Advanced stochastic processes: Part II

Jan A. Van Casteren

JAN A. VAN CASTEREN

ADVANCED STOCHASTIC PROCESSES: PART II

Advanced stochastic processes: Part II
4th edition

ISBN 978-87-403-3071-7

CONTENTS

CHAPTER 4

Stochastic differential equations

Some pertinent topics in the present chapter consist of a discussion on martingale theory, and a few relevant results on stochastic differential equations in spaces of finite dimension. In particular unique weak solutions to stochastic differential equations give rise to strong Markov processes whose one-dimensional distributions are governed by the corresponding second order parabolic type differential equation. Essentially speaking this chapter is part of Chapter 1 in [**184**]. (The author is thankful to WSPC for the permission to include this text also in the present book.) In this chapter we discuss weak and strong solutions to stochastic differential equations. We also discuss a version of the Girsanov transformation.

1. Solutions to stochastic differential equations

Basically, the material in this section is taken from Ikeda and Watanabe [**81**]. In Subsection 1.1 we begin with a discussion on strong solutions to stochastic differential equations, after that, in Subsection 1.2 we present a martingale characterization of Brownian motion. We also pay some attention to (local) exponential martingales: see Subsection 1.3. In Subsection 1.4 the notion of weak solutions is explained. However, first we give a definition of Brownian motion which starts at a random position.

4.1. Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with filtration $(\mathcal{F}_t)_{t\geqslant 0}$. A d-dimensional Brownian motion is a almost everywhere continuous adapted process $\{B(t) = (B_1(t), \dots, B_d(t)) : t \geqslant 0\}$ such that for $0 < t_1 < t_2 < \cdots < t_n < \infty$ and for C any Borel subset of $\left(\mathbb{R}^d\right)^n$ the following equality holds:

$$\begin{aligned}&\mathbb{P}\left[\left(B\left(t_1\right) - B(0), \dots, B\left(t_n\right) - B(0)\right) \in C\right]\\ &= \int \cdots \int_C p_{0,d}\left(t_n - t_{n-1}, x_{n-1}, x_n\right) \cdots p_{0,d}\left(t_2 - t_1, x_1, x_2\right) p_{0,d}\left(t_1, 0, x_1\right)\\ &\qquad dx_1 \dots dx_n. \end{aligned} \tag{4.1}$$

This process is called a d-dimensional Brownian motion with initial distribution μ if for $0 < t_1 < t_2 < \cdots < t_n < \infty$ and every Borel subset of $\left(\mathbb{R}^d\right)^{n+1}$ the following equality holds:

$$\begin{aligned}&\mathbb{P}\left[\left(B(0), B\left(t_1\right), \dots, B\left(t_n\right)\right) \in C\right]\\ &= \int \cdots \int_C p_{0,d}\left(t_n - t_{n-1}, x_{n-1}, x_n\right) \cdots p_{0,d}\left(t_2 - t_1, x_1, x_2\right) p_{0,d}\left(t_1, x_0, x_1\right)\end{aligned}$$

$$d\mu\left(x_0\right)\, dx_1 \dots dx_n. \tag{4.2}$$

For the definition of $p_{0,d}\left(t,x,y\right)$ see formula (4.26). By definition a filtration $\left(\mathcal{F}_t\right)_{t\geqslant 0}$ is an increasing family of σ-fields, i.e. $0 \leqslant t_1 \leqslant t_2 < \infty$ implies $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$. The process of Brownian motion $\{B(t) : t \geqslant 0\}$ is said to be adapted to the filtration $\left(\mathcal{F}_t\right)_{t\geqslant 0}$ if for every $t \geqslant 0$ the variable $B(t)$ is $\mathcal{F}_t$-measurable. It is assumed that the $\mathbb{P}$-negligible sets belong to $\mathcal{F}_0$.

1.1. Strong solutions to stochastic differential equations. In this section we discuss strong or pathwise solutions to stochastic differential equations. We also show that if the stochastic differential equation in (4.110) possesses unique pathwise solutions, then it has unique weak solutions. We begin with a formal definition.

4.2. Definition. The equation in (4.110) is said to have unique pathwise solutions, if for any Brownian motion $\{(B(t) : t \geqslant 0), (\Omega, \mathcal{F}, \mathbb{P})\}$ and any pair of $\mathbb{R}^d$-valued adapted processes $\{X(t) : t \geqslant 0\}$ and $\{X'(t) : t \geqslant 0\}$ for which

$$X(t) = x + \int_0^t \sigma\left(s, X(s)\right) dB(s) + \int_0^t b\left(s, X(s)\right)\, ds \quad \text{and} \tag{4.3}$$

$$X'(t) = x + \int_0^t \sigma\left(s, X'(s)\right) dB(s) + \int_0^t b\left(s, X'(s)\right)\, ds \tag{4.4}$$

it follows that $X(t) = X'(t)$ $\mathbb{P}$-almost surely for all $t \geqslant 0$. If for any given Brownian motion $(B(t))_{t\geqslant 0}$ the process $(X(t))_{t\geqslant 0}$ is such that for $\mathbb{P}$-almost all $\omega \in \Omega$ the equality

$$X(t,\omega) = x + \int_0^t \sigma\left(s, X(s,\omega)\right) dB(s,\omega) + \int_0^t b\left(s, X(s,\omega)\right)\, ds$$

is true, then $t \mapsto X(t)$ is called a *strong solution*.

Strong solutions are also called *pathwise solutions*. In order to facilitate the proof of Theorem 4.4 we insert the following lemma.

4.3. Lemma. *Let γ be a positive real number. Then the following inequality holds:*

$$\sum_{n=0}^{\infty} \frac{\gamma^{n/2}}{\sqrt{n!}} \leqslant \frac{1}{2}\left(\sqrt{\gamma} + \sqrt{\gamma+4}\right) \exp\left(\frac{1}{8}\left(\sqrt{\gamma} + \sqrt{\gamma+4}\right)^2 - \frac{1}{2}\right). \tag{4.5}$$

Since $\sqrt{\gamma} + \sqrt{\gamma+4} \leqslant 2\sqrt{\gamma+2}$, the inequality in (4.5) implies:

$$\sum_{n=0}^{\infty} \frac{\gamma^{n/2}}{\sqrt{n!}} \leqslant \sqrt{\gamma+2}\exp\left(\frac{1}{2}\left(\gamma+1\right)\right) < \infty. \tag{4.6}$$

We will use the finiteness of the sum rather than the precise estimate.

Proof of Lemma 4.3. Let $\delta > 0$ be a positive number. Then we have by the Cauchy-Schwarz inequality

$$\begin{aligned}\left(\sum_{n=0}^{\infty}\frac{\gamma^{n/2}}{\sqrt{n!}}\right)^2 &= \left(\sum_{n=0}^{\infty}\frac{\gamma^{n/2}}{(\delta+\gamma)^{n/2}}\frac{(\delta+\gamma)^{n/2}}{\sqrt{n!}}\right)^2 \\ &\leqslant \sum_{n=0}^{\infty}\frac{\gamma^{n}}{(\delta+\gamma)^{n}}\sum_{n=0}^{\infty}\frac{(\delta+\gamma)^{n}}{n!} = \frac{\delta+\gamma}{\delta}e^{\delta+\gamma}. \qquad (4.7)\end{aligned}$$

The choice $\delta = \frac{1}{2}\left(-\gamma + \sqrt{\gamma(\gamma+4)}\right)$ yields the equalities

$$\delta+\gamma = \frac{1}{4}\left(\sqrt{\gamma}+\sqrt{\gamma+4}\right)^2 - 1, \quad \text{and} \quad \frac{\delta+\gamma}{\delta} = \frac{1}{4}\left(\sqrt{\gamma}+\sqrt{\gamma+4}\right)^2,$$

and so the result in (4.5) follows and completes the proof of Lemma 4.3. ☐

A version of the following result can be found in many books on stochastic differential equations: see e.g. [**81, 139, 148**].

4.4. THEOREM. *Let $\sigma_{j,k}(s,x)$ and $b_j(s,x)$, $1 \leqslant j, k \leqslant d$ be continuous functions defined on $[0,\infty) \times \mathbb{R}^d$ such that for all $t > 0$ there exists a constant $K(t)$ with the property that*

$$\sum_{j,k=1}^{d} \left|\sigma_{j,k}(s,x) - \sigma_{j,k}(s,y)\right|^2 + \sum_{j=1}^{d} \left|b_j(s,x) - b_j(s,y)\right|^2 \leqslant K(t)^2 \left|x-y\right|^2 \quad (4.8)$$

for all $0 \leqslant s \leqslant t$, and all x, $y \in \mathbb{R}^d$. Fix $x \in \mathbb{R}^d$, and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $(\mathcal{F}_t)_{t \geqslant 0}$. Moreover, let $\{B(t) : t \geqslant 0\}$ be a Brownian motion on the filtered probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$. Then there exists an $\mathbb{R}^d$-valued process $\{X(t) : t \geqslant 0\}$ such that, for all $0 < T < \infty$, $\sup_{0<t\leqslant T} \mathbb{E}\left[|X(t)|^2\right] < \infty$, and such that

$$X(t) = x + \int_0^t \sigma(s, X(s))\, dB(s) + \int_0^t b(s, X(s))\, ds, \quad t \geqslant 0. \quad (4.9)$$

This process is pathwise unique in the sense of Definition 4.2.

The techniques in the proof below are very similar to a method to prove the following version of Gronwall's inequality: see e.g. [**68**]. Let f, g, $h : [0,T] \to \mathbb{R}$ be continuous functions such that $f(t) \leqslant g(t) + \int_0^t h(s) f(s)\, ds$, $0 \leqslant t \leqslant T$. If $h \geqslant 0$, then by induction with respect to k it follows that

$$f(t) \leqslant g(t) + \sum_{j=1}^{k} \int_0^t \frac{\left(\int_s^t h(\rho)\, d\rho\right)^{j-1}}{(j-1)!} g(s)\, ds + \int_0^t \frac{\left(\int_s^t h(\rho)\, d\rho\right)^{k}}{k!} h(s) f(s)\, ds,$$

and hence

$$f(t) \leqslant g(t) + \int_0^t g(s) \exp\left(\int_s^t h(\rho)\, d\rho\right) ds.$$

Let $C\left([0,T], L^2\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d\right)\right)$ be the space of all continuous $L^2\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d\right)$-valued functions supplied with the norm:

$$\|X\| = \sup_{0\leqslant t\leqslant T} \left(\mathbb{E}\left[|X(t)|^2\right]\right)^{1/2}, \quad X \in C\left([0,T], L^2\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d\right)\right).$$

Define the operator $T : C\left([0,T], L^2\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d\right)\right) \to C\left([0,T], L^2\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d\right)\right)$ by the formula

$$TX(t) = x + \int_0^t \sigma(s, X(s))\, dB(s) + \int_0^t b(s, X(s))\, ds.$$

Then the argumentation in the proof below shows that T is a mapping from $C\left([0,T], L^2\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d\right)\right)$ to $C\left([0,T], L^2\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d\right)\right)$ indeed, and that T has a unique fixed point X which is a pathwise solution to the equation in (4.9).

PROOF. *Existence.* Fix $0 < T < \infty$. Put $X_0(s) = x$, $0 \leqslant s \leqslant t$, and, for $n \geqslant 1$, $0 < t \leqslant T$,

$$X_{n+1}(t) = x + \int_0^t b(s, X_n(s))\, ds + \int_0^t \sigma(s, X_n(s))\, dB(s). \quad (4.10)$$

By (4.10) we see, for $n \geqslant 1$ and $0 < t \leqslant T$,

$$\begin{aligned} X_{n+1}(t) - X_n(t) = &\int_0^t \left(b\left(s, X_n(s)\right) - b\left(s, X_{n-1}(s)\right)\right) ds \\ &+ \int_0^t \left(\sigma\left(s, X_n(s)\right) - \sigma\left(s, X_{n-1}(s)\right)\right) dB(s). \end{aligned} \tag{4.11}$$

By assumption there exists functions $s \mapsto K_j(s)$ and $s \mapsto K_{ij}(s)$, $0 \leqslant s \leqslant T$, such that for

$$\left|b_j\left(s,y\right) - b_j\left(s,x\right)\right| \leqslant K_j(s)\left|y-x\right|,\ 0 \leqslant s \leqslant T,\ x,\, y \in \mathbb{R}^d, \text{ and} \tag{4.12}$$

$$\left|\sigma_{ij}\left(s,y\right) - \sigma_{ij}\left(s,x\right)\right| \leqslant K_{i,j}(s)\left|y-x\right|,\ 0 \leqslant s \leqslant T,\ x,\, y \in \mathbb{R}^d, \tag{4.13}$$

and such that $\int_0^T \left(K_j(s)^2 + K_{i,j}(s)^2\right) ds < \infty$ for $0 \leqslant 1 \leqslant i,\, j \leqslant d$. Let the function $K(s) \geqslant 0$ be such that $K(s)^2 = \sum_{j=1}^d K_j(s)^2 + \sum_{i=1}^d \max_{1\leqslant j\leqslant d} K_{ij}(s)^2$. Then $\int_0^T K(s)^2\, ds < \infty$. Moreover, for $n \geqslant 1$ and $0 \leqslant t \leqslant T$ we infer, by using (4.11). (4.12) and (4.13), by the definition of $K(s)$, and by standard properties of stochastic integrals relative to Brownian motion, the following inequality:

$$\mathbb{E}\left[\left|X_{n+1}(t) - X_n(t)\right|^2\right] \leqslant 2\int_0^t K(s)^2 \mathbb{E}\left[\left|X_n(s) - X_{n-1}(s)\right|^2\right] ds. \tag{4.14}$$

In order to obtain (4.14) we also used an inequality of the form $\left(|a|+|b|\right)^2 \leqslant 2\left(|a|^2+|b|^2\right)$, $a,\, b \in \mathbb{R}^d$. The proofs of (4.15) and (4.18) require equalities of the form

$$\int_{s<s_1<\dots<s_j<t}\int \prod_{i=1}^j K\left(s_i\right)^2\, ds_1 \dots ds_j = \frac{\left(\int_s^t K(\rho)^2\, d\rho\right)^j}{j!},\ j \in \mathbb{N},\ j \geqslant 1.$$

By employing induction the inequality in (4.14) yields, for $1 \leqslant j \leqslant n$ and for $0 \leqslant t \leqslant T$, the inequality:

$$\begin{aligned} &\mathbb{E}\left[\left|X_{n+1}(t) - X_n(t)\right|^2\right] \\ &\leqslant 2^j \int_0^t \frac{\left(\int_s^t K(\rho)^2 d\rho\right)^{j-1}}{(j-1)!} K(s)^2 \mathbb{E}\left[\left|X_{n-j+1}(s) - X_{n-j}(s)\right|^2\right] ds. \end{aligned} \tag{4.15}$$

Since $X_0(s) = x$ the equality in (4.10) for $n = 0$ yields

$$X_1(s) - X_0(s) = \int_0^s b\left(\rho, x\right)\, d\rho + \int_0^s \sigma\left(\rho, x\right)\, dB(\rho),$$

and hence, for $0 \leqslant s \leqslant T$,

$$\mathbb{E}\left[\left|X_1(s) - X_0(s)\right|^2\right] \leqslant 2\left(\left|\int_0^s b\left(\rho,x\right)\, d\rho\right|^2 + \sum_{i,j=1}^d \int_0^s \left|\sigma_{ij}\left(\rho,x\right)\right|^2\, d\rho\right). \tag{4.16}$$

Let $A(s,x) \geqslant 0$ be such that

$$A(s,x)^2 = \sup_{0<\tau\leqslant s}\left|\int_0^\tau b\left(\rho,x\right)\, d\rho\right|^2 + \sum_{i,j=1}^d \int_0^s \left|\sigma_{ij}\left(\rho,x\right)\right|^2\, d\rho. \tag{4.17}$$

Then (4.17) together with (4.15) with $j = n$ yields

$$\begin{aligned}
\mathbb{E}\left[|X_{n+1}(t) - X_n(t)|^2\right] &\leqslant 2^n \int_0^t \frac{\left(\int_s^t K(\rho)^2 d\rho\right)^{n-1}}{(n-1)!} K(s)^2 \mathbb{E}\left[|X_1(s) - X_0(s)|^2\right] ds \\
&\leqslant 2^{n+1} \int_0^t \frac{\left(\int_s^t K(\rho)^2 d\rho\right)^{n-1}}{(n-1)!} K(s)^2 A(s,x)^2 \, ds \\
&\leqslant 2^{n+1} A(t,x)^2 \int_0^t \frac{\left(\int_s^t K(\rho)^2 d\rho\right)^{n-1}}{(n-1)!} K(s)^2 \, ds \\
&= 2^{n+1} A(t,x)^2 \int_{0<s_1<\cdots<s_n<t} \int \prod_{j=1}^n K\left(s_j\right)^2 \, ds_1 \dots ds_n \\
&= 2^{n+1} A(t,x)^2 \frac{\left(\int_0^t K(s)^2 ds\right)^n}{n!}. \qquad (4.18)
\end{aligned}$$

From Lemma (4.3) and inequality (4.6) with $\gamma = 2\int_0^t K(s)^2 \, ds$ we infer:

$$\sum_{n=0}^{\infty} \left(\mathbb{E}\left[|X_{n+1}(t) - X_n(t)|^2\right]\right)^{1/2} \leqslant 2A(t,x) \sqrt{\int_0^t K(s)^2 \, ds + 1} \, e^{\int_0^t K(s)^2 \, ds + \frac{1}{2}}. \qquad (4.19)$$

