Brownian Motion

N.J. Nielsen

1 Notation

In these notes we shall in general use standard notation. For every $n \in \mathbb{N}$ $\mathcal{B}_n$ denotes the Borel algebra on $\mathbb{R}^n$ and if $(\Omega, \mathcal{F}, P)$ is a probability space, $X \colon \Omega \to \mathbb{R}^n$ a random variable, then we let $X(P)$ denote the distribution measure (the image measure) on $\mathbb{R}^n$ of X, e.g.

$$X(P)(A) = P(X^{-1}(A)) \quad \text{for all } A \in \mathcal{B}_n. \tag{1.1}$$

If $n \in \mathbb{N}$, we let $\langle \cdot, \cdot \rangle$ denote the canonical inner product on $\mathbb{R}^n$. Hence for all $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ og alle $y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ we have

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j y_j. \tag{1.2}$$

All vector spaces which occur in these notes are assumed to be real unless otherwise stated.

Finally we let $\boldsymbol{m}_n$ denote the Lebesgue measure on $\mathbb{R}^n$ and put $\boldsymbol{m} = \boldsymbol{m}_1$.

2 Characteristic functions and the normal distribution.

We start with the following definition:

Definition 2.1 *Let $n \in \mathbb{N}$ and let μ be a Borel probability measure on $\mathbb{R}^n$. The characteristic function $\varphi_\mu \colon \mathbb{R}^n \to \mathbb{C}$ of μ is defined by:*

$$\varphi_\mu(y) = \int_{\mathbb{R}^n} e^{i\langle y, x\rangle} d\mu(x) \quad \textit{for all} y \in \mathbb{R}^n. \tag{2.1}$$

The following example shows that there is a connection between the Fourier transform and characteristic functions.

Example 2.2 *Let* $h \in L_1(\mathbb{R})$, $h \geq 0$ *with* $\int_{-\infty}^{\infty} h(x)dx = 1$. *Put*

$$\mu(A) = \int_A h(x)dx \quad \text{for all } A \in \mathcal{B}. \tag{2.2}$$

Then $\varphi_\mu(y) = \sqrt{2\pi}\hat{h}(-y)$ *for all* $y \in \mathbb{R}$ *where* $\hat{h}$ *denotes the Fourier transformen.*

The integral transformation theorem gives namely that for all $y \in \mathbb{R}$ *we have:*

$$\varphi_\mu(y) = \int_{-\infty}^{\infty} e^{iyx} d\mu(x) = \int_{-\infty}^{\infty} e^{iyx} h(x)dx = \sqrt{2\pi}\hat{h}(-y). \tag{2.3}$$

We shall need the following theorem which we state without proof (a proof can e.g. be found in [L, side 199–201]).

Theorem 2.3 *The map* $\mu \to \varphi_\mu$ *is one to one.*

A classical theorem of Bochner gives together with Theorem 2.3 that the map $\mu \to \varphi_\mu$ gives a one-to-one correspondance between Borel probability measures on $\mathbb{R}^n$ and the continuous non-negative definite functions from $\mathbb{R}^n$ to $\mathbb{R}$, taking the value 1 at zero.

Definition 2.4 *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space and let* $X \colon \Omega \to \mathbb{R}^n$ *be a random variable. The characteristic function* φ_X *of* X *is defined as* $\varphi_X = \varphi_{X(P)}$. *This gives*

$$\varphi_X(y) = \int_{\mathbb{R}^n} e^{i\langle y,x\rangle} dX(P) = \int_\Omega e^{i\langle y,X\rangle} dP \tag{2.4}$$

for ally $\in \mathbb{R}^n$.

It follows immediately from Theorem 2.3 and Definition 2.4 that two n-dimensional random variables (not necessarily defined on the same probability space) has the same distribution if and only if their characteristic functions are identical.

Let us recall that a symmetric, real$n \times n$ matrix C is called positive definite if

$$\langle Cx, x\rangle > 0 \quad \text{for all } x \in \mathbb{R}^n \setminus \{0\}. \tag{2.5}$$

We shall now define and briefly describe the normal distributed random variables in terms of charateristic functions. Vi shall mostly use the 1-dimensional case but in proofs involving independence it is often necessary also to consider multi-dimensional normal distributions. We make the following definition.

Definition 2.5 *Let* C *be a symmetric, positive definite* $n\times n$ *matrix and let* $\xi = (\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{R}^n$.

An n-dimensional random variable X *is said to be normally distributed* $N(\xi, C)$ *if* X *has a density function* f *given by*

$$f(x) = \frac{1}{(2\pi)^{n/2}(\det C)^{\frac{1}{2}}} \exp(-\frac{1}{2}\langle C^{-1}(x-\xi), x-\xi\rangle) \tag{2.6}$$

for all $x \in \mathbb{R}^n$.

Example 2.6 *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space and let* $X \colon \Omega \to \mathbb{R}$ *be a random variable,* $\xi \in \mathbb{R}$ *og* $\sigma > 0$. X *is said to be normally distributed* $N(\xi, \sigma^2)$ *if* X *has a density function* f *given by*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{1}{2\sigma^2}(x-\xi)^2). \tag{2.7}$$

Using the next lemma it is relatively easy to see that $EX = \xi$ *og* $VX = \sigma^2$.

To justify Definition 2.5 we have to show that the function f given by (2.6) (or (2.7) in the 1-dimensional case) actually is a density function, e.g. that $f \geq 0$ og $\int_{\mathbb{R}^n} f \, d\boldsymbol{m}_n = 1$. For this we we need the following lemma which is probably well known from earlier courses in mathematics

Lemma 2.7 *We have the following formula:*

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx = \sqrt{2\pi}. \tag{2.8}$$

By substitution in the integral we get:

$$\int_{-\infty}^{\infty} \exp(-\frac{1}{2\sigma^2}(x-\xi)^2) dx = \sigma\sqrt{2\pi}. \tag{2.9}$$

This Lemma give immediately that the function f given by (2.7) is a density function . In order to prove that this is also the case in higher dimensions we need a bit of linear algebra.

Theorem 2.8 *The function* f *given by* (2.6) *is a density function.*

Proof: Let us for simplicity assume that $\xi = 0$. Since C and therefore also C^{-1} is symmetric, $\mathbb{R}^n$ has an orthonormal basis $(e_j)_{j=1}^n$ consisting of eigenvectors for C^{-1}. Let λ_j be the eigenvalue of C^{-1} corresponding to e_j. Since C^{-1} er positive definit, it follows that $\lambda_j > 0$ for all $1 \leq j \leq n$. For every $x \in \mathbb{R}^n$ we have

$$x = \sum_{j=1}^n \langle x, e_j \rangle e_j \tag{2.10}$$

and therefore

$$C^{-1}x = \sum_{j=1}^n \lambda_j \langle x, e_j \rangle e_j \tag{2.11}$$

and

$$\langle C^{-1}x, x \rangle = \sum_{j=1}^n \lambda_j \langle x, e_j \rangle^2. \tag{2.12}$$

Let $U\colon \mathbb{R}^n \to \mathbb{R}^n$ be defined by

$$Ux = (\langle x, e_j\rangle)_{j=1}^n \quad \text{for all } x \in \mathbb{R}^n. \tag{2.13}$$

