3. Basic things from Markov processes and martingales

We will now start generalising the Brownian motion to a larger class of processes that can be restricted to live inside of given domain. For this purpose we will introduce first the Markov processes. We will be refining the defintion and properties of Markov processes more refined ways later on, but let's start with a simpler terms first (as always).

3.1. Markov process.

First we define the Markov property.

If the set of times $T = \mathbb{N}$ and S is countable, then the Markov process is a *Markov chain*. In this case we can define the Markov property in a form

$$\begin{aligned}\mathbf{P}\left(X_{n+m} \in A \,|\, \mathscr{H}_n \right) &= \sum_{j\in A} \mathbf{P}\left(X_{n+m} = j \,|\, \mathscr{H}_n \right) \\ &= \sum_{j\in A} \mathbf{P}\left(X_{n+m} = j \,|\, X_n \right) = \mathbf{P}\left(X_{n+m} \in A \,|\, X_n \right)\end{aligned}$$

for every $A \subset S$. Here $\mathscr{H}_n = \sigma(\{ X_m : \ m \leq n \})$ is the *history*[16] of the process X at time n. We will need a general concept of that captures the essential part of the history. This property is that the history increases with time. Everything that is history-like will be called *filtration*.

3.1. **Definition.** A family $(\mathscr{F}_t; t \in T)$ indexed by time is a *filtration*, if $\mathscr{F}_t \subset \mathscr{F}$ is a sub-σ-algebra for every $t \in T$ and $\mathscr{F}_s \subset \mathscr{F}_t$ when ever $s \leq t$. If there is no confusion of the set of timest T, we will denote a filtration by $(\mathscr{F}_t)$.

Now we can define the Markov property in general.

3.2. **Definition.** A stochastic process (X_t) has a *Markov property with respect to a filtration* $(\mathscr{F}_t)$, if for every time instances $t \geq s$ and every $A \in \mathscr{S}$ we have

$$\mathbf{P}\left(X_t \in A \,|\, \mathscr{F}_s \right) = \mathbf{P}\left(X_t \in A \,|\, X_s \right) \quad \text{a.s.} \tag{3.3}$$

During this course we will almost explicitly be considering the so called *time stationary* Markov processes, so let's define that.

3.4. **Definition.** A stochastic processes (X_t) has a *time stationary Markov property with respect to a filtration* $(\mathscr{F}_t)$ if for every time instances $t \geq s$, for every $x \in S$ and for every $A \in \mathscr{S}$ we have

$$\mathbf{P}_x\left(X_t \in A \,|\, \mathscr{F}_s \right) = \mathbf{P}_{X_s}\left(X_{t-s} \in A \right) \quad \text{a.s.} \tag{3.5}$$

[16]loosely, everything we can talk about the process X before and at time n

3.6. *Remark.* The notation $\mathbf{P}_x(A)$ should be read as $\mathbf{P}\left(A \,|\, X_0 = x \right)$. This way the time stationary Markov property reads as

$$\mathbf{P}\left(X_t \in A \,|\, X_0 = x, \mathscr{F}_s \right) = \mathbf{P}\left(X_{t-s} \in A \,|\, X_0 = y \right)|_{y=X_s} \quad \text{a.s.}$$

This way the subscript should be read as *the starting point* of the process. So the time stationary Markov property means that we can think that we restart the time and treat the current point $X_s(\omega)$ as a new starting point.

Now the first attempt to a definition of a Markov process.

3.7. **Preliminary definition.** By a Markov process $\mathcal{M}$ we mean a triplet $\mathcal{M} = (X_t, \mathscr{F}_t, \{\mathbf{P}_x\})$, where (X_t) is a stochastic process, which has *a time stationary* Markov property with respect to a filtration $(\mathscr{F}_t)$ and with respect to every probability measure $\mathbf{P}_x$, when $x \in S$. We will also require that for every $x \in S$ we have $\mathbf{P}_x\left(X_0 = x \right) = 1$ and that the mapping $x \mapsto \mathbf{P}_x\left(A \right)$ is $\mathscr{S}$-measurable for every $A \in \sigma(\mathscr{F}_t; t \geq 0)$.

