in Propositon 7.1.*

**Definition 7.6.** *A Gaussian Hilbert space* $W$ *is a closed subspace of* $L^2(\Omega, \mathcal{F}, \mathbb{P})$ *of centered Gaussian Random Variables.*

**Definition 7.7.** *A Gaussian process* $\Phi$ *indexed by a separable Hilbert space* $H$ *is an isomorphism* $\Phi : H \to W$ *where* $W$ *is a Gaussian Hilbert space, such that*

$$\mathbb{E}(\Phi(h)\Phi(g)) = \langle h, g \rangle_H, \text{ for all } h, g \in H.$$

## 7.1 Construction of Brownian Motion

Here is a sketch of the construction of a Brownian Motion.
Let $H = L^2([0,\infty), dx)$ and let $(\varphi_n)_{n\in\mathbb{N}}$ be an ONB of $H$. Let now $\xi_1, \xi_2, \dots$ be a sequence of i.i.d. $\mathcal{N}(0,1)$ random variables.

Let $W := \overline{\operatorname{span}(\xi_1, \xi_2, \dots)}$, then $W$ is a Gaussian Hilbert space. If we now require that $\Phi : H \to W$ linear satisfies

$$\Phi(\varphi_n) = \xi_n, \text{ for all } n \in \mathbb{N}$$

Then by linearty of $\Phi$ and the fact that every element $h \in H$ can be written as $h = \sum_{n=1}^{\infty} \alpha_n \varphi_n$ for some $\alpha_i \in \mathbb{R}$, we get that

$$\Phi(h) = \sum_{n=1}^{\infty} \alpha_n \xi_n$$

Consequently $\Phi$ is an isomorphism. Moreover we have that

$$\mathbb{E}(\Phi(h)\Phi(g)) = \mathbb{E}\left(\sum_{n=1}^{\infty} \alpha_n \xi_n \sum_{k=1}^{\infty} \beta_k \xi_k\right) = \sum_{i=1}^{\infty} \alpha_i \beta_i \mathbb{E}(\xi_i^2) = \sum_{i=1}^{\infty} \alpha_i \beta_i$$

Analogeously, since $\varphi_n$ form an ONB of $L^2$ we get that

$$\langle h, g\rangle_{L^2} = \int_0^{\infty} \sum_{n=1}^{\infty} \alpha_n \varphi_n(x) \sum_{k=1}^{\infty} \beta_k \varphi_k(x) dx = \sum_{i=1}^{\infty} \alpha_i \beta_i$$

Thus $\Phi$ is a Gaussian process indexed by the Hilbert Space $L^2$. We also notice that $\mathbb{E}(\Phi(h)) = 0$ and consequently

$$\mathbb{E}((\Phi(h))^2) = \operatorname{Var}(\Phi(h)) = \|h\|_H^2 = \sum_{n=1}^{\infty} \alpha_n^2 < \infty$$

Where the finiteness of the above sum comes from the fact that $\|\cdot\|_{H^2}$ is a well-defined norm on $H = L^2$. This entails that we have

$$\Phi(h) = \lim_{N\to\infty} \underbrace{\sum_{n=1}^{N} \alpha_n \xi_n}_{\sim \mathcal{N}(0, \sum_{n=1}^{N} \alpha_n^2)} \sim \mathcal{N}\left(0, \sum_{n=1}^{\infty} \alpha_n^2\right)$$

Using our work done above, we can set $X_t := \Phi(1_{[0,t]})$ and following our calculations above, it is obvious to conclude that $X_t$ is a Gaussian process. Moreover we have

$$\begin{aligned}\mathbb{E}(X_t X_s) = \mathbb{E}\left(\Phi(1_{[0,t]})\Phi(1_{[0,s]})\right) &= \langle 1_{[0,t]}, 1_{[0,s]}\rangle_{L^2} \\ &= \int_0^{\infty} 1_{[0,t]}(x) 1_{[0,s]}(x) ds = t \wedge s\end{aligned}$$

We have thus shown that the process $X_t = \Phi(1_{[0,t]})$ satisfies 2) in Proposition 7.1 and consequently also 1). We are not done yet with the construction, because we need a continuous process.

