Stochastic Processes

Winter Term 2018/2019

Lecture held by Prof. Volker Betz

Typeset in LaTeX by Jonas Lenz
Errors can be reported at jlenz@mathematik.tu-darmstadt.de

Version of November 22, 2018

Contents

Motivation and Planned Topics

Laws of nature are encoded by differential equations. Some are encoded by ODEs, e.g. classical mechanics

$$\partial_t^2 x(t) = -a(x(t))\partial_t x(t) + F(t, x(t))$$

where a is a friction coefficients and F describes the force. Others by PDEs, e.g. electrodynamics or quantum mechanics. The aim of this course is to describe ODEs with noise, for example with random force. The simplest form of an ODE is

$$\partial_t X_t = b(t, X_t) + \sigma(t, X_t)\xi_t \tag{1}$$

where b is a deterministic force (drift), σ denotes the diffusion coefficient (tells how important the noise is) and as the most important part ξ_t is the white noise.

First Question: What Noise?

The "natural" approach is a random Fourier series:

$$\xi_t^{(N)} := \sum_{k=0}^{N} Y_k \cos(kt) + \sum_{k=1}^{N} Z_k \sin(kt) \quad \text{for} \quad t \in [0, 2\pi)$$

with $Y_k, Z_k \sim \mathcal{N}(0,1)$. The limit $\lim_{N\to\infty} \xi_t^{(N)}$ does *not* exist[1] and is called "white noise".

[1] The limit exists as a tempered distribution but not as an integrable function.

Second Question: Can we fix this?

A possible solution might be to integrate in time to improve convergence. For this, we define

$$B_t^{(N)} := \int_0^t \xi_s^{(N)} \,\mathrm{d}s = tY_0 + \sum_{k=1}^{N} \frac{1}{k}(Y_k \sin(kt) - Z_k(\cos(kt) - 1)).$$

Soon, we will see that this limit exists. It is called “Brownian motion”. Assume in (1) that $\sigma(t, y) \equiv \sigma_0$ is constant. Then we can compute

$$X_t - X_0 = \int_0^t \partial_s X_s \,\mathrm{d}s \overset{(1)}{=} \int_0^t b(s, X_s) \,\mathrm{d}s + \sigma_0 B_t.$$

This equation at least makes sense but we don’t know whether we can solve it. If σ is not constant, a term $\int_0^t \sigma(s, X_s)\xi_s \,\mathrm{d}s$ appears. In order to give sense to this integral, we will introduce Itô-integrals or rough integrals.

Third Question: What to read?

Apart from these fabulous lecture notes, the book “Brownian Motion” by Schilling and Partsch was recommended.

1 Brownian Motion

1.1 Definition: Let (E, Σ) be a measurable space and T a set. A collection of (E, Σ)-valued random variables (RVs) $\mathbf{X} = (X_t)_{t \in T}$ is called *E-valued stochastic process (SP) with index set T*. $\Diamond$

1.2 Example: We introduce three particularly interesting special cases of SPs.

(a) An SP with $(E, \Sigma) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and $T = \mathbb{R}_{\geq 0}$ is called *$\mathbb{R}^d$-valued, continuous-time SP*.

(b) For $E = \{-1, 1\}$, $\Sigma = \mathcal{P}(E)$ and $T = \mathbb{Z}^d$, the SP is called *spin system*.

(c) If E is countable and $T = \mathbb{N}_0$, we speak of a *time discrete SP*. $\Diamond$

From a dynamical point of view, X_t is a t-dependent quantity that changes with time. This perspective is suitable for the comprehension of the first and third example. From a global point of view, an SP is a single RV with values in the space $\Omega = E^T = \{f : T \to E\}$. In the first example, this means to consider the whole *path* $(X_t)_{t \geq 0}$ as one object. In the second example, each spin configuration in $\{-1, 1\}^{\mathbb{Z}^d}$ is an element of Ω. This raises some questions: What is the “right” σ-algebra on Ω? Does it even exist?

1.3 Definition: Let $\mathbf{X} = (X_t)_{t \in T}$ be an SP where the state space E is a group, e.g. $E = \mathbb{R}^d$, and $T \subseteq \mathbb{R}$. The family $(X_{s,t})_{s,t \in T}$ with $X_{s,t} \coloneqq X_t - X_s$ is called the *increment process of* $\mathbf{X}$ or *set of increments*.

An SP has *independent increments* if for all $n \in \mathbb{N}$ and all $s_1 < t_1 \leq s_2 < t_2 \leq \cdots \leq s_n < t_n$ with $s_i, t_i \in T$, the RVs $(X_{s_i,t_i})_{1 \leq i \leq n}$ are independent.

An SP has *stationary increments* if for all $n \in \mathbb{N}$, all $r \in T$ and all $s_1 < t_1 \leq s_2 < t_2 \leq \cdots \leq s_n < t_n$ with $s_i, t_i \in T$, we have

$$(X_{s_i,t_i})_{i=1,\dots,n} \sim (X_{s_i+r,t_i+r})_{i=1,\dots,n},$$

i.e. that the increments are equal in distribution regardless at which time we started looking. $\Diamond$

1.4 Definition: An $\mathbb{R}^d$-valued SP $\mathbf{B} = (B_t)_{t\in\mathbb{R}_{\geq 0}}$ is called *Brownian Motion* (BM) if

(B1) $B_0(\omega) = 0$ for almost all $w \in \Omega$,

(B2) $\mathbf{B}$ has independent increments,

(B3) $\mathbf{B}$ has stationary increments,

(B4) the increments are normally distributed, i.e.

$$B_{s,t} \coloneqq B_t - B_s \sim B_{t-s} \sim \mathcal{N}\left(0, (t-s)\mathrm{I}_d\right),$$

and

(B5) the map $t \mapsto B_t(\omega)$ is continuous for all $\omega \in \Omega$. $\diamondsuit$

In (B4) we require continuity of all paths. For many settings, it would be more natural to require this only for almost-all paths. However, the set of continuous paths is usually not measurable and it is often easier to work exclusively with the continuous paths. It does not make much difference for the theory but we will consistently work with this definition.

1.5 Remark: Checking the requirements (B0) – (B4) in view of $B_t = \int_0^t \xi_s \,\mathrm{d}s$:

(B1) $\int_0^0 \xi_s \,\mathrm{d}s = 0$

(B2) $B_t - B_s = \int_s^t \xi_r \,\mathrm{d}r$ and $\xi_r \perp\!\!\!\perp (\xi_s)_{s\neq r} \quad \forall r$

(B3) The distribution of ξ_r does not depend on r.

(B4) Central limit theorem and Riemann approximation.

(B5) The map $t \mapsto \int_0^t f_s \,\mathrm{d}s$ is continuous for all "sensible" functions f, in particular for $f = \xi$. $\diamondsuit$

1.6 Definition: The *Gaussian measure* $\mathcal{N}\left(m, \sigma^2\right)$ with mean m and variance σ^2 is the probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with Lebesgue density

$$g_{m,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-m)^2\right). \qquad \diamondsuit$$

1.7 Proposition: Let $X \sim \mathcal{N}\left(m, \sigma^2\right)$. Then:

(a) $\mathbb{E}\left(X\right) = m$, $\mathbb{V}\left(X\right) = \sigma^2$

(b) We have the *Gaussian tail estimate*

$$\frac{1}{\sqrt{2\pi}}\frac{C}{C^2+1}\mathrm{e}^{-\frac{C^2}{2}} \le \mathbb{P}(X-m \ge C\sigma) \le \frac{1}{\sqrt{2\pi}}\frac{1}{C}\mathrm{e}^{-\frac{C^2}{2}},$$

for all $C>0$, $\sigma>0$.

(c) For $(m_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}$, $m\in\mathbb{R}$, $(\sigma_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}_{\ge 0}$ and $\sigma\in\mathbb{R}_{\ge 0}$ we have that $(m_k,\sigma_k)\to(m,\sigma)$ if and only if $\mathcal{N}(m_k,\sigma_k^2) \xrightarrow{\mathrm{d}} \mathcal{N}(m,\sigma^2)$. ♢

1.8 Definition: An $\mathbb{R}^d$-valued RV X is called *d-dimensional Gaussian* if for all linear functionals $L\colon \mathbb{R}^d\to\mathbb{R}$ there are m, σ^2 with $LX\sim\mathcal{N}(m,\sigma^2)$. Explicitly: If $\mathbf{X}=(X^1,\dots,X^d)$, this means that for all $a_1,\dots,a_d\in\mathbb{R}$ there are m, σ^2 such that $\sum_{i=1}^d a_iX^i\sim\mathcal{N}(m,\sigma^2)$. ♢

1.9 Example: (a) If $X^1,\dots,X^d$ are independent 1-dimensional Gaussian, then $\mathbf{X}=(X^1,\dots,X^d)$ is d-dimensional Gaussian.

(b) *Warning:* Without independence, this is not true in general. Consider $X^1\sim\mathcal{N}(0,1)$ and

$$X^2(\omega)=\begin{cases}-X^1(\omega), & \text{if } |X^1(\omega)|\le 1,\\ +X^1(\omega), & \text{if } |X^1(\omega)|>1.\end{cases}$$

Then, $X^2\sim\mathcal{N}(0,1)$ (to check this, compute $\mathbb{P}(X^2<c)$ for all $c\in\mathbb{R}$) but (X^1,X^2) is not Gaussian as $|X^1(\omega)-X^2(\omega)|\le 2$ for all $\omega\in\Omega$ and $|X^1(\omega)-X^2(\omega)|\not\equiv 0$, which implies that X^1-X^2 is not Gaussian. ♢

1.10 Exercise: Are there pairwise independent $X^1,\dots,X^d\sim\mathcal{N}$ such that $\mathbf{X}=(X^d,\dots,X^d)$ is not Gaussian? ♢

1.11 Proposition: A real RV X is $\mathcal{N}(m,\sigma^2)$-distributed if and only if its characteristic function is given by

$$\varphi_X(u) \stackrel{*}{=} \mathrm{e}^{\mathrm{i}um}\mathrm{e}^{-\frac{1}{2}\sigma^2u^2} = \exp\left(\mathrm{i}um-\frac{1}{2}\sigma^2u^2\right). \qquad ♢$$

Proof. Recall that $\varphi_X(u)=\mathbb{E}(\mathrm{e}^{\mathrm{i}uX})$ uniquely determines the distribution of X. So it is enough to show $(*)$ for $X\sim\mathcal{N}(m,\sigma^2)$. Since we have the regularity $\varphi_{X+m}(u)=\mathbb{E}(\mathrm{e}^{\mathrm{i}u(X+m)})=\mathrm{e}^{\mathrm{i}um}\varphi_X(u)$, it suffices to consider the case $m=0$.

By the Lebesgue differentiation theorem we have

$$\frac{\mathrm{d}}{\mathrm{d}u}\varphi_X(u)=\frac{1}{\sqrt{2\pi\sigma^2}}\int \mathrm{i}x\mathrm{e}^{\mathrm{i}ux}\mathrm{e}^{-\frac{x^2}{2\sigma^2}}\,\mathrm{d}x$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int \mathrm{i}(\mathrm{i}u\mathrm{e}^{\mathrm{i}ux})\sigma^2 \mathrm{e}^{-\frac{x^2}{2\sigma^2}}\,\mathrm{d}x$$
$$= -u\sigma^2\varphi_X(u),$$

and $\varphi_X(0) = 1$. Hence, $h(u) = \ln(\varphi_X(u))$ solves the ODE

$$h'(u) = \frac{\varphi_X'(u)}{\varphi_X(u)} = -u\sigma^2$$
$$h(0) = 0,$$

which implies $h(u) = -\frac{1}{2}u^2\sigma^2$. □

1.12 Corollary: For $X \sim \mathcal{N}(0, \sigma^2)$ and $J \in \mathbb{C}$ we have $\mathbb{E}\left(\mathrm{e}^{JX}\right) = \mathrm{e}^{\sigma^2\frac{J^2}{2}}$. ◊

Proof. This follows by analytic continuation of the previous proposition. □

1.13 Theorem: Let X be d-dimensional Gaussian.

(a) The distribution of $\mathbf{X}$ is uniquely determined by

$$\mathbf{m} = \mathbb{E}(\mathbf{X}) = \left(\mathbb{E}(X^i)\right)_{1\leq i\leq d} \in \mathbb{R}^d,$$

the *mean vector* of $\mathbf{X}$, and

$$C = (C_{ij})_{1\leq i,j\leq d} \quad \text{with} \quad C_{ij} = \operatorname{Cov}(X^i, X^j),$$

the *covariance matrix* of $\mathbf{X}$. We write $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, C)$.