From (4.19) it easily follows that there exists an adapted $\mathbb{R}^d$-valued process $(X(t))_{0\leqslant t\leqslant T}$ in $L^2\left(\Omega, \mathfrak{F}_T, \mathbb{P}; \mathbb{R}^d\right)$ such that

$$\lim_{n\to\infty} \mathbb{E}\left[|X_n(t) - X(t)|^2\right] = 0. \qquad (4.20)$$

From (4.19) it also follows that this convergence also holds $\mathbb{P}$-almost surely. The latter can be seen as follows. Fix $\eta > 0$. Then the probability of the event $\{\limsup_{n\to\infty} |X_n(t) - X(t)| > \eta\}$ can be estimated as follows:

$$\begin{aligned}
\mathbb{P}\left[\limsup_{n\to\infty} |X_n(t) - X(t)| > \eta\right] &\leqslant \inf_{m\in\mathbb{N}} \mathbb{P}\left[\bigcup_{n=m}^{\infty} \{|X_n(t) - X(t)| > \eta\}\right] \\
&\leqslant \inf_{m\in\mathbb{N}} \mathbb{P}\left[\bigcup_{n_1>n_2\geqslant m}^{\infty} \{|X_{n_1}(t) - X_{n_2}(t)| > \eta\}\right] \\
&\leqslant \inf_{m\in\mathbb{N}} \mathbb{P}\left[\left\{\sum_{n=m}^{\infty} |X_{n+1}(t) - X_n(t)| > \eta\right\}\right] \\
&\leqslant \inf_{m\in\mathbb{N}} \frac{1}{\eta} \mathbb{E}\left[\sum_{n=m}^{\infty} |X_{n+1}(t) - X_n(t)|\right] \\
&\leqslant \inf_{m\in\mathbb{N}} \frac{1}{\eta} \sum_{n=m}^{\infty} \mathbb{E}\left[|X_{n+1}(t) - X_n(t)|\right] = 0. \qquad (4.21)
\end{aligned}$$

The final equality is a consequence of Lemma 4.3 together with (4.19) and the inequality $\mathbb{E}\left[|X_{n+1}(t) - X_n(t)|\right] \leqslant \left(\mathbb{E}\left[|X_{n+1}(t) - X_n(t)|^2\right]\right)^{1/2}$. Since $\eta > 0$ is arbitrary in (4.21) we infer that $\lim_{n\to\infty} X_n(t) = X(t)$ ($\mathbb{P}$-almost surely). This

$\mathbb{P}$-almost sure convergence (as $n \to \infty$) also implies that we may take pointwise limits in (4.10) to obtain:

$$X(t) = x + \int_0^t b(s, X(s))\, ds + \int_0^t \sigma(s, X(s))\, dB(s). \tag{4.22}$$

The equality in (4.22) shows the existence of pathwise or strong solutions to the equation in (4.9).

Uniqueness. Let $(X_1(t))_{0\leqslant t\leqslant T}$ and $(X_2(t))_{0\leqslant t\leqslant T}$ be two solutions to the stochastic differential equation in (4.9). By using a stopping time argument we may assume that $\sup_{0\leqslant s\leqslant T} |X_2(s) - X_1(s)|$ is $\mathbb{P}$-almost surely bounded. Then

$$\begin{aligned} X_2(t) - X_1(t) &= \int_0^t \left(b(s, X_2(s)) - b(s, X_1(s))\right) ds \\ &\quad + \int_0^t \left(\sigma(s, X_2(s)) - \sigma(s, X_1(s))\right) dB(s). \end{aligned} \tag{4.23}$$

As in the proof of (4.15) with $j = n$ and (4.18) it then follows that

$$\begin{aligned} &\mathbb{E}\left[|X_2(t) - X_1(t)|^2\right] \leqslant 2^n \int_0^t \frac{\left(\int_s^t K(\rho)^2 d\rho\right)^{n-1}}{(n-1)!} K(s)^2 \mathbb{E}\left[|X_2(s) - X_1(s)|^2\right] ds \\ &\leqslant 2^n \sup_{0<s<t} \mathbb{E}\left[|X_2(s) - X_1(s)|^2\right] \frac{\left(\int_0^t K(\rho)^2 d\rho\right)^n}{n!}. \end{aligned} \tag{4.24}$$

Since the right-hand side of (4.24) tends to 0 as $n \to \infty$ we see that $X_2(t) = X_1(t)$ $\mathbb{P}$-almost surely. So uniqueness follows.

The proof of Theorem 4.4 is complete now. ☐

1.2. A martingale characterization of Brownian motion. The following result we owe to Lévy.

4.5. Theorem. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space with filtration (or reference system)* $(\mathcal{F}_t)_{t \geqslant 0}$. *Suppose* $\mathcal{F}$ *is the* σ*-algebra generated by* $\cup_{t\geqslant 0}\mathcal{F}_t$ *augmented with the* $\mathbb{P}$*-zero sets, and suppose* $\mathcal{F}_t$ *is continuous from the right:* $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$ *for all* $t \geqslant 0$. *Let* $\{M(t) = (M_1(t), \dots, M_d(t)) : t \geqslant 0\}$ *be an* $\mathbb{R}^d$*-valued local* $\mathbb{P}$*-almost surely continuous martingale with the property that the quadratic covariation processes* $t \mapsto \langle M_i, M_j \rangle (t)$ *satisfy*

$$\langle M_i, M_j \rangle (t) = \delta_{i,j} t, \quad 1 \leqslant i, \, j \leqslant d. \tag{4.25}$$

Then $\{M(t) : t \geqslant 0\}$ *is* d*-dimensional Brownian motion with initial distribution given by* $\mu(B) = \mathbb{P}[M(0) \in B]$, $B \in \mathcal{B}_{\mathbb{R}^d}$, *the Borel field of* $\mathbb{R}^d$.

It follows that the finite-dimensional distributions of the process $t \mapsto M(t)$ are given by:

$$\begin{aligned}&\mathbb{P}\left[M(t_1) \in B_1, \dots, M(t_n) \in B_n\right] \\ &= \int \left(\int_{B_1} \cdots \int_{B_n} p_{0,d}(t_n - t_{n-1}, x_{n-1}, x_n) \cdots p_{0,d}(t_2 - t_1, x_1, x_2)\, p_{0,d}(t_1, x, x_1) \right. \\ &\qquad \left. dx_n \cdots dx_1 \right) d\mu(x).\end{aligned}$$

Here $p_{0,d}(t, x, y)$ is the classical Gaussian kernel:

$$p_{0,d}(t, x, y) = \frac{1}{\left(\sqrt{2\pi t}\right)^d} \exp\left(-\frac{|x-y|^2}{2t} \right). \tag{4.26}$$

4.6. Remark. There is even a nicer result which says the following. Let X be a continuous $\mathbb{R}^d$-valued process with stationary independent increments. Then, there exist unique $b \in \mathbb{R}^d$ and $\Sigma \in \mathbb{R}^{d^2}$ such that $X(t) - X(0)$ is a (b, Σ)-Brownian motion. This means that $X(t)$ is a Gaussian (or multivariate normal) vector such that $\mathbb{E}[X(t)] = bt$ and

$$\mathbb{E}\left[(X_{j_1}(t) - b_{j_1}t)(X_{j_2}(t) - b_{j_2}t) \right] = t\Sigma_{j_1,j_2}.$$

For the one-dimensional case the reader is referred to Breiman [**35**]. For the higher dimensional case, see, *e.g.*, Lowther [**116**].

Proof of Theorem 4.5. Let $\xi \in \mathbb{R}^d$ be arbitrary. First we show that it suffices to establish the equality:

$$\mathbb{E}\left[e^{-i\langle \xi, M(t)-M(s)\rangle} \bigm| \mathcal{F}_s \right] = e^{-\frac{1}{2}|\xi|^2 (t-s)}, \quad t > s \geqslant 0. \tag{4.27}$$

For suppose that (4.27) is true for all $\xi \in \mathbb{R}^d$. Observe that (4.27) implies $\mathbb{E}\left[e^{-i\langle \xi, M(t)-M(s)\rangle} \right] = e^{-\frac{1}{2}|\xi|^2 (t-s)}$. Then, by standard approximation arguments, it follows that the variable $M(t) - M(s)$ is $\mathbb{P}$-independent of $\mathcal{F}_s$. In other words the process $t \mapsto M(t)$ possesses independent increments. Since the Fourier transform of the function $y \mapsto p_{0,d}(t-s, 0, y)$ is given by

$$\int_{\mathbb{R}^d} e^{-i\langle \xi, y\rangle} p_{0,d}(t-s, 0, y)\, dy = e^{-\frac{1}{2}|\xi|^2 (t-s)}$$

it also follows that the distribution of $M(t)-M(s)$ is given by

$$\mathbb{P}\left[M(t)-M(s)\in B\right]=\int_B p_{0,d}\left(t-s,0,y\right)dy. \tag{4.28}$$

Moreover, for $0<t_1<\cdots<t_n$ we also have

$$\begin{aligned}&\mathbb{P}\left[M(0)\in B_0,\, M\left(t_1\right)-M(0)\in B_1,\dots,M\left(t_n\right)-M\left(t_{n-1}\right)\in B_n\right]\\&=\mathbb{P}\left[M(0)\in B_0\right]\mathbb{P}\left[M\left(t_1\right)-M(0)\in B_1\right]\cdots\mathbb{P}\left[M\left(t_n\right)-M\left(t_{n-1}\right)\in B_n\right]\\&=\int_{B_0}\int_{B_1}\cdots\int_{B_n}p_{0,d}\left(t_1,0,y_1\right)\cdots p_{0,d}\left(t_n-t_{n-1},0,y_n\right)d\mu\left(y_0\right)dy_1\cdots dy_n.\end{aligned}$$

Here $B_0,\dots,B_n$ are Borel subsets of $\mathbb{R}^d$. Hence, if B is a Borel subset of $\underbrace{\mathbb{R}^d\times\cdots\times\mathbb{R}^d}_{n+1\text{times}}$, then it follows that

$$\begin{aligned}&\mathbb{P}\left[\left(M(0),M\left(t_1\right)-M(0),\dots,M\left(t_n\right)-M\left(t_{n-1}\right)\right)\in B\right]\\&=\underset{B}{\int\cdots\int}p_{0,d}\left(t_1,0,y_1\right)\cdots p_{0,d}\left(t_n-t_{n-1},0,y_n\right)d\mu\left(y_0\right)dy_1\cdots dy_n.\end{aligned} \tag{4.29}$$

Next we compute the joint distribution of $\left(M(0),M\left(t_1\right),\dots,M\left(t_n\right)\right)$ by employing (4.29). Define the linear map $\ell:\mathbb{R}^d\times\cdots\times\mathbb{R}^d\to\mathbb{R}^d\times\cdots\times\mathbb{R}^d$ by

$$\ell\left(x_0,x_1,\dots,x_n\right)=\left(x_0,x_1-x_0,x_2-x_1,\dots,x_n-x_{n-1}\right).$$

Let B be a Borel subset of $\mathbb{R}^d\times\cdots\times\mathbb{R}^d$. By (4.29) we get

$$\begin{aligned}&\mathbb{P}\left[\left(M(0),\dots,M\left(t_n\right)\right)\in B\right]\\&=\mathbb{P}\left[\ell\left(M(0),\dots,M\left(t_n\right)\right)\in\ell\left(B\right)\right]\\&=\mathbb{P}\left[\left(M(0),M\left(t_1\right)-M(0),\dots,M\left(t_n\right)-M\left(t_{n-1}\right)\right)\in\ell\left(B\right)\right]\\&=\underset{\ell(B)}{\int\dots\int}p_{0,d}\left(t_1,0,y_1\right)\cdots p_{0,d}\left(t_n-t_{n-1},0,y_n\right)d\mu\left(y_0\right)dy_1\cdots dy_n\end{aligned}$$

(change of variables: $\left(y_0,y_1,\dots,y_n\right)=\ell\left(x_0,x_1,\dots,x_n\right)$)

$$=\underset{B}{\int\cdots\int}p_{0,d}\left(t_1,x_0,x_1\right)\cdots p_{0,d}\left(t_n-t_{n-1},x_{n-1},x_n\right)\,d\mu\left(x_0\right)dx_1\cdots dx_n. \tag{4.30}$$

In order to complete the proof of Theorem 4.5 from equality (4.30) it follows that it is sufficient to establish the equality in (4.27). Therefore, fix $\xi\in\mathbb{R}^d$ and $t>s\geqslant 0$. An application of Itô's lemma to the function $x\mapsto e^{-i\langle\xi,x\rangle}$ yields

$$\begin{aligned}&e^{-i\langle\xi,M(t)\rangle}-e^{-i\langle\xi,M(s)\rangle}\\&=-i\sum_{j=1}^d\xi_j\int_s^t e^{-i\langle\xi,M(\tau)\rangle}dM_j(\tau)-\frac{1}{2}\sum_{j,k=1}^d\xi_j\xi_k\int_s^t e^{-i\langle\xi,M(\tau)\rangle}d\left\langle M_j,M_k\right\rangle(\tau)\end{aligned}$$

(formula (4.25))

$$=-i\sum_{j=1}^d\xi_j\int_s^t e^{-i\langle\xi,M(\tau)\rangle}dM_j(\tau)-\frac{1}{2}\left|\xi\right|^2\int_s^t e^{-i\langle\xi,M(\tau)\rangle}\,d\tau. \tag{4.31}$$

Hence, from (4.31) it follows that

$$\begin{aligned} & e^{-i\langle\xi,M(t)-M(s)\rangle}-1 \qquad (4.32)\\ &= -i\sum_{j=1}^d \xi_j\int_s^t e^{-i\langle\xi,M(\tau)-M(s)\rangle}dM_j(\tau)-\frac{1}{2}\left|\xi\right|^2\int_s^t e^{-i\langle\xi,M(\tau)-M(s)\rangle}\,d\tau. \end{aligned}$$

Since the processes

$$t\mapsto\int_s^t e^{-i\langle\xi,M(\tau)-M(s)\rangle}\,dM_j(\tau),\ t\geqslant s,\ 1\leqslant j\leqslant d,$$

are local martingales, we infer by (possibly) using a stopping time argument that

$$\mathbb{E}\left[e^{-i\langle\xi,M(t)-M(s)\rangle}\bigm|\mathcal{F}_s\right]=1-\frac{1}{2}\left|\xi\right|^2\int_s^t\mathbb{E}\left[e^{-i\langle\xi,M(\tau)-M(s)\rangle}\bigm|\mathcal{F}_s\right]d\tau. \qquad (4.33)$$

Next, let $v(t)$, $t\geqslant s$, be given by

$$v(t)=\int_s^t\mathbb{E}\left[e^{-i\langle\xi,M(\tau)-M(s)\rangle}\bigm|\mathcal{F}_s\right]d\tau.$$

Then $v(s)=0$, and (4.33) implies

$$v'(t)+\frac{1}{2}\left|\xi\right|^2v(t)=1. \qquad (4.34)$$

From (4.34) we infer

$$\frac{d}{dt}\left(e^{\frac{1}{2}(t-s)|\xi|^2}v(t)\right)=\left(\frac{1}{2}\left|\xi\right|^2v(t)+v'(t)\right)e^{\frac{1}{2}(t-s)|\xi|^2}=e^{\frac{1}{2}(t-s)|\xi|^2}. \qquad (4.35)$$

The equality in (4.35) implies:

$$e^{\frac{1}{2}(t-s)|\xi|^2}v(t)-v(s)=\frac{2}{\left|\xi\right|^2}\left(e^{\frac{1}{2}(t-s)|\xi|^2}-1\right),$$

and thus we see

$$v'(t)+\frac{1}{2}v(s)e^{-\frac{1}{2}(t-s)|\xi|^2}=e^{-\frac{1}{2}(t-s)|\xi|^2} \qquad (4.36)$$

Since $v(s)=0$ (4.36) results in

$$\mathbb{E}\left[e^{-i\langle\xi,M(\tau)-M(s)\rangle}\bigm|\mathcal{F}_s\right]=v'(t)=e^{-\frac{1}{2}(t-s)|\xi|^2}. \qquad (4.37)$$

The equality in (4.37) is the same as the one in (4.27). By the above arguments this completes the proof of Theorem 4.5. ☐

As a corollary to Theorem 4.5 we get the following result due to Lévy.

4.7. Corollary. *Let $\{M(t):t\geqslant 0\}$ be a continuous local martingale in $\mathbb{R}$ such that the process $t\mapsto M(t)^2-t$ is a local martingale as well. Then the process $\{M(t):t\geqslant 0\}$ is a Brownian motion with initial distribution given by $\mu(B)=\mathbb{P}\left[M(0)\in B\right]$, $B\in\mathcal{B}_{\mathbb{R}}$.*

Proof. Since $M(t)^2-t$ is a local martingale, it follows that the quadratic variation process $t\mapsto\langle M,M\rangle(t)$ satisfies $\langle M,M\rangle(t)=t$, $t\geqslant 0$. So the result in Corollary 4.7 follows from Theorem 4.5. ☐

The following result contains a d-dimensional version of Corollary 4.7.