We are however in a good position to complete the construction. We will consistently use $\Phi$ for the same isomorphism as constructed above. Moreover, we notice that is enough to construct a Brownian Motion on an interval $[0,1]$, by continuity and glueing the functions together we can then extend this to a Brownian Motion on the real line $[0,\infty)$.

Consider the Haar function, for this we first define the mother wavelet function as

$$\Psi(t) := \begin{cases} 1, & \text{if } 0 \le t \le 1/2 \\ -1, & \text{if } 1/2 \le t \le 1 \\ 0, & \text{else} \end{cases}$$

We then define for all $n \in \mathbb{N}_0$ and $k = 0, 1, \ldots, 2^n - 1$ the Haar function as

$$\Psi_{n,k}(t) = 2^{n/2}\Psi(2^n t - k), \text{ where } t \in [0,1]$$

This function is supported on the dyadic intervals on $[0,1]$, that is on $I_{n,k} = [k2^{-n}, (k+1)2^{-n}]$ for $k = 0, \ldots, 2^n - 1$ and $n \in \mathbb{N}_0$. Thus we easily conclude that for all $t \in [0,1]$ we have

$$\begin{aligned}\left|\int_0^t \Psi_{n,k}(x)dx\right| &\le \int_0^t |2^{n/2}\Psi(2^n x - k)|dx = 2^{n/2}\int_{k2^{-n}}^{(k+1)2^{-n}} |\Psi(2^n x - k)|dx \\ &\le 2^{n/2}\int_{k2^{-n}}^{(k+1)2^{-n}} dx = 2^{n/2}2^{-n} = 2^{-n/2}\end{aligned}$$

Furthermore, it is possible to show that $\{1, \Psi_{n,k}\}_{n\ge 0, k=0,\ldots,2^n-1}$ form an ONB of $L^2([0,1], dx)$. We define now a process by

$$Z_n(t) := \sum_{k=0}^{2^n-1} \xi_{n,k}\int_0^t \Psi_{n,k}(x)dx = \sum_{k=0}^{2^n-1} \xi_{n,k}\langle 1_{[0,t]}, \Psi_{n,k}\rangle_{L^2}$$

$$\text{where } (\xi_{n,k})_{\substack{k=0,\ldots,2^n-1 \\ n\ge 0}} \text{ is an array of i.i.d. } \mathcal{N}(0,1)$$

Then $Z_n$ is continuous for all $n \ge 0$. We can estimate

$$\begin{aligned}\max_{t\in[0,1]} |Z_n(t)| &\le \sum_{k=0}^{2^n-1} |\xi_{n,k}| \left|\int_0^t \Psi_{n,k}(x)dx\right| \\ &\le \max_{k=0,\ldots,2^n-1} |\xi_{n,k}| \sum_{k=0}^{2^n-1} \left|\int_{k2^{-n}}^{(k+1)2^{-n}} \Psi_{n,k}(x)dx\right| \\ &\le \max_k |\xi_{n,k}| \left|\int_0^1 \Psi_{n,k}(x)dx\right| \le \max_k |\xi_{n,k}| 2^{-n/2}\end{aligned}$$

Since all the $\xi_{n,k}$ are i.i.d. $\mathcal{N}(0,1)$ we obtain for arbitrary $\lambda > 0$

$$\mathbb{P}(\max_k |\xi_{n,k}| > \lambda) \leq 2^n \mathbb{P}(|\mathcal{N}| > \lambda) \leq 2^n \frac{e^{-\lambda^2/2}}{\lambda}$$

