Lectures on 1.4. and 3.4. 2008

2.6.2. *Histories and change of measure.* Let $Q \prec\prec \mathbb{P}$, $Y = \frac{dQ}{d\mathbb{P}}$, and $\mathbb{F} = (F_n)_{n\geq 0}$ is a history, $F_0 = \{\emptyset, \Omega\}$. Put $y_n = \mathbb{E}_{\mathbb{P}}[Y|F_n]$. Then $(y_n, F_n)_{n\geq 0}$ is a uniformly integrable martingale.
Let be $Q_n = Q|F_n$ the restriction of the measure Q to the σ-algebra F_n and respectively $\mathbb{P}_n = \mathbb{P}|F_n$. Obviously we have $Q_n \prec\prec \mathbb{P}_n$, and let $z_n = \frac{dQ_n}{d\mathbb{P}_n}$ be the corresponding R-N- derivatives. We shall show that $z_n = y_n$ $\mathbb{P}$-almost surely, $n \geq 0$. To show this, let $F \in F_n$ and we obtain

$$Q_n(F) = Q(F) = \int_F Y d\mathbb{P} = \int_F y_n d\mathbb{P},$$

where the last equality follows from the definition of the conditional expectation. On the other hand

$$\int_F y_n d\mathbb{P} = Q_n(F) = \int_F z_n d\mathbb{P},$$

and because the R-N- derivative is $\mathbb{P}$- almost surely unique, and we have proved

Theorem 2.14. *Let $Q \prec\prec \mathbb{P}$, $Y = \frac{dQ}{d\mathbb{P}}$ and $y_n = \mathbb{E}_{\mathbb{P}}[Y|F_n]$. Then $(y_n, F_n)_{n\geq 0}$ is uniformly integrable martingale and $y_n = \frac{dQ_n}{d\mathbb{P}_n}$.*

Theorem 2.15. *Let $Q \prec \mathbb{P}$ on $(\Omega, \mathcal{F})$ and $Y = \frac{dQ}{d\mathbb{P}}$ is Radon-Nikodym-derivative. Then*

- $X \in \mathbb{L}^1(Q) \Leftrightarrow XY \in L^1(\mathbb{P})$
- *If $X \in \mathbb{L}^1(Q)$, then*

$$\mathbb{E}_Q X = \mathbb{E}_P(YX). \tag{2.27}$$

Proof If $X \geq 0$ is a simple random variable, $X = \sum_{k=1}^n a_k I_{A_k}$, then we get directly from definitions that

$$\begin{aligned} \mathbb{E}_Q X &= \sum_k a_k Q(A_k) = \sum_k a_k \int (Y I_{A_k})\, d\mathbb{P} \\ &= \int \left(Y \left(\sum_k a_k I_{A_k} \right) \right) d\mathbb{P} = \mathbb{E}_{\mathbb{P}}(YX). \end{aligned}$$

If $X \geq 0, X \in \mathbb{L}^1(Q)$ and X^n is an increasing sequence of simple random variables and $X = \lim_n X^n$, then $Y^n = YX^n$ is an increasing sequence of random variable and $YX = \lim_n Y^n$.
From the definition of expectation and MCT we obtain

$$\mathbb{E}_Q X = \lim_n \mathbb{E}_Q X^{(n)} = \lim_n \mathbb{E}_{\mathbb{P}} Y^{(n)} = \mathbb{E}_{\mathbb{P}}(YX).$$

Hence, if $X \geq 0$, then $X \in \mathbb{L}^1(Q)$ if and only if $YX \in \mathbb{L}^1(\mathbb{P})$ and (2.27) is valid.
The general case follows $X = X^+ - X^-$. $\square$

Distributions of real line. Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathbb{P}_X$ be the distribution of the random variable X: $\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$, where B is a Borel set .
Let now $Q \prec\prec \mathbb{P}$. Then, if Q_X is the distribution of random variable X with respect to Q, then $Q_X \prec\prec \mathbb{P}_X$: if $\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)) = 0$, then also $Q_X(B) = Q(X^{-1}(B)) = 0$.
Let us assume that $\mathbb{P}_X(B) = \int_B f(x)\mu(dx)$ and $Q_X(B) = \int_B g(x)\mu(dx)$, where μ is a sigma- finite measure on real line. Let us compute $y = \frac{dQ_X}{d\mathbb{P}_X}$: because $Q_X(B) = \int_B g(x)\mu(dx)$ and $Q_X(B) = \int_B y(x)f(x)\mu(dx)$; we observe that if we take $y = \frac{g}{f} I_{\{f>0\}}$ we obtain that

$$Q_X(B) = \int_B g(x)\mu(dx)$$

$$= \int_B \frac{g(x)}{f(x)} I_{\{>0\}}(x) f(x)\mu(dx) = \int_B y(x)f(x)\mu(dx).$$

Because Radon-Nikodym- derivative is unique, we get that $y(x) = \frac{g(x)}{f(x)}$.
In statistics $y = \frac{g}{f}$ is likelihood ratio.

2.7. The space $\mathcal{M}^2(\mathbb{F}, \mathbb{P})$.

2.7.1. *Square integrable martingales.* Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We say that $(M_n, F_n)_{\geq 1}$ is a *square integrable martingale*, if $(M.\mathbb{F})$ is a martingale and $M_k \in L^2(\mathbb{P})$ for all $k \in \mathbb{N}$.
Let X_n, $n \geq 1$,be a process; put $\Delta X_k = X_k - X_{k-1}$.