3.8. *Remark.* Since the previous definition is bit "heavy", we will just call (X_t) as a Markov process and only if *really needed*, we will specify the filtration and the probability measures we are using. So only when it is necessary to separate the process and the triplet $\mathcal{M}$, we will use the above vocabulary.

As you might have known (or guessed), Brownian motion is a Markov process.

3.9. **Theorem.** *The Brownian motion is (a time stationary) Markov process with respect to its history $\mathscr{H}_t = \sigma(\{ B_s : s \leq t \})$.*

Proof. This just generalises the one previous computation and for simplicity, let's just prove it for 1-dimensional case. We defined $\mathbf{P}_x\left(A \right) = \mathbf{P}\left(A \,|\, B_0 = x \right)$, so clearly $\mathbf{P}_x\left(B_0 = x \right) = 1$. We know that the Brownian motion has independent increments which means that

$$B(t) - B(s) \perp\!\!\!\perp \mathscr{H}_s$$

Since

$$\mathbf{P}_y\left(B(t) \leq x \,|\, \mathscr{H}_s \right) = \sum_{j \in \mathbb{Z}} \mathbf{P}_y\left(j2^{-n} \leq B(s) < (j+1)2^{-n}, B(t) \leq x \,|\, \mathscr{H}_s \right),$$

we can estimate

$$\begin{aligned}\mathbf{P}_y\left(B(t) \leq x \,|\, \mathscr{H}_s \right) &\leq \sum_{j \in \mathbb{Z}} \mathbf{P}_y\left(B(s) \in I(j,n), B(t) - B(s) \leq x - j2^{-n} \,|\, \mathscr{H}_s \right) \\ &= \sum_{j \in \mathbb{Z}} [\, B(s) \in I(j,n) \,] \mathbf{P}\left(B(t) - B(s) \leq x - j2^{-n} \right)\end{aligned}$$

Since $B(t) - B(s) \sim B(t-s) - B(0)$, we have

$$\begin{aligned}\mathbf{P}_y\left(\, B(t) \le x \,|\, \mathscr{H}_s \,\right) &\le \sum_{j\in\mathbb{Z}} [\, B(s) \in I(j,n)\,] \mathbf{P}_{j2^{-n}}\left(\, B(t-s) \le x \,\right) \\ &= \mathbf{P}_{B^n(s)}\left(\, B(t-s) \le x \,\right)\end{aligned}$$

where

$$B^{(n)}(s) = \sum_j [\, B(s) \in I(j,n)\,] j2^{-n} \le B(s)$$

When $n \to \infty$, the random variables $B^{(n)}(s) \uparrow B(s)$ almost surely. Therefore, by continuity of distribution function[17] we obtain $\mathbf{P}_y\left(\, B(t) \le x \,|\, \mathscr{H}_s \,\right) \le \mathbf{P}_{B(s)}\left(\, B(t-s) \le x \,\right)$. Estimating from above gives the converse inequality. The remaining requirement of $x \mapsto \mathbf{P}_x\left(\, A \,\right)$ is $\mathscr{B}(\mathbb{R}_+)$-measurable is left as an excercise. □

We notice that it really was the Markov property we used in constructing the Brownian motion in the first place. We can, therefore, state and split the actual calculation as an excercise that

3.10. *Remark.* Let (X_t) be a Markov process. Suppose that for every bounded and measurable function f, for every $x \in S$ and for every time instance $t \ge 0$ we know the values

$$P_t^{(X)} f(x) := \mathbf{E}_x\, f(X_t)\,.$$

Then the mappings $(P_t^{(X)})$ form a semigroup[18] on bounded and measurable functions and that the mappings $(P_t^{(X)})$ determine uniquely the finite dimensional marginals[19] of the process (X_t) given $\{X_0 = x\}$ for any $x \in S$.