Let now $\lambda_n = n2^{-n/2}$ then we obtain by our estimate on the previous page that $\{\max_{t\in[0,1]} |Z_n(t)| > \lambda_n\} \subset \{\max_k |\xi_{n,k}| 2^{-n/2} > \lambda_n\}$ and consequentaly

$$\begin{aligned}\mathbb{P}(\max_{t\in[0,1]} |Z_n(t)| > n2^{-n/2}) &\leq \mathbb{P}(\max_k |\xi_{n,k}| 2^{n/2} > n2^{n/2}) = \mathbb{P}(\max_k |\xi_{n,k}| > n) \\ &\leq 2^n \frac{e^{-n^2/2}}{n}\end{aligned}$$

By the ratio test, one can easily show that this upper estimate is summable, thus by Borel-Cantelli Lemma we get that eventually

$$\max_{t\in[0,1]} |Z_n(t)| \leq n2^{-n/2}$$

Again, by the ratio test, this expression is summable and thus we get that

$$\sum_{n=0}^{\infty} \max_{t\in[0,1]} |Z_n(t)| < \infty$$

Since all the $Z_n$ are continuous, and the convergence above is uniformly over all $t \in [0,1]$ this entails that $\sum Z_n(t)$ converges almost surely to a continuous function on $[0,1]$.

We define now

$$B(t) := t\xi_{0,0} + \sum_{n=1}^{\infty} \sum_{k=1}^{2^n-1} \xi_{n,k} \langle 1_{[0,t]}, \Psi_{n,k} \rangle$$

Since $\{1, \Psi_{n,k}\}_{n\geq 0, k=0,\ldots,2^n-1}$ is an ONB of $L^2([0,1], dx)$ we can write every element $h \in L^2$ as $h = \sum_{\alpha\in A} \langle x_\alpha, x \rangle x_\alpha$. We thus have

$$\begin{aligned}\Phi(1_{[0,t]}) &= \Phi\left(\sum_{n=0}^{\infty} \sum_{k=0}^{2^n-1} \langle 1_{[0,t]}, \Psi_{n,k} \rangle \Psi_{n,k}\right) \\ &= t\xi_{0,0} + \sum_{n=1}^{\infty} \sum_{k=1}^{2^n-1} \langle 1_{[0,t]}, \Psi_{n,k} \rangle \Phi\left(\Psi_{n,k}\right) \\ &= t\xi_{0,0} + \sum_{n=1}^{\infty} \sum_{k=1}^{2^n-1} \langle 1_{[0,t]}, \Psi_{n,k} \rangle \xi_{n,k} = B(t)\end{aligned}$$

And we have already shown, that $\Psi(1_{[0,t]})$ satisfies ii) and therefore also i), we conclude that $B(t)$ is a Brownian Motion on $[0,1]$.

**Definition 7.8.** *The law of a Brownian Motion on* $\mathcal{C} := \{\omega : [0,\infty) \to \mathbb{R}$ *continuous*$\}$ *will be denoted by* $\mathbb{P}_0$. *We denote by* $\mathbb{P}_x$ *the image of* $\mathbb{P}_0$ *by the translation map* $\Theta : \mathcal{C} \to \mathcal{C},\ \omega \mapsto \omega + x$ *for all* $x \in \mathbb{R}$.

Brownian Motions have interesting scaling properties as the next Lemma shows.

**Lemma 7.1.** *Let* $B = (B_t)_{t\geq 0}$ *be a standard Brownian Motion, then we have*

1. *For any* $\alpha \in \mathbb{R}$, $B_{\alpha^2 t} \sim \alpha B_t$.
2. *For any* $s > 0$ *we have that* $X_t = B_{s+t} - B_s$ *is a Brownian Motion independent of* $\mathcal{F}_s^0 = \sigma(B_t, 0 \leq t \leq s)$.