4.8. THEOREM. *Let $\{M(t) = (M_1(t), \dots, M_{d'}(t)) : t \geqslant 0\}$ be a continuous local martingale with covariation process given by*

$$\langle M_j, M_k \rangle (t) = \int_0^t \Phi_{j,k}(s) ds, \ 1 \leqslant j, \, k \leqslant d'. \tag{4.38}$$

Let the $d' \times d$-matrix process $\{\chi(t) : t \geqslant 0\}$ be such that $\chi(t)\Phi(t)\chi(t)^* = I$, where I is the $d \times d$ identity matrix. Put $B(t) = \int_0^t \chi(s) \, dM(s)$. This integral should be interpreted in Itô sense. Then the process $t \mapsto B(t)$ is d-dimensional Brownian motion. Put $\Psi(t) = \Phi(t)\chi(t)^*$, and suppose that $\Psi(t)\chi(t) = I$, the $d' \times d'$ identity matrix. Then $M(t) - M(0) = \int_0^t \Psi(s) \, dB(s)$.

4.9. REMARK. Since

$$\chi(t) \left(\Phi(t)\chi(t)^*\chi(t) - I \right) = \left(\chi(t)\Phi(t)\chi(t)^* - I \right) \chi(t) = 0$$

we see that the second equality in $\Psi(t)\chi(t) = \Phi(t)\chi(t)^*\chi(t) = I$ is only possible if we assume $d = d'$. Of course here we take the dimensions of the null and range space of the matrix $\chi(t)$ into account.

Proof of Theorem 4.8. Fix $1 \leqslant i,\, j \leqslant d$. We shall calculate the quadratic covariation process

$$\begin{aligned}\left\langle B_i, B_j\right\rangle(t) &= \left\langle \sum_{k=1}^{d'} \int_0^{(\cdot)} (\chi(s))_{i,k}\, dM_k(s), \sum_{l=1}^{d'} \int_0^{(\cdot)} (\chi(s))_{j,l}\, dM_l(s) \right\rangle(t) \\ &= \sum_{k=1}^{d'} \sum_{l=1}^{d'} \int_0^t (\chi(s))_{i,k} (\chi(s))_{j,l}\, \Phi(s)_{k,l} ds \\ &= \int_0^t (\chi(s)\Phi(s)\chi(s)^*)_{i,j}\, ds = t\delta_{i,j}. \end{aligned} \tag{4.39}$$

From Theorem 4.5 and (4.39) we see that the process $t \mapsto B(t)$ is a Brownian motion. This proves the first part of Theorem 4.8. Next we calculate

$$\int_0^t \Psi(s)\, dB(s) = \int_0^t \Psi(s)\chi(s)\, dM(s) = \int_0^t dM(s) = M(t) - M(0), \tag{4.40}$$

which completes the proof of Theorem 4.8. ☐

1.3. Exponential local martingales. Let $t \mapsto N(t)$, $0 \leqslant t \leqslant T$, be a continuous (local) martingale with variation process $t \mapsto \langle N, N\rangle(t)$, $0 \leqslant t \leqslant T$. In this subsection we discuss local martingales of the form $t \mapsto e^{-Z(t)} = 1 + \int_0^t e^{-Z(s)}\, dN(s)$, $t \geqslant 0$, where $Z(t) = N(t) + \frac{1}{2}\langle N, N\rangle(t)$. Such processes are called *exponential local martingales*. The following proposition serves as a preparation for Proposition 4.12. It also has some interest of its own.

4.10. Proposition. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space endowed with a filtration $(\mathcal{F}_t)_{0\leqslant t\leqslant T}$, and let $M = (M(t))_{0\leqslant t\leqslant T}$ and $N = (N(t))_{0\leqslant t\leqslant T}$ be two local continuous martingales with $M(0) = N(0) = 0$. Put $Z(t) = N(t) + \frac{1}{2}\langle N, N\rangle(t)$, and assume that $\mathbb{E}\left[e^{-Z(t)}\right] = 1$ for all $0 \leqslant t \leqslant T$. Then the following assertions are true.*

(a) *The process $t \mapsto e^{-Z(t)}$, $0 \leqslant t \leqslant T$, is a martingale;*
(b) *The process $t \mapsto e^{-Z(t)}\left(M(t) + \langle N, M\rangle(t)\right)$ is a local martingale;*
(c) *The process $t \mapsto M(t) + \langle N, M\rangle(t)$ is a local martingale relative to the filtration $(\mathcal{F}_t)_{0\leqslant t\leqslant T}$ supplied with the measure $\mathbb{Q}_N : \mathcal{F}_T \to [0,1]$ defined by $\mathbb{Q}_N(A) = \mathbb{E}\left[e^{-Z(T)}\mathbf{1}_A\right]$, $A \in \mathcal{F}_T$.*

The measure $\mathbb{Q}_N$ can be called a *risk neutral measure*. Observe that, by assertion (a), $\mathbb{Q}_N(A) = \mathbb{E}\left[e^{-Z(t)}\mathbf{1}_A\right]$ whenever A belongs to $\mathcal{F}_t$ with $0 \leqslant t \leqslant T$. Let τ_n be the stopping time defined by

$$\tau_n = \inf\left\{ s > 0 : |N(s \wedge T)| + \frac{1}{2}\langle N, N\rangle(s \wedge T) > n \right\},$$

and set $Z_n(t) = Z(t \wedge \tau_n)$. Then the processes $t \mapsto e^{-Z_n(t)}$, $0 \leqslant t \leqslant T$, $n \in \mathbb{N}$, are martingales. It follows that $\mathbb{E}\left[e^{-Z_n(t)}\right] = 1$, for all $0 \leqslant t \leqslant T$, and for all $n \in \mathbb{N}$. By Fatou's lemma we infer that

$$\mathbb{E}\left[e^{-Z(t)}\right] = \mathbb{E}\left[\lim_{n\to\infty} e^{-Z_n(t)}\right] \leqslant \liminf_{n\to\infty} \mathbb{E}\left[e^{-Z_n(t)}\right] \leqslant 1. \tag{4.41}$$

In fact we have a stronger result. It says that an exponential local martingale is a super-martingale.

4.11. Theorem. *Let the process* $t \mapsto e^{-Z(t)}$, $0 \leqslant t \leqslant T$, *be a continuous local martingale. In general, this process is a super-martingale relative to the filtration* $(\mathcal{F}_t)_{0t\leqslant t\leqslant T}$ *with* $\mathcal{F}_t = \sigma\left(N(s) : 0 \leqslant s \leqslant t\right)$. *Consequently, if* $\mathbb{E}\left[e^{-Z(t)}\right] = 1$ *for all* $0 \leqslant t \leqslant T$, *then the process* $t \mapsto e^{-Z(t)}$, $0 \leqslant t \leqslant T$, *is a martingale.*

Proof. This result can be seen as follows. Let $0 \leqslant t_1 < t_2$, and choose the sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ as above. Then, for $A \in \mathcal{F}_{t_1\wedge\tau_m}$, we have

$$\begin{aligned}\mathbb{E}\left[e^{-Z(t_2)}\mathbf{1}_A\right] &= \mathbb{E}\left[\lim_{n\to\infty} e^{-Z(t_2\wedge\tau_n)}\mathbf{1}_A\right]\\ &\leqslant \liminf_{n\to\infty}\mathbb{E}\left[e^{-Z(t_2\wedge\tau_n)}\mathbf{1}_A\right] = \mathbb{E}\left[e^{-Z(t_1\wedge\tau_m)}\mathbf{1}_A\right] \end{aligned}\tag{4.42}$$

In (4.42) we employed Doob's optional sampling Theorem 1.23 for martingales. From (4.42) it follows that:

$$\mathbb{E}\left[e^{-Z(t_2)} \,\middle|\, \mathcal{F}_{t_1\wedge\tau_m}\right] \leqslant e^{-Z(t_1\wedge\tau_m)}. \tag{4.43}$$

Since the event $\{\tau_m > t_1\}$ belongs to $\mathcal{F}_{t_1\wedge\tau_m}$, from (4.43) we infer

$$\begin{aligned}\mathbb{E}\left[e^{-Z(t_2)}\mathbf{1}_{\{\tau_m>t_1\}} \,\middle|\, \mathcal{F}_{t_1}\right] &= \mathbb{E}\left[e^{-Z(t_2)}\mathbf{1}_{\{\tau_m>t_1\}} \,\middle|\, \mathcal{F}_{t_1\wedge\tau_m}\right]\\ &\leqslant e^{-Z(t_1\wedge\tau_m)}\mathbf{1}_{\{\tau_m>t_1\}} = e^{-Z(t_1)}\mathbf{1}_{\{\tau_m>t_1\}}. \end{aligned}\tag{4.44}$$

The first equality in (4.44) is a consequence of the fact that, if an event A belongs to $\mathcal{F}_{t_1}$, then $A \cap \{\tau_m > t_1\}$ belongs to $\mathcal{F}_{t_1\wedge\tau_m}$. With $A \in \mathcal{F}_{t_1}$ of the form $A = \{(N(s_1), \dots, N(s_n)) \in B\}$, where $0 \leqslant s_1 < \dots < s_n \leqslant t_1$ and B is a Borel subset of $\mathbb{R}^n$, the identities

$$\begin{aligned}\mathbf{1}_{\{\tau_m>t_1\}}\mathbf{1}_A &= \mathbf{1}_{\{\tau_m>t_1\}}\mathbf{1}_B\left(N(s_1), \dots, N(s_n)\right)\\ &= \mathbf{1}_{\{\tau_m>t_1\}}\mathbf{1}_B\left(N(s_1\wedge\tau_m), \dots, N(s_n\wedge\tau_m)\right)\end{aligned}$$

are self-explanatory. Fix $1 \leqslant j \leqslant n$. Right-continuity implies that the equality $N(s_j\wedge\tau_m) = \lim_{k\to\infty} N\left(\left(2^{-k}\left\lceil 2^k(s_j\wedge\tau_m)\right\rceil\right)\wedge s_j\right)$ holds, and hence the variable $N(s_j\wedge\tau_m)$ is $\mathcal{F}_{s_j\wedge\tau_m}$-measurable. Consequently, the event $A \cap \{\tau_m > t_1\}$ belongs to $\mathcal{F}_{t_1\wedge\tau_m}$. Theorem 2.42 about π-λ systems then implies that for all $A \in \mathcal{F}_{t_1}$ the event $\{\tau_m > t_1\} \cap A$ belongs to the σ-field $\mathcal{F}_{t_1\wedge\tau_m}$. In the left-hand side and the far right-hand side of (4.44) we let $m \to \infty$ to obtain

$$\mathbb{E}\left[e^{-Z(t_2)} \,\middle|\, \mathcal{F}_{t_1}\right] \leqslant e^{-Z(t_1)}, \quad \mathbb{P}\text{-almost surely.} \tag{4.45}$$

By definition, the inequality in (4.45) shows that the process $t \mapsto e^{-Z(t)}$ is a super-martingale. If $0 \leqslant t_1 < t_2 \leqslant T$, and if $\mathbb{E}\left[e^{-Z(t_2)}\right] = \mathbb{E}\left[e^{-Z(t_1)}\right]$, then (4.45) implies that

$$\mathbb{E}\left[e^{-Z(t_2)} \,\middle|\, \mathcal{F}_{t_1}\right] = e^{-Z(t_1)}, \quad \mathbb{P}\text{-almost surely.} \tag{4.46}$$

This completes the proof of Theorem 4.11. □

Proof of 4.10. (a) An application of Itô's formula and employing the equality $\langle Z, Z\rangle(t) = \langle N, N\rangle(t)$ yields:

$e^{-Z(t)}$

$$\begin{aligned}
&= e^{-Z(0)} - \int_0^t e^{-Z(\rho)}\, dZ(\rho) + \frac{1}{2}\int_0^t e^{-Z(\rho)}\, d\left\langle Z, Z\right\rangle(\rho) \\
&= e^{-Z(0)} - \int_0^t e^{-Z(\rho)}\, dN(\rho) - \frac{1}{2}\int_0^t e^{-Z(\rho)}\, d\left\langle N, N\right\rangle(\rho) + \frac{1}{2}\int_0^t e^{-Z(\rho)}\, d\left\langle N, N\right\rangle(\rho) \\
&= e^{-Z(0)} - \int_0^t e^{-Z(\rho)}\, dN(\rho). \qquad (4.47)
\end{aligned}$$

From the equalities in (4.47) it follows that the process $t \mapsto e^{-Z(t)}$, $0 \leqslant t \leqslant T$, is a local martingale. In view of the assumption that $\mathbb{E}\left[e^{-Z(t)}\right] = 1$ for all $0 \leqslant t \leqslant T$ it follows that the process in (a) is a genuine martingale: see Theorem 4.11.

(b) Again we apply Itô's lemma, now to the function $(x, y) \mapsto e^{-x}y$. Then we obtain:

$$\begin{aligned}
&e^{-Z(t)}\left(M(t) + \left\langle N, M\right\rangle(t)\right) \\
&= -\int_0^t e^{-Z(\rho)}\left(M(\rho) + \left\langle N, M\right\rangle(\rho)\right)\, dZ(\rho) + \int_0^t e^{-Z(\rho)}\left(dM(\rho) + d\left\langle N, M\right\rangle(\rho)\right) \\
&\quad + \frac{1}{2}\int_0^t e^{-Z(\rho)}\left(M(\rho) + \left\langle N, M\right\rangle(\rho)\right)\, d\left\langle Z, Z\right\rangle(\rho) \\
&\quad - \int_0^t e^{-Z(\rho)} d\left\langle Z, M + \left\langle N, M\right\rangle\right\rangle(\rho). \qquad (4.48)
\end{aligned}$$

By applying the equalities $\left\langle Z, Z\right\rangle = \left\langle N, N\right\rangle$ and $\left\langle Z, M + \left\langle N, M\right\rangle\right\rangle = \left\langle N, M\right\rangle$ to the equality in (4.48) we obtain

$$\begin{aligned}
&e^{-Z(t)}\left(M(t) + \left\langle N, M\right\rangle(t)\right) \\
&= -\int_0^t e^{-Z(\rho)}\left(M(\rho) + \left\langle N, M\right\rangle(\rho)\right)\, dN(\rho) \\
&\quad - \frac{1}{2}\int_0^t e^{-Z(\rho)}\left(M(\rho) + \left\langle N, M\right\rangle(\rho)\right)\, d\left\langle N, N\right\rangle(\rho) \\
&\quad + \int_0^t e^{-Z(\rho)}\left(dM(\rho) + d\left\langle N, M\right\rangle(\rho)\right) \\
&\quad + \frac{1}{2}\int_0^t e^{-Z(\rho)}\left(M(\rho) + \left\langle N, M\right\rangle(\rho)\right)\, d\left\langle N, N\right\rangle(\rho) - \int_0^t e^{-Z(\rho)} d\left\langle N, M\right\rangle(\rho) \\
&= -\int_0^t e^{-Z(\rho)}\left(M(\rho) + \left\langle N, M\right\rangle(\rho)\right)\, dN(\rho) + \int_0^t e^{-Z(\rho)}\, dM(\rho). \qquad (4.49)
\end{aligned}$$

Being the sum of two stochastic integrals with respect to (local) martingales the equality in (4.49) implies that the process in (b) is a local martingale.

(c) By using a stopping time argument we may and do assume that the process $t \mapsto M(t) + \left\langle N, M\right\rangle(t)$ is bounded and so it belongs to $L^1\left(\Omega, \mathcal{F}_T, \mathbb{Q}_N\right)$. Let $0 \leqslant t_1 < t_2 \leqslant T$, and put

$$Y\left(t_1\right) = \mathbb{E}_{\mathbb{Q}_N}\left[M\left(t_2\right) + \left\langle N, M\right\rangle\left(t_2\right) \,\middle|\, \mathcal{F}_{t_1}\right].$$

Then the stochastic variable $Y(t_1)$ is $\mathcal{F}_{t_1}$-measurable and, for all bounded $\mathcal{F}_{t_1}$-measurable variables G we have

$$\mathbb{E}\left[e^{-Z(T)}\left(M(t_2)+\langle N,M\rangle(t_2)\right)G\right]=\mathbb{E}\left[e^{-Z(T)}Y(t_1)\,G\right]. \tag{4.50}$$

Since the process $t\mapsto e^{-Z(t)}$, $0\leqslant t\leqslant T$, is a $\mathbb{P}$ martingale, the equality in (4.50) implies:

$$\mathbb{E}\left[e^{-Z(t_2)}\left(M(t_2)+\langle N,M\rangle(t_2)\right)G\right]=\mathbb{E}\left[e^{-Z(t_1)}Y(t_1)\,G\right]. \tag{4.51}$$

From assertion (b) together with our stopping time argument we see that the process $t\mapsto e^{-Z(t)}\left(M(t)+\langle N,M\rangle(t)\right)$ is a $\mathbb{P}$-martingale. From (4.51) we then infer:

$$\mathbb{E}\left[e^{-Z(t_1)}\left(M(t_1)+\langle N,M\rangle(t_1)\right)G\right]=\mathbb{E}\left[e^{-Z(t_1)}Y(t_1)\,G\right] \tag{4.52}$$

for all bounded $\mathcal{F}_{t_1}$-measurable variables G. So finally we get, $\mathbb{P}$-almost surely,

$$e^{-Z(t_1)}Y(t_1)=e^{-Z(t_1)}\left(M(t_1)+\langle N,M\rangle(t_1)\right),$$

and hence,

$$Y(t_1)=M(t_1)+\langle N,M\rangle(t_1),\quad \mathbb{P}\text{-almost surely.}$$

This shows assertion (c) and completes the proof of Proposition 4.10. ☐

A combination of Proposition 4.10 and Lévy's characterization of Brownian motion in $\mathbb{R}^d$ yields the following result.