How does the previous remark work for the Brownian motion ? Now

$$\begin{aligned}P_t^{(B)} f(x) = \mathbf{E}_x\, f(B_t) \; = \mathbf{E}_0\, f(x + B_t) \; &= \int_{\mathbb{R}^d} f(x+y) q_t(y)\,\mathrm{d}y \\ &= \int_{\mathbb{R}^d} f(y) q_t(y-x)\,\mathrm{d}y\end{aligned}$$

where q_t is the density function[20] of the d-dimensional Gaussian random variable $B_t \sim \mathfrak{N}(0, tI)$. So the function

$$p_t(x,y) := q_t(y-x) = (2\pi t)^{-d/2} \exp\left(-\tfrac{1}{2}|x-y|^2/t\right). \qquad (3.11)$$

determines uniquely the Brownian motion. Let's give names to these functions.

[17] the distribution function of a $\mathbb{R}_+$-valued random variable X is $F(x) = \mathbf{P}\left(\, X \le x \,\right)$

[18] by semigroup we mean that $P_t^{(X)} P_s^{(X} f(x) = P_{t+s}^{(X)} f(x)$ ja $P_0^{(X)} f(x) = f(x)$

[19] i.e. the measures $(A_1, \ldots, A_n) \mapsto \mathbf{P}_x\left(\, \forall j : X(t_j) \in A_j \,\right)$

[20] the density function is the Radon–Nikodym derivative of the distribution with respect to the Lebesgue measure, if it happens exists

3.12. **Definition.** Let (X_t) be a Markov process and let μ be a measure on the state space S. The mappings $(P_t^{(X)})$ are called as *transition probability operators*, if every $P_t^X f$ is bounded and measurable, when ever f is bounded and measurable. The function $p_t^{(X)}(x,y)$, that satisfies the condition

$$P_t^{(X)} f(x) = \int_S p_t^{(X)}(x,y) f(y) \mu(\,\mathrm{d}y)$$

for every bounded and measurable f, for every $t > 0$ and for μ-a.e.. points $x \in S$ and is measurable in x, is called as the *transition probability density* with respect to measure μ.

For Brownian motion the function p_t in equation (3.11) is, therefore, the transition probability density of the Brownian motion with respect to Lebesgue measure.

We notice that the mapping p_t is a C^∞-function,[21] when $t > 0$ with respect to to all the variables x, y and t. Moreover,

$$\begin{aligned}
\partial_t p_t(x,y) &= p_t(x,y)\partial_t \log p_t(x,y) = (-\tfrac{1}{2}t^{-1}d + \tfrac{1}{2}|x-y|^2 t^{-2})p_t(x,y), \\
\partial_{x_j} p_t(x,y) &= p_t(x,y)\partial_{x_j} \log p_t(x,y) = -(x_j - y_j)t^{-1}p_t(x,y), \\
\partial^2_{x_j} p_t(x,y) &= p_t(x,y)\left((x_j - y_j)^2 t^{-2} + \partial_{x_j}\ \log p_t(x,y)\right) \\
&= 2p_t(x,y)\left(\tfrac{1}{2}(x_j - y_j)^2 t^{-2} - \tfrac{1}{2}t^{-1}\right) \\
\triangle_x p_t(x,y) &= 2p_t(x,y)\sum_{j=1}^d \left(\tfrac{1}{2}(x_j - y_j)^2 t^{-2} - \tfrac{1}{2}t^{-1}\right) = 2\partial_t p_t(x,y).
\end{aligned}$$

Thus, we have found out that the mapping p_t satisfies the *heat equation* $\partial_t p_t(x,y) = \frac{1}{2}\triangle_x p_t(x,y) = \frac{1}{2}\triangle_y p_t(x,y)$ for every $t > 0$ and for every $x, y \in \mathbb{R}^d$. Furthermore, we notice that (Excercise) that if f is continuous and bounded, then

$$\lim_{t\to 0} P_t^{(B)} f(x) = f(x)$$

The mapping p_t is therefore called as *the heat kernel*, since it satisfies the following *singular* initial value problem for heat equation

$$\begin{cases} \partial_t p_t &= \frac{1}{2}\triangle_y p_t, \quad \text{when } t > 0 \\ p_0 &= \delta_x, \end{cases}$$

where δ_x is the *Dirac point mass* at the point $x \in \mathbb{R}^d$. One good way of understanding this singular equation is to approach it directly via the transition probability operators, since by the Lebesgue dominated convergence the

[21]i.e. infinitely differentiable

function $u_t = P_t^{(B)} f$ also is a solution to an initial value problem for the heat equation

$$\begin{cases} \partial_t u_t &= \frac{1}{2}\triangle_y u_t, \quad \text{when } t > 0 \\ u_0 &= f. \end{cases}$$

So, if we are given a transition probability density, we can construct the Markov process starting from that. If we don't have a density, it is still extremely helpful to talk about *transition functions* that will still provide the transition probability operators as integrals.