U is an isometry since for all $x \in \mathbb{R}^n$ we have

$$\|Ux\|^2 = \sum_{j=1}^n \langle x, e_j\rangle^2 = \|\sum_{j=1}^n \langle x, e_j\rangle e_j\|^2 = \|x\|^2,$$

and since $(e_j)_{j=1}^n$ is a basis U is onto as well. We now get:

$$\begin{aligned}
\int_{\mathbb{R}^n} \exp(-\frac{1}{2}\langle C^{-1}x, x\rangle) d\boldsymbol{m}_n(x) &= \int_{\mathbb{R}^n} \exp(-\frac{1}{2}\sum_{j=1}^n \lambda_j \langle x, e_j\rangle^2) d\boldsymbol{m}_n(x) \qquad (2.14)\\
&= \int_{\mathbb{R}^n} \exp(-\frac{1}{2}\sum_{j=1}^n \lambda_j (Ux)_j^2) d\boldsymbol{m}_n(x)\\
&= \int_{\mathbb{R}^n} \exp(-\frac{1}{2}\sum_{j=1}^n \lambda_j x_j^2) d\boldsymbol{m}_n(x)\\
&= \prod_{j=1}^n \int_{-\infty}^{\infty} \exp(-\frac{1}{2}\lambda_j x_j^2) dx_j\\
&= (2\pi)^{n/2} \frac{1}{\sqrt{\prod_{j=1}^n \lambda_j}}\\
&= (2\pi)^{n/2} \frac{1}{\sqrt{\det C^{-1}}} = (2\pi)^{n/2}\sqrt{\det C},
\end{aligned}$$

where we in the third equality have used that the Lebesgue measure $\boldsymbol{m}_n$ is rotation invariant (or said in another way: we transform the integral by U which has a Jacobian with absolute value 1) and in the fifth equality have used Lemma 2.7. □

The proof of the next theorem is left to the reader:

Theorem 2.9 *Let $n \in \mathbb{N}$ and let ξ og $C = (c_{jk})$ be as in Definition 2.5. If $X = (X_j)_{j=1}^n$ is an n-dimensional random variable, normally distributed $N(\xi, C)$, then*

$$\begin{aligned}
EX &= \xi \qquad (2.15)\\
Cov(X_j, X_k) &= E(X_j - \xi_j)(X_k - \xi_k) = c_{jk} \qquad 1 \le j, k \le n. \qquad (2.16)
\end{aligned}$$

We now wish to compute the characteristic function of a normally distributed random variable. We need the following lemma:

Lemma 2.10 *For every $y \in \mathbb{R}$ we have*

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-iy)^2} dx = \sqrt{2\pi}. \tag{2.17}$$

Proof: The proof requires complex function theory and we shall only give a sketch.

For every $N \in \mathbb{N}$ we let R_N denote the rectangle in the complex plane determined by the pointsbestemt ved $-N, N, -N - iy$ og $N - iy$. Since the function $e^{-\frac{1}{2}z^2}$ is holomorphic in $\mathbb{C}$ we get that

$$\int_{\partial R_N} e^{-\frac{1}{2}z^2} dz = 0, \tag{2.18}$$

where we have integrated along the boundary boundary of the rectangle in the positive direction (anticlockwise). We now write the intgral as the sum of the four line integrals along the sides of the rectangle. It is readily checked that the integrals along the vertical sides go to 0 for $N \to \infty$. Together with (2.18) this gives

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-iy)^2} dx = \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx = \sqrt{2\pi}. \tag{2.19}$$

□

We start by finding the characteristic function for a one-dimensional normally distributed random variable.

Theorem 2.11 *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X \colon \Omega \to \mathbb{R}$ be a random variable. X is normally distributed $N(\xi, \sigma^2)$ if and only if*

$$\varphi_X(y) = e^{-\frac{1}{2}\sigma^2 y^2} \cdot e^{iy\xi} \quad \text{for all } y \in \mathbb{R}. \tag{2.20}$$

Proof: Let first X be normally distibuted $N(0,1)$. We then get that

$$\begin{aligned} \varphi_X(y) &= \int_{-\infty}^{\infty} e^{iyx} dX(P)(x) \qquad (2.21)\\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iyx} e^{-\frac{1}{2}x^2} dx \\ &= e^{-\frac{1}{2}y^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iyx} e^{-\frac{1}{2}x^2} \cdot e^{\frac{1}{2}y^2} dx \\ &= e^{-\frac{1}{2}y^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-iy)^2} dx = e^{-\frac{1}{2}y^2}. \end{aligned}$$

Assume next that X is normally distributed $N(\xi, \sigma^2)$ and put $Z = \frac{X-\xi}{\sigma}$. Since Z is normally distributed $N(0,1)$ and $X = \xi + \sigma Z$, we get from the above

$$\begin{aligned} \varphi_X(y) &= \int_{-\infty}^{\infty} e^{iyX(w)} dP(w) \qquad (2.22)\\ &= e^{i\xi y} \int_{-\infty}^{\infty} e^{iy\sigma Z} dP \\ &= e^{i\xi y} \varphi_Z(\sigma y) = e^{i\xi y} \cdot e^{-\frac{1}{2}y^2\sigma^2}. \end{aligned}$$

Assume now that φ_X er af formen (2.20) and let Y be normally distributed $N(\xi, \sigma^2)$. From the already proved it follows that $\varphi_Y = \varphi_X$ and hence we conclude from Theorem 2.3 that X has the same distribution as Y. □

We will now show the theorem analogous to Theorem 2.11 for multi-dimensional normal distributions.

Theorem 2.12 *Let $(\Omega, \mathcal{F}, P)$ be a probability space, $n \in \mathbb{N}$ and $X\colon \Omega \to \mathbb{R}^n$ a random variable. If $\xi \in \mathbb{R}^n$ and C is a symmetric, positive definite $n \times n$ matrix, then X is normally distributed $N(\xi, C)$ if and only if*

$$\varphi_X(y) = e^{i\langle \xi, y\rangle} \exp(-\frac{1}{2}\langle Cy, y\rangle) \quad \text{for all} y \in \mathbb{R}^n. \tag{2.23}$$

Proof: Let first X be normally distributed $N(0, C)$ and $(e_j)_{j=1}^n$, $(\lambda_j)_{j=1}^n$ and U be chosen as in Theorem 2.8. Using (2.12) we find that

$$\sqrt{\det C}(2\pi)^{n/2}\varphi_X(y) = \int_{\mathbb{R}^n} \exp(i\langle y, x\rangle) \exp(-\frac{1}{2}\langle C^{-1}x, x\rangle) d\boldsymbol{m}_n(x) \tag{2.24}$$

$$= \int_{\mathbb{R}^n} \exp(i\sum_{j=1}^n \langle x, e_j\rangle\langle y, e_j\rangle) \exp(-\frac{1}{2}\sum_{j=1}^n \lambda_j\langle x, e_j\rangle^2) d\boldsymbol{m}_n(x)$$

$$= \exp(-\frac{1}{2}\sum_{j=1}^n \frac{1}{\lambda_j}\langle y, e_j\rangle^2) \int_{\mathbb{R}^n} \prod_{j=1}^n \exp(-\frac{1}{2}(\sqrt{\lambda_j}\langle x, e_j\rangle - i\frac{1}{\sqrt{\lambda_j}}\langle y, e_j\rangle)^2) d\boldsymbol{m}_n(x)$$

$$= \exp(-\frac{1}{2}\langle Cy, y\rangle) \int_{\mathbb{R}^n} \prod_{j=1}^n \exp(-\frac{1}{2}(\sqrt{\lambda_j}(Ux)_j - i\frac{1}{\sqrt{\lambda_j}}\langle y, e_j\rangle)^2) d\boldsymbol{m}_n(x)$$

$$= \exp(-\frac{1}{2}\langle Cy, y\rangle) \int_{\mathbb{R}^n} \prod_{j=1}^n \exp(-\frac{1}{2}(\sqrt{\lambda_j}x_j - i\frac{1}{\sqrt{\lambda_j}}\langle y, e_j\rangle)^2) d\boldsymbol{m}_n(x)$$

$$= \exp(-\frac{1}{2}\langle Cy, y\rangle) \prod_{j=1}^n \int_{-\infty}^{\infty} \exp(-\frac{1}{2}(\sqrt{\lambda_j}x_j - i\frac{1}{\sqrt{\lambda_j}}\langle y, e_j\rangle)^2 dx_j$$

$$= \exp(-\frac{1}{2}\langle Cy, y\rangle)(2\pi)^{n/2}\sqrt{\det C}.$$

In the calculations above we have used Lemma 2.7 and Lemma 2.10.