*Proof.* Exercise! □

**Definition 7.9.** *We define* $\mathcal{F}_s^0 = \sigma(B_t : 0 \leq t \leq s)$ *and* $\mathcal{F}_s^+ = \cap_{t>s}\mathcal{F}_t^0$ *the right-continuous extension. In particular we call* $\mathcal{F}_0^+$ *the germ-sigma-algebra.*

**Theorem 7.2** (Markov Property). *For any bounded measurable* $F : \mathcal{C} \to \mathbb{R}$ *and any* $t > 0$ *the Markov Property holds, i.e. we have*

$$\mathbb{E}_x(F \circ \Theta_t \mid \mathcal{F}_t^+) = \mathbb{E}_{\omega_t}(F),\ x \in \mathbb{R}$$

*where* $(\Theta_t)_{t\geq 0}$ *is the shift on* $\mathcal{C}$, *i.e. for any* $s \in \mathbb{R}$ *we have*

$$\Theta_s : \begin{cases} \mathcal{C} & \longrightarrow \mathcal{C} \\ \omega & \longmapsto \theta_s(\omega) = (\omega_{t+s})_{t\geq 0} \end{cases}$$

**Remark 7.3.** Markov's property for BM states that under $\mathbb{P}_x$, the conditional law of $\Theta_t\omega$ given $\mathcal{F}_t^+$ is $\mathbb{P}_{\omega_t}$.

A consequence of the Markov's property for BM is given by Blumenthal's 0/1 law, it establishes that the germ sigma algebra is in fact a trivial sigma algebra.

**Theorem 7.3** (Blumenthal's 0/1 law). *For any* $x \in \mathbb{R}$ *and any* $A \in \mathcal{F}_0^+$ *we have* $\mathbb{P}_x(A) \in \{0,1\}$, *i.e. the germ-sigma-algebra* $\mathcal{F}_0^+$ *is trivial.*

**Theorem 7.4** (Strong Markov Property). *Let* $F : \mathcal{C} \to \mathbb{R}$ *be bounded measurable,* $\tau$ *be a stopping time with respect to* $(\mathcal{F}_t^0)_{t\geq 0}$, *then for all* $x \in \mathbb{R}$ *with* $\mathbb{P}_x(\tau < \infty) = 1$ *we have*

$$\mathbb{E}_x(F \circ \Theta_\tau \mid \mathcal{F}_\tau^+) = \mathbb{E}_{\omega_\tau}(F)$$

**Remark 7.4.** This means that under $\mathbb{P}_x(\cdot \mid \tau < \infty)$ the process $t \mapsto \omega_{t+\tau}$ has law $\mathbb{P}_{\omega_\tau}$ and is indepent from $\mathcal{F}_\tau^+ = \cap_{\epsilon>0}\mathcal{F}_{\tau+\epsilon}^0$.

**Proposition 7.2** (Reflection principle)**.** *Let $a > 0$ and $b < a$ be arbitrary, then for any $t > 0$ we have*

$$\mathbb{P}_0\left(\max_{s\le t}\omega_s > a,\ \omega_t < b\right) = \mathbb{P}_0(\omega_t > 2a-b) = \int_{2a-b}^{\infty}\frac{e^{-x^2/(2t)}}{\sqrt{2\pi t}}dx$$

**Remark 7.5.**

1. The Reflection principle gives us the joint law of the supremum of the process and the process itself.

2. The last equality was just stated for the sake of completeness, indeed $\omega_t$ is a Gauss Random variable with mean zero and variance $t$ at time $t$.

3. For any $t > 0$, let us denote $M_t = \max_{s\le t}\omega_s$, then $M_t$ and $M_t - \omega_t$ have the same law as $(\omega_t)_{t\ge 0}$.

Here is the Reflection principle as presented by Norbert Wiener.