3.13. **Definition.** A function $(t, x, B) \mapsto P_{t,x}(B)$ on $[0, \infty) \times S \times \mathscr{S}$ is called a *transition function* if

i) for all $t \geq 0$ and $x \in S$, the function $B \mapsto P_{t,x}(B)$ is a probability measure on $(S, \mathscr{S})$ and $P_{0,x} = \delta_x$

ii) for all $t \geq 0$ and $B \in \mathscr{S}$, the function $x \mapsto P_{t,x}(B)$ is $\mathscr{S}$-measurable

iii) for all $t \geq 0$, $s \geq 0$, $x \in S$ and $B \in \mathscr{S}$ it satisfies

$$P_{t+s,x}(B) = \int_S P_{t,x}(\,\mathrm{d}y) P_{s,y}(B)$$

The last condition *iii*) is called *Chapman–Kolmogorov* equation. For Brownian motion we notice that

$$P_{t,x}^{(B)}(A) = \int_A p_t(x, y)\,\mathrm{d}y$$

is the transition function. Also we notice that the transition probability operator can be expressed with the transition function

$$P_t^{(B)} f(x) = \int_{\mathbb{R}^d} f(y) P_{t,x}(\,\mathrm{d}y).$$

So, even if we don't necessarily have the transition probability density, we still are going to have something almost nice.

3.14. **Lemma.** *Suppose $(P_{t,x})$ is a transition function. Then there exists a Markov process $\mathcal{M} = (X_t, \mathscr{H}_t, \mathbf{P}_x)$ such that $P_{t,x}(A) = \mathbf{P}_x\left(X_t \in A \right)$ for every $t \geq 0$ and every $A \in \mathscr{S}$.*

Proof. We leave the constrution of $\mathbf{P}_x$ and the process (X_t) and the condition that $\mathbf{P}_x(X_0 = x) = 1$ as an excercise. These follow from the Kolmogorov extension theorem.[22]. So we only need to define the finite dimensional marginals of $\mathbf{P}_x$, which are given by

$$\mathbf{P}_x(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n) = \int_{A_1} P_{t_1,x}(\,\mathrm{d}y_1) \ldots \int_{A_n} P_{t_n - t_{n-1}, y_{n-1}}(\,\mathrm{d}y_n)$$

[22]Recall we assume that S is a Borel subset, which is needed for the Kolmogorov extension theorem

for every $t_1 < t_2 < \cdots < t_n$ and every $A_j \in \mathscr{S}$.

We will still need to verify the measurability $x \mapsto \mathbf{P}_x(A)$ for every $A \in \mathscr{F}_\infty$ and the Markov property, i.e.

$$\mathbf{P}_x\left(A, X_t \in B\right) = \mathbf{E}_x\left[A\right]\mathbf{P}_{X_s}\left(X_{t-s} \in B\right)$$

holds for every x, for every $0 \le s < t$, for every $A \in \mathscr{F}_s$ and for every $B \in \mathscr{S}$.

Since we cannot directly describe *all the elements* in σ-algebras $\mathscr{F}_\infty$ and $\mathscr{F}_s$, we will first show something less, by proving it for a *certain special* sets that we can describe: $A = \{X_{t_j} \in C_j, j = 1, \ldots, n\}$, where $n \in \mathbb{N}$, time instances $0 \le t_1 < \cdots < t_n$ and $C_j \in \mathscr{S}$. Let's denote the family of these kinds of sets by $\mathscr{A}_\infty$ and if, in addition, $t_n \le t$, then by $\mathscr{A}_t$. We assume that we know[23] that $\mathscr{F}_\infty = \sigma(\mathscr{A}_\infty)$ and $\mathscr{F}_t = \sigma(\mathscr{A}_t)$. We also consider known that both $\mathscr{A}_\infty$ and $\mathscr{A}_t$ are *algebras.*[24]

When we know the claimed measurability and Markov property for these sets, we will show that the claims hold for the full σ-algebras as well.