4.12. Proposition. *Let the $\mathbb{R}^d$-valued process $s \mapsto c(s)$ be an adapted process which is predictable relative to Brownian motion $(B(t))_{t\geqslant 0}$. Put $N(t) = \int_0^t c(s)\,dB(s)$, and*

$$Z(t) = N(t) + \frac{1}{2}\langle N, N\rangle(t) = \int_0^t c(s)\,dB(s) + \frac{1}{2}\int_0^t |c(s)|^2\,ds, \quad t \geqslant 0.$$

Suppose that for all $t > 0$ the equality $\mathbb{E}\left[e^{-Z(t)}\right] = 1$ holds. Then the process $(W(t))_{t\geqslant 0}$, defined by $W(t) = B(t) + \int_0^t c(s)\,ds$ is Brownian motion relative to the measure $A \mapsto \mathbb{Q}_N(A)$, $A \in \mathcal{F}_T$, as defined in Proposition 4.10.

Proof. An application of Proposition 4.10 with $M(t) = B_j(t)$ shows that the process

$$\begin{aligned} W_j(t) &= B_j(t) + \int_0^t c_j(s)\,ds = B_j(t) + \left\langle \int_0^{(\cdot)} c(s)\,dB(s), B_j \right\rangle(t) \\ &= B_j(t) + \langle N, B_j\rangle(t) \end{aligned}$$

is a local $\mathbb{Q}_N$-martingale. Moreover, $\langle W_{j_1}, W_{j_2}\rangle(t) = \delta_{j_1,j_2}t$. From Theorem 4.5 we see that the process $t \mapsto W(t)$ is a $\mathbb{Q}_N$-Brownian motion. This completes the proof of Proposition 4.12. □

It will be very convenient to introduce Hermite polynomials $(h_k(x))_{k\in\mathbb{N}}$, and to establish some of their properties. In the context of stochastic calculus they also play a central role. The Hermite polynomial $h_k(x)$ is defined by

$$h_k(x) = (-1)^k e^{\frac{1}{2}x^2}\left(\frac{d}{dx}\right)^k\left(e^{-\frac{1}{2}x^2}\right). \tag{4.53}$$

For $k \in \mathbb{N}$, $x \in \mathbb{R}$, $a > 0$, we write

$$H_k(x,a) = a^{k/2}h_k\left(\frac{x}{\sqrt{a}}\right).$$

Then we have $H_0(x,a) = 1$, $H_1(x,a) = x$, $H_2(x,a) = x^2 - a$, $H_3(x,a) = x^3 - ax$. The Hermite polynomials satisfy the following recurrence relation:

$$h_{k+2}(x) - xh_{k+1}(x) + (k+1)h_k(x) = 0,\ k \geqslant 0, \tag{4.54}$$

and therefore

$$H_{k+2}(x,a) - xH_{k+1}(x,a) + (k+1)aH_k(x,a) = 0,\ k \geqslant 0. \tag{4.55}$$

The equality in (4.54) can be proved by induction and the definition of h_k in (4.53). From the definition of $h_{k+1}(x)$ it follows that $h'_{k+1}(x) = xh_{k+1}(x) - h_{k+2}(x)$, and so, by (4.54) we see

$$h'_{k+1}(x) = (k+1)h_k(x),\ k \geqslant 0. \tag{4.56}$$

From (4.54) and (4.56) we infer

$$h_{k+2}(x) - xh_{k+1}(x) + h'_{k+1}(x) = 0,\ k \geqslant 0,$$

and hence

$$h_{k+1}(x) - xh_k(x) + h_k'(x) = 0,\ k \geqslant 0. \tag{4.57}$$

By differentiating the equality in (4.57) and again using (4.56) we obtain the following differential equation:

$$h_k''(x) - xh_k'(x) + kh_k(x) = 0,\ k \geqslant 0. \tag{4.58}$$

In the following proposition we collect some of their properties.

4.13. Proposition. *For τ, $x \in \mathbb{R}$ and $a > 0$ the following identities are true:*

$$e^{\tau x - \frac{1}{2}\tau^2 a} = \sum_{k=0}^{\infty} \frac{\tau^k}{k!} H_k(x, a), \tag{4.59}$$

$$e^{\tau x - \frac{1}{2}\tau^2} = \sum_{k=0}^{\infty} \frac{\tau^k}{k!} H_k(x, 1) = \sum_{k=0}^{\infty} \frac{\tau^k}{k!} h_k(x), \tag{4.60}$$

$$\frac{\partial}{\partial x} H_{k+1}(x, a) = (k+1) H_k(x, a), \quad \textit{and} \quad \frac{1}{2}\frac{\partial^2}{\partial x^2} H_k(x, a) + \frac{\partial}{\partial a} H_k(x, a) = 0. \tag{4.61}$$

Proof. Let the sequence $\left(\widetilde{h}_k(x)\right)_{k\in\mathbb{N}}$ be such that, for all x and $\tau \in \mathbb{C}$, the equality

$$e^{\tau x - \frac{1}{2}\tau^2} = \sum_{k=0}^{\infty} \frac{\tau^k}{k!} \widetilde{h}_k(x) \tag{4.62}$$

holds. Then

$$\begin{aligned}
\widetilde{h}_k(x) &= \left(\frac{\partial}{\partial \tau}\right)^k \left(e^{\tau x - \frac{1}{2}\tau^2}\right)\Big|_{\tau=0} = e^{\frac{1}{2}x^2} \left(\frac{\partial}{\partial \tau}\right)^k \left(e^{-\frac{1}{2}(\tau - x)^2}\right)\Big|_{\tau=0} \\
&= (-1)^k \left(\frac{\partial}{\partial x}\right)^k \left(e^{-\frac{1}{2}(\tau - x)^2}\right)\Big|_{\tau=0} = (-1)^k \left(\frac{d}{dx}\right)^k \left(e^{-\frac{1}{2}x^2}\right) \\
&= h_k(x).
\end{aligned} \tag{4.63}$$

The equality in (4.63) implies the identity in (4.60). By a correct scaling ($\tau\sqrt{a}$ replaces τ, and $\dfrac{x}{\sqrt{a}}$ replaces x) the equality in (4.59) follows from (4.60) and the definition of $H_k(x, a)$. The equalities in (4.61) follow from (4.56) and from (4.58) respectively. Altogether this completes the proof of Proposition 4.13. ☐

In the following proposition the process $t \mapsto M(t)$, $t \in [0, T]$, is a martingale on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Its quadratic variation process is denoted by $t \mapsto \langle M, M\rangle(t)$, $t \in [0, T]$.

4.14. Proposition. *The following identities hold:*

$$\begin{aligned}
\frac{H_{k+1}\left(M(t), \langle M, M\rangle(t)\right)}{(k+1)!} &= \int_0^t \frac{H_k\left(M(s), \langle M, M\rangle(s)\right)}{k!}\, dM(s) \\
&= \int_{0<s_1<\dots<s_{k+1}<t} \mathbf{1} dM(s_1) \dots dM(s_{k+1}).
\end{aligned} \tag{4.64}$$

In addition, the following equalities hold as well:

$$
\begin{aligned}
& e^{\tau M(t)-\frac{1}{2}\tau^2\langle M,M\rangle(t)} \\
&= 1+\tau\int_0^t e^{\tau M(s)-\frac{1}{2}\tau^2\langle M,M\rangle(s)}\,dM(s) \\
&= 1+\sum_{k=1}^{\ell-1}\tau^k\int_{0<s_1<\dots<s_k<t}\!\!\int \mathbf{1}dM\left(s_1\right)\dots dM\left(s_k\right) \\
&\qquad + \tau^\ell\int_{0<s_1<\dots<s_\ell<t}\!\!\int e^{\tau M(s_1)-\frac{1}{2}\tau^2\langle M,M\rangle(s_1)}\,dM\left(s_1\right)\dots dM\left(s_\ell\right) \qquad (4.65) \\
&= \sum_{k=0}^{\ell-1}\frac{\tau^k}{k!}H_k\left(M(t),\langle M,M\rangle(t)\right) \\
&\quad + \tau^\ell\int_{0<s_1<\dots<s_\ell<t}\!\!\int e^{\tau M(s_1)-\frac{1}{2}\tau^2\langle M,M\rangle(s_1)}\,dM\left(s_1\right)\dots dM\left(s_\ell\right) \\
&= \sum_{k=0}^{\ell}\frac{\tau^k}{k!}H_k\left(M(t),\langle M,M\rangle(t)\right) \\
&\quad + \tau^\ell\int_{0<s_1<\dots<s_\ell<t}\!\!\int \left(e^{\tau M(s_1)-\frac{1}{2}\tau^2\langle M,M\rangle(s_1)}-\mathbf{1}\right)dM\left(s_1\right)\dots dM\left(s_\ell\right). \qquad (4.66)
\end{aligned}
$$

Please notice that in the equalities in (4.64) through (4.66) the order of integration has to be respected: first we integrate with respect $dM\left(s_1\right)$, then with respect to $dM\left(s_2\right)$ and so on.

PROOF OF PROPOSITION 4.14. These equalities follow from Itô's formula and the equalities in Proposition 4.13. Itô's lemma is applied to the functions $(x,a) \mapsto H_{k+1}(x,a)$, and $(x,a) \mapsto e^{\tau x - \frac{1}{2}\tau^2 a}$ with $x = M(s)$, and $a = \langle M, M\rangle(s)$. In particular the equalities in (4.61) are relevant. This completes the proof of Proposition 4.14. $\square$

In the following proposition we collect some equalities in case we consider an exponential martingale $t \mapsto e^{M(t) - \frac{1}{2}\langle M,M\rangle(t)}$ in case the process $t \mapsto \langle M,M\rangle(t)$ is deterministic.

4.15. PROPOSITION. *Let* $t \mapsto M(t)$, $0 \leqslant t \leqslant T$, *be a martingale on* $(\Omega, \mathcal{F}, \mathbb{P})$ *with the property that the variation process* $t \mapsto \langle M,M\rangle(t)$, $0 \leqslant t \leqslant T$, *is deterministic. The the following identities are true:*

$$\begin{aligned}
&\mathbb{E}\left[\int_{0<s_1<\cdots<s_{k_1}<t}\int dM(s_1)\dots dM(s_{k_1}) \cdot \int_{0<\rho_1<\cdots<\rho_{k_2}<t}\int dM(\rho_1)\dots dM(\rho_{k_2})\right] \\
&= \mathbb{E}\left[\frac{H_{k_1}(M(t),\langle M,M\rangle(t))}{k_1!}\frac{H_{k_2}(M(t),\langle M,M\rangle(t))}{k_2!}\right] \\
&= \frac{(\langle M,M\rangle(t))^{k_1}}{k_1!}\delta_{k_1,k_2}, \quad \textit{and} \qquad (4.67)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{0<s_1<\cdots<s_\ell<t}\int e^{M(s_1)-\frac{1}{2}\langle M,M\rangle(s_1)}\,dM(s_1)\dots dM(s_\ell)\right|^2\right] \\
&= \int_0^t e^{\langle M,M\rangle(s)}\frac{(\langle M,M\rangle(t) - \langle M,M\rangle(s))^{\ell-1}}{(\ell-1)!}\,d\langle M,M\rangle(s) \\
&= e^{\langle M,M\rangle(t)} - \sum_{j=0}^{\ell-1}\frac{(\langle M,M\rangle(t))^j}{j!}. \qquad (4.68)
\end{aligned}$$

PROOF. Let the predictable processes $s \mapsto F_1(s)$ and $s \mapsto F_2(s)$ be such that the quantities $\mathbb{E}\left[\int_0^T |F_1(s)|^2\,d\langle M,M\rangle(s)\right]$ and $\mathbb{E}\left[\int_0^T |F_2(s)|^2\,d\langle M,M\rangle(s)\right]$ are finite. Then we have

$$\mathbb{E}\left[\int_{t_1}^{t_2} F_1(s)\,dM(s)\cdot\int_{t_1}^{t_2} F_2(s)\,dM(s)\right] = \mathbb{E}\left[\int_{t_1}^{t_2} F_1(s)F_2(s)\,d\langle M,M\rangle(s)\right], \qquad (4.69)$$

for $0 \leqslant t_1 < t_2 \leqslant T$. By repeatedly employing the equality in (4.69) and using the fact that the process $s \mapsto \langle M,M\rangle(s)$ is deterministic we infer, for $1 \leqslant k_1 < k_2$, and $0 < t \leqslant T$, with $\ell = k_2 - k_1$,

$$\begin{aligned}
&\mathbb{E}\left[\int_{0<s_1<\cdots<s_{k_1}<t}\int dM(s_1)\dots dM(s_{k_1}) \cdot \int_{0<\rho_1<\cdots<\rho_{k_2}<t}\int dM(\rho_1)\dots dM(\rho_{k_2})\right] \\
&= \mathbb{E}\left[\int_{0<s_1<\cdots<s_{k_2}<t}\int dM(s_1)\dots dM(s_\ell)\;d\langle M,M\rangle(s_{\ell+1})\dots d\langle M,M\rangle(s_{k_2})\right]
\end{aligned}$$

$$= \mathbb{E}\left[\int_{0<s_1<\dots<s_\ell<t}\int \frac{\left(\langle M,M\rangle(t)-\langle M,M\rangle(s_\ell)\right)^{k_1}}{k_1!}\,dM(s_1)\dots dM(s_\ell)\right]=0. \tag{4.70}$$

If in (4.70) $k_1 = k_2$, and so $\ell = 0$, then we obtain

$$\mathbb{E}\left[\left|\int_{0<s_1<\dots<s_{k_1}<t}\int dM(s_1)\dots dM(s_{k_1})\right|^2\right] = \frac{\left(\langle M,M\rangle(t)\right)^{k_1}}{k_1!}. \tag{4.71}$$

The equalities in (4.70) and (4.71) show the equalities in (4.67). The proof of the equalities requires an induction argument. For $\ell = 1$ we have

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_0^t e^{M(s_1)-\frac{1}{2}\langle M,M\rangle(s_1)}\,dM(s_1)\right|^2\right]\\
&=\int_0^t \mathbb{E}\left[e^{2M(s)-\langle M,M\rangle(s)}\right]d\langle M,M\rangle(s)\\
&=\int_0^t \mathbb{E}\left[e^{2M(s)-\frac{1}{2}\langle 2M,2M\rangle(s)}\right]e^{\langle M,M\rangle(s)}\,d\langle M,M\rangle(s)\\
&=\int_0^t e^{\langle M,M\rangle(s)}\,d\langle M,M\rangle(s) = e^{\langle M,M\rangle(t)}-1.
\end{aligned} \tag{4.72}$$

The equalities in (4.72) imply those in (4.68) for $\ell = 1$. The second equality follows by partial integration and induction with respect to ℓ. The first equality in (4.68) can be obtained by an argument which is very similar to the proof of the equality in (4.67) with $k_1 = k_2 = \ell$. The details are left to the reader.

This completes the proof of Proposition 4.15. □

4.16. Corollary. *Let the hypotheses and notation be as in Proposition 4.14. Then*

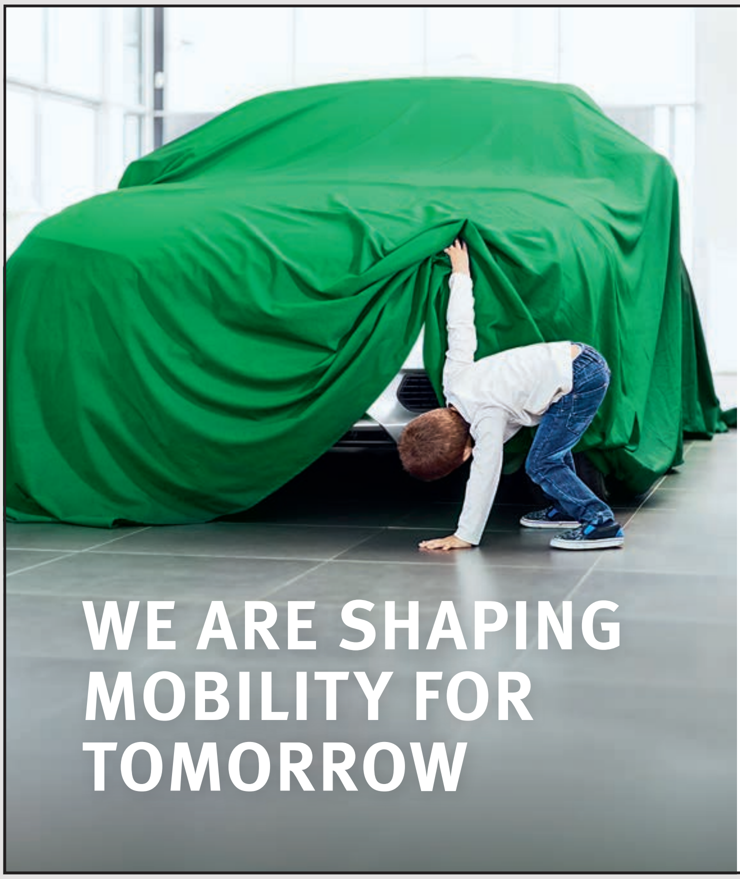
$$\lim_{\ell\to\infty}\tau^\ell\int_{0<s_1<\dots<s_\ell<t}\int e^{\tau M(s_1)-\frac{1}{2}\tau^2\langle M,M\rangle(s_1)}\,dM(s_1)\dots dM(s_\ell)=0, \tag{4.73}$$

$\mathbb{P}$-*almost surely. If the limit in (4.73) is in fact an* L^1*-limit, then the process* $t \mapsto e^{\tau M(t)-\frac{1}{2}\tau^2\langle M,M\rangle(t)}$ *is a martingale. In particular, it then follows that* $\mathbb{E}\left[e^{\tau M(t)-\frac{1}{2}\tau^2\langle M,M\rangle(t)}\right]=1$*; compare with the inequality in (4.41) and with Theorem 4.11.*

If the process $t \mapsto \langle M,M\rangle(t)$ *is real-valued and deterministic, then the limit in (4.73) is an* L^2*-limit, and so also an* L^1*-limit.*

Proof. Equality (4.73) in Corollary 4.16 follows from the equality in (4.59) in Proposition 4.13 with $x = M(t)$ and $a = \langle M,M\rangle(t)$ together with the equalities in (4.65) and (4.61). The assertion about the L^1-convergence also follows from these arguments. The only topic that requires some is the one about the situation where the process $t \mapsto \langle M,M\rangle(t)$ is deterministic. In this case the terms in the sum in (4.65) are orthogonal in $L^2(\Omega,\mathcal{F}_T,\mathbb{P})$, and this sum converges in L^2-sense to $e^{\tau M(t)-\frac{1}{2}\tau^2\langle M,M\rangle(t)}$. These assertions follow from the identities (4.67) and (4.68) in Proposition 4.15. This completes the proof of Corollary 4.16. ☐

The previous results, i.e. Proposition 4.14 and Corollary 4.16 are applicable if the martingale $M(t)$ is of the form $M(t) = \int_0^t h(s) \cdot dW(s)$, where $t \mapsto W(t)$ is standard Brownian motion. Then $\langle M, M \rangle (t) = \int_0^t |h(s)|^2 \, ds$. If $s \mapsto h(s)$ is deterministic, then in (4.73) we have L^2-convergence. These martingales play a role in the martingale representation theorem: see Theorem 4.21.