Lemma 2.5 (Energyequation, Pythagoras). *Let $(X, \mathbb{F}) = (X_k, F_k)_{k\geq 0}$ be a square integrable martingale. Then*

$$\begin{aligned} \mathbb{E}X_n^2 &= \mathbb{E}X_0^2 + \sum_{k=1}^n \mathbb{E}(X_k - X_{k-1})^2 \\ &= \mathbb{E}X_0^2 + \sum_{k=1}^n \mathbb{E}(\Delta X_k)^2. \end{aligned} \tag{2.28}$$

Proof Let $j < k$. Because X is square integrable, we get that $X_j X_k \in L^1(\mathbb{P})$. Since X is a martingale, then by exercise 18.3., problem 6 we get

$$\mathbb{E}[X_j(X_k - X_{k-1})|F_{k-1}] = X_j \mathbb{E}[X_k - X_{k-1}|F_{k-1}] = X_j \cdot 0 = 0;$$

which implies $\mathbb{E}(X_j(X_k - X_{k-1})) = 0$. On the other hand, any sequence Y_n can be written as a telescopic sum

$$Y_n = Y_0 + \sum_{k=1}^n \Delta Y_k,$$

and so we write X_n^2 as

$$X_n^2 = X_0^2 + 2X_0 \sum_{k=1}^n (\Delta X_k) + \sum_{k=1}^n (\Delta X_k)^2 + \sum_{i\neq k} (\Delta X_j \Delta X_k). \tag{2.29}$$

But for $j \geq 1$ we have $\mathbb{E}(X_0(\Delta X_j)) = 0$, and $\mathbb{E}(\Delta X_j \Delta X_k) = 0$, for $j \neq k$. After these observations we obtain (2.28) by taking expectations on the left and right hand side of (2.29). $\square$

The proof of lemma 2.5 was based on the fact that a square integrable martingale has orthogonal increments. Next we give another proof, using the properties of martingale transforms.
Recall the Abel summation formula for two sequences of numbers $a = (a_k)_{0\leq k\leq n}$ and $b = (b_k)_{0\leq k\leq n}$. The Abel summation formula is the identity

$$a_m b_m = a_0 b_0 + \sum_{k=1}^{m} a_{k-1}\Delta b_k + \sum_{k=1}^{m} b_k \Delta a_k. \tag{2.30}$$

Let us utilize the following notation:

$$[a,b]_m = \sum_{k=1}^{m} \Delta a_k \Delta b_k.$$

Now we can write the Abel summation in a symmetric form as follows:

$$a_m b_m = a_0 b_0 + \sum_{k=1}^{m} a_{k-1}\Delta b_k + \sum_{k=1}^{m} b_{k-1}\Delta a_k + [a,b]_m. \tag{2.31}$$

Often one says that the formula (2.31) is the *integration by parts formula* in the discrete time.
Let us return to the proof of the equality (2.28) in Lemma 2.5 Integration by parts formula gives

$$X_n^2 = X_0^2 + 2\sum_{k=1}^{n} X_{k-1}\Delta X_k + [X,X]_n,$$

where $Y_n \doteq \sum_{k=1}^n X_{k-1}\Delta X_k$ is a martingale transform and , the process $C_k \doteq X_{k-1}$ is predictable. Moreover, we get that $X_{k-1}\Delta X_k \in L^1(\mathbb{P})$, because X is a square integrable martingale. Hence $Y = (Y_k)_{1\leq k\leq n}$ is a martingale. Note also that $\mathbb{E}Y_m = 0$, and so

$$\mathbb{E}X_n^2 = \mathbb{E}X_0^2 + \mathbb{E}[X,X]_n,$$

and this the equality (2.29) with our new notation.
Denote by $\mathcal{M}^2(\mathbb{F},\mathbb{P})$ the space of square integrable martingales, bounded in $L^2(\mathbb{P})$: $M \in \mathcal{M}^2(\mathbb{F},\mathbb{P})$ if and only if $\sup_n \mathbb{E}M_n^2 < \infty$.

Theorem 2.16. *Let $(M,\mathbb{F}) = (M_n, F_n)_{n\geq}$ be a square integrable martingale. Then*

$$M \in \mathcal{M}^2 \Leftrightarrow \sum_{k=1}^{\infty} \mathbb{E}\left(\Delta M_k\right)^2 < \infty. \tag{2.32}$$

Moreover, if $M \in \mathcal{M}^2$, then there exists $M_\infty = \lim_n M_n$ $\mathbb{P}$- almost surely. Moreover $M_\infty = L^2(\mathbb{P}) - \lim_n M_n$ and

$$\mathbb{E}M_\infty^2 = \mathbb{E}M_0^2 + \mathbb{E}[M,M]_\infty. \tag{2.33}$$

Proof By Lemma 2.5 $n \to \mathbb{E}M_n^2$ is increasing in n [or directly, because M^2 is a submartingale]. With Fatou lemma we get $\mathbb{E}M_\infty^2 = \lim_n \mathbb{E}M_2^2$ and hence

$$\mathbb{E}M_\infty^2 = \lim_n \mathbb{E}M_n^2 = \mathbb{E}M_0^2 + \sum_{k=1}^{\infty} \mathbb{E}(\Delta M_k)^2,$$

and this gives (2.32).

We have always $\mathbb{E}|M_n| \leq \left(\mathbb{E}M_n^2\right)^{\frac{1}{2}}$, and so from the assumption $M \in \mathcal{M}^2$ it follows that $\sup_n \mathbb{E}|M_n| < \infty$. Martingale convergence theorem implies that there exists M_∞ such that $M_\infty = \lim_n M_n$ $\mathbb{P}$- almost surely [and also in $\mathbb{L}^1(\mathbb{P})$, since the martingale M is uniformly integrable, because it is bounded in $L^2(\mathbb{P})$].