**Theorem 7.5** (Reflection Principle)**.** *Let $(W_t)_{t\ge0}$ be a Wiener process (i.e. a standard Brownian Motion), and let $a > 0$ be a treshold (also called a crossing point), then we have*

$$\mathbb{P}_0\left(\sup_{0\le t\le s} W_t \ge a\right) = 2\mathbb{P}(W_t \ge a)$$

## 7.2 Sample path properties of a Brownian Motion

We start with the important time Inversion Formula for the Brownian Motion.

**Proposition 7.3** (Time Inversion)**.** *Let $(B_t)_{t\ge0}$ be a standard Brownian Motion (i.e. it has law $\mathbb{P}_0$), then $(tB_{1/t})_{t\ge0}$ is also a standard Brownian Motion (i.e. also has law $\mathbb{P}_0$).*

*Proof.* Let us define the process $Z_t := tB_{1/t}$. Then clearly $(Z_t)_{t\ge0}$ is a Gaussian process which is continuous for all $t > 0$ (we need to pay seperate attention the time $t = 0$). Moreover we have for all $s, t \in \mathbb{R}_0^+$,

$$\mathbb{E}(Z_tZ_s) = ts\mathbb{E}(B_{1/t}B_{1/s}) = ts\min\left(\frac{1}{t}, \frac{1}{s}\right) = \min(t,s)$$

It remains to show that $\lim_{t\to0} Z_t = 0$, i.e. $Z_t$ is continuous at $t = 0$. We notice that by Kolmogorov's LLN we have $\lim_{n\to\infty}\frac{B_n}{n} = 0$ almost surely. Showing that $\lim_{t\to0} Z_t = 0$ is equivalent to showing that for all $\epsilon > 0$ we have

$$\mathbb{P}\left(\limsup_{t\to\infty}\frac{|B_t|}{t} > \epsilon\right) = 0$$

By the Reflection principle we obtain for all $\lambda > 0$ that

$$\mathbb{P}\left(\sup_{0\le s\le t} B_s \ge \lambda\right) = 2\mathbb{P}(B_t \ge \lambda) = \mathbb{P}(|B_t| \ge \lambda) \le 2e^{-\lambda^2/(2t)}$$

Specializing this bound on $\lambda = \sqrt{n}$ for a $n \in \mathbb{N}_{\geq \nVdash}$ gives a summable upper bound. Thus by Borel-Cantelli Lemma we obtain that

$$\sup_{n\le s\le n+1} B_s \le \sqrt{n}, \text{ almost surely, eventually.}$$

This establishes that

$$\limsup_{s\to\infty} \frac{B_s}{s} = 0, \text{ almost surely.}$$

Just replacing $B_s$ with $-B_s$ gives the statement about the liminf, we conclude that $\lim_{s\to\infty} B_s/s = 0$. $\square$

**Corollary 7.1** (0/1 Law)**.** *Let* $\tau = \bigcap_{t>0} \sigma(\omega_s : s \ge t)$ *(tail* $\sigma$*-algebra), if* $A \in \tau$*, then* $\mathbb{P}_x(A) \in \{0, 1\}$ *for all* $x \in \mathbb{R}$*.*

The next Proposition entails that the behaviour of a Brownian Motion is rather wild around its origin (conveniently picked to be at zero) and oscilating in general for large values of $t$.

**Proposition 7.4.** *We have* $\mathbb{P}_0$*-almost surely that*

$$\limsup_{t\to 0} \frac{\omega_t}{\sqrt{t}} = +\infty \text{ and } \liminf_{t\to 0} \frac{\omega_t}{\sqrt{t}} = -\infty$$

*It also follows that*

$$\limsup_{t\to \infty} \frac{\omega_t}{\sqrt{t}} = +\infty \text{ and } \liminf_{t\to \infty} \frac{\omega_t}{\sqrt{t}} = -\infty$$

*Proof.* Let $a > 0$ be arbitrary, by symmetry it is enough to establish that

$$\mathbb{P}_0\Big(\underbrace{\limsup_{t\to 0} \frac{\omega_t}{\sqrt{t}} > a}_{\in \mathcal{F}_0^+}\Big) = 1$$