1.4. Weak solutions to stochastic differential equations. In the following theorem the symbols $\sigma_{i,j}$, $1 \leqslant i \leqslant d$, $1 \leqslant j \leqslant n$, and b_j, $1 \leqslant j \leqslant d$, stand for real-valued locally bounded Borel measurable functions defined on $[0,\infty) \times \mathbb{R}^d$. The matrix $(a_{i,j}(s,x))_{i,j=1}^d$ is defined by

$$a_{j,k}(s,x) = \sum_{k=1}^{n} \sigma_{i,k}(s,x)\sigma_{j,k}(s,x) = (\sigma(s,x)\sigma^*(s,x))_{i,j}\,.$$

For $s \geqslant 0$, the operator $L(s)$ is defined on $C^2\left(\mathbb{R}^d\right)$ with values in the space of locally bounded Borel measurable functions:

$$L(s)f(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}\left(s,x\right) D_iD_jf\left(x\right) + \sum_{j=1}^{d} b_j(s,x)D_jf(x), \quad f \in C^2\left(\mathbb{R}^d\right). \tag{4.74}$$

In Theorem 4.17 and its proof we use the following notation $E_N(s,x)$ is the orthogonal projection from $\mathbb{R}^n$ onto the zero-space of the matrix $\sigma(s,x)$ viewed as an operator from $\mathbb{R}^n$ to $\mathbb{R}^d$ and $E_R(s,x)$ is the orthogonal projection from $\mathbb{E}^n_R$ onto the range-space of the matrix $\sigma(s,x)$. Formally these operators can be written as

$$\begin{aligned} E_N(s,x) &= I - \sigma\left(s,x\right)^* \left(\sigma(s,x)\sigma(s,x)^*\right)^{-1}\sigma(s,x) \\ &= I - \lim_{\varepsilon\downarrow 0}\sigma\left(s,x\right)^* \left(\sigma(s,x)\sigma(s,x)^* + \varepsilon I\right)^{-1}\sigma(s,x), \quad \text{and} \\ E_R(s,x) &= \sigma\left(s,x\right)^* \left(\sigma(s,x)\sigma(s,x)^*\right)^{-1}\sigma(s,x) \\ &= \lim_{\varepsilon\downarrow 0}\sigma\left(s,x\right)^* \left(\sigma(s,x)\sigma(s,x)^* + \varepsilon I\right)^{-1}\sigma(s,x). \end{aligned}$$

The following theorem shows the close relationship between weak solutions and solutions to the martingale problem. A filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ satisfies the standard conditions if $\mathcal{F}_0$ contains the $\mathbb{P}$-null sets.

4.17. THEOREM. *Let* $(\Omega,\mathcal{F},\mathbb{P})$ *be a probability space with a right-continuous filtration* $(\mathcal{F}_t)_{t\geqslant 0}$ *satisfying the standard conditions and determined by a Brownian motion* $\{W(t) = (W_1(t),\dots,W_n(t)) : t \geqslant 0\}$ *on* $(\Omega,\mathcal{F},\mathbb{P})$. *Suppose that*

$$\{X(t) = (X_1(t),\dots,X_d(t)) : t \geqslant 0\}$$

is a d-dimensional continuous adapted process. Then the following assertions are equivalent:

(i) *For every* $f \in C^2\left(\mathbb{R}^d\right)$ *the process*

$$t \mapsto f\left(X(t)\right) - f\left(X(0)\right) - \int_0^t L(s)f\left(X(s)\right)\,ds \tag{4.75}$$

is a local martingale.

(ii) *The processes*

$$t \mapsto M_j(t) = X_j(t) - \int_0^t b_j\left(s,X(s)\right)\,ds,\ t \geqslant 0, \quad 1 \leqslant j \leqslant d, \tag{4.76}$$

are local martingales with covariation processes

$$t \mapsto \langle M_i, M_j \rangle (t) = \int_0^t a_{i,j}\left(s, X(s)\right) ds, \ t \geqslant 0, \quad 1 \leqslant i,\, j \leqslant d. \tag{4.77}$$

(iii) *There exists a Brownian motion* $\{B(t) : t \geqslant 0\}$ *starting at 0 such that*

$$X(t) = X(0) + \int_0^t b\left(s, X(s)\right) ds + \int_0^t \sigma\left(s, X(s)\right) dB(s), \quad t \geqslant 0. \tag{4.78}$$

Notice that the covariation of the processes $t \mapsto \int_0^t \sigma\left(s, X(s)\right) dW(s)$ and of $t \mapsto \int_0^t E_N\left(s, X(s)\right) dW(s)$ vanish. The same is true for the processes $t \mapsto \int_0^t E_R\left(s, X(s)\right) dW(s)$ and $t \mapsto \int_0^t E_N\left(s, X(s)\right) dW(s)$. Observe that the covariation of the processes $t \mapsto M(t)$ and $t \mapsto \int_0^t E_N\left(s, X(s)\right) dW(s)$ vanish if the Brownian motion $t \mapsto W(t)$ is independent of the process $t \mapsto M(t)$. This can be accomplished by enlarging the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with an independent copy $\left(\Omega^W, \mathcal{F}^W, \mathbb{P}^W\right)$. On the probability space $\left(\Omega \times \Omega^W, \mathcal{F} \otimes \mathcal{F}^W, \mathbb{P} \times \mathbb{P}^W\right)$ the new variables $\widetilde{X}(t)$ and $\widetilde{W}(t)$ are given by

$$\widetilde{X}(t)\left(\omega, \omega'\right) = X(t)(\omega) \ \text{ and } \ \widetilde{W}(t)\left(\omega, \omega'\right) = W(t)\left(\omega'\right), \ \left(\omega, \omega'\right) \in \Omega \times \Omega^W$$

respectively.

Examples of (Feller) semigroups can be manufactured by taking a continuous function $\varphi : [0, \infty) \times E \to E$ with the property that $\varphi\left(s + t, x\right) = \varphi\left(t, \varphi\left(s, x\right)\right)$, for all s, $t \geqslant 0$ and $x \in E$. Then the mappings $f \mapsto P(t)f$, with $P(t)f(x) = f\left(\varphi\left(t, x\right)\right)$ defines a semigroup. It is a Feller semigroup if $\lim_{x \to \triangle} \varphi\left(t, x\right) = \triangle$. An explicit example of such a function, which does not provide a Feller-Dynkin semigroup on $C_0\left(\mathbb{R}\right)$ is given by $\varphi(t, x) = \dfrac{x}{\sqrt{1 + \frac{1}{2}tx^2}}$ (example due to V. Kolokoltsov [**96**], and [**95**]). Put $u(t, x) = P(t)f(x) = f\left(\varphi(t, x)\right)$. Then $\dfrac{\partial u}{\partial t}(t, x) = -x^3 \dfrac{\partial u}{\partial x}(t, x)$. In fact this (counter-)example shows that solutions to the martingale problem do not necessarily give rise to Feller-Dynkin semigroups. These are semigroups which preserve not only the continuity, but also the fact that functions which tend to zero at $\triangle$ are mapped to functions with the same property. However, for Feller semigroups we only require that continuous functions with values in $[0, 1]$ are mapped to continuous functions with the same properties. Therefore, it is not needed to include a hypothesis like (4.79) which reads as follows: for every $(\tau, s, t, x) \in [0, T]^3 \times E$, $\tau < s < t$, the following equality holds:

$$\mathbb{P}_{\tau,x}\left[X(t) \in E\right] = \mathbb{P}_{\tau,x}\left[X(t) \in E,\ X(s) \in E\right]. \tag{4.79}$$

Nadirashvili [**128**] constructs an elliptic operator in a bounded open domain $U \subset \mathbb{R}^d$ with a regular boundary such that the martingale problem is not uniquely solvable. More precisely the result reads as follows. Consider an elliptic operator $L = \sum_{j,k=1}^{d} a_{j,k}^2 \frac{\partial^2}{\partial x_j \partial x_k}$, where $a_{j,k} = a_{j,k}$ are measurable functions on $\mathbb{R}^d$ such that

$$c^{-1} |\xi|^2 \leqslant \sum_{j,k=1}^{d} a_{j,k}\xi_j\xi_k \leqslant c\,|\xi|^2 , \quad \xi \in \mathbb{R}^d,$$

for some ellipticity constant $c \geqslant 1$. There exists a diffusion $(X(t), \mathbb{P}_x)$ corresponding to the operator L which can be defined as a solution to the martingale problem $\mathbb{P}\left[X(0) = x\right] = 1$, $f\left(X(t)\right) - f\left(X(0)\right) - \int_0^t f\left(X(s)\right) ds$ is a $\mathbb{P}_x$-martingale for all $f \in C^2\left(\mathbb{R}^d\right)$. Nadirashvili is interested in non-uniqueness in the above martingale problem and in non-uniqueness of solutions to the Dirichlet problem $Lu = 0$ in Ω, the unit ball in $\mathbb{R}^d$, $u = g$ on $\partial\Omega$, where $\Omega \subset \mathbb{R}^d$ is a bounded domain with smooth boundary and $g \in C^2\left(\partial\Omega\right)$. In particular, so-called good solutions u to the Dirichlet problem are investigated. A good solution is a function u which is the limit of a subsequence of solutions u_n, $n \in \mathbb{N}$, to the equation $L^n u_n = \sum_{j,k=1}^{d} a_{j,k}^n \frac{\partial^2 u_n}{\partial x_j \partial x_k} = 0$ in Ω, $u_n = g$ on $\partial\Omega$, where the operators L^n are elliptic with smooth coefficients $a_{j,k}^n$ and a common ellipticity constant c such that $a_{j,k}^n \to a_{j,k}$ almost everywhere in Ω as $n \to \infty$. The main result is the following theorem: There exists an elliptic operator L of the above form defined in the unit ball $B_1 \subset \mathbb{R}^d$, $d \geqslant 3$, and there is a function $g \in C^2\left(\partial B_1\right)$ such that the formulated Dirichlet problem has at least two good solutions. An immediate consequence is non-uniqueness of solutions to the corresponding martingale problem.

The following corollary easily follows from Theorem 4.17. It establishes a close relationship between unique weak solutions to stochastic differential equations and unique solutions to the martingale problem. For the precise notion of "unique weak solutions" see Definition 4.19 below. This result should also be compared with Proposition 3.43, where the connection with (strong) Markov processes is explained.

4.18. Corollary. *Let the notation and hypotheses be as in Theorem 4.17. Put* $\Omega = C\left([0,\infty), \mathbb{R}^d\right)$, *and* $X(t)(\omega) = \omega(t)$, $t \geqslant 0$, $\omega \in \Omega$. *Fix* $x \in \mathbb{R}^d$. *Then the following assertions are equivalent:*

(i) *There exists a unique probability measure* $\mathbb{P}$ *on* $\mathcal{F}$ *such that the process*

$$f\left(X(t)\right) - f\left(X(0)\right) - \int_0^t L(s) f\left(X(s)\right) ds$$

is a $\mathbb{P}$*-martingale for all* C^2*-functions* f *with compact support, and such that* $\mathbb{P}\left[X(0) = x\right] = 1$.

(ii) *The stochastic integral equation*

$$X(t) = x + \int_0^t \sigma\left(s, X(s)\right) dB(s) + \int_0^t b\left(s, X(s)\right) ds \tag{4.80}$$

has unique weak solutions.

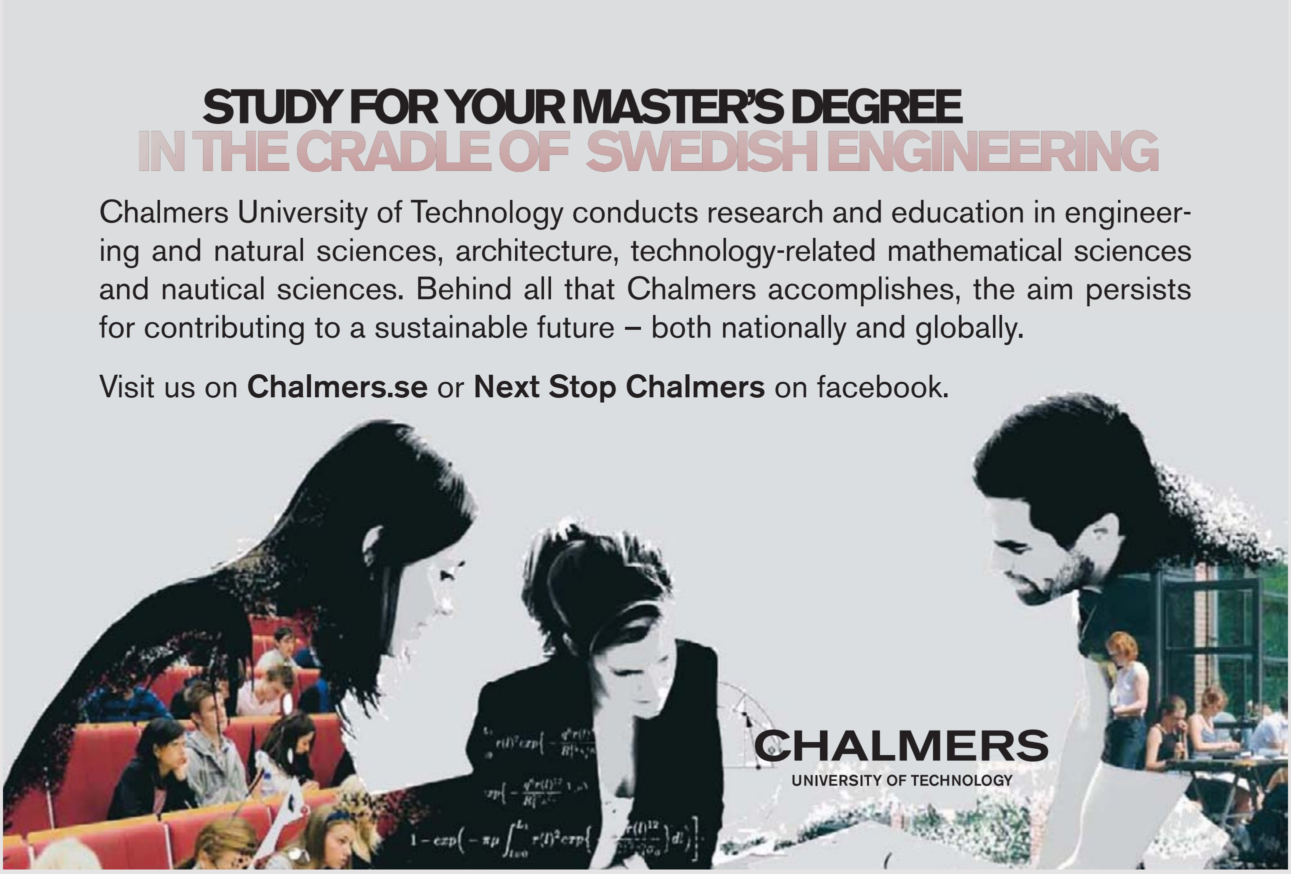
4.19. Definition. The equation in (4.80) is said to have unique weak solutions on the interval $[0, T]$, also called unique distributional solutions, provided that the finite-dimensional distributions of the process $X(t)$, $\leqslant t \leqslant T$, which satisfy (4.80) do not depend on the particular Brownian motion $B(t)$ which occurs in (4.80). This is the case if and only if for any pair of Brownian motions

$$\{(B(t) : T \geqslant t \geqslant 0), (\Omega, \mathcal{F}, \mathbb{P})\} \quad \text{and} \quad \{(B'(t) : T \geqslant t \geqslant 0), (\Omega', \mathcal{F}, \mathbb{P}')\}$$

and any pair of adapted processes $\{X(t) : T \geqslant t \geqslant 0\}$ and $\{X'(t) : T \geqslant t \geqslant 0\}$ for which

$$\begin{aligned} X(t) &= x + \int_0^t \sigma(s, X(s))\, dB(s) + \int_0^t b(s, X(s))\, ds \quad \text{and} \\ X'(t) &= x + \int_0^t \sigma(s, X'(s))\, dB'(s) + \int_0^t b(s, X'(s))\, ds \end{aligned}$$

the finite-dimensional distributions of the process $\{X(t) : T \geqslant t \geqslant 0\}$ relative to $\mathbb{P}$ coincide with the finite-dimensional distributions of $\{X'(t) : T \geqslant t \geqslant 0\}$ relative to $\mathbb{P}'$.