Next we shall show that $M_\infty = \mathbb{L}^2(\mathbb{P}) - \lim_n M_n$. By the energy equality we obtain

$$\mathbb{E}(M_{n+r} - M_n)^2 = \sum_{k=n+1}^{n+r} \mathbb{E}(\Delta M_k)^2 \leq \sum_{k=n+1}^{\infty} \mathbb{E}(\Delta M_k)^2.$$

By the Fatou lemma

$$\begin{aligned}\mathbb{E}(M_\infty - M_n)^2 &= \mathbb{E}\liminf_r (M_{n+r} - M_n)^2 \\ &\leq \liminf_r \mathbb{E}(M_{n+r} - M_n)^2 \leq \sum_{k=n+1}^{\infty} \mathbb{E}(\Delta M_k)^2.\end{aligned}$$

If $n \to \infty$, then $\sum_{k=n+1}^{\infty} \mathbb{E}(\Delta M_k)^2 \to 0$ because the whole series converges, so $M_\infty = \mathbb{L}^2(\mathbb{P}) - \lim_n M_n$. Moreover,

$$\mathbb{E}M_\infty^2 = \lim_n \mathbb{E}M_n^2 = \mathbb{E}M_0^2 + \lim_n \mathbb{E}[M,M]_n$$

and we have proved the equality (2.33). □

The next corollary is useful, when one applies it to sequences of martingales.

Corollary 2.6. *Let $N, M \in \mathcal{M}^2(\mathbb{F}, \mathbb{P})$ such that $N_0 = M_0$. Then*

$$\mathbb{E}[M-N, M-N]_\infty \leq \mathbb{E}\left((M-N)_\infty^*\right)^2 \leq 4\mathbb{E}[M-N, M-N]_\infty. \tag{2.34}$$

Proof If $M, N \in \mathcal{M}^2$ and $N_0 = M_0$, then also $U = M - N \in \mathcal{M}^2$ and $U_0 = 0$. From the equality $\mathbb{E}U_n^2 = \mathbb{E}[U,U]_n$, and from the Doob maximal inequality we get that

$$\mathbb{E}[U,U]_n = \mathbb{E}U_n^2 \leq \mathbb{E}\left((U)_n^*\right)^2 \leq 4\mathbb{E}U_n^2 = 4\mathbb{E}[U,U]_n.$$

We obtain the inequalities in (2.34) from this by letting $n \to \infty$. □

2.8. **Doob decomposition.** The next theorem, again due to Doob, is very useful, and the proof is very simple[6].

Theorem 2.17 (Doob decomposition)**.** *Let $(X_n, F_n)_{n\geq 0}$ be an integrable stochastic process. Then there exists a unique predictable integrable process $A = (A_n, F_n)_{n\geq 0}$ such that*

$$M_n \doteq -X_0 + X_n - A_n \tag{2.35}$$

is martingale. The process $(X_n, F_n)_{n\geq 0}$ is a submartingale if and only if A is non-decreasing: $A_n \leq A_{n+1}$.

[6]The corresponding continuous time version of this theorem is called Doob-Meyer decomposition, and it is one of the most important theorems in stochastic analysis. The proof of the continuous time result is difficult.

Proof Note that due to the telescopic formula one can define discrete time processes by defining their jumps. Define $M_0 = 0$ and if $k \geq 1$ then

$$\Delta M_k \doteq \Delta X_k - \mathbb{E}[\Delta X_k | F_{k-1}];$$

it is easy to see that $\Delta M_k \in \mathbb{L}^1(\mathbb{P})$, and $\mathbb{E}[\Delta M_k | F_{k-1}] = 0$, and so ΔM_k is a martingale difference sequence. Hence $M_n = \sum_{k=1}^n \Delta M_k$ is a martingale. In order to have the equality (2.35), we put $A_0 = 0$ and

$$A_n = \sum_{k=1}^{n} \mathbb{E}[\Delta X_k | F_{k-1}];$$

clearly $A_n \in \mathbb{L}^1(\mathbb{P})$ and $A_n \in F_{n-1}$.
Let us show that the representation (2.35) is unique. So assume that we have another representation $X_n = X_0 + \tilde{M}_n + \tilde{A}_n$, with a martingale $\tilde{M}$ and predictable $\tilde{A}$. Then we have

$$\begin{aligned} \mathbb{E}[\Delta X_k | F_{k-1}] &= \mathbb{E}[\Delta \tilde{M}_k + \Delta \tilde{A}_k | F_{k-1}] = \Delta \tilde{A}_k \\ &= \mathbb{E}[\Delta M_k + \Delta A_k | F_{k-1}] = \Delta A_k. \end{aligned}$$

Hence $\Delta A_k = \Delta \tilde{A}_k$ $\mathbb{P}$- almost surely; this in turn implies that $A_n = \tilde{A}_n$ $\mathbb{P}$-almost surely, and then we get that $\tilde{M}_n = M_n$ $\mathbb{P}$- almost surely.
If X is a submartingale, then $\mathbb{E}[X_k | F_{k-1}] \geq X_{k-1}$, and so $\Delta A_k \geq 0$; this shows that the process A in non-decreasing.
Conversely, if A is non-decreasing, then

$$\begin{aligned} \mathbb{E}[X_k | F_{k-1}] &= \mathbb{E}[X_0 + M_k + A_k | F_{k-1}] = X_0 + M_{k-1} + A_k \\ &\geq X_0 + M_{k-1} + A_{k-1} = X_{k-1}. \end{aligned}$$

We have proved the theorem. □

Remark 2.4. *The predictable process A from the Doob decomposition is called the compensator of X.*