(b) If C is invertible, then the distribution of $\mathbf{X}$ has a Lebesgue-density which is given by

$$\mathbb{P}(X \in \mathrm{d}\mathbf{x}) = \frac{1}{(2\pi)^{\frac{d}{2}}}\frac{1}{(\det C)^{\frac{1}{2}}} \exp\Big(-\frac{1}{2}\left\langle \mathbf{x}-\mathbf{m}, C^{-1}(\mathbf{x}-\mathbf{m})\right\rangle\Big)\,\mathrm{d}\mathbf{x}\,. \quad ◊$$

Proof. (a) Assume $\mathbf{X}$, $\mathbf{Y}$ are d-dimensional Gaussian with mean $\mathbf{m}$ and covariance matrix C. Let $\mathbf{a} \in \mathbb{R}^d$ be arbitrary and set $Z := \sum_{i=1}^d a_i X^i$ and $W := \sum_{i=1}^d a_i Y^i$. Then, Z and W are 1-dimensional Gaussian with $\mathbb{E}(Z) = \mathbb{E}(W) = \langle \mathbf{a}, \mathbf{m}\rangle$ and

$$\mathbb{V}(Z) = \mathbb{V}(W) = \langle a, Ca\rangle\,. \tag{$*$}$$

So,

$$\varphi_{\mathbf{X}}(a) = \mathbb{E}\left(\mathrm{e}^{\mathrm{i}\langle \mathbf{a},\mathbf{X}\rangle}\right) = \mathrm{e}^{\mathrm{i}\langle \mathbf{a},\mathbf{m}\rangle}\mathrm{e}^{-\frac{1}{2}\langle a,Ca\rangle} = \varphi_{\mathbf{Y}}(a)$$

holds for all $a \in \mathbb{R}^d$, which implies $\mathbf{X} \sim \mathbf{Y}$.

(b) By $(*)$, C must be positive semidefinite. If C is invertible, C must be positive definite. Hence, the density is well-defined. To check that it is the right one, compute its characteristic function (remains as an exercise). $\square$

1.14 Proposition: Let $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, C)$ be a d-dimensional Gaussian random variable and $A \in \mathbb{R}^{n \times d}$. Then, $A\mathbf{X} \sim \mathcal{N}(A\mathbf{m}, ACA^*)$ where A^* denotes the transpose of A. $\Diamond$

Proof. The proof remains as an exercise. $\square$

1.15 Proposition: Let $\mathbf{X} \sim \mathcal{N}(\mathbf{m}, C)$. Then $X^1, \ldots, X^d$ are independent random variables if and only if $C_{ij} = 0$ for all $i \neq j$, i.e. the pairs X^i, X^j are uncorrelated. $\Diamond$

Proof. The implication "$\Rightarrow$" always holds (if the variances exist). For the other direction, let $Y^1, \ldots, Y^d$ be independent with $Y^i \sim \mathcal{N}(m_i, C_{ii})$. Then by 1.13, $\mathbf{X} \sim \mathbf{Y}$, which implies that the X^i are independent. $\square$

1.16 Definition: Let $(X_t)_{t \in T}$ be an E-valued stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The set of *finite dimensional distributions* (*fdd*) of $\mathbf{X}$ is the family of probability measures

$$\{p_{t_1,\ldots,t_n} \mid t_1, \ldots, t_n \in T; t_i \neq t_j \text{ if } i \neq j; n \in \mathbb{N}\}$$

where $p_{t_1,\ldots,t_n} = \mathbb{P} \circ (X_{t_1}, \ldots, X_{t_n})^{-1}$ is the image of $\mathbb{P}$ under $(X_{t_1}, \ldots, X_{t_n})$. In order words, $p_{t_1,\ldots,t_n}(A_1 \times \cdots \times A_n) = \mathbb{P}(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n)$ for all "good" sets $A_1, \ldots, A_n$. $\Diamond$

1.17 Example: Let $T = \mathbb{N}$, $E = \mathbb{Z}$ and $(X_n)_{n \in \mathbb{N}}$ be a simple random walk, this is to say $X_n = \sum_{i=1}^n Z_i$ with Z_i i.i.d., $\mathbb{P}(Z_i = \pm 1) = \frac{1}{2}$. Then

$$p_{1,4,17}(A \times B \times C) = \mathbb{P}(X_1 \in A, X_4 \in B, X_{17} \in C). \qquad \Diamond$$

1.18 Proposition: Let $\mathbf{X}$ be as in 1.16. Then its fdd fulfil the *consistency conditions* that for all $t_1, \ldots, t_n \in T$, $C_1, \ldots, C_n \in \mathcal{E}$, $\sigma \in \mathrm{S}_n$ it holds that

(C1) $p_{t_1,\ldots,t_n}(C_1 \times \cdots \times C_n) = p_{t_{\sigma(1)},\ldots,t_{\sigma(n)}}(C_{\sigma(1)} \ldots, C_{\sigma(n)})$ and

(C2) $p_{t_1,\ldots,t_n}(C_1 \times \cdots \times C_{n-1} \times E) = p_{t_1,\ldots,t_{n-1}}(C_1 \times \cdots \times C_{n-1})$. $\Diamond$

Proof. This remains as an easy exercise. $\square$

1.19 Definition: An $\mathbb{R}^d$-valued process $(\mathbf{X}_t)_{t \in T}$ is called *Gaussian process* if all its fdd are Gaussian measures. $\Diamond$

1.20 Remark: (a) Explicitly, $p_{t_1,\dots,t_n}$ is Gaussian on $\mathbb{R}^{dn}$.

(b) (1.9 b) shows that there are processes where X_t is Gaussian for all $t \in T$ but where $\mathbf{X}$ is not a Gauss process. Take for example $T = \{1, 2\}$, $E = \mathbb{R}$, $X_1 = X^1$ and $X_2 = X^2$. Morale: The one-dimensional distributions are not enough to make a process Gaussian!

(c) If $\mathbf{X}$ is a Gauss process, its fdd are fully determined by the mean and covariance functions

$$\begin{aligned} T &\to \mathbb{R}^d, & t &\mapsto \mathbb{E}\left(\mathbf{X}_t\right) \\ T^2 &\to \mathbb{R}^{d\times d}, & (s,t) &\mapsto \mathrm{Cov}(\mathbf{X}_s, \mathbf{X}_t). \end{aligned}$$

This follows immediately from Theorem 1.13 $\diamondsuit$

1.21 Theorem: (a) An $\mathbb{R}^d$-valued Brownian motion $\mathbf{B}$ is a Gaussian process with $\mathbb{E}\left(\mathbf{B}_t\right) = 0$ for all t and

$$\mathrm{Cov}(\mathbf{B}_s, \mathbf{B}_t) = \mathbb{E}\left(\mathbf{B}_s \otimes \mathbf{B}_t\right) = \mathbb{E}\left(\left(B_s^i B_t^j\right)_{i,j=1,\dots,d}\right) = \min\{s,t\} \cdot \mathrm{I}_d$$

(b) Conversely, any Gaussian process with the mean and covariance functions from (a) is a Brownian motion if it fulfills (B4). $\diamondsuit$

Proof. (a) Let $t_1, \dots, t_n \in \mathbb{R}_0^+$ with $t_1 < \dots < t_n$. Then with $t_0 = 0$ and $\mathbf{D}_i = \mathbf{B}_{t_i} - \mathbf{B}_{t_{i-1}}$

$$\begin{aligned} &\left(\mathbf{B}_{t_1}(\omega), \dots, \mathbf{B}_{t_n}(\omega)\right)^\top = \\ &A\left(B_{t_1}(\omega) - B_0(\omega), B_{t_2}(\omega) - B_{t_1}(\omega), \dots, B_{t_n}(\omega) - B_{t_{n-1}}(\omega)\right)^\top \end{aligned}$$

with the lower triangle matrix

$$A = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$$

holds. By (B1),(B3) and (1.9), $(B_{t_i} - B_{t_{i-1}})_{1\le i\le n} \sim \mathcal{N}\left(0, C\right)$ with $C_{ij} = \delta_{ij}(t_i - t_{i-1})$. By (1.13), $p_{t_1,\dots,t_n} \sim \mathcal{N}\left(0, ACA^*\right)$ which implies that $\mathbf{B}$ is a Gaussian process. Now, we compute the covariance and assume $s < t$. Then we have

$$\begin{aligned} \mathrm{Cov}(B_s, B_t) &= \mathbb{E}\left(B_s \otimes B_t\right) = \mathbb{E}\left(B_s \otimes (B_t - B_s)\right) + \mathbb{E}\left(B_s \otimes B_s\right) \\ &= s\mathrm{I}_d = \min\{s,t\}\mathrm{I}_d. \end{aligned}$$

(b) We check that (B0)-(B2) hold, as (B3),(B4) holdn by assumption. (B0) follows from $\mathbb{V}(B_0) = 0$ and $\mathbb{E}(B_0) = 0$. For (B1) and (B2) let $0 < t_1 < \cdots < t_n$. The covariance matrix $(B_{t_1}, \ldots, B_{t_n})$ is

$$M = (m_{ij})_{i,j\in\mathbb{N}} = (t_{\min\{i,j\}})_{i,j\in\mathbb{N}}$$

and with A as in a). Then, $(B_{t_1} - B_0, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}})$ has covariance matrix

$$M' = A^{-1}M\left(A^{-1}\right)^* = \mathrm{diag}(t_1, t_2 - t_1, \ldots, t_n - t_{n-1}).$$

which implies that (B1) and (B2) holds. $\square$

1.22 Proposition: Let $\mathbf{B}^1, \ldots, \mathbf{B}^d$ be 1-dimensional Brownian motions and let the $(\mathbf{B}^i)_{i=1,\ldots,d}$ be independent (as stochastic processes). Then $\left(B_t^1, \ldots, B_t^d\right)_{t\geq 0}$ is a d-dimensional Brownian motion. Conversely, the coordinate processes $(B_t^i)_{t\geq 0}$ of a d-dimensional Brownian motion are independent 1-dimensional Brownian motions. $\Diamond$

Proof. This is remains as an exercise or can be found in Section 2.3 of Schilling/Partzsch. $\square$

1.23 Proposition: Let $\mathbf{B}$ be a 1-dimensional Brownian motion. Then its fdd are given by

$$p_{t_1,\ldots,t_n}(A_1 \times \cdots \times A_n) = \mathbb{P}\left(B_{t_1\in A_1,\ldots,B_{t_n}\in A_n}\right)$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{\left[\prod_{j=1}^n (t_j - t_{j-1})\right]} \int_{A_1\times\cdots\times A_n} \exp\left(-\frac{1}{2}\sum_{j=1}^n \frac{(x_j - x_{j-1})^2}{t_j - t_{j-1}}\right) \mathrm{d}x \qquad (1.1)$$

for all $0 = t_0 < t_1 < \cdots < t_n$, $A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R}), n \in \mathbb{N}$ with $x_0 = 0$ and $x = (x_1, \ldots, x_n)$. $\Diamond$

Proof. Referring to Thm 1.20 this remains as an exercise. $\square$

1.24 Proposition: The family of fdd given in the previous proposition is consistent in the sense of (1.17),(C1) and (C2). $\Diamond$

So, Brownian motions have a chance to exist. We now that it does. Nevertheless, this will take a while.