Proof of Theorem 4.17. (i) $\Longrightarrow$ (ii) With $f_j(x_1, \dots, x_d) = x_j$, $1 \leqslant j \leqslant d$, assertion (i) implies that the process

$$M_j(t) = X_j(t) - \int_0^t b_j(s, X(s))\, ds = f_j(X(t)) - \int_0^t L(s) f_j(X(s))\, ds \tag{4.81}$$

is a local martingale. We will show that the processes

$$\left\{ M_i(t) M_j(t) - \int_0^t a_{i,j}(s, X(s))\, ds : \, t \geqslant 0 \right\}, \quad 1 \leqslant i,\, j \leqslant d,$$

are local martingales as well. To this end fix $1 \leqslant i,\, j \leqslant d$, and define the function $f_{i,j} : \mathbb{R}^d \to \mathbb{R}$ by $f_{i,j}(x_1, \dots, x_d) = x_i x_j$. From (i) it follows that the process

$$\left\{ X_i(t) X_j(t) - \int_0^t \left(a_{i,j}(s, X(s)) + b_i(s, X(s)) X_j(s) + b_j(s, X(s)) X_i(s) \right) ds \right\}$$

is a local martingale. For brevity we write

$$\begin{aligned}
\alpha_{i,j}(s) &= a_{i,j}(s, X(s)), \quad \beta_j(s) = b_j(s, X(s)), \quad \beta_i(s) = b_i(s, X(s)), \\
M_i(s) &= X_i(s) - \int_0^s \beta_i(\tau)\, d\tau, \quad M_j(s) = X_i(s) - \int_0^s \beta_j(\tau)\, d\tau, \\
M_{i,j}(s) &= X_i(s) X_j(s) - \int_0^s \left(\beta_i(\tau) X_j(\tau) + \beta_j(\tau) X_i(\tau) + \alpha_{i,j}(\tau) \right) d\tau.
\end{aligned} \tag{4.82}$$

Then the processes M_i and $M_{i,j}$ are local martingales. Moreover, we have

$$\begin{aligned}
&\left(M_i(t) + \int_0^t \beta_i(s)\, ds \right) \left(M_j(t) + \int_0^t \beta_j(s)\, ds \right) = X_i(t) X_j(t) \\
&= \int_0^t \left(\beta_i(\tau) X_j(\tau) + \beta_j(\tau) X_i(\tau) + \alpha_{i,j}(\tau) \right) d\tau + M_{i,j}(t) \\
&= \int_0^t \left(\beta_i(\tau) \left(X_j(\tau) - M_j(\tau) \right) + \beta_j(\tau) \left(X_i(\tau) - M_i(\tau) \right) + \alpha_{i,j}(\tau) \right) d\tau \\
&\qquad + \int_0^t \left(\beta_i(\tau) M_j(\tau) + \beta_j(\tau) M_i(\tau) \right) d\tau + M_{i,j}(t) \\
&= \int_0^t \beta_i(\tau) \left(X_j(\tau) - M_j(\tau) \right) d\tau + \int_0^t \beta_j(\tau) \left(X_i(\tau) - M_i(\tau) \right) d\tau \\
&\qquad + \int_0^t \alpha_{i,j}(\tau)\, d\tau + \int_0^t \left(\beta_i(\tau) M_j(\tau) + \beta_j(\tau) M_i(\tau) \right) d\tau + M_{i,j}(t) \\
&= \int_0^t \beta_i(\tau) \int_0^\tau \beta_j(s)\, ds\, d\tau + \int_0^t \beta_j(\tau) \int_0^\tau \beta_i(s)\, ds\, d\tau \\
&\qquad + \int_0^t \alpha_{i,j}(\tau)\, d\tau + \int_0^t \left(\beta_i(\tau) M_j(\tau) + \beta_j(\tau) M_i(\tau) \right) d\tau + M_{i,j}(t) \\
&= \iint_{0<s<\tau<t} \beta_i(\tau) \beta_j(s)\, d\tau\, ds + \iint_{0<\tau<s<t} \beta_i(\tau) \beta_j(s)\, d\tau\, ds \\
&\qquad + \int_0^t \alpha_{i,j}(\tau)\, d\tau + \int_0^t \left(\beta_i(\tau) M_j(\tau) + \beta_j(\tau) M_i(\tau) \right) d\tau + M_{i,j}(t)
\end{aligned}$$

$$\begin{aligned}&= \int_0^t \beta_i(\tau)\,d\tau \int_0^t \beta_j(s)\,ds + \int_0^t \alpha_{i,j}(s)\,ds + M_{i,j}(t)\\ &\quad + \int_0^t \left(\beta_i(s)M_j(s) + \beta_j(s)M_i(s)\right)\,ds. \end{aligned} \tag{4.83}$$

Consequently, from (4.83) we see

$$\begin{aligned}&M_i(t)M_j(t) - \int_0^t \alpha_{i,j}(s)\,ds\\ &= M_{i,j}(t) - \int_0^t \left(\beta_i(s)\left(M_j(t) - M_j(s)\right) + \beta_j(s)\left(M_i(t) - M_i(s)\right)\right)\,ds. \end{aligned} \tag{4.84}$$

It is readily verified that the processes

$$\begin{aligned}\int_0^t \beta_i(s)\left(M_j(t) - M_j(s)\right)\,ds &= \int_0^t \int_0^\tau \beta_i(s)\,ds\,dM_j(\tau) \quad \text{and}\\ \int_0^t \beta_j(s)\left(M_i(t) - M_i(s)\right)\,ds &= \int_0^t \int_0^\tau \beta_j(s)\,ds\,dM_i(\tau)\end{aligned}$$

are local martingales. It follows that the process

$$\left\{M_i(t)M_j(t) - \int_0^t \alpha_{i,j}(s)\,ds : t \geqslant 0\right\}$$

is a local martingale. So that the covariation process $\langle M_i, M_j\rangle$ is given by $\langle M_i, M_j\rangle(t) = \int_0^t \alpha_{i,j}(s)\,ds$.

(ii) $\Longrightarrow$ (iii) This implication follows from an application of Theorem 4.8 with $\Phi_{i,j}(t) = a_{i,j}(t, X(t))$ and $\chi(t) = \sigma(t, X(t))^* \left(\sigma(t, X(t))\,\sigma(t, X(t))^*\right)^{-1}$, and where $B(t) = \int_0^t \chi(s)\,dM(s)$, provided that null-space of the matrix $\sigma(t, X(t))$ is $\{0\}$. Observe that then $\langle B_{j_1}, B_{j_2}\rangle(t) = t\delta_{j_1,j_2}$, because

$$\chi(t)\Phi(t)\chi(t)^* = \sigma(t, X(t))^* \left(\sigma(t, X(t))\,\sigma(t, X(t))^*\right)^{-1} \sigma(t, X(t)) = I.$$

In addition, the process $t \mapsto B(t)$ is a Brownian motion and for $t > 0$ we have

$$\begin{aligned}&X(t) - X(0) - \int_0^t b(s, X(s))\,ds - \int_0^t \sigma(s, X(s))\,dB(s)\\ &= X(t) - X(0) - \int_0^t b(s, X(s))\,ds - \int_0^t \sigma(s, X(s))\,\chi(s)\,dM(s)\\ &= X(t) - X(0) - \int_0^t b(s, X(s))\,ds - \int_0^t dM(s)\\ &= X(t) - X(0) - \int_0^t b(s, X(s))\,ds - M(t) + M(0) = 0, \end{aligned} \tag{4.85}$$

where in the very final step of (4.85) we invoked the equalities in (4.76).

If the matrix process $\sigma(t, X(t))\,\sigma^*(t, X(t))$ is not invertible, then we proceed as follows. We have a martingale $M(s)$, $0 \leqslant s \leqslant t$, on $(\Omega, \mathcal{F}_t, \mathbb{P})$ with the properties of assertion (ii). We introduce the matrix processes $\widetilde{\psi}_\varepsilon(s)$, $\varepsilon > 0$, $E_R(s)$, and $E_N(s)$ as follows

$$\begin{aligned}
\widetilde{\psi}_\varepsilon(s) &= \sigma^*\left(s, X(s)\right)\left(\sigma\left(s, X(s)\right)\sigma^*\left(s, X(s)\right) + \varepsilon I\right)^{-1} \\
E_R(s) &= \lim_{\varepsilon\downarrow 0} \sigma^*\left(s, X(s)\right)\left(\sigma\left(s, X(s)\right)\sigma^*\left(s, X(s)\right) + \varepsilon I\right)^{-1}\sigma\left(s, X(s)\right), \quad \text{and} \\
E_N(s) &= I - E_R(s).
\end{aligned}$$

The matrix $E_R(s)$ can be considered as an orthogonal projection on the range of the matrix $\sigma^*\left(s, X(s)\right)\sigma\left(s, X(s)\right)$, and $E_N(s)$ as an orthogonal projection on its null space. More precisely,

$$E_R(s)\sigma^*\left(s, X(s)\right) = \sigma^*\left(s, X(s)\right), \quad \text{and} \quad \sigma\left(s, X(s)\right)E_N(s) = 0.$$

The matrix process $V_N(\tau)$ is to be determined, but it satisfies $V_N(\tau)V_N^*(\tau) = E_N(\tau) = E_N(\tau)V_N(\tau)V_N^*(\tau)E_N^*(\tau)$. In terms of these processes we define the following process:

$$\begin{aligned}
B(s) &= \lim_{\varepsilon\downarrow 0}\int_0^s \widetilde{\psi}_\varepsilon(\tau)\,dM(\tau) + \int_0^s V_N(\tau)\,dW(\tau) \\
&= \int_0^s V_R(\tau)\,dW(\tau) + \int_0^s V_N(\tau)\,dW(\tau).
\end{aligned} \tag{4.86}$$

Here, we applied the martingale representation theorem to get the matrix process $V_R(\tau)$: see Theorem 4.22. It will turn out that

$$V_R(\tau)V_R^*(\tau) = E_R(\tau) = E_R(\tau)V_R(\tau)V_R^*(\tau)E_R(\tau).$$

Next we will prove that the process $s \mapsto B(s)$ is a Brownian motion, and that $M(s) - M(0) = \int_0^s \sigma\left(\tau, X(\tau)\right)\,dB(\tau)$. Put

$$B_\varepsilon(s) = \int_0^s \widetilde{\psi}_\varepsilon(\tau)\,dM(\tau) + \int_0^s V_N(\tau)\,dW(\tau). \tag{4.87}$$

Then the restriction of the matrix process $V_R^*(\tau)$ to the range of the orthogonal projection $E_R^*(\tau)$ is extended to a linear mapping $V^*(\tau)$ to the whole space $\mathbb{R}^n$ in such a way that $V^*(\tau)E_R(\tau)x = V_R^*(\tau)x$, $x \in \mathbb{R}^n$, and such that $V_N^*(\tau) := V^*(\tau)E_N(\tau)$ is an orthogonal transformation from the range of $E_N(\tau)$ onto the range of $I - V_R^*(\tau)E_R(\tau)V_R(\tau)$ which is the null-space of $E_R(\tau)V_R(\tau)$. If we define $V(\tau) = E_R(\tau)V_R(\tau)$ on the range of $V_R^*(\tau)E_R(\tau)$ and as the adjoint $V_N(\tau)$ of $V_N^*(\tau)$ on the range of $I - V_R^*(\tau)E_R(\tau)V_R(\tau)$. Then we see that $V(\tau)V^*(\tau) = I$. Such an orthogonal map can be obtained via a Gramm-Schmidt procedure applied the finite sequences $\left(I - V_R^*(\tau)E_R(\tau)V_R(\tau)\right)e_k$, and $E_N(\tau)e_k$ $1 \leqslant k \leqslant n$, where the vectors e_k, $1 \leqslant k \leqslant n$, form an ortho-normal basis for $\mathbb{R}^n$. In addition we get

$$E_R(\tau)V_R(\tau)V_N^*(\tau)E_N(\tau) = E_R(\tau)V(\tau)V^*(\tau)E_N(\tau) = E_R(\tau)E_N(\tau) = 0. \tag{4.88}$$

Moreover, we have

$$\begin{aligned}
&\left\langle B_{\varepsilon,j_1}, B_{\varepsilon,j_2}\right\rangle(s) \\
&= \sum_{k_1,k_2=1}^{d}\sum_{\ell=1}^{n}\int_0^s \widetilde{\psi}_{\varepsilon,j_1,k_1}(\tau)\widetilde{\psi}_{\varepsilon,j_1,k_1}(\tau)\sigma_{k_1,\ell}\left(\tau, X(\tau)\right)\sigma_{k_2,\ell}\left(\tau, X(\tau)\right)\,d\tau
\end{aligned}$$

$$+\sum_{k=1}^{n}\sum_{k_1=1}^{d}\int_0^s \widetilde{\psi}_{\varepsilon,j_1,k_1}(\tau)V_{N,j_2,k}(\tau)\,d\left\langle M_{k_1},W_k\right\rangle(\tau)$$
$$+\sum_{k=1}^{n}\sum_{k_1=1}^{d}\int_0^s \widetilde{\psi}_{\varepsilon,j_2,k_1}(\tau)V_{N,j_1,k}(\tau)\,d\left\langle M_{k_1},W_k\right\rangle(\tau)$$
$$+\sum_{k=1}^{n}\int_0^s V_{N,j_1,k}(\tau)V_{N,j_2,k}(\tau)\,d\tau$$

(employ the equality $M(s)-M(0)=\int_0^s\sigma\left(\tau,X(\tau)\right)\,dB(\tau)$)

$$=\int_0^s\left(\widetilde{\psi}_{\varepsilon}(\tau)\sigma\left(\tau,X(\tau)\right)\sigma^*\left(\tau,X(\tau)\right)\widetilde{\psi}_{\varepsilon}^*(\tau)\right)_{j_1,j_2}\,d\tau+\int_0^s\left(V_N(\tau)V_N^*(\tau)\right)_{j_1,j_2}\,d\tau$$
$$+\sum_{k=1}^{n}\sum_{k_1=1}^{d}\sum_{\ell=1}^{d}\int_0^s \widetilde{\psi}_{\varepsilon,j_1,k_1}(\tau)V_{N,j_2,k}(\tau)\sigma_{k_1,\ell}\left(\tau,X(\tau)\right)\,d\left\langle B_\ell,W_k\right\rangle(\tau)$$
$$+\sum_{k=1}^{n}\sum_{k_1=1}^{d}\sum_{\ell=1}^{d}\int_0^s \widetilde{\psi}_{\varepsilon,j_2,k_1}(\tau)V_{N,j_1,k}(\tau)\sigma_{k_1,\ell}\left(\tau,X(\tau)\right)\,d\left\langle B_\ell,W_k\right\rangle(\tau). \tag{4.89}$$

Notice that for $\varepsilon\downarrow 0$ the first term in (4.89) converges to the covariation of the processes

$$\left(\int_0^s E_R(\tau)\,dB(\tau)\right)_{j_1}=\left(\int_0^\tau V_R(\tau)\,dW(\tau)\right)_{j_1}$$
$$=\left(\int_0^\tau E_R(\tau)V_R(\tau)\,dW(\tau)\right)_{j_1},\quad\text{and}$$
$$\left(\int_0^s E_R(\tau)\,dB(\tau)\right)_{j_2}=\left(\int_0^\tau V_R(\tau)\,dW(\tau)\right)_{j_2}.$$

On the one-hand this covariation is equal to $\int_0^s (E_R(\tau))_{j_1,j_2}\, d\tau$, but on the other hand, by (4.86), it is also equal to $\int_0^\tau (V_R(\tau)V_R^*(\tau))_{j_1,j_2}\, d\tau$. As a consequence we infer $(E_R(\tau))_{j_1,j_2} = (V_R(\tau)V_R^*(\tau))$, and hence $E_R(\tau) = (V_R(\tau)V_R^*(\tau))_{j_1,j_2}$. So the procedure which we explained above can be carried out to obtain the matrix process $\tau \mapsto V(\tau)$ which satisfies, e.g., $V(\tau)V^*(\tau) = I$. From (4.89) we infer by continuity and the definition of $E_R(\tau)$ that

$$\begin{aligned}
\langle B_{j_1}, B_{j_2}\rangle (s) &= \lim_{\varepsilon\downarrow 0} \langle B_{\varepsilon,j_1}, B_{\varepsilon,j_2}\rangle (s) \\
&= \int_0^s (E_R(\tau)E_R^*(\tau))_{j_1,j_2}\, d\tau + \int_0^s (E_N(\tau)E_N^*(\tau))_{j_1,j_2}\, d\tau \\
&\quad + \left\langle \left(\int_0^{(\cdot)} E_R(\tau)\, dB(\tau)\right)_{j_1}, \left(\int_0^{(\cdot)} V_N(\tau)\, dW(\tau)\right)_{j_2} \right\rangle (s) \\
&\quad + \left\langle \left(\int_0^{(\cdot)} E_R(\tau)\, dB(\tau)\right)_{j_2}, \left(\int_0^{(\cdot)} V_N(\tau)\, dW(\tau)\right)_{j_1} \right\rangle (s) \\
&= \int_0^s (E_R(\tau)E_R^*(\tau))_{j_1,j_2}\, d\tau + \int_0^s (E_N(\tau)E_N^*(\tau))_{j_1,j_2}\, d\tau \\
&\quad + \left\langle \left(\int_0^{(\cdot)} E_R(\tau)V_R(\tau)\, dW(\tau)\right)_{j_1}, \left(\int_0^{(\cdot)} V_N(\tau)\, dW(\tau)\right)_{j_2} \right\rangle (s) \\
&\quad + \left\langle \left(\int_0^{(\cdot)} E_R(\tau)V_R(\tau)\, dW(\tau)\right)_{j_2}, \left(\int_0^{(\cdot)} V_N(\tau)\, dW(\tau)\right)_{j_1} \right\rangle (s)
\end{aligned}$$

(the covariations of the processes $\int_0^t V_R(\tau)\, dW(\tau)$ and $\int_0^t V_N(\tau)\, dW(\tau)$ vanish: see (4.88))

$$= \int_0^s (E_R(\tau)E_R^*(\tau) + E_N(\tau)E_N^*(\tau))_{j_1,j_2}\, d\tau = \delta_{j_1,j_2} s. \tag{4.90}$$

In the final step of (4.90) we used the fact that the processes $E_R(\tau)$ and $E_N(\tau)$ are orthogonal projections such that $E_R(\tau) + E_N(\tau) = I$. From Lévy's theorem 4.5 it follows that the process $s \mapsto B(s)$, $0 \leqslant s \leqslant t$, is a Brownian motion.