By Blumenthal's 0/1 law, we just need to show that the above probability is strictly positive (since it can either be 0 or 1). However, we easily see that

$$\begin{aligned}
\mathbb{P}_0\left(\limsup_{\to 0} \frac{\omega_t}{\sqrt{t}} > a\right) &= \lim_{\epsilon\to 0} \mathbb{P}_0\left(\sup_{0\le t\le\epsilon} \frac{\omega_t}{\sqrt{t}} > a\right) \\
&\ge \lim_{\epsilon\to 0} \mathbb{P}_0\left(\frac{\omega_\epsilon}{\sqrt{\epsilon}} > a\right) = \mathbb{P}(\mathcal{N}(0,1) > a) > 0
\end{aligned}$$

$\square$

**Definition 7.10.** *Let $f$ be a real or complex-valued function on $\mathbb{R}$ and let $0 < \alpha \leq 1$. We say that $f$ is $\alpha$-Hölder continuous or that it satisfies a Hölder condition when there exists a real constant $C > 0$ such that*

$$|f(x) - f(y)| \leq C|x - y|^\alpha, \text{ for all } x, y \in \mathbb{R}.$$

**Remark 7.6.**

1. We say $f$ is (locally) $\alpha$-Hölder continuous at $t \in \mathbb{R}$ if there exists a $\delta > 0$ such that for all $|x - y| < \delta$ the above inequality is satisfied.

2. For $\alpha = 1$ we obtain the Lipschitz-Continuity. The case $\alpha = 0$ isn't interesting because it only gives us bounded but not necessarily continuous functions. Moreover it can be shown that for $\alpha > 1$ only constant functions satisfy the Hölder condition.

3. We have the following chain of inclusions for all $\alpha \in (0, 1]$:
   Continuously differentiable $\subset$ Lipschitz continuous $\subset$ $\alpha$-Hölder continuous $\subset$ uniformly continuous $\subset$ continuous.

**Proposition 7.5.** *For any $\alpha > 1/2$, $\mathbb{P}_0$ almost surely, the sample paths of $\omega$ are nowhere $\alpha$-Hölder-continuous. In particular, $\mathbb{P}_0$ almost-surely, the function $t \mapsto \omega_t$ is nowhere differentiable.*

*Proof.* Let $C > 0$ and $\alpha > 1/2$. For any natural number $N \geq 1$ we define the set

$$A_N := \{t \in [0, 1] : |\omega_s - \omega_t| \leq C|t - s|^\alpha, \text{ for all } s : |s - t| \leq 1/N\}$$

We want to show that $\mathbb{P}_0(A_N) = 0$ for all $N \geq 1$. In fact, we have that $A_N$ is an increasing sequence of events, so we only need to prove that

$$\lim_{N\to\infty} \mathbb{P}_0(A_N) = 0$$

Let $n \in \mathbb{N}$ and for all $k = 0, 1, \ldots, 2nN$ define

$$Z_{N,k} := \max_{j=1,\ldots,n} \left| \frac{\omega_{j+k}}{2nN} - \frac{\omega_{j+k-1}}{2nN} \right|$$

Let us now define the set

$$D_N := \left\{ k : Z_{N,k} \leq \frac{2C}{N^\alpha} \right\}$$

We claim that that $A_n \subset D_n$, indeed for $\omega \in A_N$ we set $k = \lfloor 2nNt \rfloor$, then the triangle inequality gives us

$$\left| \frac{\omega_{j+k+1}}{2nN} - \frac{\omega_{j+k}}{2nN} \right| \leq \left| \frac{\omega_{j+k+1}}{2nN} - \omega_t \right| + \left| \omega_t - \frac{\omega_{j+k}}{2nN} \right| \leq \frac{C}{N^\alpha} + \frac{C}{N^\alpha} = \frac{2C}{N^\alpha}$$

which shows that $\omega \in D_N$, consequentally we have that $\mathbb{P}_0(A_N) \leq \mathbb{P}_0(D_N)$.