If X is normally distributed $N(\xi, C)$, then $Z = X - \xi$ is normally distributed $N(0, C)$ so we find that

$$\varphi_X(y) = \int_\Omega e^{i\langle y, X\rangle} dP = \int_\Omega e^{i\langle y, \xi + Z\rangle} dP = e^{i\langle \xi, y\rangle}\varphi_Z(y) = e^{i\langle \xi, y\rangle} e^{-\frac{1}{2}\langle Cy, y\rangle} \tag{2.25}$$

$y \in \mathbb{R}^n$.

The rest of the proof of the theorem follows from Theorem 2.3. □

In the following we shal call an n-dimensional random variable X normally distributed if there exist a $\xi \in \mathbb{R}^n$ and a non-negative definite matrix C so that φ_X satisfies (2.23). Hence we do not any longer require that C is invertible. This generalization of normally distributed random variables is important when we consider linear combinations and limits of (usual)) normally distributed random variables. We can for example consider a constant $c \in \mathbb{R}$ as a normally distributed random variable variabel with mean value c og variance 0.

We shall now characterize n-dimensional normally distributed random variables in terms of their coordinate functions.

Theorem 2.13 *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X_j \colon \Omega \to \mathbb{R}$ be random variables. If $X = (X_1, X_2, \dots, X_n) \colon \Omega \to \mathbb{R}^n$, then the following two statements are equivalen:*

(i) X is normally distributed.

(ii) Every linear combination of the X_j's is normally distributed..

Proof: (i) $\implies$ (ii). Let X be normally distributed $N(\xi, C)$, let $t_1, t_2, \dots, t_n \in \mathbb{R}$ and put $Y = \sum_{j=1}^n t_j X_j$. We shall show that Y is normally distributed. If $t = (t_1, t_2, \dots, t_n)$, then clearly at $\langle t, X \rangle = Y$. With this observation we get for every $y \in \mathbb{R}$:

$$\begin{aligned} \varphi_Y(y) &= \int_{-\infty}^{\infty} \exp(iyY) dP = \int_{-\infty}^{\infty} \exp(i\langle yt, X \rangle) dP \\ &= \varphi_X(yt) = e^{iy\langle \xi, t \rangle} \exp(-\frac{1}{2} y^2 \langle Ct, t \rangle). \end{aligned} \tag{2.26}$$

Theorem 2.11 now gives that Y is normally distributed with mean value $\langle \xi, t \rangle$ og variance$\langle Ct, t \rangle$.

(ii) $\implies$ (i). For every $i \le j$, $k \le n$ we put

$$c_{jk} = E(X_j - EX_j)(X_k - EX_k) \tag{2.27}$$

and let $C = (c_{kj})$, $\xi = (EX_1, EX_2, \dots, EX_n)$. We shal show that X is normally distributed $N(\xi, C)$.

Let $y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ be arbitrary. Since $\langle y, X \rangle = \sum_{j=1}^n y_j X_j$,it follows from our assumptions that $\langle y, X \rangle$ is normally distributed with $E\langle y, X \rangle = \langle y, EX \rangle$ og variance

$$\begin{aligned} 0 \le V(\langle y, X \rangle) &= E(\langle y, X - EX \rangle)^2 \\ &= E(\sum_{j=1}^n y_j (X_j - EX_j))^2 \\ &= \sum_{j=k} y_j y_k c_{jk} = \langle Cy, y \rangle. \end{aligned} \tag{2.28}$$

(2.28) gives in particular that C er non-negative definite.

Hence using Theorem 2.11 we get

$$\varphi_X(y) = \int_\Omega \exp(i\langle y, X\rangle)dP = \varphi_{\langle y,X\rangle}(1) = \exp(i\langle y, \xi\rangle - \frac{1}{2}\langle Cy, y\rangle), \tag{2.29}$$

which according to Theorem 2.12 implies that X is normally distributed $N(\xi, C)$. $\square$

If $\mathcal{N}$ is a family of random variables on a probability space $(\Omega, \mathcal{F}, P)$, then the elements in $\mathcal{N}$ are called uncorelated if $\text{cov}(X, Y) = E(X - EX)(Y - EY) = 0$ for all $X, Y \in \mathcal{N}$. It is immediate that if $\mathcal{N}$ is independent, then the elements in $\mathcal{N}$ are uncorelated. We shall now see that the converse holds for certain sets of normally distributed random variables.

Theorem 2.14 *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X_j \colon \Omega \to \mathbb{R}$ be random variables so that $X = (X_1, X_2, \ldots, X_n) \colon \Omega \to \mathbb{R}^n$ is normally distributed. If X_j's are uncorelated, then they are independent..*

Proof: Lad $\xi \in \mathbb{R}^n$ and let C be the covariance matrix of X. From our assumptions it follows that it is a diagonal matrix with $\sigma_j^2 = (V(X_j)$ in the diagonal. If f denotes the density function of X and f_j denotes the density function of X_j, $1 \le j \le n$,then we have

$$f(x_1, x_2, \ldots, x_n) = (2\pi)^{-n/2} \prod_{j=1}^{n} (\frac{1}{\sigma_j}) \exp(-\frac{1}{2} \sum_{j=1}^{n} (\frac{x_j}{\sigma_j})^2) = \prod_{j=1} f_j(x_1, x_2, \ldots, x_n), \tag{2.30}$$

where we have used Theorem 2.13 to conclude that X_j is normally distributed for every $1 \le j \le n$.

(2.30) shows that $\{X_1, X_2, \ldots, X_n\}$ is an independent set. $\square$

If we combine Theorem 2.13 with the above, we get

Theorem 2.15 *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\mathcal{M} \subseteq L_2(P)$ be a subspace so that every $Y \in \mathcal{M}$ is normally distributed. If $\mathcal{N} \subseteq \mathcal{M}$ is a subset the elements of which are uncorelated, then $\mathcal{N}$ is independent.*

Proof: Let $n \in \mathbb{N}$ and $X_1, X_2, \ldots, X_n \in \mathcal{N}$ and put $X = (X_1, X_2, \ldots, X_n)$. Since every linear combination of the X_j's belongs to $\mathcal{M}$, X normally distributed according to Theorem 2.13. Since the X_j's are uncorelated, they are independent according to Theorem 2.14. This shows that $\mathcal{N}$ is an independent set.. $\square$

The above gives the following very useful corollary.