1.25 Definition: Let $(E, \mathcal{E})$ be a measurable space and T a set.

i) The map

$$pi_t \colon E^T \to E$$
$$(e_s)_{s\in T} \mapsto e_t$$

is called *coordinate projection* to the t-th coordinate. When we identify E^T with $\{f \colon T \to E\}$ then $\pi_t(e) = e_t$ is the point evaluation of the function e at point t.

ii) The σ-algebra $\mathcal{E}^{\otimes T}$ is the smalles σ-algebra on E^T so that all maps π_t are $\mathcal{E}^{\otimes T}$-$\mathcal{E}$-measurable.

iii) The measurable space $(E^T, \mathcal{E}^{\otimes T})$ is the *canonical space* for E valued stochastic processes with index set T.

iv) If $\Omega_0 \subset E^T$ is any subset (not necessarily measurable), the σ-algebra

$$\mathcal{E}^{\otimes T} \cap \Omega_0 \coloneqq \{A \cap \Omega_0 \colon A \in \mathcal{E}^{\otimes T}\}$$

is called the *trace* of $\mathcal{E}^{\otimes T}$ on Ω_0. The measurable space $(\Omega_0, \mathcal{E}^{\otimes T} \cap \Omega_0)$ is the canonical space for E-valued process with sample paths in Ω_0.

$\diamondsuit$

1.26 Example: $E = \mathbb{R}^d$, $T = \mathbb{R}_0^+$ and $\Omega = E^T = \{\omega \colon \mathbb{R}_0^+ \to \mathbb{R}^d\}$, $\pi_t(\omega) = \omega(t)$, Ω=space of all „paths" $t \to \omega(t)$. Write $X_t(\omega) = \pi_t(\omega) = \omega(t)$. We consider $\Omega \cap C_0(\mathbb{R}^d) = \{\omega \in C(\mathbb{R}_0^+, \mathbb{R}^d), \omega(0) = 0\}$ and $\mathcal{F} = \mathcal{B}(\mathbb{R}^d)^{\otimes \mathbb{R}_0^+} \cap C_0(\mathbb{R}^d)$. Then $(C_0(\mathbb{R}^d), \mathcal{F})$ is the canonical measurable space for a stochastic process with continuous paths. $\diamondsuit$

1.27 Remark: The metric of *local uniform convergence* on $C_0(\mathbb{R}^d)$ is given by

$$\rho \colon C_0 \times C_0 \to \mathbb{R}_0^+$$
$$(f, g) \mapsto \sum_{n=1}^{\infty} \min\{1, \sup_{0\le t\le n} |f(t) - g(t)|\} 2^{-n}.$$

The Borel-σ-algebra $\mathcal{B}(C_0)$ on C_0 is the smallest σ-algebra on C_0 such that all ρ-open sets are measurable. We have $\mathcal{B}(C_0) = \mathcal{B}(\mathbb{R}^d)^{\otimes \mathbb{R}_0^+} \cap C_0$. $\diamondsuit$

Proof. This remains as an exercise. $\square$

1.28 Lemma: Let $(E, \mathcal{E})$ be a measurable space, T a set and $A \subseteq E^T$. Then $A \in \mathcal{E}^{\otimes T}$ if and only if there exists $I \subseteq T$ countable with $A \in \{\pi_t \colon t \in I\}$. $\diamondsuit$

Proof. This remains as an exercise (on some exercise sheet). □

Recall the following:

1.29 Theorem (Thm 3.29 from Probability Theory, Winter Term 17/18)**:** Let $(\Omega, \mathcal{F})$ be a Borel space and $\mathbb{P}$ a probability measure on $(\Omega, \mathcal{F})$. Then for each σ-algebra $\mathcal{G} \subseteq \mathcal{F}$, a map $\mu\colon \Omega \times \mathcal{F} \to [0,1]$ with the properties

(i) $\mu(\cdot, A)$ is $\mathcal{F}$-measurable for all $A \in \mathcal{F}$

(ii) $\mu(\omega, \cdot)$ is a probability measure for all $\omega \in \Omega$

(iii) Moreover, $\mu(\omega, \cdot)$ is a conditional probability of A given $\mathcal{G}$, i.e. $\mu(\omega, A) = \mathbb{P}(A \mid \mathcal{G})(\omega)$ for $\mathbb{P}$-almost all $\omega \in \Omega$.

exists. μ is called *regular conditional probability.* ◊

Proof. Will be uploaded in the notes (to be done later). □

1.30 Lemma: Let $(\Omega, \mathcal{F})$ be a Borel space, $\mu\colon \Omega \times \mathcal{F} \to [0,1]$ with properties (1.28)(i),(ii) [called a *probability kernel*]. Then there exists a $\mathcal{U}[0,1]$-RV Y and an $\mathcal{F} \otimes \mathcal{B}([0,1])$-measurable function $f\colon \Omega \times [0,1] \to \Omega$ with

$$\mu(\omega, A) = \mathbb{P}(f(\omega, Y) \in A) = \int_0^1 \mathbb{1}_{\{f(\omega,\cdot)\in A\}}(u)\ \mathrm{d}u$$

for all $\omega \in \Omega$, $A \in \mathcal{F}$. ◊

Proof. This remains as an exercise or can be found in Kallenberg [Foundations of Modern Probability, 3.22]. □

1.31 Theorem: Let $(E, \mathcal{E})$ be Borel. For each $n \in \mathbb{N}$ let $\mathbb{P}_n$ be a probability measure on $(E^n, \mathcal{E}^{\otimes n})$ and assume *consistency*, i.e. for all $n \in \mathbb{N}$ and all $A \in \mathcal{E}^{\otimes n}$ it holds that

$$\mathbb{P}_{n+1}(A \times E) = \mathbb{P}_n(A).$$

Let $(\Omega, \mathcal{F}, \mathbb{P}) = \Big([0,1]^{\mathbb{N}}, \mathcal{B}([0,1])^{\otimes\mathbb{N}}, \mathcal{U}([0,1])^{\otimes\mathbb{N}}\Big)$. Then there exist random variables $X_i\colon \Omega \to E$, $i \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ and all $A \in \mathcal{E}^{\otimes n}$ it holds that

$$\mathbb{P}_n(A) = \mathbb{P}((X_1, \ldots, X_n) \in A).$$

◊

Proof. 1. Fix $n \in \mathbb{N}$. Then $(E^{n+1}, \mathcal{E}^{\otimes n+1})$ is Borel as a product of Borel spaces (exercise!). We set $\mathcal{G}_n \coloneqq \sigma(\{A_1 \times \cdots \times A_n \times E \colon A_i \in \mathcal{E}\})$ and $\mu_n \colon E^{n+1} \times \mathcal{E}^{\otimes n+1} \to [0,1]$ as in (1.28), i.e. $\mu_n(\mathbf{x}, A) = \mathbb{P}_{n+1}(A \mid \mathcal{G}_n)(x)$ almost surely with respect to $\P_{n+1}$ for all $A \in \mathcal{E}^{\otimes n+1}$. Since $\mathbf{x} \mapsto \mu_n(\mathbf{x}, A)$ is measurable with respect to $\mathcal{G}_n$, it depends only on $x_1, \ldots, x_n$ and not on x_{n+1}. Write

$$\tilde{\mu_n}((x_1, \ldots, x_n), A) = \mu_n(\mathbf{x}, A).$$

Note that $\mu_0(\mathbf{x}, A) = \mathbb{P}_1(A)$ does not depend on x_1. 3. By (1.29), there exist functions $f_n \colon E^n \times [0,1] \to E$ with $\mu_n(\mathbf{x}, A) = \mathbb{P}(f(x_1, \ldots, x_n, Y_{n+1}) \in A)$. In particular

$$\mu_0(\mathbf{x}, A) = \tilde{\mu_0}(A) = \mathbb{P}(f_0(Y_1) \in A).$$

Now, we will proceed by induction. Put $X_1 = f(Y_1)$. Assume that $(X_1, \ldots, X_n)$ have been constructed. We set $X_{n+1}(\omega) \coloneqq f_n((X_1, (\omega), \ldots, X_n(\omega), Y_{n+1}(\omega))$. Since $\mu_n(\mathbf{x}, A) = \tilde{\mu}_n(x_1, \ldots, x_n), A) = \mathbb{P}(f_{n+1}(x_1, \ldots, x_n, Y_{n+1}) \in A)$, we find that for all $A_1, \ldots, A_{n+1} \in \mathcal{E}$, it holds that

$$\begin{aligned}\mathbb{P}(X_i \in A_i \forall i \leq n+1) &= \mathbb{E}\left(\mathbb{P}(X_{n+1} \in A_{n+1} \mid \mathcal{G}_n) \prod_{i=1}^n \mathbb{1}_{\{x_i \in A_i\}}\right) \\ &= \mathbb{E}\left(\mathbb{P}(f_n((X_1, \ldots, X_n), Y_{n+1}) \in A_{n+1} \mid \mathcal{G}_n) \prod_{i=1}^n \mathbb{1}_{\{X_i \in A_i\}}\right) \eqqcolon (*).\end{aligned}$$

Since $Y_{n+1} \amalg (X_1, \ldots, X_n)$, Proposition 3.23 from Probability Theory implies that

$$\begin{aligned}\mathbb{P}(f_n((X_1, \ldots, X_n), Y_{n+1}) \in A_{n+1} \mid \mathcal{G}_n)(\omega) &= \mathbb{P}(f_n((X_1(\omega), \ldots, X_n(\omega)), Y_{n+1}) \in A_{n+1}) \\ &= \tilde{\mu_n}((X_1(\omega), \ldots, X_n(\omega)), A_{n+1})\end{aligned}$$

holds $\mathbb{P}$-almost surely. So using the image measure we have

$$\begin{aligned}(*) &= \mathbb{E}\left((\tilde{\mu_n}((X_1, \ldots, X_n), A_{n+1}) \prod_{i=1}^n \mathbb{1}_{\{X_i \in A_i\}}\right) \\ &= \int \tilde{\mu_n}((x_1, \ldots, x_n), A_{n+1}) \prod_{i=1}^n \mathbb{1}_{\{x_i \in A_i\}} \mathbb{P}_n(\mathrm{d}\mathbf{x}) \\ &= \int \tilde{\mu_n}((x_1, \ldots, x_n), A_{n+1}) \prod_{i=1}^n \mathbb{1}_{\{x_i \in A_i\}} \mathbb{P}_{n+1}(\mathrm{d}\mathbf{x}) \\ &= \mathbb{E}\left(\mathbb{P}_{n+1}(A_{n+1} \mid \mathcal{G}_n) \prod_{i=1}^n \mathbb{1}_{A_i}\right)\end{aligned}$$

$$= \mathbb{E}\left(\prod_{i=1}^{n+1} \mathbb{1}_{A_i}\right) = \mathbb{P}_{n+1}(A_1, \times \cdots \times A_{n+1}).$$

Then the claim follows by induction. $\square$

1.32 Theorem (Kolmogorov 1932)**:** Let $(E, \mathcal{E})$ be Borel, T a set. Let $\{p_{t_1,\ldots,t_n} : t_1, \ldots, t_n \in T, n \in \mathbb{N}\}$ be a family of probability measures, that are consistent, i.e. fulfill (C1),(C2) of (1.17). Then there exists a probability measure $\mathbb{P}$ on $(E^T, \mathcal{E}^{\otimes T})$ such that

$$p_{t_1,\ldots,t_n}(A) = \mathbb{P}((\Pi_{t_1}, \ldots, \Pi_{t_n}) \in A)$$

holds for all $A \in \mathcal{E}^{\otimes T}$ and all $t_1, \ldots, t_n, n \in \mathbb{N}$. $\Diamond$

Proof. Let $A \in \mathcal{E}^{\otimes T}$. By Lemma 1.27, there exists a countable subset $I \subset T$ with $A \in \sigma(\Pi_t \colon t \in I)$. We write $A)B \times E^{T\setminus I}$ for some $B \in \mathcal{E}^{\otimes I}$. By the previous theorem, there exists a unique probability measure $\mathbb{P}_I$ on $\mathcal{E}^{\otimes I}$ with

$$p_{t_1,\ldots,t_n}(A_1 \times \cdots \times A_n) = \mathbb{P}_I(\Pi_{t_i} \in A_i \forall i \leq n)$$

for all $A_1, \ldots, A_n, t_1, \ldots, t_n$ and $n \in \mathbb{N}$. $\square$

Define $\mathbb{P}(A) = \mathbb{P}_I(B)$ if $A = B \times E^{T\setminus I}$ for some countable $I \subseteq I$ and some $B \in \mathcal{E}^{\otimes I}$. By consistency, $\mathbb{P}$ is well-defined and finitely additive. For σ-additivity, let $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{E}^{\otimes T}$ be disjoint. Then for each n there exists a countable set $I_n \subseteq T$ and $B_n \in \mathcal{E}^{\otimes T_n}$ with $A_n = B_n \times E^{T\setminus I_n}$. Set $I = \bigcup_{n\in\mathbb{N}} I_n$ the we have $A_n = \tilde{B}_n \times E^{T\setminus I}$ for all n and the $\tilde{B}_n$ are in $\mathcal{E}^{\otimes I}$ and are disjoint. Thus,

$$\mathbb{P}\left(\cup_{n\in\mathbb{N}} A_n\right) = \mathbb{P}_I\left(\cup_{n\in\mathbb{N}} \tilde{B}_n\right) = \sum_{n=1}^{\infty} \mathbb{P}_I(\tilde{B}_n) = \sum_{i=1}^{n} \mathbb{P}(A_1),$$

and we are done.