Example 2.8. *We try to illustrate several aspects of the obtained results in this example.*
Let $S_0 = 0$ and $S_n = \sum_{k=1}^n \xi_k$, where square integrable random variables ξ_k are independent, with $\mathbb{E}\xi_k = 0$ and $Var(\xi_k) = \sigma_k^2 > 0$. Let $F_n = F_n^S = \sigma\{S_k : k \leq n\}$. We already know from Example 2.2. that (S_k, F_k), $k \geq 0$, is a square integrable martingale. We have that $[S,S]_n = \sum_{k=1}^n \xi_k^2$, then from the condition (2.33) we obtain that $S \in \mathcal{M}^2$ if and only if

$$\mathbb{E}[S,S]_\infty = \mathbb{E} \sum_{k=1}^{\infty} \xi_k^2 = \sum_{k=1}^{\infty} \sigma_k^2 < \infty.$$

The integration by parts formula gives

$$S_n^2 = S_0^2 + 2 \sum_{k=1} S_{k-1} \Delta S_k + [S,S]_n,$$

and hence the process

$$S_n^2 - [S,S]_n = S_0^2 + 2 \sum_{k=1}^{n} S_{k-1} \Delta S_k$$

is a martingale, because $\sum_{k=1}^{n} S_{k-1}\Delta S_k$ *is a martingale transform,* $S_{k-1}\Delta S_k \in \mathbb{L}^1(\mathbb{P})$ *and hence* $\mathbb{E}[S_{k-1}\Delta S_k|F_{k-1}] = S_{k-1}\cdot 0 = 0$. *[It is not difficult to check that if* M *is a square integrable martingale, then* $M^2 - [M,M]$ *is a martingale. This generalizes the more concrete discussion with sums of independent centered random variables.]*
Let us compare the decomposition $S^2 - [S,S]$ *to the Doob decomposition: By Jensen inequality* $(S_n^2, F_n)_{n\geq 0}$ *is a submartingale, and so according to the Doob decomposition*

$$S_n^2 = S_0^2 + M_n + A_n,$$

$A = (A_n, F_n)_{n\geq 0}$ *is predictable and increasing, and* M *is a martingale.*
As in the proof of the Doob decomposition, we first try to figure out, how the process A *looks like.*
Because $\Delta S_k^2 = 2S_{k-1}\Delta S_k + (\Delta S_k)^2$ *we obtain that*

$$\begin{aligned}\Delta A_k &= \mathbb{E}[\Delta S_k^2|F_{k-1}] = \mathbb{E}[2S_{k-1}\Delta S_k + (\Delta S_k)^2|F_{k-1}] \\ &= 2S_{k-1}\mathbb{E}[\Delta S_k|F_{k-1}] + \mathbb{E}[\xi_k^2|F_{k-1}] = \sigma_k^2,\end{aligned}$$

where we have used the following information: $\Delta S_k = \xi_k$, $\mathbb{E}[\xi_k|F_{k-1}] = \mathbb{E}\xi_k = 0$, *because* $\xi_k \perp\!\!\!\perp F_{k-1}$ *and by the same reason* $\mathbb{E}[\xi_k^2|F_{k-1}] = \mathbb{E}[\xi_k^2] = \sigma_k^2$. *From this we will get that*

$$A_n = \sum_{k=1}^{n} \sigma_k^2.$$

Later we use the notation: $A = < S, S >$ *[for general martingales* M *we will also use the notation* $< M.M >$ *for the compensator of* M^2*.].*
Let us study the Doob maximal inequality in this example. From the Doob maximal inequality we get

$$\mathbb{E}S_n^2 \leq \mathbb{E}(S_n^*)^2 \leq 4\mathbb{E}S_n^2.$$

But the process $S^2 - [S,S]$ *is a martingale, and hence* $\mathbb{E}S_n^2 = \mathbb{E}[S,S]_n$, *and we can write the maximal inequality in the form*

$$\mathbb{E}[S,S]_n \leq \mathbb{E}(S_n^*)^2 \leq 4\mathbb{E}[S,S]_n. \tag{2.36}$$

But the process $S^2 - < S, S >$ *is also a martingale, and so* $\mathbb{E}S_n^2 = \mathbb{E} < S, S >_n = \sum_{k=1}^{n}\sigma_k^2$. *We get that*

$$\mathbb{E} < S, S >_n = \sum_{k=1}^{n}\sigma_k^2 \leq \mathbb{E}(S_n^*)^2 \leq 4\sum_{k=1}^{n}\sigma_k^2 = 4\mathbb{E} < S, S >_n . \tag{2.37}$$

With deterministic times we see no difference in the inequalities. But let τ *be a finite stopping time. Moreover, assume that* $\sigma_k^2 = \sigma^2$. *Then* $< S, S >_n = n\sigma^2$. *Doobs theorem on optimal stopping tells us that*

$$\mathbb{E}S_\tau^2 = \mathbb{E}[S,S]_\tau = \mathbb{E} < S, S >_\tau = \sigma^2\mathbb{E}\tau.$$

Now we can write the inequality (2.37) in the form

$$\sigma^2\mathbb{E}\tau \leq \mathbb{E}(S_\tau^*)^2 \leq 4\sigma^2\mathbb{E}\tau.$$

Let finally τ *be a stopping time such that* $\mathbb{P}(\tau < \infty) = 1$. *Then* $\tau \wedge n$ *is a bounded stopping time and we get*

$$\sigma^2\mathbb{E}\tau\wedge n \leq \mathbb{E}(S_{\tau\wedge n}^*)^2 \leq 4\sigma^2\mathbb{E}\tau\wedge n.$$

If we let $n \to \infty$, then we will finally get

$$\sigma^2 \mathbb{E}\tau \le \mathbb{E}(S_\tau^*)^2 \le 4\sigma^2 \mathbb{E}\tau.$$

From this we get that $S_\tau^* \in \mathbb{L}^2(\mathbb{P})$ if and only if $\mathbb{E}\tau < \infty$. Similarly, from the inequalities (2.36) we obtain the inequalities

$$\mathbb{E}[S,S]_\tau \le \mathbb{E}(S_\tau^*)^2 \le 4\mathbb{E}[S,S]_\tau.$$

This is less informative that the previous one in our case.