Since for $j = 1, \dots, n$ the random variables

$$\frac{\omega_{j+1+1}}{2nN} - \frac{\omega_{j+k}}{2nN}$$

are i.i.d. $\mathcal{N}(0, 1/2nN)$ it follows that $Z_{N,k}$ are identically distributed for all $k = 0, 1, \dots, 2nN$. Thus we can easily establish the bound

$$\mathbb{P}_0(D_N) \leq 2nN\mathbb{P}_0\left(Z_{N,k} \leq \frac{2C}{N^\alpha}\right) = 2nN\mathbb{P}_0\left(\left|\mathcal{N}\left(0, \frac{1}{2nN}\right)\right| \leq \frac{2C}{N^\alpha}\right)^n$$

Using the trivial bound $\mathbb{P}(|\mathcal{N}(0,1)| \leq \lambda) \leq \lambda$ for all $\lambda > 0$ we find that

$$\begin{aligned}\mathbb{P}_0(D_N) &\leq 2nN\mathbb{P}\left(\frac{1}{\sqrt{2nN}}|\mathcal{N}(0,1)| \leq \frac{2C}{N^\alpha}\right)^n \\ &= 2nN\mathbb{P}(|\mathcal{N}(0,1)| \leq C_n N^{1/2-\alpha})^n \leq 2nC_n^n N^{n\left(\frac{1}{2}-\alpha\right)+1}\end{aligned}$$

Since $\alpha > \frac{1}{2}$, we have that $(1/2 - \alpha) < 0$, choosing $n \in \mathbb{N}$ large enough we can guarantee that $n(0.5 - \alpha) + 1 < 0$. We conclude that

$$\lim_{N\to\infty} \mathbb{P}_0(D_N) = 0, \text{ and thus } \lim_{N\to\infty} \mathbb{P}_0(A_N) = 0$$

□

**Remark 7.7.** For any $0 < \alpha < \frac{1}{2}$ one can show, that $\mathbb{P}_0$ almost surely, the sample path of $\omega$ are in fact $\alpha$-Hölder continuous at every point. However, no statement is possible for the case $\alpha = 1/2$.

**Proposition 7.6** (Law of iterated logarithm). *Let $h_t := \sqrt{2t \log|\log(t)|}$ for any $t > 0$, then $\mathbb{P}_0$ almost surely we have*

$$\begin{aligned}\limsup_{t\to 0} \frac{\omega_t}{h_t} &= 1 \\ \liminf_{t\to 0} \frac{\omega_t}{h_t} &= -1 \\ \limsup_{t\to\infty} \frac{\omega_t}{h_t} &= 1\end{aligned}$$

*Proof.* Exercise! □

### 7.3 Martingales derived from Brownian Motion

Let $B$ be a standard Brownian Motion

**Proposition 7.7.** *Let $(B_t)_{t\geq 0}$ be a standard Brownian Motion, then the following processes are martingales:*

1. $(B_t^2 - t)_{t\geq 0}$
2. $M_t^\alpha = \exp\left(\alpha B_t - \frac{\alpha^2}{2}t\right)$ *for all* $t \geq 0$ *with expected value* $1$.
3. $\left(\frac{d}{d\alpha}\right)^k M^\alpha_{t\,|\alpha=0}$, *for any* $k \in \mathbb{N}$

*Proof.* Exercise! □

**Theorem 7.6** (Skorokhod's Theorem)**.** *Let $\xi$ be a random variable such that $\mathbb{E}(\xi) = 0$ and $\mathbb{E}(\xi^2) = \sigma$. Then there exists a stopping time $\tau$ such that $\mathbb{E}(\tau) = \sigma^2$ and*

$$B_\tau \stackrel{d}{=} \xi.$$

In order to prove Skorokhod's Theorem we have another relevant exercise, it is a special case of Skorokho'ds Theorem for centered Bernoulli random variables.