1.33 Corollary: An $\mathbb{R}^d$-valued stochastic process fulfilling (B0)-(B3) from (1.4) exists. Explicitly, there exists a unique probability measure $\mathcal{W}$ on $\left(\left(\mathbb{R}^d\right)^{\mathbb{R}_0^+}, \left(\mathcal{B}(\mathbb{R}^d)\right)^{\otimes\mathbb{R}_0^+}\right)$ such that the random variables

$$\begin{aligned} B_t \colon \Omega &\to \mathbb{R}^d \\ \omega &\mapsto B_t(\omega) \coloneqq \omega(t) \end{aligned}$$

fulfill (B0)-(B3). $\mathcal{W}$ is called *pre-Wienermeasure* $\Diamond$

It remains to see (B4). Note that the statement $\mathcal{W}(C(\mathbb{R}^+, \mathbb{R}^d)) = 1$ makes no sense, as $C(\mathbb{R}^+, \mathbb{R}^d) \notin \mathcal{F}$. We need to be more careful.

1.34 Definition: Let $D \subseteq \mathbb{R}_0^+$ be open and $\alpha > 0$. A function $f \colon \mathbb{R}_0^+ \to \mathbb{R}^d$ is called *α-Hoelder continuous on D* if the Hoelder-seminorm is finite. f is locally α-HC on D if it is Hoelder-continuous on D if it is Hoelder-continuous on each $D \cap [0, n]$. We write $f \in C^\alpha(D)$ or $f \in C^\alpha_{loc}(D)$ ◊

1.35 Remark:

a) If D has no cluster points, any function on D is an element of C^α_{loc}.

b) Usually, D is a dense subset of some interval $[a, b] \subseteq \mathbb{R}_0^+$.

c) In the case of b), $f \in C^\alpha(D)$ can be uniquely extended to $[a, b]$ by the usual extension. Uniqueness follows from the continuity.

d) If D is dense and $f \in C^\alpha_{loc}(D)$ for some $\alpha > 1$, then f is constant.

e) $f \in C^1_{loc}(\mathbb{R})$ if and only if f is locally Lipschitz.

f) $\|\cdot\|_{D,\alpha}$ is only a seminorm, as it does not detect constants.

g) The map $\|\cdot\|_{D,\alpha}$ is $\mathcal{B}(\mathbb{R}^d)^{\otimes \mathbb{R}_0^+}$-$\mathcal{B}([0,\infty])$-measurable if D is countable.

◊

1.36 Theorem: Let $(X_t)_{t \in [0,T]}$ be an $\mathbb{R}^d$-valued stochastic process. Assume that there exist $q \geq 2$, $\beta > \frac{1}{q}$, $C < \infty$ such that

$$\mathbb{E}\left(|X_{s,t}|^q\right) \leq C\,|t-s|^{\beta q} \tag{1.2}$$

holds for all $s, t \in [0, T]$ with $|t - s| < \frac{1}{2}$. Let $D \coloneqq \{k \cdot 2^{-n} \colon k, n \in \mathbb{N}\} \cap [0, T]$ (dyadic rationals). Then

$$\mathbb{E}\left(\|X\|^q_{D,\alpha}\right) < \infty$$

holds for all $\alpha \in [0, \beta - \frac{1}{q})$. ◊

Proof. Let $D_n \coloneqq \{k \cdot 2^{-n}\} \cap [0, T]$. Then we have $D = \bigcup_{n \in \mathbb{N}} D_n$. We define

$$K_n(\omega) \coloneqq \max\{|X_{t,t+2^{-n}}(\omega)| : t \in D_n\}.$$

Then by (1.2) it holds that

$$\mathbb{E}\left(K_n^q\right) \leq \mathbb{E}\left(\sum_{t \in D_n} |X_{t,t+2^{-n}}|^q\right)$$

$$\leq |D_n| C(2^{-n})^{\beta q} \leq T 2^n C(2^{-n})^{\beta q}$$
$$= CT(2^{-n})^{\beta q - 1}.$$

Now let $s < t$, $s, t \in D$. If $|t - s| > \frac{1}{2}$ we pick $t_1, \ldots, t_m \in D$ with $t_0 = s$, $t_m = t$ and $|t_i - t_{i-1}| < \frac{1}{2}$ for all i. Since $|X_t - X_s| \leq \sum_{i=1}^m \left|X_{t_i} - X_{t_{i-1}}\right|$, we find

$$\frac{|X_t - X_s|}{|t-s|^\alpha} \leq \sum_{i=1}^n \frac{\left|X_{t_i} - X_{t_{i-1}}\right|}{|t_i - t_{i-1}|^\alpha}$$
$$\leq m \sup\{\frac{|X_t - X_s|}{|t-s|} : s, t \in D, |t-s| < \frac{1}{2}\}.$$

So it suffices to consider $|t - s| \leq \frac{1}{2}$. Then there exists $j \in \mathbb{N}$ with $2^j < t - s \leq 2^{-j-1}$ and $N \in \mathbb{N}$ so that $s, t \in D_N$. We define $A_j \colon D_j \cap (s, t)$, then $|D_j| \in \{1, 2\}$. Pick $t_j^- := \min A_j$ as well as $t_j^+ := \max A_j$ and set $A_{j+1}^- \cap [s, t_j^-)$ then $\left|A_{j+1}^-\right| \in \{0, 1\}$. Pick $t_{j+1}^- := \min\{t_j^{-1}, \inf A_{j+1}^-\}$. Analogously for t_{j+1}^+ via A_{j+1}^+ with min and inf replaced by max and sup. Inductively, $A_{j+l}^- := D_{j+l} \cap [s, t_{j+l-1}^-)$, $t_{j+l}^- := \min\{t_{j+l-1}, \inf A_{j+l}^-$, $A_{j+l}^+ := D_{j+l} \cap (t_{j+l-1}^+, t]$, $t_{j+l}^+ := \max\{t_{j+l-1}^+, \sup A_{j+l}^+\}$. This will stop when $j+l = N$ with $t_N^- = s, t_N^+ = t$. Now,

$$|X_{s,t}| \leq \sum_{l=0}^{N-j} \left|X_{t_{j+l}^-}(\omega) - X_{t_{j+l-1}^-}(\omega)\right| + \sum_{l=0}^{N-j} \left|X_{t_{j+l}^+}(\omega) - X_{t_{j+l-1}^+}(\omega)\right|$$

Each term is either equal to 0 or the difference of some $\left|X_{t_{j+l}^\pm} - X_{t_{j+l}^\pm + 2^{-(j+l)}}\right|$. So, $(**) \leq 2 \sum_{l=0}^{N-j} Kj + l(\omega) \leq 2 \sum_{l=j}^\infty K_l(\omega)$. Since we assumed $|t - s| > 2^{-j}$, we get

$$\frac{|X_{s,t}(\omega)|}{|t-s|^\alpha} \leq 2^{j\alpha} \cdot 2 \sum_{l=j}^\infty K_l(\omega)$$
$$\leq 2 \sum_{l=0}^\infty 2^{l\alpha} K_l(\omega).$$

The right hand side above does not depend on s and t. Thus,

$$\mathbb{E}\left(\|X\|_{D,\alpha}^q\right)^{\frac{1}{q}} \leq \mathbb{E}\left((2T)^q (2 \sum_{l=0}^\infty 2^{l\alpha} K_l)^q\right)^{\frac{1}{q}}$$
$$= 4T \left|\sum_{l=0}^\infty 2^{l\alpha} K_l\right|_{L^q} \leq 4T \sum_{l=0}^\infty 2^{l\alpha} |K_l|_{L^q}$$

$$\leq 2CT \sum_{l=0}^{\infty} 2^{l\alpha} 2^{-l\frac{\beta q-1}{q}}$$

where we used $(*)$ in the last step. For $\alpha < \beta - \frac{1}{q}$ this sum is finite. $\square$

1.37 Corollary: Let $(X_t)_{t\in\mathbb{R}_0^+}$ be an $\mathbb{R}^d$-valued stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$ which fulfills (1.2) for all T.

a) (Kolmogorov-Chentsov-Than) There exists a stochastic process $(\tilde{X}_t)_{t\in\mathbb{R}_0^+}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ such that

   i) for all $\omega \in \Omega$ the map $t \mapsto \tilde{X}_t(\omega)$ is continuous, we say that $\tilde{X}$ has continuous paths.

   ii) for all $\in \mathbb{R}_0^+$ we have $\mathbb{P}(X_t = \tilde{X}_t$, this is to say $\tilde{X}$ is a *version* of X.

b) On the canonical space $(C(\mathbb{R}_0^+, \mathbb{R}^d), \mathcal{F}_{can})$ there exists a probability measure $\mathbb{P}_0$ so that with

$$Y_t(\hat{\omega}) = \hat{\omega}(t) \quad \forall \hat{\omega} \in C(R_0^+, \mathbb{R}^d)$$

the processes (X_t) and (Y_t) have the same fdd.

$\diamondsuit$

Proof. By Thm 1.36 we have $\mathbb{E}\left(\|X\|_{D_T,\alpha}^q\right) < \infty$ for some $\alpha > 0, q \geq 2$ and all $D_T = D_{\mathbb{R}_0^+} \cap [0,T]$. Hence $\mathbb{P}(\|X\|_{D_T,\alpha} < \infty) = 1$ for all T and therefore $\mathbb{P}(X \in C_{loc}^\alpha(D)) = 1$. The set $\{\omega \in \Omega \colon X(w) \in C_{loc}^\alpha(D)) =: \Omega_0$ depends on countably many indices and is thus measurable. Define

$$\tilde{X}_t(\omega) := \begin{cases} \lim_{t_n \to t} X_{t_n}(\omega), & \text{for } t_n \in D, w\omega \in \Omega_0, \\ 0, & \text{otherwise }. \end{cases}$$

$\tilde{X}_t(\omega)$ is independent of the approximating sequence and $t \mapsto \tilde{X}_t(\omega)$ is continuous for all $\omega \in \Omega$. Hence (i) holds. For (ii), let

$$\begin{aligned} Z_t(\omega) &:= X_t(\omega) - \tilde{X}_t(\omega) \\ Z_{t_n} &:= X_t(\omega) - \tilde{X}_{t_n}(\omega) \end{aligned}$$

for $t_n \in D$, $t_n \to t$. Then $Z_{t_n}(\omega) \to Z_t(\omega)$ on Ω_0, i.e. almost surely, and by (1.2)

$$\mathbb{P}(|Z_{t_n}| > \varepsilon) \leq \frac{1}{\varepsilon^q}\mathbb{E}\left(|X_{t_n} - X_t|^q\right)$$

$$\leq C\frac{1}{\varepsilon^q}|t_n - t|^{\beta q} \to 0,$$

$Z_{t_n} \to 0$ in probability and hence $Z_t = 0$ almost surely.