Let's start with the measurability. So, we want to show that

$$x \mapsto \mathbf{E}_x\, f_1(X_{t_1}) \ldots f_n(X_{t_n})$$

is $\mathscr{S}$-measurable for every n, for every sequence of times $0 \le t_1 < \cdots < t_n$ and for every bounded and measurable f_j which makes the claim bit more general, but easier to prove. The orginal claim would then follow by choosing $f_j = [C_j]$.

Suppose we know that

$$\mathbf{E}_x\, f_1(X_{t_1}) \ldots f_n(X_{t_n}) = P_{t_1} g_{n,1}(x)$$

when

$$g_{n,j} = f_j P_{s_j} g_{n,j+1}, \quad s_j = t_{j+1} - t_j \quad \text{ja} \quad g_{n,n} = f_n.$$

We will (again) omit the details of this and leave it to excercises. Now since f_j is measurable for every j and $P_{s_j} g$ is measurable for measurable g for every j, therefore every $g_{n,j}$ is measurable by induction and hence the claimed measurability follows.

Simililarly the Markov property claim is

$$\mathbf{P}_x\left(X_{t_j} \in C_j, j \le n\right) = \mathbf{E}_x\left[X_{t_j} \in C_j, j < n\right]\mathbf{P}_{X_{t_{n-1}}}\left(X_{t_n - t_{n-1}} \in C_n\right)$$

This follows from the identity

$$P_{t_1} g^{(1)}_{n,1} = P_{t_1} g^{(2)}_{n-1,1}$$

[23] and this is quite easy excercise type material

[24] just additive, not necessarily σ-additive

where

$$g^{(1)}_{n,j} = [\,C_j\,]P_{s_j} g^{(1)}_{n,j+1}, \quad \text{and} \quad g^{(1)}_{n,n} = [\,C_n\,]$$

and

$$g^{(2)}_{n-1,j} = [\,C_j\,]P_{s_j} g^{(2)}_{n-1,j+1}, \quad \text{and} \quad g^{(2)}_{n-1,n-1} = [\,C_{n-1}\,]P_{s_{n-1}}[\,C_n\,].$$

So, we notice that the claim follows if

$$g^{(1)}_{n,n-1} = g^{(2)}_{n-1,n-1}.$$

But this is immediate, since

$$g^{(2)}_{n-1,n-1} = [\,C_{n-1}\,]P_{s_{n-1}} g^{(1)}_{n,n} = g^{(1)}_{n,n-1}.$$

In the end of the proof, we will show that the measurability can be extended from the algebra $\mathscr{A}_\infty$ to the σ-algebraa $\mathscr{F}_\infty$. This follows from the Dynkin's π-λ Theorem, as we will soon see.

Denote

$$\mathscr{G} = \{\, A \in \mathscr{F}_\infty : \ x \mapsto \mathbf{P}_x(A) \text{ is measurable } \}$$

so we are now *expressing the wanted claim* as $\mathscr{G} = \mathscr{F}_\infty$. We already know that $\mathscr{A}_\infty \subset \mathscr{G} \subset \mathscr{F}_\infty = \sigma(\mathscr{A}_\infty)$. If we manage to show that $\mathscr{G}$ is a *Dynkin system*[25] then the Dynkin's π-λ Theorem says that $\sigma(\mathscr{A}_\infty) \subset \mathscr{G}$, which proves the claim.

First of all, $S \in \mathscr{G}$, since $S \in \mathscr{A}_\infty$. Moreover, if $A \in \mathscr{G}$, then $x \mapsto 1 - \mathbf{P}_x(A) = \mathbf{P}_x(A^C)$ is measurable and thus, $A^C \in \mathscr{G}$. Let $(A_i) \subset \mathscr{G}$ and suppose $A_i \cap A_j = \emptyset$ for every $i \neq j$ and denote the union of these sets by A. By σ-additivity

$$x \mapsto \mathbf{P}_x(A) = \sum_j \mathbf{P}_x(A_j)$$

which is measurable as a sum of measurable functions. So $A \in \mathscr{G}$ and we have verified that $\mathscr{G}$ is a Dynkin system.