Summary. Let $(M_n, F_n)_{n\ge 0}$ be a square integrable martingale.
From the integration by parts formula we obtain that the process $M^2 - [M,M]$ is a martingale, where $[M,M]$ is an increasing process given by $[M,M]_n = \sum_{k=1}^n (\Delta M_k)^2$. In the case of $\mathbb{P}(\tau < \infty) = 1$ Doobs maximal inequality and stopping equality give the following inequalities

$$\mathbb{E}[M,M]_\tau \le \mathbb{E}(M_\tau^*)^2 \le 4\mathbb{E}[M,M]_\tau. \tag{2.38}$$

Because the process M^2 is a submartingale, the Doob decomposition gives that there exists an increasing integrable process $A = (A_n, F_n)_{n\ge 1}$ such that $M^2 - A$ is a martingale. The process A is the compensator of M^2, and we denote it by $< M, M >$. From Doobs theorem on optimal stopping it follows that $\mathbb{E}M_\tau^2 = \mathbb{E} < M, M >_\tau$, when τ is a bounded stopping time. We can now write predictable version of the Doobs maximal inequality (2.38):

$$\mathbb{E} < M, M >_\tau \le \mathbb{E}(M_\tau^*)^2 \le 4\mathbb{E} < M, M >_\tau . \tag{2.39}$$

In the example 2.8 we gave an example from a situation, where the predictable maximal inequalities (2.39) are more informative than the inequalities (2.38).
The processes $M^2 - [M,M]$ and $M^2 - < M, M >$ are martingales, the process $[M,M] - < M, M >$ is a martingale, too:

$$[M,M] - < M, M > = M^2 - < M, M > -(M^2 - [M,M]).$$

Hence the process $< M, M >$ is the compensator of $[M,M]$..

Remark 2.5. *The inequalities (2.38) and (2.39) hold, when $\mathbb{E}(M_\tau^*)^2 = \infty$: if this is the case, then $\mathbb{E}[M,M]_\tau = \infty$ and $\mathbb{E} < M, M >_\tau = \infty$.*

The process $[M,M]_n = \sum_{k=1}^n (\Delta M_k)^2$ is sometimes called the energy process of M. If M is a square integrable process, then the process $< M, M >$ exists, and it called the predictable energy process. Martingales $M \in \mathcal{M}^2$ have finite energy.

3. Brownian motion and continuous time martingales

3.1. **Basic facts on continuous time processes.**

Definition 3.1. *Let $X = (X_t)_{t\ge 0}$ be a collection of random variables; then X is a continuous time stochastic process. The map [with ω fixed] $t \mapsto X_t(\omega)$ is the path of the process X [sometimes also a trajectory of X]. The process X is continuous, if it has continuous paths almost surely. The process X has D-paths, if it has right-continuous paths: $X_{t+}(\omega) = X_t(\omega)$ with left-hand limits: there exists $X_{t-}(\omega) = \lim_{s\uparrow t} X_s(\omega)$. The jump of the process X at time t is $\Delta X_t(\omega) = X_t(\omega) - X_{t-}(\omega)$.*

A discrete time process $Y = (Y_k)_{k\geq 0}$ can be embedded in the continuous time as follows. Define a continuous time process X by $X_t = Y_{\lfloor t \rfloor}$. Then the process X has D- paths and the sequence of the jumps ΔX satisfies $\Delta X_n = Y_n$. Hence we can interpret in this way all the discrete time processes as continuous time processes with simple D- paths.
Processes with D- paths have several useful properties, which are given in the next theorem.

Theorem 3.1. *Let $X = (X_t)_{t\geq 0}$ be a process with D- paths.*

a) *The expressions $X_t^* = \sup_{s\leq t} |X_s|$, X_∞^*, $\sup_{s\leq t} X_s$ ... are random variables.*
b) *Let $T > 0$ and $\pi_n = \{t_k : k = 0, \dots k(n)\}$ a partition of the interval $[0, T]$, $0 = t_0 < t_1 \cdots < t_{k(n)} = T$ and $|\pi_n| = \max_{\pi_n}(t_k - t_{k-1})$. Let $X_t^{(n)} = X_{t_k}$ when $t \in [t_{k-1}, t_k)$ be a discretization of X and let $|\pi_n| \to 0$, as $n \to \infty$; then $X_t = \lim_n X_t^{(n)}$.*
c) *With ω fixed the set*

$$\Delta X(\omega) \neq 0 = \{t : \Delta X_t(\omega) \neq 0\}$$

is numerable.

Proof Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a function with $f(t+) = f(t)$ and with left-hand limits $f(t-)$. Put $f^*(t) = \sup_{s\leq t} |f(s)|$ and $\tilde{f}^*(t) = \sup\{|f(q)| : q \in [0, t] \cap Q \cup \{t\}\}$. Clearly $\tilde{f}^*(t) \leq f^*(t)$. Assume now that $\tilde{f}^* < f^*$, that is $f^*(t) - \tilde{f}^*(t) = \epsilon > 0$. From the definition we obtain that there exists $t_0 \in [0, t)$ such that $|f(t_0)| > f^*(t) - \frac{\epsilon}{2}$. But $t_0 < t$ and by the right-continuity of f $|f(t_0)| = \lim_{q_k \downarrow t_0} |f(q_k)$, and this gives $\tilde{f}^*(t) \geq f(t_0)$, which is in conflict with the assumption that $f^*(t) - \tilde{f}^*(t) = \epsilon > 0$. Apply now these facts to the functional X_t^*. We have shown that $X_t^* = \tilde{X}_t^*$, where $\tilde{X}_t^* = \sup\{|X_q| : q \in [0, t] \cap Q \cup \{t\}\}$, and the functional $\tilde{X}_t^*$ is a random variable. The other claims of part a) are shown similarly.
The claim b) follows directly form the fact that the process X has D- paths.
Let f be a D-function and let $T > 0$. Consider the set