The map

$$\begin{aligned} F\colon \Omega &\to \Omega_0 \\ \omega &\mapsto (\tilde{X}_t(\omega))_{t\in\mathbb{R}_0^+} \end{aligned}$$

is measurable. Then $\mathbb{P}_0 = \mathbb{P} \circ F^{-1}$. □

1.38 Theorem: a) There exists a stochastic process satisfying B0-B4, i.e. a Brownian motion.

b) There exists a probability measure $\mathcal{W}$ on $(C_0, \mathcal{F})$ such that under $\mathcal{W}$ the projection $\omega \mapsto \pi_t(\omega) = B_t(\omega) = \omega(t)$ form a Brownian motion. The space $(C_0, \mathcal{F}, \mathcal{W})$ is called *Wiener space.*

◊

Proof. (i) By 1.22 we can restrict to 1-d BM. By Cor 1.33 a process $\tilde{B}_t)$ with B0-B3 exists. By B3

□

2 Properties of Brownian Motion

Brownian Motions have a lot of the general properties an SP might have. Hence, we can do a lot with Brownian Motions we can not do in a more general setting.

Invariance properties

The statements are far more interesting than the proofs, which are too easy to help in understanding the properties. We can do all kinds of funny things to Brownian Motionand obtain Brownian Motions again.

When we just transform the values of a Brownian Motion, we can get new Brownian Motions quite easily.

2.1 Proposition (Orthogonal invariance): Let $\mathbf{B}$ be a Brownian Motion$^{\mathrm{d}}$and U a d by d orthogonal matrix then $(U\mathbf{B}_t)_{t\in T}$ is a Brownian Motion$^{\mathrm{d}}$. In particular, $-\mathbf{B} = (-\mathbf{B}_t)_{t\in T}$ is a Brownian Motion$^{\mathrm{d}}$. ◊

Proof. With $UU^* = \mathrm{I}_d$ we have using Proposition 1.14 that

$$(U\mathbf{B})_{s,t} \sim \mathcal{N}\left(0, (t-s)\mathrm{I}_d\right)$$

holds true. The other properties are clear. □

Instead of transforming the values, we can also shift the process in time. This is an important property which to understand will be quite useful.

2.2 Proposition (Time shift invariance): Let $\mathbf{B}$ be a Brownian Motion$^{\mathrm{d}}$ and $a \in \mathbb{R}_{\geq 0}$. Then $(\mathbf{B}_{t+a} - \mathbf{B}_a)_{t\geq 0}$ is a Brownian Motion$^{\mathrm{d}}$. ◊

Proof. The proof remains as an exercise. □

The next property is very fundamental and central to our understanding of Brownian Motions. The path a Brownian Motiontakes is entirely independent of the path taken so far except for the value the Brownian Motionnow starts at. If we forget where we are we essentially delete all memory of the past. For all relevant purposes, the value we take at the current point in time is all we know about what has happened so far.

2.3 Proposition (Memoryless property, elementary Markov property)**:** Let $(\mathbf{B}_t)$ be a Brownian Motion[d] and $a \in \mathbb{R}_{\geq 0}$. The the SPs $(\mathbf{B}_t)_{0 \leq t \leq a}$ and $(\mathbf{W}_t)_{t \geq 0} :- (\mathbf{B}_{t+a} - \mathbf{B}_t)_{t \geq 0})$ are independent. In particular,

$$\mathbb{E}\left(F\left((\mathbf{W}_t)_{t\geq 0}\right) \mid \sigma(\mathbf{B}_s \colon s \leq a)\right) = \mathbb{E}\left(F\left(\mathbf{B}_t\right)_{t \geq 0}\right))$$

for any measurable function $F \colon \mathrm{C}_0 \to \mathbb{R}$. ◊

Proof. Idea: fdd and intersection stable generator of the σ-algebras Transform the fdd with

$$A = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$$

to get another int stable generator. Do this for both processes.

$$W_{t_i} - W_{t_{i-1}} = B_{t_i + a} - B_{t_{i-1}+a}$$

Use (B1). P-Theo: Indep of int-stable generator means indep of sigma-alg. □

A diffusive rescaling is a rescaling of a process where we rescale the time, e.g. let it run twice as fast, where we also transform the values to offset the rescaling. If we were to double the speed of a linear function and then were to divide the values by 2 we would obtain the original function again. In case of Brownian Motionwe obviously will not be able to recover the exact function but we can get a process of identical distribution.

2.4 Proposition (Invariance under diffusive rescaling)**:** ◊

Proof. Exercise □

Imagine ink in water. If we let time run 100 times as fast the ink will seem to disperse at (merely) ten times the speed.

The function recovered exactly by this rescaling is the square root. This means that a Brownian Motionlooks "locally like a square root except it doesn't". Brownian Motionvery rough and does not look like sqrt at all. (It only looks like sqrt on small and large scales. –.–) Sqrt leaves 0 very fast but gets slower over time.

If we stop a Brownian Motionat some point in time and let it play backwards we obtain a Brownian Motionagain.

2.5 Proposition (Time reversal symmetry): ◊

Proof. Exercise □

Another way in which Brownian Motions look like the sqrt:

$$\sqrt{t} = t\sqrt{\frac{1}{t}}$$

relates behaviour at zero to behaviour at infinity How a bm looks at infinity how it looks at zero a very strong symmetry

The involution is not a Brownian Motionon the original probability space since there will be paths that are not continuous in zero. However, nearly all paths are continuous and we can simply throw away the remaining points.

2.6 Proposition (Time involution invariance): ◊

Proof. The hard part of the proof is continuity at zero

Something that is often very useful when dealing with SP: Write out definition of the limit in quantors over countable sets. Intersection with $\mathbb{Q}$ is sufficient since we have continuity outside of zero

$\mathcal{F} \cap \Omega_0 := \{F \cap \Omega_0 : F \in \mathcal{F}\}$ □

We now change our perspective. So far, we took values and time and did things to that. We took a function from R to R and transformed it in some way. All of these can be written as operators on function spaces to obtain measure-preserving maps.

2.7 Remark (Invariances of Wiener measure): ◊

Think of C_0 like a big $\mathbb{R}^n$. Many maps will not preserve the measure but some do.

Martingale properties of Brownian Motion

In probability theory often only martingales on discrete sets are defined but it works just as well on continuous sets.

2.8 Definition: The third condition is the one that makes a martingale a martingale

sub/super harmonic functions

sub-martingale sup-martingale ◊

2.9 Proposition (2.10)**:** BM is an $\mathcal{F}_t^B$ martingale ◊

Later we will see that $\mathcal{F}^B$ is sometimes to small. Replacing it by a larger filtration can in principle destroy the martingale property.

2.10 Definition (admissble filtration)**:** ◊

Now, we check that a BM is a martingale wrt an admissible martingale.

2.11 Proposition: BM is an $\mathcal{F}_t$ martingale. ◊

Proof. same as for 2.10 as we used only independence of increments. □

2.12 Proposition: (B_t) BM $\mathcal{F}_t$ admissible. Then the following SPs are martingales.

i) $M_t := |B_t|^2 - dt$

ii) $M_t^v := \mathrm{e}^{(v,B_t)-\frac{t}{2}|v_R|^2+\frac{t}{2}|v_I|^2}$ for all $v = v_R + \mathrm{i}V_I \in \mathbb{C}^d$. (for real ξ this may look like a Fourier transform).

iii) For $d = 1$, $M_t^n := t^{\frac{n}{2}} H_n(t^{-\frac{1}{2}} B_t)$ where H_n is the n-th Hermite polynomial, i.e. , $H_n(x) = (-1)^{\frac{1}{2}} \mathrm{e}^{-\frac{x^2}{2}} \partial_x^n (\mathrm{e}^{-\frac{x^2}{2}})$. These Hermite polynomials are an orthogonal basis in a weighted L^2 space, i.e. ,

$$\int H_n(x) H_m(x) \mathrm{e}^{-\frac{x^2}{2}} \,\mathrm{d}x = \delta_{n,m}$$

when weighted by the correct factor (probably some $\frac{1}{\sqrt{2\pi}}$).

◊

Proof. i) It suffices to consider the case $d = 1$ since

$$|B_t|^2 - dt = \sum_{j=1}^{d} \left(\left| B_t^j \right|^2 - t \right).$$

We need to introduce increments in order to use the properties on an admissible filtration.

$$\mathbb{E}\left(B_t^2 \mid \mathcal{F}_s\right) = \mathbb{E}\left((B_t - B_2)^2 + 2B_sB_t - B_s^2 \mid \mathcal{F}_s\right)$$

and use linearity and independence.

ii) First need to check integrability, $\mathbb{E}\left(|M_t^v| \infty\right)$ by Corollary 1.12 Now, $\mathbb{E}\left(M_t^v \mid \mathcal{F}_s\right) = \mathrm{e}^{-\frac{t}{2}|v|^2} \mathrm{e}^{(v,B_2)} \mathbb{E}\left(\mathrm{e}^{(v,B_t - B_s)} \mid \mathcal{F}_s\right)$ and use what we calculated in 1.12.

iii) exercise: expand $M_t^\alpha := \mathrm{e}^{\alpha B_t} \mathrm{e}^{-\alpha^2 \frac{t}{2}}$ and sort by powers of α.

□

All martingales so far were of the form $M_t = f(t, B_t)$ for some function f. But there are more general martingales. In order to construct those we need following important lemma. It is a classical theorem in theory of PDEs.

2.13 Lemma: The transition density

$$p_t(x) = \frac{1}{(2\pi t)^{\frac{d}{2}}} \exp(-\frac{|x|^2}{2t})$$

of BM^d solves the *(scaled) heat equation*, i.e. ,

$$\partial_t p_t(x) = \frac{1}{2} \Delta p_t(x)$$

for all $t > 0$ and $x \in \mathbb{R}^d$. ◊

Proof. We consider the Fourier transform of the density as then it will be easier to calculate derivatives

$$\hat{p}_t(k) := \int \mathrm{e}^{\mathrm{i}(k,x)} p_t(x) \, \mathrm{d}x = \mathrm{e}^{-\frac{1}{2}|k|^2 t}$$

Clearly $\partial_t \hat{p}_t(k) = -\frac{1}{2}|k|^2 \hat{p}_t(k)$ and by Dominated convergence we may change derivative and integration

$$\int e^{i(k,x)} \partial_t p_t(x) \, dx = -\frac{1}{2} \int |k|^2 e^{i(k,x)} p_t(x) \, dx$$

(...) and conclude by integration by parts 2 times in all coordinates. In order to conclude the desired equality we use that Fourier transform is an isomorphism which shows that both terms coincide almost everywhere and as they are continuous they coincide everywhere. □

Main Theorem for today

2.14 Theorem: Let (B_t) be a Brownian Motion, $\mathcal{F}_t$ admissible. Let $f \in C([0,\infty) \times \mathbb{R}^d, \mathbb{R}) \cap C((0,\infty) \times \mathbb{R}^d, \mathbb{R})$ and assume that there exists $C < \infty$ and a locally bounded function $t \mapsto c(t)$ with

$$\max\{\partial_t f(t,x), \partial_{x_i}^j f(t,x), f(t,x) \colon 1 \leq i \leq d, j = 1, 2\} \leq c(t) e^{C|x|}.$$

technical assumption, most functions do satisfy this Define for $t \geq 0$

$$M_t^f := f(t, B_t) - f(0, B_0) - \int_0^t (Lf)(r, B_r) \, dr$$

where $(Lf)(t,x) = \partial_t f(t,x) + \frac{1}{2}\Delta f(t,x)$. Then $(M_t^f)_{t\geq 0}$ is an $\mathcal{F}_t$ martingale. It is called the *fundamental martingale* wrt f. ◊

For the proof we note the following lemma, which we have already seen in the discrete case.