Extending the Markov property is done with the same method, so let's omit that. □

3.15. *Remark.* We will later on update our idea of Markov processes and this is still preliminary.

[25]a family $\mathscr{G}$ is a Dynkin system on the set S, if $S \in \mathscr{G}$, $S^C \in \mathscr{G}$ and it is closed with respect to countable unions of *disjoint* sets $(A_i) \subset \mathscr{G}$ i.e. $\bigcup A_i \in \mathscr{G}$, if $A_i \cap A_j = \emptyset$ for every $i \neq j$

3.2. **Martingales.** Before returning back to Markov processes and specific type Markov processes we will jump to another property of a Brownian motion namely the martingale proporety.

Where as the Markov proporety gives a method to describe the future evolution based on simply knowing the "current" state, the martingale property tells that what we *expect to happen* in future, given the history, is that nothing really changes.

So we define

3.16. **Definition.** Let (X_t) be a real valued stochastic process and $(\mathscr{F}_t)$ a filtration. We say that X is a *martingale* with respect to filtration $(\mathscr{F}_t)$ (or $(\mathscr{F}_t)$-martingale), if $\mathbf{E}\,|X|_t < \infty$ for every $t \in T$, process X is $(\mathscr{F}_t)$-adpated[26] and

$$X_t = \mathbf{E}\,(X_s \,|\, \mathscr{F}_t) \tag{3.17}$$

for every time instance $s > t \in T$. If the equality $=$ is replaced in condition (3.17) by inequality $\leq$ (respectively $\geq$), then this process is called as a *submartingale* (respectively *supermartingale*).

3.18. *Remark.* Note that this means that for martingales the mean function $\mathbf{E}\,X_t$ is constant, for submartingales it is an increasing function and for supermartingales it is a decreasing function.

First some examples of martingales (namely, $1D$-Brownian motion and some of its functions).

3.19. **Theorem.** *Brownian motion is a martingale with respect to its history.*

Proof. So we know already that B is adapted and integrable. And since it has independent increments,

$$\mathbf{E}\,(B_s \,|\, \mathscr{H}_t) = \mathbf{E}\,(B_s - B_t + B_t \,|\, \mathscr{H}_t) = B_t + \mathbf{E}\,(B_s - B_t) = B_t.$$

□

3.20. **Example.** A process $Y_t := B_t^2$ is clearly not a martingale, since, for example, $t \mapsto \mathbf{E}\,Y_t = t$ is not a constant but it is increasing function. However, it is a submartingale with respect to the history of Brownian motion, since it's clearly adapted, integrable and moreover,

$$\mathbf{E}\,\left(B_s^2 \,|\, \mathscr{H}_t\right) = \mathbf{E}\,\left((B_s - B_t)^2 + 2B_sB_t - B_t^2 \,|\, \mathscr{H}_t\right) \geq 2B_t\mathbf{E}\,(B_s \,|\, \mathscr{H}_t) - B_t^2$$

where we just estimated the positive term away and hence $Y_t \leq \mathbf{E}\,(Y_s \,|\, \mathscr{H}_t)$.

[26]adapted means X_t is $\mathscr{F}_t$-measurable for every t

Moreover, the process $X_t := Y_t - t$ is a *martingale* with respect to the history of Brownian motion. Again, X is adapted and $\mathbf{E}\,|X|_t \le t + \mathbf{E}\,Y_t = 2t < \infty$. Now we have to actually compute the omitted term in the previous calculation, but by independence of increments

$$\mathbf{E}\,\Big((B_s - B_t)^2 \,|\, \mathscr{H}_t\Big) = \mathbf{E}\,(B_s - B_t)^2 = s - t.$$

Collecting these, we therefore have,

$$\mathbf{E}\,(Y_s \,|\, \mathscr{H}_t) = (s-t) + Y_t = s + X_t$$

or $\mathbf{E}\,(X_s \,|\, \mathscr{H}_t) = X_t$.