In order to finish the proof of the implication (ii) $\Longrightarrow$ (iii) we still have to prove the equality $M(s) - M(0) = \int_0^s \sigma(\tau, X(\tau))\, dB(\tau)$ and to construct the matrix process $V_N(\tau)$ in such a way that the covariation of the processes $M(t)$ and $\int_0^t V_N(\tau)\, dW(\tau)$ vanishes. For brevity we write $\sigma(\tau) = \sigma(\tau, X(\tau))$. Then by definition and standard calculations with martingales we obtain:

$$\begin{aligned}
&M(s) - M(0) - \int_0^s \sigma(\tau)\, dB_\varepsilon(\tau) \\
&= M(s) - M(0) - \int_0^s \sigma(\tau)\widetilde{\psi}_\varepsilon(\tau)\, dM(\tau) - \int_0^s \sigma(\tau)\, V_N(\tau)\, dW(\tau) \\
&= \int_0^s \left(I - \sigma(\tau)\sigma^*(\tau)\left(\sigma(\tau)\sigma^*(\tau) + \varepsilon I\right)^{-1}\right) dM(\tau)
\end{aligned}$$

$$= \varepsilon \int_0^s \left(\sigma(\tau)\sigma^*(\tau) + \varepsilon I\right)^{-1} dM(\tau). \tag{4.91}$$

From (4.91) together with the fact that covariation process of the local martingale $M(s)$ is given by $\int_0^s \sigma(\tau)\sigma^*(\tau)\, d\tau$, it follows that the covariation matrix of the local martingale

$$M(s) - \int_0^s \sigma\left(\tau\right)\, dB_\varepsilon(\tau)$$

is given by

$$\varepsilon^2 \int_0^s \left(\sigma(\tau)\sigma^*(\tau) + \varepsilon I\right)^{-1} \sigma(\tau)\sigma^*(\tau) \left(\sigma(\tau)\sigma^*(\tau) + \varepsilon I\right)^{-1}\, d\tau. \tag{4.92}$$

In addition, we have in spectral sense:

$$0 \leqslant \varepsilon^2 \left(\sigma(\tau)\sigma^*(\tau) + \varepsilon I\right)^{-1} \sigma(\tau)\sigma^*(\tau) \left(\sigma(\tau)\sigma^*(\tau) + \varepsilon I\right)^{-1} \leqslant \frac{\varepsilon}{4} I, \tag{4.93}$$

and thus in L^2-sense we have

$$M(s) - M(0) - \int_0^s \sigma(\tau)\, dB(\tau) = L^2\text{-}\lim_{\varepsilon\downarrow 0} \left(M(s) - M(0) - \int_0^s \sigma(\tau) B_\varepsilon(\tau)\right) = 0. \tag{4.94}$$

The equality in (4.94) completes the proof of the implication (ii) $\longrightarrow$ (iii).

(iii) $\Longrightarrow$ (i) Let $f : \mathbb{R}^d \to \mathbb{R}$ be a twice continuously differentiable function. By Itô's lemma we get

$$\begin{aligned}
& f\left(X(t)\right) - f\left(X(0)\right) - \int_0^t L(s) f\left(X(s)\right)\, ds \\
&= \int_0^t \nabla f\left(X(s)\right) \cdot dX(s) + \frac{1}{2}\sum_{i,j=1}^d \int_0^t D_i D_j f\left(X(s)\right) d\left\langle X_i, X_j\right\rangle(s) \\
&\quad - \int_0^t L(s) f\left(X(s)\right) ds \\
&= \sum_{j=1}^d \int_0^t b_j\left(s, X(s)\right) D_j f\left(X(s)\right)\, ds \\
&\quad + \frac{1}{2}\sum_{i,j=1}^d \sum_{k=1}^d \int_0^t \sigma_{i,k}\left(s, X(s)\right) \sigma_{j,k}\left(s, X(s)\right) D_i D_j f\left(X(s)\right)\, ds \\
&\quad + \int_0^t \nabla f\left(X(s)\right) \sigma\left(s, X(s)\right) dB(s) - \int_0^t L(s) f\left(X(s)\right) ds \\
&= \int_0^t \nabla f\left(X(s)\right) \sigma\left(s, X(s)\right)\, dB(s).
\end{aligned} \tag{4.95}$$

The final expression in (4.95) is a local martingale. Hence (iii) implies (i).

This completes the proof of Theorem 4.8. ☐

4.20. Remark. The implication (ii) $\Longrightarrow$ (i) in Theorem 4.17 can also be proved directly by using Itô calculus. Suppose that the local martingales $t \mapsto M_j(t)$,

$1 \leqslant j \leqslant d$, are defined as in assertion (ii) with covariation processes as in (4.77). Let f be a C^2-function defined on $\mathbb{R}^d$. Then we have:

$$\begin{aligned}
& f\left(X(t)\right) - f\left(X(0)\right) - \int_0^t L(s) f\left(X(s)\right)\, ds \\
&= \int_0^t \nabla f\left(X(s)\right)\, dX(s) + \frac{1}{2}\sum_{i,j=1}^d \int_0^t D_i D_j f\left(X(s)\right)\, d\left\langle X_i, X_j\right\rangle(s) \\
&\quad - \int_0^t L(s) f\left(X(s)\right)\, ds \\
&= \int_0^t \nabla f\left(X(s)\right)\, dM(s) + \int_0^t \nabla f\left(X(s)\right) b\left(s, X(s)\right)\, ds \\
&\quad + \frac{1}{2}\sum_{i,j=1}^d \int_0^t D_i D_j f\left(X(s)\right)\, d\left\langle M_i, M_j\right\rangle(s) - \int_0^t L(s) f\left(X(s)\right)\, ds \\
&= \int_0^t \nabla f\left(X(s)\right)\, dM(s) + \int_0^t \nabla f\left(X(s)\right) b\left(s, X(s)\right)\, ds \\
&\quad + \frac{1}{2}\sum_{i,j=1}^d \int_0^t D_i D_j f\left(X(s)\right) a_{i,j}\left(s, X(s)\right)\, ds - \int_0^t L(s) f\left(X(s)\right)\, ds \\
&= \int_0^t \nabla f\left(X(s)\right)\, dM(s). \qquad (4.96)
\end{aligned}$$

Assertion (i) is a consequence of equality (4.96).

2. A martingale representation theorem

In this section we formulate and prove the martingale theorem based on an n-dimensional Brownian motion. Proofs are, essentially speaking, taken from [**138**]. Let $(W(s))_{0\leqslant s<\infty}$ be standard Brownian motion in $\mathbb{R}^n$, and let $\mathcal{F}_t$ be the σ-field generated by $(W(s))_{0\leqslant s\leqslant t}$ augmented with the $\mathbb{P}$-null sets. For $h \in L^{\infty}([0,T];\mathbb{R}^n)$ we write $X_h(t) := e^{\int_0^t h(s)\cdot dW(s)-\frac{1}{2}\int_0^t |h(s)|^2\,ds}$.

4.21. Theorem. *Let Ψ_T be the subspace of $L^2(\Omega,\mathcal{F}_T,\mathbb{P})$ spanned by the exponentials $X_h(T) := e^{\int_0^T h(s)\cdot dW(s)-\frac{1}{2}\int_0^T |h(s)|^2\,ds}$, $h \in L^{\infty}_{\text{simple}}([0,T];\mathbb{R}^n)$. Then Ψ_T is dense in the space $L^2(\Omega,\mathcal{F}_T,\mathbb{P})$.*

In Theorem 4.21 the space $L^{\infty}_{\text{simple}}([0,T];\mathbb{R}^n)$ consists of those $\mathbb{R}^n$-valued functions $h \in L^{\infty}([0,T];\mathbb{R}^n)$ which can be written in the form

$$h(s) = \sum_{k=1}^{N} \mathbf{1}_{(t_{k-1},t_k]}(s)\left(\sum_{j=k}^{N}\lambda_j\right) = \sum_{j=1}^{N}\mathbf{1}_{(0,t_j]}(s)\lambda_j,\ 0\leqslant s\leqslant T,\ N\in\mathbb{N}, \tag{4.97}$$

where, for any $N \in \mathbb{N}$, $0 = t_0 < t_1 < \cdots < t_N = T$ is an arbitrary partition of the interval $[0,T]$, and where $(\lambda_j)_{1\leqslant j\leqslant N}$ are arbitrary vectors in $\mathbb{R}^n$. Observe that, for such functions h, $\int_0^T h(s)\cdot dW(s) = \sum_{j=1}^N \lambda_j\cdot W(t_j)$. Also notice that, by Itô's lemma, $X_h(T) = 1 + \int_0^T X_h(s)h(s)\cdot dW(s)$, $h \in L^{\infty}([0,T];\mathbb{R}^n)$. In the proof of Theorem 4.21 the following notation is employed. The symbol $C_0^{\infty}\left(\mathbb{R}^{n\times N}\right)$ stands for the vector space of those C^{∞}-functions φ defined on all real $n\times N$ matrices λ with the property that all functions of the form

$$\lambda \mapsto \left(1+\|\lambda\|_{\text{HS}}^2\right)^m D_{j,k}^{\alpha_{j,k}}\varphi(\lambda),\quad m\in\mathbb{N},\ 1\leqslant j\leqslant N,\ 1\leqslant k\leqslant n,\ \alpha_{j,k}\in\mathbb{N},$$

are bounded. Here $D_{j,k}^{\alpha_{j,k}}$ stands for the derivative of order $\alpha_{j,k}$ relative to the variable $\lambda_{j,k}$. The symbol $\|\lambda\|_{\text{HS}}$ stands for the Hilbert-Schmidt norm of the matrix λ; that is

$$\|\lambda\|_{\text{HS}}^2 = \sum_{j=1}^{N}\sum_{k=1}^{n}|\lambda_{j,k}|^2,\ \lambda = (\lambda_{j,k})_{1\leqslant j\leqslant N,1\leqslant k\leqslant n}.$$

Functions of the form $\lambda \mapsto \exp\left(-\frac{1}{2}\|\lambda\|_{\text{HS}}^2\right)$ belong to the space $C_0^{\infty}\left(\mathbb{R}^{n\times N}\right)$. Observe that $C_0^{\infty}\left(\mathbb{R}^{n\times N}\right)$ constitutes a dense subspace of $C_0\left(\mathbb{R}^{n\times N}\right)$, i.e. the space of complex-valued continuous functions which tend to 0 at ∞ equipped with the supremum norm.

Proof of Theorem 4.21. This statement is true if there exists no $g \in L^2(\Omega,\mathcal{F}_T,\mathbb{P})$, which is perpendicular to all $X(T) \in \Psi_T$. We start by assuming that there is a $g \in L^2(\Omega,\mathcal{F}_T,\mathbb{P})$ such that g is orthogonal to all variables $X(T) \in \Psi_T$. This orthogonality means that $\mathbb{E}[X_h(T)g] = 0$, for all $h \in L^{\infty}_{\text{simple}}([0,T];\mathbb{R}^n)$. Or, what is the same,

$$\int_{\Omega} e^{\int_0^T h(s)\cdot dW(s)-\frac{1}{2}\int_0^T h(s)^2\,ds} g\,d\mathbb{P} = 0,\quad \text{for all } h\in L^{\infty}_{\text{simple}}([0,T];\mathbb{R}^n). \tag{4.98}$$

The equalities in (4.98) are equivalent to

$$e^{-\frac{1}{2}\int_0^T h(s)^2\,ds}\int_\Omega e^{\int_0^T h(s)\cdot dW(s)(w)} g\,d\mathbb{P} = 0, \quad \text{for all } h \in L^\infty_{\text{simple}}\left([0,T]\,;\mathbb{R}^n\right),$$

which amounts to

$$\int_\Omega e^{\int_0^T h(s)\cdot dW(s)} g\,d\mathbb{P} = 0, \quad \text{for all } h \in L^\infty_{\text{simple}}\left([0,T]\,;\mathbb{R}^n\right), \tag{4.99}$$

By taking h as in (4.97), we see that for all $\lambda = (\lambda_1,\dots,\lambda_N) \in (\mathbb{R}^n)^N = \mathbb{R}^{n\times N}$ and for all $(t_1,\dots,t_N) \in [0,T]^N$ with $0 = t_0 < t_1 < \cdots < t_N = T$, the following equality holds: $\int_\Omega e^{\sum_{j=1}^N \lambda_j\cdot W(t_j)} g\,d\mathbb{P} = 0$. Next, put $G(\lambda) = \int_\Omega e^{\sum_{j=1}^N \lambda_j\cdot W(t_j)} g\,d\mathbb{P}$. The function $\lambda \mapsto G(\lambda)$ is real analytic on $\mathbb{R}^{n\times N}$, and thus has an analytic extension to the complex space $\mathbb{C}^{n\times N}$: $G(z) := \int_\Omega e^{\sum_{j=1}^N z_j\cdot W(t_j)} g\,d\mathbb{P}$ for all $z \in \mathbb{C}^{n\times N}$. Here $z_j \cdot W_j(t) = \sum_{k=1}^n z_{j,k}W_k(t)$, $z_j = (z_{1,k},\dots,z_{n,k}) \in \mathbb{C}^n$. Since $G(\lambda) = 0$ for $\lambda \in \mathbb{R}^{n\times N}$, it follows that $G(z) = 0$ for $z \in \mathbb{C}^{n\times N}$. However, for $\varphi \in C_0^\infty\left(\mathbb{R}^{n\times N}\right)$, and with

$$\widehat{\varphi}\left(y_1,\dots,y_N\right) = \underbrace{\int_{\mathbb{R}^n}\dots\int_{\mathbb{R}^n}}_{N\ \text{times}} e^{-i\sum_{j=1}^N y_j\cdot x_j}\varphi\left(x_1,\dots,x_N\right)\,dx_N\dots dx_1,$$

where $(y_1,\dots,y_N) \in \mathbb{R}^{n\times N}$, we see that

$$\mathbb{E}\left[\varphi\left(W\left(t_1\right),\dots,W\left(t_N\right)\right)g\right] = \int_\Omega \varphi\left(W\left(t_1\right),\dots,W\left(t_N\right)\right)g\,d\mathbb{P}$$

(inverse Fourier transform)

$$= \int_\Omega \left(\frac{1}{(2\pi)^{n\times N}}\int_{\mathbb{R}^{n\times N}} e^{i\sum_{j=1}^N W(t_j)\cdot y_j}\widehat{\varphi}(y)\,dy\right) g\,d\mathbb{P}$$

(Fubini's theorem)

$$\begin{aligned} &= \frac{1}{(2\pi)^{n\times N}}\int_{\mathbb{R}^{n\times N}}\int_\Omega e^{i\sum_{j=1}^N W(t_j)\cdot y_j} g\,d\mathbb{P}\widehat{\varphi}(y)\,dy \\ &= \frac{1}{(2\pi)^{n\times N}}\int_{\mathbb{R}^{n\times N}} G(iy)g\widehat{\varphi}(y)\,dy = 0. \end{aligned} \tag{4.100}$$

From the monotone class theorem, and the fact the space $C_0^\infty\left(\mathbb{R}^{n\times N}\right)$ is dense in $C_0\left(\mathbb{R}^{n\times N}\right)$ for the uniform topology, it follows that the equality in (4.100), i.e. the equality

$$\mathbb{E}\left[\varphi\left(W\left(t_1\right),\dots,W\left(t_N\right)\right)g\right] = 0 \tag{4.101}$$

can only be true for all $\varphi \in C_0^\infty\left(\mathbb{R}^{n\times N}\right)$, and for all $(t_1,\dots,t_N) \in (0,\infty)^N$, $0 < t_1 < \cdots < t_N = T$, for all $N \in \mathbb{N}$, provided that $\mathbb{E}\left[Fg\right] = 0$ for all bounded $\mathcal{F}_T$-measurable random variables F. Consequently, $\mathbb{E}\left[X_h(T)g\right] = 0$ for all $h \in L^\infty_{\text{simple}}\left([0,T]\,;\mathbb{R}^n\right)$ if and only if the random variable $g \in L^2\left(\Omega,\mathcal{F}_T,\mathbb{P}\right)$ is identically 0. This completes the proof of Theorem 4.21. ☐

The following theorem is known as the Itô representation theorem.