2.15 Lemma (Doob's maximal inequality)**:** Let (M_t) be a nonnegative submartingale with continuous paths. Then for all $t \geq 0$ and $p > 1$ we have

$$\mathbb{E}\left(\sup_{s\leq t} M_s^p\right) \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left(M_t^p\right)$$

◊

Proof. exercise, use discrete Martingales and 4.39 from Probability Theory. □

Proof of Theorem 2.14. Fix $\overline{\omega}$ and calculate

$$\mathbb{E}\left(M_t^f \mid \mathcal{F}_s\right)(\overline{\omega}) = f(s, B_s(\overline{\omega})) - f(0, B_0(\overline{\omega})) \\ - \int_0^t (Lf)(r, B_r(\overline{\omega}))\ \mathrm{d}r + \mathbb{E}\left(f(t, B_t) - f(s, B_s) - \int_s^t (Lf)(r, B_r)\ \mathrm{d}r \mid \mathcal{F}_s\right)(\overline{\omega}) = (*).$$

By Proposition 2.23 of PT we know that if $X \amalg Y$ then

$$\mathbb{E}\left(h(X, Y) \mid \sigma(Y)\right)(\overline{\omega}) = \mathbb{E}\left(h(X, Y(\overline{\omega})\right).$$

Applying this with $X = (B_r - B_s)_{r \geq s}$, $Y(s) = (B_r)_{r \leq s}$ and

$$h(x, y) = f(t, B_{s,t} - B_s) - f(s, B_s) - \int_s^t (Lf)(r, B_{r,s} - B_s)\ \mathrm{d}r$$

we obtain from above

$$(*) = M_s^f(\overline{\omega}) + \mathbb{E}\left(f(t, B_{s,t} - B_s(\overline{\omega})) - f(s, B_s(\overline{\omega}) - \int_s^t (Lf)(r, B_{r,s} - B_s(\overline{\omega}))\ \mathrm{d}r\right. \\ = M_s^f(\overline{\omega}) + \mathbb{E}\left(\tilde{f}_{\overline{\omega}}(t-s, B_{t-s}) - \tilde{f}_{\overline{\omega}}(0,0) - \int_0^{t-s} (L\tilde{f}_{\overline{\omega}}(r, B_{r-s})\ \mathrm{d}r\right)$$

with $\tilde{f}_{\overline{\omega}}(u, x) = f(u+s, x - B_s(\overline{\omega}))$. We used that L commuters with shifts in the parameter space. The function $\tilde{f}_{\overline{\omega}}$ also fulfills of Thm 2.15, so the theorem will be proved if we can show $\mathbb{E}\left(M_t^f\right) = 0$ for all allowed f, all t. Let $\varepsilon > 0$ be arbitrary. Then by Fubini and the assumptions on f we have

$$\begin{aligned}\mathbb{E}\left(\int_\varepsilon^t (Lf)(s, B_s)\ \mathrm{d}s\right) &= \int_\varepsilon^t \mathbb{E}\left((Lf)(s, B_s)\right)\ \mathrm{d}s \\ &= \int_\varepsilon^t \int p_s(x)(Lf)(s, x)\ \mathrm{d}x\ \mathrm{d}s \\ &= \int \mathrm{d}x \int_\varepsilon^t p_s(x)(\partial_s f)(s, x)\ \mathrm{d}s + \int_\varepsilon^t \mathrm{d}s \int p_s(x)(\frac{1}{2}\Delta f)(s, x) = (...)\end{aligned}$$

Hence, we get $\mathbb{E}\left(M_t^f\right) = \mathbb{E}\left(f(\varepsilon, B_\varepsilon) - f(0, B_0)\right) - \mathbb{E}\left(\int_0^\varepsilon (Lf)(s, B_s)\ \mathrm{d}s\right)$ for all t and $\varepsilon > 0$. It remains to show $\mathbb{E}\left(M_t^f\right) \to 0$ as $\varepsilon \to 0$. We estimate the last summand using Jensen's formula and obtain

$$\begin{aligned}\left|\mathbb{E}\left(\int_0^\varepsilon (Lf)(s, B_s)\ \mathrm{d}s\right)\right| &\leq \mathbb{E}\left(\int_0^\varepsilon |Lf(s, B_s)|\ \mathrm{d}s\right) \\ &\leq \int_0^\varepsilon c(s)\mathbb{E}\left(\mathrm{e}^{|B_s|}\right)\ \mathrm{d}s \to 0\end{aligned}$$

as $\varepsilon < 0$ as the integrand can be bounded uniformly by $\mathbb{E}\left(\mathrm{e}^{|B_\varepsilon|}\right)$ multiplied by some constant, since $\mathrm{e}^{|B_s|}$ is a submartingale. For the other expression, Doobs maximal inequality will be the key tool. We have

$$\begin{aligned}\mathbb{E}\left(\sup_{s\leq\varepsilon}|f(s,B_s)|\right) &\leq \sup_{s\leq\varepsilon} c(s)\mathbb{E}\left(\sup_{s\leq\varepsilon}\left(\mathrm{e}^{\frac{1}{2}C|B_s|^2}\right)\right)\\ &\leq \sup_{s\leq\varepsilon} c(s)\cdot 4\cdot\mathbb{E}\left(\mathrm{e}^{C|B_s|}\right).\end{aligned}$$

By the continuity of f and $t\mapsto B_t$ we have

$$\lim_{\varepsilon\to 0} f(\varepsilon, B_\varepsilon(\omega)) = f(0,0)$$

for all ω. The dominated convergence theorem then yields $\mathbb{E}\left(f(\varepsilon,B_\varepsilon)\right)\to f(0,0)$ as $\varepsilon\to 0$. □

2.16 Example:

a) Let $f(x,t)=x^2$. Then $Lf=1$ and we have $M_t^f = B_t^2 - t$ (see 2.12 a)).

b) Let $f(x,t)=x^3$. Then $Lf=3x$ and $M_t^f = B_t^3 - 3\int_0^t B_s\,\mathrm{d}s$.

c) Let $f\colon\mathbb{R}^d\to\mathbb{R}$ be harmonic, i.e. $\Delta f=0$. Then $Lf=0$ and hence $(f(B_t))_t$ is a martingale. For example for $d=2$ and $f(x,y)=\mathrm{e}^x\sin(y)$ then for any two-dimensional Brownian motion also $(\mathrm{e}^{B_t{}^{(1)}}\sin(B_t^{(2)}))_{t\geq 0}$ is a martingale.

d) Let $\mathbf{B_t} = B_t^{(1)} + \mathrm{i}B_t^{(2)}$ be a complex Brownian motion where $B_t^{(j)}$ are independent BMs. Let h be an entire function with not too much growth at ∞ Then $(h(\mathbf{B_t}))_{t\geq 0}$ is a martingale. The reason for this is that the real as well as the imaginary part of h are harmonic (due to the Cauchy-Riemann differential equations).

◊

Stopping times

2.17 Definition: Let $(\mathcal{F}_t)$ be a filtration. A *random time* is a $\mathbb{R}_0^+\cup\{\infty\}$-valued random variable. A *stopping time* with respect to $(\mathcal{F}_t)$ is a random time τ with $\{\tau\leq t\}\in\mathcal{F}_t$ for all $t\in\mathbb{R}_0^+$. ◊

2.18 Example: Let (X_t) be an $(\mathcal{F}_t)$-adapted stochastic process with continuous paths and $F \subseteq \mathbb{R}^d$ a closed set. Then the random variable

$$\tau_F \coloneqq \inf\{s \geq 0 \colon X_s \in F\}$$

is a stopping time called the *hitting time of* F. $\Diamond$

Proof. Define $d(x, F) = \inf\{|x - y| : y \in F\}$ which is continuous in x. By definition we have $\tau_F(\omega) \leq t$ if and only if $\inf\{s \geq 0 \colon X_s(\omega) \in F\} \leq t$. As F is continuous and (X_t) has continuous paths this is equivalent to the existence of $x \in F, s \leq t$ such that $X_s = x$. Using the continuity of the distance function an equivalent formulation is

$$\inf\{d(X_s(\omega), F), s \in [0, t]\} = 0.$$

Due to the continuity we can restrict ourselves to $s \in [0, t9 \cap \mathbb{Q}$. Since $\{\omega \colon \inf\{d(X_s(\omega), F), s \in [0, t]\} = 0\} \in \mathcal{F}_t$, the claim follows. $\square$

Obviously, this proof does not work for open sets, so we need a larger filtration.

2.19 Definition: Let $\mathbf{X}$ be a stochastic process. The filtration (check this!)

$$\mathcal{F}^{\mathbf{X}}_{t+} \coloneqq \bigcap_{u>t} \mathcal{F}^{\mathbf{X}}_u = \bigcap_{n=1}^{\infty} \mathcal{F}^{\mathbf{X}}_{t+\frac{1}{n}}$$

is called the *right-continuous completion* of $(\mathcal{F}^{\mathbf{X}}_t)$. Any filtration with the property $\mathcal{F}_t = \bigcap_{n\geq 1} \mathcal{F}_{t+\frac{1}{n}}$ is called *right-continuous.* $\Diamond$

Now we can formulate an analogous result for open sets.

2.20 Lemma: Let $\mathbf{X}$ be a stochastic process with right-continuous paths and let $U \subseteq \mathbb{R}^d$ be open. Then

$$\tau_U \coloneqq \inf\{s \geq 0\} \colon X_s \in U\}$$

is an $(\mathcal{F}_{t+})$-stopping time. $\Diamond$

Proof. We have

$$\begin{aligned}
\{\tau_U \leq t\} &= \bigcap_{n\in\mathbb{N}} \{\tau_U < t + \frac{1}{n}\} \\
&= \bigcap_{n\in\mathbb{N}} \bigcup_{j<t+\frac{1}{n}} \{X_j \in U\} \\
&= \bigcap_{n\in\mathbb{N}} \bigcup_{j\in[0,\frac{1}{n})\cap\mathbb{Q}} \{X_j \in U\} \in \mathcal{F}_{t+}.
\end{aligned}$$

$\square$

2.21 Example: Let (B_t) be a two-dimensional Brownian motion and $\tau_b := \inf\{t \geq 0 \colon B_t = b\}$ for $b \in \mathbb{R}^2$ be the *first passage time of* b. Then $\mathbb{P}(\tau_b < \infty) = 1$ for all b. In fact, $t \mapsto B_t(\omega)$ infinitely often for almost all ω. This follows from the next proposition. $\Diamond$

2.22 Proposition: Let (B_t) be a one-dimensional Brownian motion. Then

$$\mathbb{P}\left(\limsup_{t\to\infty} \frac{1}{\sqrt{t}} B_t = \infty\right) = \mathbb{P}\left(\liminf_{t\to\infty} \frac{1}{\sqrt{t}} B_t = \infty\right) = 1.$$

Moreover, for all $t_0 \in \mathbb{R}_0^+$ we have

$$\mathbb{P}\left(\omega \colon B \text{ is not differentiable at } t_0\right) = 1.$$

$\Diamond$

Proof. It is clear that

$$\begin{aligned}\limsup_{t\to\infty} \frac{1}{\sqrt{t}} B_t &\geq \limsup_{n\in\mathbb{N}, n\to\infty} \frac{1}{\sqrt{n}} B_n \\ &= \limsup_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} B_{k,k-1}\end{aligned}$$

holds true where the $B_{k,k-1}$ are i.i.d. random variables with variance 1. By [PT, (2.50a)] we have $\mathbb{P}\left(\limsup_{n\to\infty} \frac{1}{\sqrt{n}} B_n = \infty\right) = 1$ which implies that the first two statements hold. For the third, note that by the time shift invariance it is enough to consider $t_0 = 0$. By (2.6) we have

$$\begin{aligned}\mathbb{P}\left(\limsup_{t\to\infty} \frac{1}{\sqrt{t}} B_t = \infty\right) &= \mathbb{P}\left(\limsup_{t\to\infty} \sqrt{t} B_{\frac{1}{t}} = \infty\right) \\ &= \mathbb{P}\left(\limsup_{t\to\infty} \frac{1}{\sqrt{t}} t B_{\frac{1}{t}} = \infty\right) = 1.\end{aligned}$$

The same calculation can be done for $\liminf$. Hence, $\frac{1}{t}(B_t - B_0)$ does not converge almost surely. $\square$

2.23 Remark: It follows that $\mathbb{P}(t \to B_t \text{ is differentiable at any } t \in \mathbb{Q}_0^+) = 0$. With more work, one can show

$$\mathbb{P}(\lambda(\{t \in \mathbb{R}_0^+ \colon B \text{ differentiable at } t\}) = 0) = 1.$$

With even more work, one can show

$$\mathbb{P}(\exists t \geq 0 \colon B \text{ differentiable at } t) = 0.$$

$\Diamond$

Now, we return to considering stopping times.