$$\Delta_k^T(f) = \{t : |\Delta(f)(t)| > \frac{1}{k}, t \leq T\}.$$

We claim that the set $\Delta_k(f)$ is finite. To prove this, assume that this is not the case. Then we can find a sequence $t_k \in \Delta_k^T(f)$, $k \geq 1$. Because the interval $[0, T]$ is bounded, then there a subsequence $t_{n(k)}$ such exists $t_0 = \lim t_{n(k)} \in [0, T]$. By going to a further subsequence t_q we can assume that t_q converges to the point t_0 either from the left or from the right. If for example $t_q \downarrow t_0$, then the function f is not right-continuous at the point t_0. If $t_q \uparrow t_0$, then the function f does not have left limit at the point t_0. But f is continuous from the right and has left-hand limits at every point. So $\Delta_k^T(f)$ is finite.
So for fixed ω the sets $\Delta_k^T(X)(\omega)$ are finite. We obtain that the set

$$\Delta X(\omega) = \cup_{T\in\mathbb{N}} \cup_{k\geq 1} \Delta_k^T(X)(\omega)$$

is at most numerable, and we have proved part c). □

The next definition gives several ways to identify in what sense two processes are the same.

Definition 3.2. *Let X, Y be stochastic processes.*

- *The processes X and Y have the same finite dimensional distributions, if for all $t_1,\dots,t_k$, $k \ge 1$, $B_1,\dots,B_k \in \mathbb{B}_{\mathbb{R}}$:*

$$\mathbb{P}(X_{t_1} \in B_1,\dots,X_{t_k} \in B_k) = \mathbb{P}(Y_{t_1} \in B_1,\dots,Y_{t_k} \in B_k). \tag{3.1}$$

- *The process X is a version of a process Y if $t \ge 0$ $\mathbb{P}(X_t = Y_t) = 1$.*
- *The processes X are Y indistinguishable, if $\mathbb{P}(X_t = Y_t \ \forall t \ge 0) = 1$.*

Clearly indistinguishably of X and Y implies that they are versions of each other. More facts of this type in the next theorem.

Theorem 3.2. *Let X, Y be stochastic processes.*

- *If the processes X and Y are versions of each other, then they have the same finite dimensional distributions.*
- *If the processes X and Y are versions of each other with D- paths, then they are indistinguishable.*

Proof Let $t_1,\dots,t_k$, $k \ge 1$ and

$$A = \{\omega : X_{t_1}(\omega) = Y_{t_1}(\omega),\dots,X_{t_k} = Y_{t_k}\}.$$

If the processes X and Y are versions of each other, then $\mathbb{P}(A) = 1$, and so

$$\begin{aligned}\mathbb{P}(X_{t_1} \in B_1,\dots,X_{t_k} \in B_k) &= \mathbb{P}((X_{t_1} \in B_1,\dots,X_{t_k} \in B_k) \cap A)\\ &= \mathbb{P}((Y_{t_1} \in B_1,\dots,Y_{t_k} \in B_k) \cap A)\\ &= \mathbb{P}(Y_{t_1} \in B_1,\dots,Y_{t_k} \in B_k),\end{aligned}$$

since, for example we have that

$$\mathbb{P}((X_{t_1} \in B_1,\dots,X_{t_k} \in B_k) \cap A^c) \le \mathbb{P}(A^c) = 0.$$

Hence the processes X and Y have the same finite dimensional distributions. We assume the the processes X and Y are versions of each other with D-paths. Put $B = \{\omega :\ \text{exists } t \ge 0 \text{ with } X_t(\omega) \ne Y_t(\omega)\}$. If $\omega \in B$, then there exists a $q \in Q$ such that $X_q(\omega) \ne Y_q(\omega)$. On the other hand, for a fixed $q \in Q$ $\mathbb{P}(\omega : X_q(\omega) \ne Y_q(\omega), q \ge 0, q \in Q) = 0$, and hence $\mathbb{P}(B) = 0$. We have shown that X and Y are indistinguishable. □

Example 3.1 (Brownan motion). *Stochastic process $W = (W_t)_{t \ge 0}$ is a Brownian motion, if $W_0 = 0$*

- *For all $0 \le t_0 < t_1 < \dots < t_n$ the increments $W_{t_k} - W_{t_{k-1}}$, $k = 1,\dots,n$ are independent.*
- *The increment has a normal distribution: $W_t - W_s \sim N(0, t-s)$ when $t > s \ge 0$.*
- *The paths of the process W are continuous.*

This definition is axiomatic, and at this point we do not know, in what sense Brownian motion exists. We will discuss this on Tuesday, 8.4.
The name comes from Robert Brown who studied the movement of small particles in water, and heuristically described the path properties of Brownian motion having values in three dimensional space. Bachelier obtained first mathematical results on Brownian motion in his PhD thesis in 1900. He also

used Brownian motion to model stock price movements at the Paris stock exchange! In 1905 Einstein used Brownian motion as a model in physics. Note that this was about twenty five years earlier than the axiomatization of probability theory by Kolmogorov.

Next we show that the paths of Brownian motion are rather wild. This fact explains why we must work for a while to be able to define stochastic integrals with respect to Brownian motion.