Corollary 2.16 *Let $\mathcal{M} \subseteq L_2(P)$ be as in Theorem 2.15 and let $\mathcal{N} \subseteq \mathcal{M}$ be a subset the elements of which all have mean value 0. Then $\mathcal{N}$ is independ if and only if it is an orthogonal set.*

We also want to mention the folowing well known theorem.

Theorem 2.17 *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $X_j\colon \Omega \to \mathbb{R}$ independent, normally distributed random variables. Then the following statements hold:*

(i) $X = (X_1, X_2, \dots, X_n)$ is normally distributed.

(ii) Every linear combination of the X_j's is normally distributed.

Proof: (i) follow by direct calculation, and (ii) follows from (i) by using Theorem 2.13. □

We conclude this section with the following theorem which will be very useful for us in the sequel:

Theorem 2.18 *Let $(X_k) \subseteq L_2(P)$ be a sequence of normally distributed random variables. If $X_k \to X$ i $L_2(P)$, then X is normally distributed.*

Proof: For every $k \in \mathbb{N}$ we put $\xi_k = EX_k$ og $\sigma_k^2 = V(X_k) = \|X_k\|_2^2 - (EX_k)^2$, $\xi = EX$ and $\sigma^2 = V(X)$. For evey $k \in \mathbb{N}$ we get

$$|\xi_k - \xi| = |EX_k - EX| \le E|X_k - X| \le \|X_k - X\|_2 \tag{2.31}$$

and

$$|\|X_k\|_2 - \|X\|_2| \le \|X_k - X_m\|_2. \tag{2.32}$$

Since $X_k \to X$ i $L_2(P)$, (2.30) and (2.32) show that $\xi_k \to \xi$ and $\sigma_k^2 \to \sigma^2$. We will show that X is normally distributed $N(\xi, \sigma^2)$. For every $y \in \mathbb{R}$ we find

$$\begin{aligned} |\varphi_X(y) - \varphi_{X_k}(y)| &\le \int_\Omega |e^{iyX} - e^{iyX_k}|dP \qquad (2.33) \\ &\le |y| \int_\Omega |X_k - X|dP \le |y| \|X_k - X\|_2 \to 0 \quad \text{for } k \to \infty \end{aligned}$$

so that $\varphi_{X_k}(y) \to \varphi_X(y)$ for all $y \in \mathbb{R}$.

Since X_k is normally distributed $N(\xi_k, \sigma_k^2)$, Theorem 2.11 gives that for every $y \in \mathbb{R}$ we have

$$\varphi_{X_k}(y) = \exp(iy\xi_k - \frac{1}{2}\sigma_k^2 y^2) \to \exp(iy\xi - \frac{1}{2}\sigma^2 y^2). \tag{2.34}$$

This shows that

$$\varphi_X(y) = \exp(iy\xi - \frac{1}{2}\sigma^2 y^2) \quad \text{for alle } y \in \mathbb{R}, \tag{2.35}$$

so that X is normally distributed $N(\xi, \sigma^2)$. □

Remark: In Theorem 2.18it can easily happen that $\sigma = 0$ so that X is a constant. By combining Theorem 2.18 with Theorem 2.13 a multi-dimensional version of Theorem 2.18is readily obtained . This version can however also easily be proved directly in a similar manner as above.

3 Brownian Motion

In this section we shall show the existence of the very important stochastic process knoas Brownian Motion and invistigate its basic properties.

The process is named after the Scottish botanist Robert Brown who in 1827 observed that when he suspended pollen grains in water, then they moved in an apparent random manner where they all the time changed directions. Besides discovering this phenomenon he has not contributed to the development of the mathematical theory for this process

The first attempt to give a mathematical definition of the process was done in 1900 by the French mathematician Louis Bachelier who was interested in modelling fluctuations in prices in financial markets and Albert Einstein who in 1905 gave a mathematical model of the phenomenon observed by Brown. The first mathematical rigorous treatment of this model was given by Norbert Wiener in 1923 and therefore the process is also often called the Wiener process.

Since then Brownian motion has played an important role in pure mathematics, applied mathematics and physics where in particular it is used in Einsteins Relativity Theory.

In mathematical financing Brownian motion has played a dominant role since it is the generating process of almost all models in this field. This is also the case for the Black-Scholes model, which as is well resulted in the Nobelprize in economy for its creators.

In the following we let $(\Omega, \mathcal{F}, P)$ be a fixed probability space.

Westart with the following definition:

Definition 3.1 *A stochastic process in continuous time is a family* $(X_t)_{t\geq 0}$ *of real random variables defined on a probability space* $(\Omega, \mathcal{F}, P)$.

Given a stochastic process $(X_t)_{t\geq 0}$ we often only consider X_t for t i et interval $[0, R]$.

We shall also need the following definitions:

Definition 3.2 *Let*$(\mathcal{F}_t)_{t\geq 0}$ *be a family of sub-*σ*-algebras of* $\mathcal{F}$ *so that* $\mathcal{F}_s \subseteq \mathcal{F}_t$ *for alls* $\leq t$. *A stochastic process* $(X_t)_{t\geq 0}$ *is called adapted if* X_t *er* $\mathcal{F}_t$*-measurable for every* $t \geq 0$.

Definition 3.3 *Let* $(\mathcal{F}_t)$ *be as in Definition 3.2 and let* $(X_t) \subseteq L_1(P)$ *be an* $(\mathcal{F}_t)$*-adapted process.* (X_t) *is called a submartingale if*

$$X_s \leq E(X_t|\mathcal{F}_s) \quad \text{for all } s < t. \tag{3.1}$$

If there for all $s < t$ *is equality in* (3.1), *then* (X_t) *is called a martingale.* (X_t) *is said to be a supermartingale if* $(-X_t)$ *is a submartingale.*

In this course we shall later discuss the theory of martingales in greater detail.

The next definition is important:

Definition 3.4 *A proceess (X_t) on $(\Omega, \mathcal{F}, P)$ is called continuous if the function $t \to X_t(\omega)$ is continuous for a.e. $t.\omega \in \Omega$.*

A process (Y_t) is said to have a continuous version if there exists a continuous process (X_t) so that $P(X_t = Y_t) = 1$ for all$t \geq 0$. If (X_t) is a process on $(\Omega, \mathcal{F}, P)$, then the functions $t \to X_t(\omega)$, $\omega \in \Omega$ are called the paths of the process.

Now it is the time to define the Brownian motion.

Definition 3.5 *A real stochastic process (B_t) is called a Brownian motion starting at 0 with mean value ξ and variance σ^2 if the following conditions are satisfied:*

(i) $P(B_0 = 0) = 1$

(ii) $B_t - B_s$ *is normally distributed* $N((t-s)\xi, (t-s)\sigma^2)$ *for all* $0 \leq s < t$.

(iii) $B_{t_1}, B_{t_2} - B_{t_1}, \dots B_{t_n} - B_{t_{n-1}}$ *are (stochastically) independent for alle* $0 \leq t_1 < t_2 < t_3 < \cdots t_n$. *$(B_t)$ is called a normalized Brownian motion if $\xi = 0$ and $\sigma^2 = 1$.*

(B_t) is called a normalized Brownian motion if $\xi = 0$ and $\sigma^2 = 1$.