2.24 Proposition: Let τ, σ and $(\tau_n)_{n\in\mathbb{N}}$ with respect to the filtration $(\mathcal{F}_t)$. Then we have

a) $\{\tau < t\} \in \mathcal{F}$ for all t and if $(\mathcal{F}_t)$ is right-continuous then $\{\rho < t\} \in \mathcal{F}_t$ for all t then ρ is a stopping time.

b) $\tau + \sigma$, $\tau \wedge \sigma$, $\tau \vee \sigma$ and $\sup_{n\in\mathbb{N}} \tau_n$ are $\mathcal{F}_t$-stopping times.

c) If $(\mathcal{F}_t)$ is right-continuous then $\inf_n \tau_n$, $\liminf_{n\to\infty} \tau_n$ and $\limsup_{n\to\infty} \tau_n$ are stopping times.

◊

Proof. This remains as an exercise. □

2.25 Lemma (Approximation)**:** Let τ be a stopping time. For $n, m \in \mathbb{N}_0$ define $\tau_n(\omega) = (m+1)2^{-n}$ if $m2^{-n} \leq \tau(\omega) < (m+1)2^{-n}$ and $\tau_n(\omega) = \infty$ if $\tau(\omega) = \infty$. Then τ_n is an $(\mathcal{F}_t)$-stopping time, taking countably many values. Also $\tau_n(\omega) \searrow \tau(\omega)$ for all $\omega \in \Omega$. ◊

Proof. Also this remains as an exercise. □

2.26 Definition: Let $(\mathcal{F}_t)$ be a filtration and τ be an $(\mathcal{F}_t)$-stopping time. The σ-algebra (check this!)

$$\mathcal{F}_t \coloneqq \{A \in \mathcal{F} \colon A \cap \{\tau \leq t\} \mathcal{F}_t \forall t \geq 0\}$$

is called the *σ-algebra of τ's past.* ◊

2.27 Example: Let $C^0(\mathbb{R}_0^+, \mathbb{R}^d)$ and $\mathcal{F}_t = \sigma(\pi_s \colon s \leq t)$ then for $A \in \mathcal{F} = \sigma\left(\cup_{t\geq 0}\mathcal{F}_t\right)$, we have $A \in \mathcal{F}_t$ if and only if for all $\omega \in \Omega$, the values of $(\pi_s)_{s\leq t}$ determine whether $\omega \in A$. Moreover, we have $A \in \mathcal{F}_\tau$ if and only if for all $\omega \in \Omega$ such that $\tau(\omega) \leq t$ the values of $(\pi_s)_{s\leq t}$ determine whether $\omega \in A$. In the case that $\tau(\omega)$ there is no restriction. Hence, the first condition is the important one. For example consider closed sets $F, G, H \subseteq \mathbb{R}^d$ and their entry times τ_F, τ_G, τ_H as in (2.20). Check whether the set

$$\{\omega \colon t \mapsto \pi_t(\omega) \text{ first hits } F\text{, then } G\text{, then } H\}$$

is in $\mathcal{F}_\sigma$ where $\sigma = \tau_F, \tau_G, \tau_H$. Do similarly for

$$\{\omega \colon t \mapsto \pi_t(\omega) \text{ has not entered } G \text{ before entering } F\}.$$

◊

2.28 Proposition: Let $(\mathcal{F}_t)$ be a filtration and $\tau, (\tau_n)$ be $(\mathcal{F}_t)$-stopping times. Then

a) $\tau \in m\mathcal{F}_\tau$

b) If $(\mathcal{F}_t)$ is right-continuous, then

$$\mathcal{F}_\tau = \{A \in \mathcal{F} \colon A \cap \{\tau < t\} \in \mathcal{F}_t \forall t\}.$$

Note that in the definition we needed to consider $\{\tau \leq t\}$.

c) If $\tau_1(\omega) \leq \tau_2(\omega)$ for all ω then this ordering also holds for the corresponding σ-algebras, i.e. $\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$.

d) We have the following approximation property: If $(\mathcal{F}_\tau)$ is right-continuous and $\tau_n \searrow \tau$ then

$$\mathcal{F}_\tau = \bigcap_{n \in \mathbb{N}} \mathcal{F}_{\tau_n}.$$

e) If (X_t) is $(\mathcal{F}_t)$-adapted, $(\mathcal{F}_t)$ right-continuous and $t \mapsto X_t(\omega)$ is right-continuous, then the map

$$\omega \mapsto X_{\tau(\omega)}(\omega) \mathbb{1}_{\{\tau_(\omega) < \infty\}},$$

i.e. the stopped process, is measurable with respect to $\mathcal{F}_\tau$.

◊

Proof. a),b),c) are exercises

d) By c) we have $\mathcal{F}_\tau \subseteq \cap_{n \in \mathbb{N}} \mathcal{F}_{\tau_n}$. Conversely, for $A \in \mathcal{F}_\tau$ we have

$$A \cap \{\tau < t\} = \bigsqcup_{n \in \mathbb{N}} A \cap \{\tau_n < t\} \in \mathcal{F}_t.$$

e) Assume first that $\tau(\Omega)$ is countable. Then for all $B \in \mathcal{B}(\mathbb{R}^d)$ we have

$$\{X_\tau \in B\} \cap \{\tau \leq t\} = \bigsqcup_{k \colon t_k \leq t} \{\tau = t_k, X_{t_k} \in B\} \in \mathcal{F}_t.$$

For general τ we approximate as in (2.26), i.e. $\tau_n(\omega) \searrow \tau(w)$ for all $\omega \in \Omega$. By the right-continuity of $t \mapsto X_t(\omega)$, we have

$$X_{\tau_n(\omega)}(\omega) \to X_{\tau(\omega)}(\omega)$$

as $n \to \infty$ for all $\omega \in \Omega$. So, for all $n \geq m$ it holds that

$$X_{\tau_n} \in m\mathcal{F}_{\tau_n} \subseteq m\mathcal{F}_{\tau_m}$$

hence $X_\tau \in m\mathcal{F}_{\tau_m}$ which implies $X_\tau \in m \cap_{m\geq 0} \mathcal{F}_{\tau_m}$ which by d) is equal to $\mathcal{F}_\tau$.

□

Next, we state and prove a very important theorem. We only formulate it for submartingales but by considering the negative process it holds with reversed inequalities for supermartingales and hence with equality for martingales.

2.29 Theorem (Doobs optional stopping theorem)**:** Let (X_t) be a $(\mathcal{F}_t)$-submartingale, σ, τ be $(\mathcal{F}_t)$-stopping times.

a) The process $(X_{\tau\wedge k})_{k\in\mathbb{N}}$ is an $(\mathcal{F}_k)_{k\in\mathbb{N}}$ and an $(\mathcal{F}_{\tau\wedge k})$-submartingale.

b) If there exists $T < \infty$ with $\sigma(\omega) \leq \tau(\omega) \leq T$ for all $\omega \in \Omega$ (i.e. the stopping times are bounded) then

$$\mathbb{E}\left(X_\tau \mid \mathcal{F}_\sigma\right)(\omega) \geq X_{\sigma(\omega)}(\omega)$$

holds $\mathbb{P}$-almost surely.

◊

Proof. We need to transfer known statements from discrete martingales and this remains as an exercise. □

2.30 Remark: The boundedness restriction on the stopping times seems to be very strong. But, e.g. , the condition $\mathbb{P}(\tau < \infty) = 1$ is not enough, as the following example shows. Let $(B_t)_{t\geq 0}$ be a one-dimensional Brownian motion and $\tau = \tau_{\{3\}}$. Then by (2.23) $\mathbb{P}(\tau_3 < \infty) = 1$ and $B_{\tau_3(\omega}(\omega) = 3$ for all $\omega \in \{\tau_3 < \infty\}$. So, $3 = \mathbb{E}\left(B_{\tau_3}\right) \neq \mathbb{E}\left(B_0\right) = 0$. Then, statement b) of Theorem 2.30 with $\tau = \tau_3$ and $\sigma \equiv 0$ is false. The statement of (2.30b) does hold under several weaker conditions, see e.g. Thm. A18 in Schilling-Partzsch or Thm 1.93 in [Ligger-Cts time Markov Processes]. We focus on a useful extension that requires $\mathbb{E}\left(X_t^2\right) < \infty$ for all t. ◊

2.31 Definition: Let (X_t) be an $\mathbb{R}$-valued martingale with continuous paths and $\mathbb{E}\left(X_t^2\right) < \infty$ for all t. A stochastic process $(A_t)_{t\geq 0}$ is called *quadratic variation process*(qvp) of (X_t) if

(i) $A_0 = 0$,

(ii) $t \mapsto A_t(\omega)$ is increasing almost surely,

(iii) $(X_t^2 - A_t)_{t \geq 0}$ is a martingale.

$\Diamond$

2.32 Remark: a) Proposition 2.12 a) shows that $A_t(\omega) := t$ defines a qvp for a one-dimensional Brownian motion.

b) For discrete martingales, Thm 4.49 of [PT] shows that $(A_n)_{n \in \mathbb{N}}$ is unique and is given by some nice formula which can be found there

c) For continuous martingales, there also exists a unique stochastic process fulfilling the preceding definition. It is given by

$$A_t = \lim_{|\pi| \to \infty} \sum_{t_i \in \pi} (X_{t_i} - X_{t_{i-1}})^2$$

where π denotes some partition. The limit is independent of the approximating sequence of partitions and exists in L^2. The proof is long and can be found in [Ligget, Thm 5.3].

$\Diamond$

2.33 Theorem: Let (X_t) be a continuous martingale, (A_t) its qvp and σ, τ be stopping times with $\sigma \leq \tau$. If $\mathbb{P}(\tau < \infty) = 1$ and $\mathbb{E}\left(A_\tau\right) < \infty$, then

a) $\mathbb{E}\left(X_\tau \mid \mathcal{F}_\sigma\right) = X_\sigma$ almost surely,

b) $\mathbb{E}\left(X_\tau^2 - A_\tau \mid \mathcal{F}_\sigma\right) = \mathbf{X}_\sigma^2 - A_\sigma$ almost surely.

In particular $\mathbb{E}\left(X_\tau\right) = \mathbb{E}\left(X_0\right)$ and $\mathbb{E}\left(X_\tau^2\right) = \mathbb{E}\left(X_0^2\right) - \mathbb{E}\left(A_\tau\right)$. $\Diamond$

Proof. Let $n \geq k$. Since $\mathcal{F}_{\tau \wedge n} \supseteq \mathcal{F}_{\tau \wedge k} \supseteq \mathcal{F}_{\sigma \wedge k}$ and since $\tau \wedge n \leq n < \infty$, for any martingale (M_t) we have by the tower property

$$\begin{aligned} \mathbb{E}\left(M_{\tau \wedge n} \mid \mathcal{F}_{\sigma \wedge k}\right) &= \mathbb{E}\left(\mathbb{E}\left(M_{\tau \wedge n} \mid \mathcal{F}_{\tau \wedge k}\right) \mid \mathcal{F}_{\sigma \wedge k}\right) \\ &= \mathbb{E}\left(M_{\tau \wedge k} \mid \mathcal{F}_{\sigma \wedge k}\right) = M_{\sigma \wedge k}, \end{aligned}$$

where we have used 2.30 a) and b) in the last two steps. Applying this for $\sigma = \tau, M_t = X_t^2 - A_t$ we obtain by the calculations above

$$\begin{aligned} \mathbb{E}\left(\left(X_{\tau \wedge n} - X_{\tau \wedge k}\right)^2\right) &= \mathbb{E}\left(X_{\tau \wedge n}^2 - X_{\tau \wedge k}^2\right) \\ &= \mathbb{E}\left(A_{\tau \wedge n} - A_{\tau \wedge k}\right) \end{aligned}$$

$$= \mathbb{E}\left((A_{\tau\wedge n} - A_{\tau\wedge k}\mathbb{1}_{\{\tau\geq k\}}\right) \overset{n,k\to\infty}{\longrightarrow} 0.$$