Theorem 3.3. *Let $W = (W_t)_{t\geq 0}$ be a Brownian motion, $T > 0$ and $\pi_n = \{t_k : t_0 < t_1 \cdots t_{k(n)}\}$ are partitions of the interval $[0, T]$ such that $|\pi_n| \to 0$, as $n \to \infty$. Then*

$$\sum_{t_k\in\pi_n} (W_{t_k} - W_{t_{k-1}})^2 \stackrel{L^2(\mathbb{P})}{\to} T, \text{ when } |\pi_n| \to 0. \tag{3.2}$$

Proof Let

$$X^{(n)} = \sum_{t_k\in\pi_n} (W_{t_k} - W_{t_{k-1}})^2.$$

We must show that $X^{(n)} \stackrel{L^2(\mathbb{P})}{\to} T$, as $|\pi_n| \to 0$. We recall that it is equivalent to show that $\mathbb{E}X^{(n)} \to T$ and $\operatorname{Var}(X^{(n)}) \to 0$. Now

$$\mathbb{E}X^{(n)} = \sum_{t_k\in\pi_n} \mathbb{E}(W_{t_k} - W_{t_{k-1}})^2 = \sum_{t_k\in\pi_n} (t_k - t_{k-1}) = T.$$

Recall also the following property of the normal distribution: if $\xi \sim N(0, \sigma^2)$, then $\operatorname{Var}(\xi^2) = \mathbb{E}\xi^4 - \sigma^4 = 2\sigma^4$. Brownian motion has independent increments, and so

$$\begin{aligned} \operatorname{Var}(X^{(n)}) &= \sum_{t_k\in\pi_n} \operatorname{Var}((W_{t_k} - W_{t_{k-1}})^2) \\ &= 2\sum_{t_k\in\pi_n} (t_k - t_{k-1})^2 \leq |\pi_n|T \to 0. \end{aligned}$$

This proves the claim. □

Next we recall the definition of <u>variation</u> of a function.

Definition 3.3. *The <u>variation</u> of the function f on the interval $[0, T]$ with respect to a partition π is the number*

$$\mathcal{V}_\pi(f) = \sum_{t_k\in\pi} |f_{t_k} - f_{t_{k-1}}|,$$

and its <u>total variation</u> on the interval $[0, T]$ is the number

$$\mathcal{V}_T(f) = \sup_\pi var_\pi(f);$$

and f has <u>bounded variation</u> on the interval $[0, T]$ if $\mathcal{V}_T(f) < \infty$.

Let us show that for a Brownian motion W we have that $\mathcal{V}_T(W) = \infty$ almost surely. Let us assume that this is not the case. Let $A = \{\omega : \mathcal{V}_T(W)(\omega) < \infty\}$ and assume that $\mathbb{P}(A) > 0$. If $\omega \in A$, then

$$\sum_{t_k\in\pi_n} (W_{t_k} - W_{t_{k-1}})^2 \leq \max_k |W_{t_k} - W_{t_{k-1}}|\mathcal{V}_T(W)(\omega).$$

Brownian motion has continuous paths and a continuous function is uniformly continuous on compact intervals $[0,T]$, then $\max_k |W_{t_k} - W_{t_{k-1}}| \to 0$, as $|\pi_n| \to 0$.
Hence for ω in the set A $\sum_{t_k\in\pi_n}(W_{t_k} - W_{t_{k-1}})^2 \to 0$; but we have always $\sum_{t_k\in\pi_n}(W_{t_k} - W_{t_{k-1}})^2 \xrightarrow{\mathbb{P}} T$, as $|\pi_n| \to 0$. This in turn implies that we must have $\mathbb{P}(A) = 0$. Hence we have shown that $mathcalV_T(W) = \infty$ almost surely.

3.2. **Measurability of a process, histories and stopping times.** Recall once more the definition of a stochastic process. The collection $X = (X_t)_{t\geq 0}$ is a stochastic process, if X_t is a random variable for all $t \geq$. Here we fix t and ask measurability with respect to the ω only. Sometimes it is useful to ask measurability with respect to the pair (ω, t).

Example 3.2. *Let $X = (X_n)_{n\geq 0}$ be a stochastic process with discrete time. Put $\mathcal{N} = \mathcal{P}(\mathbb{N})$. Then $\mathcal{N}$ is a sigma-algebra on $\mathbb{N}$ [and it is the smallest sigma-algebra containing singletons $\{k\}$]. Consider the process X as a mapping $(n,\omega) \mapsto X_n(\omega)$. We see that this is a measurable mapping from the product space to real line $(\mathbb{N}\times\Omega, \mathcal{N}\otimes\mathcal{F}) \to (\mathbb{R}, \mathbb{B}_{\mathbb{R}})$:*

$$\{(n,\omega) : X_n(\omega) \in B\} = \cup_k \{k\} \times \{\omega : X_k(\omega) \in \mathbb{B}\} \in \mathcal{N}\otimes\mathcal{F}.$$

If $X = (X_t)_{t\geq}$ is a continuous time stochastic process, then we again have $\{s\} \times \{\omega : X_s(\omega) \in B\} \in \mathbb{B}_{\mathbb{R}_+} \otimes \mathcal{F}$; but this does not imply that $\{(s,\omega) : X_s(\omega) \in B\} \in \mathbb{B}_{\mathbb{R}_+} \otimes \mathcal{F}$, since the union is not anymore numerable.