3.21. **Example.** If f is any convex function and X is a martingale, then $f(X_t)$ is a submartingale if in addition we know that $\mathbf{E}\,|f(X_t)| < \infty$ for every $t \in T$. This already shows that $|B(t)|$, $B(t)_+$, $e^{B(t)}$, etc. are submartingales, which however are not martingales.

3.22. *Remark.* In the sequal we will only talk about super- tai submartingales. It is good to notice that if (X_t) is a submartingale, then $(-X_t)$ is a supermartingale and vice versa.

Martingales have so many good properties and one of the most important is the good integration theory so we will be collecting these properties during the road. In addition, the martingales are suitably well behaved and have very good limit properties.

Quite a many results concerning martingales are first formulated to discrete time martingales (like the *Upcrossing inequality* ja *Doob's maximal-L^p-inequalities*, for instance in "Stokastiset differentiaaliyhtälöt" -course material), but we will in the sequel formulate these with countable time sets as well as *right continuous processes*. This assumption on right continuity will be repeated many times later on.

Upcrossing inequality tells that a supermartingale will not *cross* a given interval unboundedly many times in a bounded time interval. We define the *upcrossing number* $U(f, T, [a, b])$ as:

$$U(f, T_F, [a, b]) = \#\{\text{upcrossings of } [a, b] \text{ a function } f \text{ makes on a set } T_F\}$$

$$\text{and} \quad U(f, T, [a, b]) = \sup\{\, U(f, T_F, [a, b]) : \; T_F \subset T \,\}$$

$$\text{when } T_F = \{t_1 < \cdots < t_d\}$$

By an upcrossing on a set T_F we mean the time instances $a_1 < b_1 < \cdots < a_n < b_n \subset T_F$, that are defined recursively as

$$a_{k+1} = \min\{\, t_j > b_k : \; f(t_j) < a \,\}, \quad b_{k+1} = \min\{\, t_j > a_{k+1} : \; f(t_j) > b \,\}$$

3.23. **Lemma** (Upcrossing inequality)**.** *Let (X_t) be supermartingale with respect to filtration $(\mathscr{F}_t)$. Let $T_0 \subset T = [0, \infty)$ be countable and dense. Then*

$$(b-a)\mathbf{E}\, U(X, T_0, [a,b]) \;\leq \sup_{t\in T_0} \mathbf{E}\,(X_t + a)^-$$

We will still introduce Joseph Leo Doob's martingale inequalities.

3.24. **Theorem** (Doob's maximal-L^p-inequalities)**.** *Let (X_t) be supermartingale with respect to filtration $(\mathscr{F}_t)$. Let $T_0 \subset T \subset [0, \infty)$ be a countable and dense. Then*

$$\lambda^p \mathbf{P}\left(\sup_{t\in T_0} |X_t| \geq \lambda\right) \leq \sup_t \mathbf{E}\, X_t^- \,.$$

Moreover, when $p > 1$, we have

$$\mathbf{E} \sup_{t\in T_0} |X_t|^p \;\leq \left(\frac{p}{p-1}\right)^p \sup_t \mathbf{E}\,|X_t|^p \,.$$

These generalize nicely to the continuous case.

3.25. **Corollary** (Doob's maximal-L^p-inequalities (part II))**.** *If (X_t) is right continuous supermartingale and $T \subset \mathbb{R}$ is an interval and $X_t^* = \sup_{s\leq t}|X_s|$, then for every $t \in T$ it holds that*

$$\lambda^p \mathbf{P}\,(\, X_t^* \geq \lambda\,) \leq \mathbf{E}\,|X_t|^p \,.$$

Furthermore, when $p > 1$, we have for every $t \in T$

$$\mathbf{E}\,(X_t^*)^p \;\leq \left(\frac{p}{p-1}\right)^p \mathbf{E}\,|X_t|^p \,.$$

With the help of these results we can say somehting about the paths of supermartingales, namely that they have limits from both directions.