The essential task of this section is of course to prove the existence of the Brownian motion, i.e. we have to show that there exists a probability space $(\Omega, \mathcal{F}, P)$ and a process (B_t) on that space so that the conditions in Definition 3.5 are satisfied. It is of course enough to show the existence of a normalized Brownian motion (B_t) for then $(\xi t + \sigma B_t)$ is a Brownian motion with mean value ξ and variance σ^2. We shall actually show a stronger result, namely that the Brownian motion has a continuous version. When we in the following talk about a Brownian motion we will always mean a normalized Brownian motion unless otherwise stated.

We will use Hilbert space theory for the construction so lets us recall some of its basic facts.

In the following $(\cdot, \cdot)$, respectively $\|\cdot\|$ will denote the inner product, respectively the norm in an arbitrary Hilbert space H. If we consider several different Hilbert spaces at the same time it is of course a slight misuse of notation to use the same symbols for the inner products and norms in these spaces but it is customary and eases the notation.

Lets us recall the polarization formula:

Lemma 3.6 *If H is a real Hilbert space, then:*

$$(x, y) = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) \quad \textit{for alle } x, y \in H. \tag{3.2}$$

If H is a complex Hilbert space, then:

$$(x, y) = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2). \tag{3.3}$$

Proof: Calculate the right hand sides! □

Definition 3.7 *Let H_1 and H_2 be Hilbert spaces. A linear map $T\colon H_1 \to H_2$ is called an isometry if $\|Tx\| = \|x\|$ for all $x \in H_1$.*

Using Lemma 3.6 we easily get:

Proposition 3.8 *Let H_1 and H_2 be Hilbert space $T\colon H_1 \to H_2$ a linear map. The following statements are equivalent:*

(i) T is an isometry.

(ii) $(Tx, Ty) = (x, y)$ for all $x, y \in H_1$.

Proof: (i) $\implies$ (ii). Use the polarization formula on (x, y) and (Tx, Ty).

(ii) $\implies$ (i). Put $y = x$. □

Let us note that it follows from Proposition 3.8 that if $T\colon H_1 \to H_2$ is an isometry and $x, y \in H_1$ are orthogonal,then also Tx og Ty are orthogonal.

The next theorem shows how to construct isometries between Hilbert spaces.

Theorem 3.9 *Let H_1 be a Hilbert space with an orthonormal basis (e_n) and let (f_n) be an orthonormal sequence in a Hilbert space H_2. THen the map $T\colon H_1 \to H_2$ defined by*

$$Tx = \sum_{n=1}^{\infty} (x, e_n) f_n \quad \textit{for all } x \in H_1 \tag{3.4}$$

is an isometry of H_1 into H_2.

Proof: We must first show that T is well-defined, that is we have to show that the series in (3.4) is covergent for all $x \in H_1$.

For this let $x \in H_1$ be arbitrary. since (e_n) is orthonormal, we have that vil $\sum_{n=1}^{\infty} |(x, e_n)|^2 = \|x\|^2 < \infty$ and hence since (f_n) is orthonormal we get that $\sum_{n=1}^{\infty}(x, e_n) f_n$ in H_2, so that T is well defined. T is clearly linear since the inner product is linear in the first variable. In addition we find:

$$\|Tx\|^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2 = \|x\|^2, \tag{3.5}$$

which shows that T is an isometry. □

The next definition is new:

Definition 3.10 *Let $(\Omega, \mathcal{F}, P)$ be a probability space. A closed subspace $\mathcal{H} \subseteq L_2(P)$ is called a Gaussian Hilbert space if every $f \in \mathcal{H}$ is normally distributed with mean value 0.*

Remark: In Definition 3.10 the zero function should be considewred as normally distributed with variance 0!

The next theorem, which is one of the main theorems shows that the existence of infinite dimensional Gaussian Hilbert spaces is equivalent to the existence of Brownian motions.

Theorem 3.11 *(i) Let $(\Omega, \mathcal{F}, P)$ be a probability space so that there exists an infinite dimensional Gaussian Hilbert space $\mathcal{H} \subseteq L_2(P)$. Then there exist isometries from $L_2(0,\infty)$ til $\mathcal{H}$ and if $T\colon L_2(0,\infty) \to \mathcal{H}$ is an arbitrary isometry and we put*

$$B_t = T(1_{[0,t]}) \quad \text{for all } t \in [0,\infty[, \tag{3.6}$$

then (B_t) is a Brownnian motion.

(ii) Let $(\Omega, \mathcal{F}, P)$ be a probability space on which there is a Brownian motion. Then there exist an infinite dimensional Gaussian Hilbert space $\mathcal{H} \subseteq L_2(P)$ and an isometry$T\colon L_2(0,\infty) \to \mathcal{H}$ so that (3.6) holds.

Proof:

(i) Since $L_2(0,\infty)$ is a separable Hilbert space, it has an orthonormal basis (f_n) and since $\mathcal{H}$ is infinite dimensional, we can find an orthonormal sequence $(g_n) \subseteq \mathcal{H}$ (Note, that all the g_n's are normally distributed $N(0,1)$ and are independent according to Corollary 2.16). From Theore 3.9 it follows that there exists an isometry S af $L_2(0,\infty)$ into $\mathcal{H}$ so that $Sf_n = g_n$ for all $n \in \mathbb{N}$.

Let now $T\colon L_2(0,\infty) \to \mathcal{H}$ be an arbitrary isometry and define (B_t) by (3.6). WE have to show that the conditions (i)–(iii) of Definition 3.5 are satisfied. Since $0 = T(0) = B_0$, it is clear that (i) holds. Next let $0 \le s < t$. Since $B_t - B_s \in \mathcal{H}$, it is normally distribute with mean value 0 and furthermore we have:

$$\int_\Omega (B_t - B_s)^2 dP = \|B_t - B_s\|_2^2 = \|T(1_{]s,t]})\|_2^2 = \|1_{]s,t]}\|_2^2 = (t-s), \tag{3.7}$$

which shows that $B_t - B_s$ has variance $(t-s)$.

Let now $0 \le t_1 < t_2 < t_3 < \cdots < t_n$. Since $\{1_{[0,t_1]}, 1_{]t_1,t_2]}, \ldots, 1_{]t_{n-1},t_n]}\}$ is an orthonormal set, it follows from Proposition 3.8 that also $\{T(1_{[0,t_1]}), T(1_{]t_1,t_2]}), \ldots, T(1_{]t_{n-1},t_n]})\} = \{B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}\}$ is an orthonormal set. Hence $B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}$ are independent by Corollary 2.16. This shows (i).

(ii). Assume that $(B_t)_{t\ge 0}$ is a Brownian motion on $(\Omega, \mathcal{F}, P)$ and put $\mathcal{H} = \overline{\text{span}}\{B_t \mid t \ge 0\} \subseteq L_2(P)$. We shall first show that $\mathcal{H}$ is a Gausian Hilbert space. Since all the B_t's have mean 0, the same is true for all elements in $\mathcal{H}$. Let first $g \in \text{span}\{B_t\}$.

We can then find $0 < t_1 < t_2 < \cdots < t_n$ and $(\alpha_j)_{j=1}^n \subseteq \mathbb{R}$ so that $g = \sum_{j=1}^n \alpha_j B_{t_j}$. If we put $\beta_j = \sum_{k=j}^n \alpha_k$ for all $1 \le j \le n$, then it is easily seen that

$$g = \sum_{j=1}^n \beta_j (B_{t_j} - B_{t_{j-1}}) \quad (B_{t_0} = 0). \tag{3.8}$$

Since $B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n - B_{t_{n-1}}}$ are independent it follows from Theorem 2.17 that g is normally distributed.