**Proposition 7.8.** *For any $\alpha \in \mathbb{R}$, let $\tau_\alpha = \inf\{t > 0 : \omega_t = \alpha\}$. If $\alpha < x < \beta$, show that $\mathbb{E}_x(\tau_\alpha \wedge \tau_\beta) < \infty$ and*

$$\mathbb{P}_x(\tau_\alpha < \tau_\beta) = \frac{\beta - x}{\beta - \alpha}$$

*Deduce that if $X$ is a centered Bernoulli random variable taking values in $\{a, \beta\}$, then there exists a stopping time $\Lambda$ with finite mean so that $X$ and $\omega_\Lambda$ have the same law under $\mathbb{P}_0$.*

*Proof.* Exercise! □

**Remark 7.8.** The strategy behind the proof of Skorokhod's Theorem is that we construct a sequence $\xi_n$ (taking finitely many values) such that $\xi_n \to \xi$ almost surely. Moreover we construct an increasing sequence of stopping times $\tau_n \uparrow \tau$ such that $\mathbb{E}(\tau) < \infty$ and

$$B_{\tau_n} \stackrel{d}{=} \xi_n$$

The initiatel step is supported the above exercise, we then proceed iteratively by segmentation.

Finally we will state Donsker's invariance Theorem. Let $\xi_1, \xi_2, \dots$ be a sequence of i.i.d. Random Variables with mean zero and variance 1. Let $S_0 := 0$ and $S_n = \xi_1 + \dots + \xi_n$ for all $n \geq 1$ be a Random Walk. We consider the continuous embedding of this Random Walk on the unit interval $[0,1]$, i.e. let

$$\widetilde{S_t^N} := \left\{ \frac{S_k + \left(t - \frac{k}{N}\right) \xi_{k+1}}{\sqrt{N}} : t \in \left[\frac{k}{N}, \frac{k+1}{N}\right], k = 0, 1 \dots, N-1 \right\}$$

**Theorem 7.7** (Donsker's invariance Theorem)**.** $\widetilde{S^N}$ *converges in law to a standard Brownian Motion as* $N \to \infty$ *(in* $\mathcal{C}$*). This means that for any continuous and bounded function* $F : \mathcal{C} \to \mathbb{R}$ *we have*

$$\mathbb{E}\left(F\left(\widetilde{S^N}\right)\right) \to \mathbb{E}_0(F), \text{ as } N \to \infty$$

*Where* $\mathcal{C} = \{\omega : [0,1] \to \mathbb{R} \mid \text{continuous and } \omega_0 = 0\}$ *is a complete (separable) metric space with respect to the* $\rho(\omega, \nu) = \sup_{t \in [0,1]} \mid \omega_t - \nu_t|$.

**Remark 7.9.** Donsker's Theorem is an invariance theorem because it doesn't depend on the initial law we pick for the i.i.d. sequence $\xi$ that builds our Random Walk.

**Example 7.1.** $F(\omega) := \sup_{t \in [0,1]} \omega_t$ *is a continuous function on* $\mathcal{C}$*, thus by Donsker's invariance Theorem we obtain that*

$$\sup_{t \in [0,1]} \widetilde{S_t^N} = \frac{1}{\sqrt{N}} \sup_{k \leq N} (S_k) \implies \sup_{t \in [0,1]} B_t \stackrel{d}{=} |\mathcal{N}(0,1)|$$

**Example 7.2.** *Let* $F(\omega) = \omega_1$ *be defined on* $\mathcal{C}$*, then we obtain the classical CLT because*

$$F(\widetilde{S^N}) = \widetilde{S_1^N} = \frac{S_N}{\sqrt{N}}$$

*and consequently*

$$\frac{S_N}{\sqrt{N}} \implies \mathcal{N}(0,1), \text{ as } N \to \infty.$$