Definition 3.4. *The process $X = (X_t)_{t\geq 0}$ is (jointly) measurable, if the mapping $(t,\omega) \mapsto X_t(\omega)$ is a measurable mapping $(\mathbb{R}_+ \times \Omega, \mathbb{B}_{\mathbb{R}_+} \otimes \mathcal{A}) \to (\mathbb{R}, \mathbb{B}_{\mathbb{R}})$.*

3.2.1. *Histories and stopping times.* A continuous time history $\mathbb{F} = (F_t)_{t\geq 0}$ is an increasing family of subsigma-algebras of the sigma-algebra $\mathcal{F}$: $s \leq t \Rightarrow F_s \subset F_t$. The process X is $\mathbb{F}$- adapted, if $X_t \in F_t$, when $t \geq 0$.
The intersection of sigma-algebras is again a sigma-algebra, and so

$$F_{t+} = \cap_{u>t} F_u$$

is a sigma-algebra and $F_t \subset F_{t+}$. We say that a history $\mathbb{F}$ is continuous from the right, if $F_{t+} = F_t$.
The history of X at time t is $F_t^X = \sigma\{X_s : s \leq t\}$.
A random variable $\tau : \Omega \to \mathbb{R}_+ \cup \{\infty\}$ is a stopping time with respect to the history $\mathbb{F}$, if $\{\tau \leq t\} \in F_t$ for $t \geq 0$. The stopped sigma-algebra is defined as in discrete time:

$$F_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in F_t\}.$$

Like in discrete time we get easily the following: if σ, τ are two stopping times with $\sigma \leq \tau$, then $F_\sigma \subset F_\tau$.

Lemma 3.1. *Let $\mathbb{F}$ be continuous from the right. Then τ is a stopping time if and only if $\{\tau < t\} \in F_t$.*

Proof is an exercise [Problem 3 on 10.4. 2008].
Process X is adapted to $\mathbb{F}$, if $X_t \in F_t$.

Example 3.3. *Let X be a $\mathbb{F}$- adapted process, which has D- paths and we assume that $\mathbb{F}$ is right-continuous. Let $A \in \mathbb{B}_{\mathbb{R}}$ be an open set. Let $\tau_A = \inf\{t \geq 0 : X_t \in A\}$, where we agree that $\inf\{\emptyset\} = \infty$. Let us show that τ_A is a stopping time.*
Put

$$B = \cup_{0 \leq r \leq t, r \in Q}\{\omega : X_r(\omega) \in A\}$$

and we claim that $B = \{\omega : \tau_A(\omega) < t\}$.
Assume first that $\omega \in B$; then $\tau_A(\omega) \leq r < t$ for some rational r, and hence $\omega \in \{\tau_A < t\}$. On the other hand, if $\omega \in \{\tau_A < t\}$, then $X_s(\omega) \in A$ for some $s < t$, and hence there is a rational number $r < t$ such that $X_r(\omega) \in A$, and hence $\{\tau_A < t\} \subset B$. This means that $\{\tau_A < t\} = B$, and so $\{\tau_A < t\} \in F_t$ and by Lemma 3.1.
A typical case is the following $K > 0$ and $\tau^K = \inf\{t \geq 0 : |X_t| > K\}$.

3.3. **Martingales.** Let $(X, \mathbb{F})$ be an integrable adapted process, and let $s < t$. If in addition $\mathbb{E}[X_t|F_s] \leq X_s$, then X is a supermartigale; if in addition $\mathbb{E}[X_t|F_s] = X_s$, the X is a martingale; and finally, if $\mathbb{E}[X_t|F_s] \geq X_s$, then X is a submartingale.

Theorem 3.4. *Let X be a process with independent increments and $X_t \in L^2(\mathbb{P})$. Put $F_t = \sigma\{X_s : s \leq t\}$. Then the process $M_t = X_t - \mathbb{E}X_t$, is a martingale and the $M_t^2 - Var(X_t)$ is a martingale.*

3.3.1. *On martingale inequalities in continuous time.* Recall Doobs L^p- inequality for a discrete time martingale $X = (X_k, F_k)_{1 \leq k \leq n}$:

$$||X_n||_p \leq ||X_n^*||_p \leq q||X_n||,$$

where $p, q > 1$ are such that $\frac{1}{p} + \frac{1}{q} = 1$.
Let $X = (X_t, F_t)_{t \leq T}$ be a continuous time martingale with D- paths. Let $q_k \in [0, T]$ be rational numbers, $k = 0, \ldots, n-1$, $0 \leq q_0 < q_1 < \cdots < q_{n-1} < T$ and $q_n = T$. If we put $Y_k^{(n)} \doteq X_{q_k}$, $G_k^{(n)} \doteq F_{q_k}$ we get a discrete time martingale, and we have the following inequality

$$||X_T||_p = ||Y_n^{(n)}||_p \leq ||(Y^{(n)}) *_n ||_p \leq q||Y_n^{(n)}|| = q||X_T||_p.$$

Because X has D- paths, increasing the rational numbers in the definition of $Y^{(n)}$, we get $(Y^{(n)})_n^* \uparrow X_T^*$, and hence the L^p-maximal inequality holds for a continuous time martingale X with D- paths.
With similar arguments one shows that upcrossing inequality holds for continuous time submartingales, and other maximal inequalities hold, too, if the processes have D-paths.

Remark 3.1. *Let (M, F) be a martingale, but not necessarily with D-paths. Let $t > 0$ and $t_n \uparrow t$. Now $Y_k = M_{t_k}$ is a uniformly integrable martingale with repsect to the history F_{t_k}, and so it converges almost surely. Hence we always have a left-hand limit $M_{t-} = \lim M_{t_n}$.*
Assume now that $s_k \downarrow t$: by using the so-called inverse martingale convergence theorem we can show that $M_{t+} = \lim_k M_{s_k}$. So martingale has always right-hand limits, too. Clearly $M_{t+} \in F_{t+}$, but not necessarily $M_{t+} \in F_t$. If the history is right-continuous, then $M_{t+} \in F_t$. One can show that in this situation the martingale M has a version with D- paths.
We omit the details.