Let now $g \in \mathcal{H}$ be arbitrary. We can then find a sequence $g_n \subseteq \operatorname{span}\{B_t\}$ so that $g_n \to g$ i $L_2(P)$. The above together with Theorem 2.18 gives that g is normally distributed and hence we have shown that $\mathcal{H}$ is a Gaussian Hilbert space

We will now construct the desired isometry. Let $\mathcal{S} \subseteq L_2(0,\infty)$ be the the subspace of $L_2(0,\infty)$ consisting of the step functions. If $f \in \mathcal{S}$, we can find $0 = t_0 < t_2 < \cdots < t_n$ og $(\alpha_j)_{j=}^n \subseteq \mathbb{R}$ so that

$$f = \sum_{j=1}^n \alpha_j 1_{]t_{j-1},t_j]}, \tag{3.9}$$

and we put

$$Sf = \sum_{j=1}^n \alpha_j (B_{t_j} - B_{t_{j-1}}). \tag{3.10}$$

It is left to the reader to show that S is a well defined linear map from $\mathcal{S}$ til $\mathcal{H}$. If f satisfies (3.9), then we have

$$\|Sf\|_2^2 = \|\sum_{j=1}^n \alpha_j (B_{t_j} - B_{t_{j-1}})\|_2^2 = \sum_{j=1}^n \alpha_j^2 \|B_{t_j} - B_{t_{j-1}}\|_2^2 = \sum_{j=1}^n \alpha_j^2 (t_j - t_{j-1}) = \|f\|_2^2 \tag{3.11}$$

which shows that $S\colon \mathcal{S} \to \mathcal{H}$ is an isometry. We shall now extend S to an isometry $T\colon L_2(0,\infty) \to \mathcal{H}$.

Let $f \in L_2(0,\infty)$. Since $\mathcal{S}$ is dense in $L_2(0,\infty)$, we can find a sequence kan vi $(f_n) \subseteq \mathcal{S}$ so that $f_n \to f$ in $L_2(0,\infty)$. If $n, m \in \mathbb{N}$, then it follows from (3.11) that

$$\|Sf_n - Sf_m\|_2 = \|S(f_n - f_m)\|_2 = \|f_n - f_m\|_2 \tag{3.12}$$

which shows that (Sf_n) is a Cauchy sequence in $\mathcal{H}$, and hence there is an $F \in \mathcal{H}$ so that $Sf_n \to F$ i $L_2(P)$. In order to define $Tf = F$ we must show that F does not depend on the chosen sequence (f_n). For this let $(h_n) \subseteq \mathcal{S}$ so that $h_n \to f$ in $L_2(0,\infty)$. From the above it follows that $\lim_{n\to\infty} Sh_n$ exists in $\mathcal{H}$. We now define the sequence $(u_k) \subseteq \mathcal{S}$ by

$$u_{2n-1} = f_n, \quad u_{2n} = h_n \quad \text{for all } n = 1, 2, \dots. \tag{3.13}$$

It is clear that $u_k \to f$ i $L_2(0,\infty)$ and again the above shows that $\lim_{h\to\infty} Su_k$ exists in $\mathcal{H}$. Since both (Sf_n) og (Sh_n) are subsequences of (Su_k), we must have:

$$F = \lim_{n\to\infty} Sf_n = \lim_{k\to\infty} Su_k = \lim_{n\to\infty} Sh_n, \tag{3.14}$$

so that F does not depend on the chosen sequence (f_n). Hence we can put stte $Tf = F$. It is readily seen that T is a linear map from $L_2(0,\infty)$ til $\mathcal{H}$ so that $Tf = Sf$ for all $f \in \mathcal{S}$.

If $f \in L_2(0,\infty)$ and $(f_n) \subseteq \mathcal{S}$ again is chosen so that $f_n \to f$ i $L_2(0,\infty)$, then it follows from (3.11) that

$$\|Tf\|_2 = \lim_n \|Sf_n\|_2 = \lim_n \|f_n\|_2 = \|f\|_2. \tag{3.15}$$

Hence T is an isometry and we get directly from (3.10) that $T(1_{[0,t]}) = B_t$ for all $t \geq 0$. □

Remark: If (B_t) is a Brownian motion, T the associated isometry and $f \in L_2(0,\infty)$, then Tf is called the Ito integral of the deterministic function f with respect to B_t and is denoted $\int f dB_t$. The term deterministic expresses that f only depends on the time variable t. In a short while we shall extend the Ito integral to certain functions $f(t,\omega)$, $t \geq 0$ og $\omega \in \Omega$. We talk about an integral because of the way the isometry T is constructed in Theortem 3.11, (ii) which resembles the way one constructs integrals. The Ito integral plays an important role in mathematical financing.

The existence of infinite dimensional Gaussian Hilbert spaces follows from the next theorem, a part of which will be shown in the appendix.

Theorem 3.12 *There exists a probability space $(\Omega, \mathcal{F}, P)$ and a sequence $(g_n) \subseteq L_2(P)$ consisting of independent random variables which are all normally distributed $N(0,1)$. $\overline{\text{span}}(g_n)$ is a Gaussian Hilbert space.*

Proof: The existence of $(\Omega, \mathcal{F}, P)$ and (g_n) will be shown in the appendix. Actually we can put $(\Omega, \mathcal{F}, P) = ([0,1], \mathcal{B}, \boldsymbol{m})$. It follows directly from Theorem 2.17 and Theorem 2.18 that $\overline{\text{span}}(g_n)$ is a Gaussian Hilbert space. □

The next theorem put our results so far together but gives also new information.

Theorem 3.13 *Let $(\Omega, \mathcal{F}, P)$ be a probability space so that there exists a sequence $(g_n) \subseteq L_2(P)$ with the properties from Theorem 3.12. Then there is a Brownian motion on $(\Omega, \mathcal{F}, P)$. More specifically: If (f_n) is an arbitrary orthonormal basis for $L_2(0,\infty)$, then the series*

$$B_t = \sum_{n=1}^{\infty} \int_0^t f_n(s)ds\, g_n \quad t \geq 0 \tag{3.16}$$

converges in $L_2(P)$ and almost surely for all $t \geq 0$. (B_t) is a Brownian motion on $(\Omega, \mathcal{F}, P)$.

Proof: It follows directly from the Theorems 3.11 and 3.12 that there exists a Brownian motion on $(\Omega, \mathcal{F}, P)$. Since (f_n) is an orthonormal basis for $L_2(0,\infty)$, there is according to Theorem 3.9 an isometry $T\colon L_2(0,\infty) \to L_2(P)$ so that $Tf_n = g_n$. T is given by

$$Tf = \sum_{n=1}^{\infty} \int_0^{\infty} f(s) f_n(s) ds\, g_n, \tag{3.17}$$

where the series converges in $L_2(P)$. It follows from Theorem 3.11 that $B_t = T(1_{[0,t]})$ is a

Brownian motion and equation (3.17) gives that

$$B_t = T(1_{[0,t]}) = \sum_{n=1}^{\infty} \int_0^t f_n(s)ds\, g_n \quad \text{for alle } t \geq 0. \tag{3.18}$$

Since the terms in this sum are independent, have mean value 0 and

$$\sum_{n=1}^{\infty} E(\int_0^t f_n(s)ds\, g_n)^2 = \|B_t\|_2^2 = t < \infty, \tag{3.19}$$

it follows from classical results in probability theory that the series (3.18) converges almost surely for every $t \geq 0$. □

We shall now prove that there is a continuous version of the Brownian motion and then we do not as so far have a free choice of the orthonormal basis (f_n) for $L_2(0,\infty)$. We construct an orthonormal basis (f_n) with the property that there is an $A \in \mathcal{F}$ with $P(A) = 1$ so that if $\omega \in A$, then the series in (3.16) converges to $B_t(\omega)$ uniformly in t on every compact subinterval of $[0,\infty[$. Since every term of the series is continuous in t, this will give that $t \to B_t(\omega)$ is continuous for all $\omega \in A$. The construction of (f_n) is based on the Haar system (an orthonormal basis for $L_2(0,1)$ explained below) with the aid of the Borel-Cantelli lemma.