3.26. **Theorem.** *Let (X_t) be a supermartingale with respect to filtration $(\mathscr{F}_t)$. Let $T_0 \subset T = [0, \infty)$ be countable and dense. Then with probability 1 it holds: for every $t \in [0, \infty)$*

$$X(t^+, \omega) = \lim_{T_0 \ni s \downarrow t} X(s, \omega)$$

the limit exists and for every $t \in (0, \infty)$

$$X(t^-, \omega) = \lim_{T_0 \ni s \uparrow t} X(s, \omega)$$

the limit exists. Moreover, for any bounded interval I, we have

$$\sup_{t\in I\cap T_0} |X_t| \;< \infty$$

almost surely.

Proof. Suppose that $t \in T$ is a time instance, that has no right (or left) limits. We will show that this event, say A, has a probability 0. Since the treatment from left and right are analogous, let's just consider the limit from right. Since the limit does not exist, we have

$$\liminf_{T_0 \ni s \downarrow t} X(s,\omega) < a < b < \limsup_{T_0 \ni s \downarrow t} X(s,\omega)$$

for some *rational* numbers $a < b$. This means also that the upcrossing number $U(f, T_0, [a,b]) = \infty$, where $f = s \mapsto X(s,\omega)$. We have to consider that these a and b may depend on the elementary event ω, so we taking this account we have

$$\mathbf{P}\,(\,A\,) \le \mathbf{P}\,(\,B(a,b) \text{ for some } \; a,b \in \mathbb{Q}\,) = \sum_{a,b \in \mathbb{Q}} \mathbf{P}\,(\,B(a,b)\,)$$

where

$$B(a,b) = \{U(X, T_0, [a,b]) = \infty\}$$

But now the Upcrossing inequality says that $\mathbf{P}\,(\,B(a,b)\,) = 0$, which implies that $\mathbf{P}\,(\,A\,) = 0$.

The second part (the boundedness) follows almost directly from Doob's maximal inequality, when $p = 1$. □

As a corollary we find that

3.27. **Corollary.** *If (X_t) is a supermartingale with respect to filtration $(\mathscr{F}_t)$ and it is right continuous, then it has left limits everywhere and it is bounded on bounded intervals, almost surely.*

Proof. This will be split as an excercise. □

Let's end the first part of martingales to the important convergence result.

3.28. **Theorem** (Martingale Convergence Theorem)**.** *Suppose (X_t) is a right continuous supermartingale, that satisfies*

$$A := \sup_t \mathbf{E}\, X_t^- \; < \infty.$$

Then there exists a random variable X_∞, such that $X_t \to X_\infty$ as $t \to \infty$ almost surely and the limit random variable X_∞ is integrable.

Proof. The condition together with Upcrossing inequality provides us

$$(b-a)\mathbf{E}\, U(X, T_0, [a,b]) \;\; \le a + A,$$

where $T_0 \subset T = \mathbb{R}$ is dense and countable. From right continuity we can deduce (but let's omit the exact details), that

$$(b-a)\mathbf{E}\, U(X, T, [a,b]) \;\; \le a + A.$$

Using the same technique as in the proof of Theorem 3.26 we can therefore deduce that

$$\mathbf{P}\left(\liminf_{t\to\infty} X_t < \limsup_{t\to\infty} X_t\right) = 0.$$

But this already says that X_∞ exists almost surely, but it might possibly be $\pm\infty$. On the other hand the Fatou's Lemma gives

$$\mathbf{E}\, X_\infty^- \le \liminf_{t\to\infty} \mathbf{E}\, X_t^- \le A.$$

Moreover, the supermartingale property gives an estimate

$$\mathbf{E}\, X_t^+ = \mathbf{E}\, X_t + \mathbf{E}\, X_t^- \le \mathbf{E}\, X_t^- + \mathbf{E}\, X_0$$

and hence Fatou's Lemma again gives

$$\mathbf{E}\, X_\infty^+ \le \liminf_{t\to\infty} \mathbf{E}\, X_t^- + \mathbf{E}\, X_0 \le A + \mathbf{E}\, X_0 < \infty$$

so the limit random variable $|X_\infty| = X_\infty^+ + X_\infty^-$ is also integrable. □

We will return again to martingales few times later on since we need the stopping times before we can introduce the Optional Stopping Theorem, which is fundamental for the later treatment.