This shows that $(X_{\tau\wedge n})_{n\in\mathbb{N}}$ is a Cauchy sequence in L^2. By the continuity of X we have $X_{\tau\wedge n} \to \mathbf{X}_\tau$ almost surely and so $X_{\tau\wedge n} \to \mathbf{X}_\tau$ in L^2 and in particular

$$X^2_{\tau\wedge n} - A_{\tau\wedge n} \to X^2_\tau - A_\tau$$

in L^1 as $n \to \infty$. So, $M_{\tau\wedge n} \to M_tau$ in L^1 both holds for $M = X$ and $M = X^2 - A$. Thus, for all k and all $n \geq k$ we have using the tower property

$$\begin{aligned}\mathbb{E}\left(M_\tau \mid \mathcal{F}_{\sigma\wedge k}\right) &= \mathbb{E}\left(\mathbb{E}\left(M_\tau \mid \mathcal{F}_{\tau\wedge n}\right) \mid \mathcal{F}_{\sigma\wedge k}\right)\\ &= M_{\sigma\wedge k}.\end{aligned}$$

By Thm 4.68 of [PT] we have

$$\mathbb{E}\left(M_\tau \mid \mathcal{F}_{\sigma\wedge k}\right) \to \mathbb{E}\left(M_\tau \mid \mathcal{F}_\sigma\right)$$

almost surely and in L^1 as $k \to \infty$ and we already know that $M_{\sigma\wedge k} \to M_\sigma$ in L^1. This shows a),b). The final statement follow by taking $\sigma \equiv 0$ and taking expectation. □

2.34 Corollary (Wald's identities)**:** Let (B_t) be a one-dimensional Brownian Motion and τ a stopping time. if $\mathbb{E}(\tau) < \infty$, then

$$\mathbb{E}\left(B_\tau\right) = 0, \mathbb{E}\left(B^2_\tau\right) = \mathbb{E}\left(\tau\right)$$

◊

Proof. $B^2_t - t$ martingale, use (2.34) note $A_{\tau(\omega)} = \tau(\omega) \Rightarrow \mathbb{E}\left(A_\tau\right) = \mathbb{E}\left(\tau\right)$. □

We could have done the following earlier or later but at some point it has to be done and this might be a good time. We now show that all the staemtens about martingales so far apply to Brownian Motion$^{\mathrm{d}}$ with the right-continuous completion filtration $(\mathcal{F}_{t+})$. First of all, we need to show that is at least admissible for the Brownian Motion$^{\mathrm{d}}$.

2.35 Lemma: The filtration $(\mathcal{F}_{t+})$ is admissible for the Brownian Motion$^{\mathrm{d}}$ $\mathbf{B}$ in the sense of Definition 2.11. ◊

Proof. a) is clear. For b) let $0 \leq t \leq u$, $F \in \mathcal{F}_{t+}$ and $f \in \mathrm{C}_b(\mathbb{R}^d)$. Then we have $F \in \mathcal{F}^B_{t+\varepsilon}$ for all $\varepsilon > 0$. Thus

$$\mathbb{E}\left(\mathbb{1}_F f(B_{u+\varepsilon} - B_{t+\varepsilon})\right) = \mathbb{E}\left(\mathbb{1}_F\right)\mathbb{E}\left(f(B_{u+\varepsilon} - B_{t+\varepsilon})\right).$$

As $B_{u+\varepsilon} - B_{t+\varepsilon} \to B_u - B_t$ pointwise as $\varepsilon \to 0$, we may apply dominated convergence and obtain $B_u - B_t \amalg F$ which shows the claim. □

As a consequence we actually can equip a Brownian motion with the right-continuous completion of its filtration.

2.36 Proposition: Let $a < x < b$, $(\mathbf{X}_t)$ a continuous real-valued martingale with $\mathbb{P}(X_0 = x) = 1$ and set $\tau_{(-\infty,a]} \wedge \tau_{[b,\infty)} = \inf\{t \geq 0 \colon \mathbf{X}_t \notin (a,b)\}$. If $\mathbb{P}(\tau < \infty) = 1$, then $\mathbb{P}(\mathbf{X}_t = a) = \mathbb{P}(\tau_{(-\infty,a]} < \tau_{[b,\infty)} = \frac{b-x}{b-a}$ and $\mathbb{P}(\mathbf{X}_t = b) = \frac{x-a}{b-a}$. ◊

Proof. Let (A_t) be the qvp of $(\mathbf{X}_t)$, $M_t \coloneqq \mathbf{X}_t^2 - A_t$. Since $\tau \wedge n \leq n$, (2.29b) gives

$$\begin{aligned} \mathbb{E}\left(M_{\tau\wedge n} \mid \mathcal{F}_0\right) &\geq M_0, \\ \mathbb{E}\left(-M_{\tau\wedge n} \mid \mathcal{F}_0\right) &\geq -M_0, \end{aligned}$$

which implies

$$\begin{aligned} \mathbb{E}\left(M_{\tau\wedge n}\right) &= \mathbb{E}\left(\mathbb{E}\left(M_{\tau\wedge n} \mid \mathcal{F}_0\right)\right) \\ &= \mathbb{E}\left(M_0\right) = x^2. \end{aligned}$$

As $0 \leq A_{\tau\wedge 1} \leq A_{\tau\wedge 2} \leq \dots$ pointwise, we may apply monotone convergence and obtain

$$\begin{aligned} \mathbb{E}\left(A_\tau\right) &= \lim_{n\to\infty} \mathbb{E}\left(A_{\tau\wedge n}\right) = \lim_{n\to\infty} \mathbb{E}\left(X_{\tau\wedge n}^2\right) - x^2 \\ &\leq \max(a^2, b^2) - x^2. \end{aligned}$$

By (2.33) we have

$$\begin{aligned} x = \mathbb{E}\left(\mathbf{X}_0\right) = \mathbb{E}\left(\mathbf{X}_\tau\right) &= \mathbb{P}(\mathbf{X}_t au = a) \cdot a + \mathbb{P}(\mathbf{X}_\tau = b) \cdot b \\ &= \mathbb{P}(\mathbf{X}_\tau = a)(a - b) + b \end{aligned}$$

and we obtain both claims. □

Combining this with (2.14) for very interesting results about Brownian Motion. In order to state them we need the following definition.

2.37 Definition: Let $\mathbf{B}$ be a Brownian Motion$^{\mathrm{d}}$ and $x \in \mathbb{R}^d$. Then the SP $(\mathbf{B}_t + x)_{t\geq 0}$ is called *Brownian Motion^d started in x*, its probability measure (path measure) is denoted by $\mathbb{P}^x$. ◊

Now we have everything ready to state our next theorem.

2.38 Theorem: Let $x \in \mathbb{R}^d$, $\mathbf{B}$ a Brownian Motion$^\mathrm{d}$ started in x. Denote $B(0,r) = \{y \in \mathbb{R}^d \colon |y| < r\}$ and $\overline{B}(0,r)$ its closure. Then for $0 < r < |x| < R$ we have

$$\mathbb{P}^x(\tau_{\overline{B}(0,r)} < \tau_{B(0,R)}) = \begin{cases} \frac{R-|x|}{R-r} & d = 1, \\ \frac{\ln R - \ln|x|}{\ln R - \ln r} & d = 2, \\ \frac{R^{2-d}-|x|^{2-d}}{R^{2-d}-r^{2-d}} & d \geq 3. \end{cases}$$

$\Diamond$

Proof. The case $d = 1$ was shown in (2.36). For $d = 2$ consider the harmonic (check this!) function

$$f \colon \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}y \qquad \mapsto \ln(y).$$

Now let

$$\tilde{f}(y) = \begin{cases} f(y), & |y| > r, \\ \mathrm{C}_b^2, & |y| \leq r, \end{cases}$$

where we extend f by any C_b^2-function. Let $g(t,y) \coloneqq \tilde{f}(y+x)$ and $\tilde{\mathbf{B}}$ a Brownian Motion^2 with $B = \tilde{B} + x$. Then

$$M_t^g = g(t, \tilde{\mathbf{B}}_t) - g(0, \tilde{\mathbf{B}}_0) - \int_0^t (Lg)(s, \tilde{\mathbf{B}}_s)\ \mathrm{d}s$$

with $L = \partial_t + \frac{1}{2}\Delta$. Then $(Lg) = \frac{1}{2}\Delta\tilde{f}(\cdot + x)$ and

$$\begin{aligned} M_t &\coloneqq M_t^g + \tilde{f}(\mathbf{B}_0) \\ &= \tilde{f}(\mathbf{B}_t) - \frac{1}{2}\int_0^t (\Delta\tilde{f})(\mathbf{B}_s)\ \mathrm{d}s \end{aligned}$$

is a continuous real-valued martingale with $M_0 = \ln|x|$. On $\{t \leq \tau_{\overline{B}(0,r)}\}$ we have $|\mathbf{B}_s| > r$ for all $0 < s < t$. Hence, we have $\Delta\tilde{f}(B_s) = 0$. We find $\mathbb{P}(\tau_{\overline{B}(0,r)} < \tau_{B(0,R)^c} = \mathbb{P}(\tau^M_{(-\infty,\ln(r)]} < \tau_{[\ln(R),\infty)})$ which by 2.36 is equal to $\frac{\ln R - \ln|x|}{\ln R - \ln r}$. Note that by (2.23) we have $\tau_{B(0,R)^c} < \infty$ almost surely and thus

$$\mathbb{P}(\tau_{[\ln(R),\infty)} < \infty) = 1.$$

For $d = 3$ check that $f \colon \mathbb{R}^d \setminus \{0\} \to \mathbb{R}, y \mapsto |y|^{2-d}$ is harmonic and proceed as above. $\square$

2.39 Corollary: For a Brownian Motion$^{\mathrm{d}}$ starting in $x \neq 0$ we have

a) $\mathbb{P}^x(\tau_{\{0\}} < \infty) = \begin{cases} 1, & d = 1, \\ 0, & d = 2. \end{cases}$

b) For all $r \in (0, |x|$ we have
$$\mathbb{P}^x(\tau_{\overline{B}(0,r)} < \infty) = \begin{cases} 1, & d \leq 2, \\ \left(\frac{|x|}{2}\right)^{2-d}, & d \geq 3. \end{cases}$$

◊

Proof. We first prove b). We have
$$\mathbb{P}(\tau_{\overline{B}(0,r)} < \infty) = \lim_{R\to\infty} \mathbb{P}^x(\tau_{\overline{B}(0,r)} < \tau_{B(0,r)^c},$$
which by the last theorem shows the claim.

For a) we consider using continuity from above
$$\begin{aligned} \mathbb{P}^x(\tau_{\{0\}} < \tau_{B(0,r)^c} &= \lim_{r\searrow 0} \mathbb{P}^x(\tau_{\overline{B}(0,r)} < \tau_{B(0,R)^c}) \\ &= \begin{cases} 1 - \frac{|x|}{R}, & d = 1, \\ 0 & d \geq 2. \end{cases} \end{aligned}$$
As $\mathbb{P}^x(\tau_{\{0\}} < \infty) = \lim_{R\to\infty} \mathbb{P}^x(\tau_{\{0\}} \leq \tau_{B(0,R)^c})$ by continuity from below, we obtain the claim. □

Note that for $d = 1$ we already saw this in 2.22.

2.40 Remark: Corollary 2.39 means that

a) For $d \geq 2$ a Brownian Motion$^{\mathrm{d}}$ never finds predetermined points. This is not surprising as $\mathbb{P}(X = a = 0$ for $X \sim \mathcal{U}[0, 1]$ and $a \in [0, 1]$.

b) Starting far away, Brownian Motion$^{\mathrm{d}}$ always arbitrary small balls (this is surprising!) but Brownian Motion$^{\mathrm{d}}$ has a (good) chance to never find them for $d \geq 3$. Similar to memorylessness (2.3) the behaviour of $\mathbf{B}$ beyond x_2 should coincide in distribution with second Brownian Motion$^{\mathrm{d}}$ started in x_2, i.e. $\mathbb{P}(\mathbf{B}$ finds $\overline{B}(0, r)$ second time$) = 1$ for $d = 2$. Iterate procedure to find $\overline{B}(0, r)$ infinitely many times $\mathbb{P}^x$-almost surely, if $d = 2$. For $d \geq 3$, every time we go back to $\overline{B}(0, r)$ we pick up a factor $(\frac{R}{r}^{2-d} < 1$. Hence, we will find $\overline{B}(0, r)$ only finitely many times $\mathbb{P}^x$-almost surely in $d \geq 3$.

$\Diamond$

Main problem: restart time is random/stopping time. For this we need stronger property then memorylessness.