In the following we let $(\tilde{h}_m)$ denote the (non-normalized) be the Haar system, defined as follows (make a picture!!):

$$\tilde{h}_1(t) = 1 \quad \text{for alle } t \in [0,1]. \tag{3.20}$$

For all $k = 0,1,2,\dots$ og $\ell = 1,2,\dots,2^k$ we put

$$\tilde{h}_{2^k+\ell}(t) = \begin{cases} 1 & \text{if} \quad t \in [(2\ell-2)2^{-k-1}, (2\ell-1)2^{-k-1}[\\ -1 & \text{if} \quad t \in [(2\ell-1)2^{-k-1}, 2\ell \cdot 2^{-k-1}[\\ 0 & \text{else.} \end{cases}$$

We norm this system in $L_2(0,1)$ and define

$$h_1 = \tilde{h}_1 \qquad h_{2^k+\ell} = 2^{k/2}\tilde{h}_{2^k+\ell} \quad \text{for alle } k = 0,1,2,\dots \text{ og } \ell = 1,2,3,\dots,2^k. \tag{3.21}$$

By direct computation we check that it is an orthonormal system and since it is easy to see that every indicator function of a dyadic interval belongs to $span(h_m)$, it follows that $span(h_m)$ is dense in $L_2(0,1)$. Therefore (h_m) is an orthonormal basis for $L_2(0,1)$. It follows from Theorem 3.13 that if $(\Omega, \mathcal{F}, P)$ is a probability space so that there exists a sequence $(g_n) \subseteq L_2(P)$ as in Theorem 3.12, then

$$B_t = \sum_{m=1}^{\infty} \int_0^t h_m(s)ds\, g_m \qquad 0 \leq t \leq 1 \tag{3.22}$$

is a Brownian motion for $t \in [0,1]$. The series converges in $L_2(P)$ and almost surely and the same is the case if we permute the terms. We should however note that the set with measure 1 on which the series converges pointwise depends on the permutation. In order not to get into difficulties with zero sets we shall fix the order of the terms in the sum. We define

$$B_t = \int_0^t h_1(s)ds\, g_1 + \sum_{k=0}^{\infty} \sum_{m=2^k+1}^{2^{k+1}} \int_0^t h_m(s)ds\, g_m \overset{def}{=} \sum_m * \int_0^t h_m(s)ds\, g_m \quad \text{for alle } 0 \le t \le 1. \tag{3.23}$$

and can now show:

Theorem 3.14 *$(B_t)_{0\le t\le 1}$ given by (3.23) is a continuous Brownian motion (on $[0,1]$).*

In the proof of the theorem we need the following lemmas:

Lemma 3.15 *For all $k \ge 0$ we have $0 \le \sum_{m=2^k+1}^{2^{k+1}} \int_0^t h_m(s)ds \le 2^{-k/2-1}$.*

Proof: For every $2^k < m \le 2^{k+1}$ we put $S_m(t) = \int_0^t h_m(s)ds$ for all $0 \le t \le 1$. If $m = 2^k + \ell$, $1 \le \ell \le 2^k$, then it follows directly from the definition of h_m, that the graph of S_m is an triangle centered in $(2\ell-1)2^{-k-1}$ and with highth $2^{-k/2-1}$. For different ℓ's these triangles do not overlap. This shows the statement. □

Lemma 3.16 *For all $k \ge 0$ we put*

$$G_k(\omega) = \max\{|g_m(\omega)| \mid 2^k < m \le 2^{k+1}\} \quad \textit{for all } \omega \in \Omega. \tag{3.24}$$

There is a subset $\tilde{\Omega} \subseteq \Omega$ with $P(\tilde{\Omega}) = 1$ so that there to every $\omega \in \tilde{\Omega}$ exists a $k(\omega)$ with the property that $G_k(\omega) \le k$ for all $k \ge k(\omega)$.

Proof: For every $x > 0$ we find

$$P(|g_m| > x\} = \sqrt{\frac{2}{\pi}} \int_x^\infty e^{-u^2/2}du \le \sqrt{\frac{2}{\pi}} \int_x^\infty \frac{u}{x} e^{-u^2/2}du = \sqrt{\frac{2}{\pi}} x^{-1} e^{-x^2/2}, \tag{3.25}$$

which gives:

$$P(G_k > k) = P(\bigcup_{m=2^k+1}^{2^k+1} (|g_m| > k) \le 2^k P(|g_1| > k) \le \sqrt{\frac{2}{\pi}} \frac{1}{k} \cdot 2^k e^{-k^2/2}. \tag{3.26}$$

Since

$$\sum_{k=1}^\infty P(G_k > k) \le \sqrt{\frac{2}{\pi}} \sum_{k=1}^\infty k^{-1} 2^k e^{-k^2/2} < \infty,$$

it follows from the Borel-Cantelli lemma that $P(G_k \le k \text{ from a certain step}) = 1$. Choosing $\tilde{\Omega}$ as this set the statement follows. □

Proof of Theorem 3.14: Let $\tilde{\Omega}$ be as in Lemma 3.16 and let $\omega \in \tilde{\Omega}$. Then there exists a $k(\omega) \geq 1$ so that $G_k(\omega) \leq k$ for alle $k \geq k(\omega)$. If $k \geq k(\omega)$ is now fixed, we find

$$\sum_{m=2^k+1}^{2^{k+1}} |\int_0^t h_m(s)ds \cdot g_m(\omega)| \leq \sum_{m=2^k+1}^{2^{k+1}} \int_0^t h_m(s)ds \cdot G_k(\omega) \leq k\, 2^{-k/2-1}. \tag{3.27}$$

for all $0 \leq t \leq 1$. Since $\sum_{k=1}^{\infty} k\, 2^{-k/2-1} < \infty$, it follows from Weierstrass' M-test that the series $\sum_{k=k(\omega)}^{\infty} \sum_{m=2^k+1}^{2^{k+1}} \int_0^t h_m(s)dsg_m(\omega)$ converges uniformly for $t \in [0,1]$. This gives that the series

$$B_t(\omega) = \sum_m * \int_0^t h_m(s)dsg_m(\omega) \tag{3.28}$$

also converges uniformly for $t \in [0,1]$ and hence $t \to B_t(\omega)$ is continuous. □

In order to find a continuous Brownian motion on $[0,\infty[$ we define the functions $h_{nm} \in L_2(0+,\infty)$ by

$$h_{nm}(t) = \begin{cases} h_m(t-n) & \text{for } t \in [n-1,n] \\ 0 & \text{else} \end{cases} \quad n \in \mathbb{N}, m \in \mathbb{N} \tag{3.29}$$

and note that $(h_{nm})_{m=1}^{\infty}$ is an orthonormal basis for $L_2(n-1,n)$ for all $n \in \mathbb{N}$ which implies that (h_{nm}) is an orthonormal basis for $L_2(0,\infty)$. We have the following theorem:

Theorem 3.17 *Let $(\Omega, \mathcal{F}, P)$ be a probability space on which there exists a sequence of $N(0,1)$-distributed random variables and let (g_{nm}) be such a sequence. Define:*

$$B_t = \sum_{n=1}^{\infty} \sum_m * \int_0^t h_{nm}(s)ds\, g_{nm} \quad \textit{for all} t \geq 0. \tag{3.30}$$

Then $(B_t)_{t\geq 0}$ is a continuous Brownian motion.

Proof: Let $n_0 \in \mathbb{N}$. For every $t \in [n_0, n_0+1]$ we find

$$B_t(\omega) - B_{n_0}(\omega) = \sum_m * \int_{n_0}^t h_{n_0m}(s)ds\, g_{n_0m}(\omega) \quad \text{for all } \omega \in \Omega. \tag{3.31}$$

By Theorem 3.14 There exists a measurable set $\tilde{\Omega}_{n_0} \subseteq \Omega$ with $P(\tilde{\Omega}_{n_0}) = 1$ so that the series (3.31) converges uniformly on $[n_0, n_0+1]$ for every $\omega \in \tilde{\Omega}_{n_0}$. This gives that $t \to B_t(\omega)$ is continuous on $[n_0, n_0+1]$ for all $\omega \in \tilde{\Omega}_{n_0}$. If we put $\tilde{\Omega} = \bigcap_{n=1}^{\infty} \tilde{\Omega}_n$, then $P(\tilde{\Omega}) = 1$ and if $\omega \in \tilde{\Omega}$, then $t \to B_t(\omega)$ will be continuous on $[0,\infty[$. □

Let now (B_t) be a Brownian motion and let for every $t \geq 0$ $\mathcal{F}_t$ denote the σ-algebra generated by $\{B_s \mid 0 \leq s \leq t\}$. We end this section with the following theorem:

Theorem 3.18 *$(B_t, \mathcal{F}_t)$ is a martingale.*

Proof: Let $0 \le s < t$. It follows directly from the definition that $B_t - B_s$ is independent of $\{B_u \mid u \le s\}$ and therefore also independent of $\mathcal{F}_s$. Hence we find

$$E(B_t|\mathcal{F}_s) = E(B_s|\mathcal{F}_s) + E(B_t - B_s|\mathcal{F}_s) = B_s + E(B_t - B_s) = B_s \tag{3.32}$$

□

4 Appendix

In this section we shall prove that there exists a sequence of independent $N(0,1)$-distributed random variables on $([0,1], \mathcal{B}, \boldsymbol{m})$. We first define the following sequence of functions (random variables) on $[0,1]$:

For all $n = 1, 2, \dots$ og $k = 1, 2, \dots, 2^{n-1}$ we put

$$u_n(t) = \begin{cases} 0 & \text{if} \quad t \in [(2k-2)2^{-n}, (2k-1)2^{-n}[\\ 1 & \text{if} \quad t \in [(2k-1)2^{-n}, 2k \cdot 2^{-n}[\end{cases} \tag{4.1}$$

and let us put $u_n(1) = 1$ for all $n \in \mathbb{N}$.

It is readily verified by induction that (u_n) is an independent sequence of random variables on $[0,1]$, each taking the values 0 and 1 with probabilities 1/2. We need the following lemma:

Lemma 4.1 *For every $t \in [0,1]$ $t = \sum_{n=1}^{\infty} u_n(t)2^{-n}$. In other words this sum is exactly the binary expansion of t.*

Proof: For $t = 1$ the lemma is obvious and for $0 \le t < 1$ it will follow from the inequality

$$0 \le t - \sum_{k=1}^{n} u_k(t)2^{-k} < 2^{-n} \quad \text{for all } t \in [0,1[\text{ and all } n \in \mathbb{N} \tag{4.2}$$

which we shall prove by induction. For $n = 1$ (4.2) is clear so let us assume that it holds for n and let $k \in \mathbb{N}$ be defined so that $k2^{-n} = \sum_{m=1}^{n} u_m(t)2^{-m}$. By the induction hypothesis $k2^{-n} \le t < (k+1)2^{-n}$. If $k2^{-n} \le t < (2k+1)2^{-n-1}$, then $u_{n+1}(t) = 0$ and the result follows. If $(2k+1)2^{-n-1} \le t < (k+1)2^{-n}$, then $u_{n+1}(t) = 1$ and hence $(2k+1)2^{-n-1} = \sum_{k=1}^{n+1} u_k(t)2^{-k}$ and again the result follows. □

We are now able to prove:

Theorem 4.2 *There exists an independent sequence (U_n) of random variables on $([0,1], \mathcal{B}, \boldsymbol{m}_1)$ so that each U_n is uniformly distributed on $[0,1]$.*

Proof:

Put $U(t) = t$ for all $t \in [0,1]$, let (A_n) be a sequence of mutually disjoint infinite subsets of $\mathbb{N}$, say $A_n = \{m_{kn} \mid k \in \mathbb{N}\}$ and put $U_n(t) = \sum_{k=1}^{\infty} u_{m_{kn}}(t) 2^{-k}$ for all $t \in [0,1]$ and all $n \in \mathbb{N}$. Since (u_n) is an independent sequence and the A_n's are mutually disjoint, (U_n) is independent as well. Since in addition the u_n's are identically distributed, it follows from Lemma 4.1 that each U_n has the same distribution as U, i.e. is uniformly distributed on $[0,1]$. □

We are now able to prove

Theorem 4.3 *Let F be a strictly increasing and continuous distribution function. On $([0,1], \mathcal{B}, \boldsymbol{m})$ there exists a sequence (X_n) of independent, F-distributed random variables.*

Proof: Let (U_n) be defined as in Theorem 4.3 and define

$$X_n(t) = F^{-1}(U_n(t)) \quad \text{for all } t \in [0,1] \text{ and all } n \in \mathbb{N} \tag{4.3}$$

Since F^{-1} is continuous, X_n is measurable for every $n \in \mathbb{N}$ and since the sequence (U_n) is independent, (X_n) is independent as well. For every $x \in \mathbb{R}$ we have:

$$\boldsymbol{m}(X_n \le x) = \boldsymbol{m}(U_n \le F(x)) = F(x) \tag{4.4}$$

□

As a corollary we obtain:

Corollary 4.4 *On $([0,1], \mathcal{B}, \boldsymbol{m})$ there exists an independent sequence (X_n), consisting of $N(0,1)$-distributed random variables.*

Proof: Use Theorem 4.3 with $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}}\, dt$ □

It is easily seen that the conditions on the distribution function F in Theorem 4.3 can be omitted. Indeed, if F is an arbitrary distribution function and we define

$$X_n(t) = \sup\{s \in \mathbb{R} \mid F(s) \le U_n(t)\} \quad \text{for all } n \in \mathbb{N} \text{ and all } t \in [0,1], \tag{4.5}$$

then it is easy to see that (X_n) is a sequence of independent F-distributed random variables.

References

[1] M. Love, *Probability Theory I*, 4th edition, springer Verlag 1977.