4.22. Theorem. *If the random variable* $X(T)$ *belongs to* $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, *then there exists a unique predictable* $\mathbb{R}^n$*-valued process* $t \mapsto F(t)$, $0 \leqslant t \leqslant T$, *for which* $\int_0^T \mathbb{E}\left[|F(s)|^2\right] ds < \infty$ *and which is such that*

$$X(T) = \mathbb{E}\left[X(T)\right] + \int_0^T F(s) \cdot dW(s). \tag{4.102}$$

In other words the space

$$\mathbb{C} + \left\{\int_0^T F(s) \cdot dW(s) : \, s \mapsto F(s) \text{ predictable and } \int_0^T \mathbb{E}\left[|F(s)|^2\right] ds < \infty\right\}$$

coincides with $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.

Proof of Theorem 4.22. Let $X(T)$ be as in Theorem 4.22. Then there exists double sequences $\left((\alpha_{j,k})_{j=1}^{N_k}\right)_{k\in\mathbb{N}}$ in $\mathbb{C}$ and $\left((h_{j,k})_{j=1}^{N_k}\right)_{k\in\mathbb{N}}$ of elements in $L^\infty_{\text{simple}}([0,T]; \mathbb{R}^n)$ such that, with $F_k(t) = \sum_{j=1}^{N_k} \alpha_{j,k} X_{h_{j,k}}(t) h_{j,k}(t)$ and with

$$\begin{aligned} X_k(T) &:= \sum_{j=1}^{N_k} \alpha_{j,k} X_{h_{j,k}}(T) = \sum_{j=1}^{N_k} \alpha_{j,k} e^{\int_0^T h_{j,k}(s)\cdot dW(s) - \frac{1}{2}\int_0^T \left|h_{j,k}(s)\right|^2 ds} \\ &= \sum_{j=1}^{N_k} \alpha_{j,k} + \int_0^T \sum_{j=1}^{N_k} \alpha_{j,k} X_{h_{j,k}}(t) h_{j,k}(t) \cdot dW(t) \\ &= \mathbb{E}\left[X_k(T)\right] + \int_0^T F_k(t) \cdot dW(t), \end{aligned} \tag{4.103}$$

we have

$$\begin{aligned} 0 &= \lim_{k\to\infty} \mathbb{E}\left[|X(T) - X_k(T)|^2\right] = \lim_{k,\ell\to\infty} \mathbb{E}\left[|X_\ell(T) - X_k(T)|^2\right] \\ &= \lim_{k,\ell\to\infty} \left\{ \left|\mathbb{E}\left[X_\ell(T) - X_k(T)\right]\right|^2 + \int_0^T \mathbb{E}\left[|F_\ell(t) - F_k(t)|^2\right] dt \right\}. \end{aligned} \tag{4.104}$$

In the third equality in (4.103) we employed Itô's lemma. From (4.104) we infer that $\mathbb{E}\left[X(T)\right] = \lim_{k\to\infty} \mathbb{E}\left[X_k(T)\right]$ and that there exists a predictable process $t \mapsto F(t)$ such that

$$\lim_{k\to\infty} \int_0^T \mathbb{E}\left[|F(t) - F_k(t)|^2\right] = 0. \tag{4.105}$$

From (4.105) we obtain

$$L^2\text{-}\lim_{k\to\infty} \int_0^T F_k(t) \cdot dW(t) = \int_0^T F(t) \cdot dW(t). \tag{4.106}$$

A combination of (4.103), (4.104) and (4.106) yields the equality in (4.102). In addition, we have $\int_0^T \mathbb{E}\left[|F(s)|^2\right] ds < \infty$, and so the existence part in Theorem 4.22 has been established now. The uniqueness part follows from the Itô isometry

$$\mathbb{E}\left[\int_0^T |F_2(s) - F_1(s)|^2 \, ds\right] = \mathbb{E}\left[\left|\int_0^T (F_2(s) - F_1(s)) \cdot dW(s)\right|^2\right] = 0,$$

if $X(T) - \mathbb{E}[X(T)] = \int_0^T F_1(s) \cdot dW(s) = \int_0^T F_2(s) \cdot dW(s)$. Altogether this completes the proof of Theorem 4.22. $\square$

Next we formulate and prove the martingale representation theorem.

4.23. THEOREM. *Let* $(M(t))_{0\leqslant t\leqslant T}$ *belong to* $\mathcal{M}^2(\Omega, \mathcal{F}_T, \mathbb{P})$. *Then there exists a unique predictable* $\mathbb{R}^n$*-valued process* $t \mapsto \zeta(t) = (\zeta_1(t), \dots, \zeta_n(t))$, $\zeta(t) \in L^2(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^n)$, *such that*

$$M(t) = M(0) + \int_0^t \zeta(s) \cdot dW(s) = M(0) + \sum_{j=1}^n \int_0^t \zeta_j(s)\, dW_j(s). \tag{4.107}$$

Of course, if in Theorem 4.23 the process $t \mapsto M(t)$, $0 \leqslant t \leqslant T$, is a martingale vector in $\mathbb{R}^d$, then we obtain a predictable matrix $Z(t) \in \mathbb{R}^{n\times d}$ such that $M(t) = M(0) + \int Z(s)^*\, dW(s)$. (This is what one needs in the context of Backward Stochastic Differential Equations or BSDEs for short.)

Proof of Theorem 4.23. Let $(M(t))_{0\leqslant t\leqslant T}$ be as in Theorem 4.23. Theorem 4.22 yields the existence of a unique predictable $\mathbb{R}^n$-valued process $t \mapsto \zeta(t) = (\zeta_1(t), \dots, \zeta_n(t))$, $\zeta(t) \in L^2(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^n)$, such that

$$M(T) = \mathbb{E}\left[M(T)\right] + \int_0^T \zeta(s)\cdot dW(s) = \mathbb{E}\left[M(T)\right] + \sum_{j=1}^n \int_0^T \zeta_j(s)\, dW_j(s). \quad (4.108)$$

Since the processes $t \mapsto M(t)$ and $t \mapsto \int_0^t \zeta_j(s)\, dW_j(s)$, $1 \leqslant j \leqslant n$, $0 \leqslant t \leqslant T$, are martingales, from (4.108) we infer

$$\begin{aligned} M(t) &= \mathbb{E}\left[M(T) \mid \mathcal{F}_t\right] = \mathbb{E}\left[\mathbb{E}\left[M(T)\right] + \int_0^T \zeta(s)\cdot dW(s) \mid \mathcal{F}_t\right] \\ &= \mathbb{E}\left[M(T)\right] + \int_0^t \zeta(s)\cdot dW(s) = \mathbb{E}\left[M(0)\right] + \int_0^t \zeta(s)\cdot dW(s) \quad (4.109) \end{aligned}$$

From (4.109) we get $M(0) = \mathbb{E}\left[M(0)\right]$, and so the representation in (4.107) follows from (4.109). The proof of Theorem 4.23 is complete now. ☐

3. Girsanov transformation

In this section we want to discuss the Cameron-Martin-Girsanov transformation or just Girsanov transformation. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $(\mathcal{F}_t)_{t\geqslant 0}$. In addition, let the process $\{B(t) : t \geqslant 0\}$ be a d-dimensional Brownian motion. Let b_j, c_j, $\sigma_{i,j}$ be Borel measurable locally bounded functions on $\mathbb{R}^d$. Suppose that the stochastic differential equation

$$X(t) = x + \int_0^t \sigma\left(s, X(s)\right) dB(s) + \int_0^t b\left(s, X(s)\right)\, ds \quad (4.110)$$

has unique weak solutions. For a precise definition of the notion of "unique weak solutions" see Definition 4.19. For more information on transformations of measures on Wiener space see e.g. Üstünel and Zakai [**177**]. In particular these observations mean that if in equation (4.111) below (for the process $Y(t)$) the process $B'(t)$ is a Brownian motion relative to a probability measure $\mathbb{P}'$, then the $\mathbb{P}'$-distribution of the process $Y(t)$ coincides with the $\mathbb{P}$-distribution of the process $X(t)$ which satisfies (4.110). Next we will elaborate on this item. Suppose that the process $t \mapsto Y(t)$ satisfies the equation (in column vectors):

$$\begin{aligned} Y(t) &= x + \int_0^t \sigma\left(s, Y(s)\right)\, dB(s) + \int_0^t \left(b\left(s, Y(s)\right) + \sigma\left(s, Y(s)\right) c\left(s, Y(s)\right)\right)\, ds \\ &= x + \int_0^t \sigma\left(s, Y(s)\right) dB'(s) + \int_0^t b\left(s, Y(s)\right) ds, \quad (4.111) \end{aligned}$$

where $B'(t) = B(t) + \int_0^t c\left(s, Y(s)\right)\, ds$. The following proposition says that relative to a martingale transformation $\mathbb{P}'$ of the measure $\mathbb{P}$ (Girsanov or Cameron-Martin transformation) the process $t \mapsto B'(t)$ is a $\mathbb{P}'$-Brownian motion. More precisely, we introduce the local martingale $M'(t)$ and the corresponding martingale measure $\mathbb{P}'$ by

$$M'(t) = \exp\left(-\int_0^t c\left(s, Y(s)\right) dB(s) - \frac{1}{2}\int_0^t \left|c\left(s, Y(s)\right)\right|^2 \, ds\right) \quad \text{and} \tag{4.112}$$

$$\mathbb{P}'\left[A\right] = \mathbb{E}\left[M'(t)\mathbf{1}_A\right], \quad A \in \mathcal{F}_t. \tag{4.113}$$

We also need the process $Z'(t)$ defined by

$$Z'(t) = -\int_0^t c\left(s, Y(s)\right) dB(s) - \frac{1}{2}\int_0^t \left|c\left(s, Y(s)\right)\right|^2 \, ds. \tag{4.114}$$

In addition, we have a need for a vector function $c_1(t, y)$ satisfying $c(t, y) = c_1(t, y)\sigma(t, y)$. Here $c(t, y)$ and $c_1(t, y)$ are considered as row vectors. This is equivalent to saying that the column vector $c(t, y)^*$ belongs to the range of $\sigma\left(s, y\right)^*$. More precisely, put

$$P(s, y) = \sigma\left(s, y\right)^* \left(\sigma\left(s, y\right)\left(\sigma\left(s, y\right)\right)^*\right)^{-1} \sigma\left(s, y\right). \tag{4.115}$$

Observe that $P(s, y)$ is well defined: it is the orthogonal projection from $\mathbb{R}^d$ on the range of $\sigma\left(s, y\right)^*$. Then the row vector $c(s, y)$ can be written in the form

$$c\left(s, y\right) = c\left(s, y\right) P\left(s, y\right) = c_1\left(s, y\right) \sigma\left(s, y\right), \tag{4.116}$$

where $c_1(s, y) = c(s, y)\sigma\left(s, y\right)^* \left(\sigma\left(s, y\right)\left(\sigma\left(s, y\right)\right)^*\right)^{-1}$. Also notice that the range of $\sigma(s, y)^*$ coincides with $\mathbb{R}^d$ provided the zero-space of $\sigma(s, y)$ is trivial. We assume that such a vector function $c_1(t, y)$ exists.

4.24. Proposition. *Suppose that the process $Y(t)$ satisfies the equation in (4.111). Let the processes $M'(t)$ and $Z'(t)$ be defined by (4.112) and (4.114) respectively. Then the following assertions are true:*

1. *The process $t \mapsto M'(t)$ is a local $\mathbb{P}$-martingale. It is a martingale provided $\mathbb{E}\left[M'(t)\right] = 1$ for all $t \geqslant 0$.*
2. *Fix $t > 0$. The variable $M'(t)$ only depends on the process $s \mapsto Y(s)$, $0 \leqslant s \leqslant t$.*
3. *Suppose that the process $t \mapsto M'(t)$ is a $\mathbb{P}$-martingale, and not just a local $\mathbb{P}$-martingale. Then $\mathbb{P}'$ can be considered as a probability measure on the σ-field generated by $\cup_{t>0}\mathcal{F}_t$.*
4. *Suppose that the process $t \mapsto M'(t)$ is a $\mathbb{P}$-martingale. Then the process $t \mapsto B'(t)$ is a Brownian motion relative to $\mathbb{P}'$.*

Proof. (1) From Itô calculus we get

$$M'(t) - M'(0) = -\int_0^t M'(s)c\left(s, Y(s)\right) dB(s),$$

and hence assertion 1 follows, because stochastic integrals with respect to Brownian motion are local martingales. Next we choose a sequence of stopping times τ_n which increase to ∞ $\mathbb{P}$-almost surely, and which are such that the processes $t \mapsto M'\left(t \wedge \tau_n\right)$ are genuine martingales. Then we see $\mathbb{E}\left[M'\left(t \wedge \tau_n\right)\right] = 1$ for all $n \in \mathbb{N}$ and $t \geqslant 0$. Fix $t_2 > t_1$. Since the processes $t \mapsto M'(t \wedge \tau_n)$, $n \in \mathbb{N}$, are $\mathbb{P}$-martingales, we see that

$$\mathbb{E}\left[M'(t_2 \wedge \tau_n) \,\middle|\, \mathcal{F}_{t_1}\right] = M'(t_1 \wedge \tau_n) \quad \mathbb{P}\text{-almost surely.} \tag{4.117}$$

In (4.117) we let $n \to \infty$, and apply Scheffé's theorem to conclude that

$$\mathbb{E}\left[M'(t_2) \,\middle|\, \mathcal{F}_{t_1}\right] = M'(t_1) \quad \mathbb{P}\text{-almost surely.} \tag{4.118}$$

The equality in (4.118) shows that the process $t \mapsto M'(t)$ is a $\mathbb{P}$-martingale provided that $\mathbb{E}[M'(t)] = 1$ for all $t \geqslant 0$. This completes the proof of assertion (1).

(2) This assertion follows from the following calculation:

$$\begin{aligned} Z'(t) &= -\int_0^t c(s, Y(s))\, dB(s) - \frac{1}{2}\int_0^t |c(s, Y(s))|^2\, ds \qquad (4.119)\\ &= -\int_0^t c(s, Y(s))\, dB'(s) + \frac{1}{2}\int_0^t |c(s, Y(s))|^2\, ds \end{aligned}$$

(the equality $c(s,y) = c_1(s,y)\sigma(s,y)$ holds)

$$\begin{aligned} &= -\int_0^t c_1(s, Y(s))\,\sigma(s, Y(s))\, dB'(s) + \frac{1}{2}\int_0^t |c(s, Y(s))|^2\, ds \\ &= -\int_0^t c_1(s, Y(s))\, d\left(\int_0^s \sigma(\tau, Y(\tau))\, dB'(\tau)\right) + \frac{1}{2}\int_0^t |c(s, Y(s))|^2\, ds \\ &= -\int_0^t c_1(s, Y(s))\, d\left(Y(s) - \int_0^s b(\tau, Y(\tau))\, d\tau\right) + \frac{1}{2}\int_0^t |c(s, Y(s))|^2\, ds. \end{aligned}$$

From (4.119), (4.112), and (4.114) it is plain that $M'(t)$ only depends on the path $\{Y(s) : 0 \leqslant s \leqslant t\}$.

(3) This assertion is a consequence of Kolmogorov's extension theorem. The measure is $\mathbb{P}'$ is well defined on $\cup_{t>0}\mathcal{F}_t$. Here we use the martingale property. By Kolmogorov's extension theorem, it extends to the σ-field generated by this union.

(4) The equality $B'(t) = B(t) + \int_0^t c\left(s, Y(s)\right)\, ds$ entails the following equality for the quadratic covariation of the processes B_i' and B_j':

$$\left\langle B_i', B_j'\right\rangle (t) = \left\langle B_i, B_j\right\rangle (t) = t\delta_{i,j}. \tag{4.120}$$

From Itô calculus we also infer

$$\begin{aligned}
&M'(t)B_i'(t)\\
&= \int_0^t M'(s)B_i'(s)\, dZ'(s) + \int_0^t M'(s)\, dB_i'(s)\\
&\quad + \frac{1}{2}\int_0^t M'(s)B'(s)\, d\left\langle Z', Z'\right\rangle (s) + \int_0^t M'(s) d\left\langle Z', B_i'\right\rangle (s)\\
&= -\int_0^t M'(s)B_i'(s)c\left(s, Y(s)\right)\, dB(s) - \frac{1}{2}\int_0^t M'(s)B_i'(s) \left|c\left(s, Y(s)\right)\right|^2 ds\\
&\quad + \frac{1}{2}\int_0^t M'(s)B_i'(s) \left|c\left(s, Y(s)\right)\right|^2 ds + \int_0^t M'(s)\, dB_i(s)\\
&\quad + \int_0^t M'(s)c_i\left(s, Y(s)\right)\, ds - \int_0^t M'(s)c_i\left(s, Y(s)\right)\, ds\\
&= -\int_0^t M'(s)B_i'(s)c\left(s, Y(s)\right)\, dB(s) + \int_0^t M'(s)\, dB_i(s).
\end{aligned} \tag{4.121}$$

Upon invoking Theorem 4.5 and employing (4.120) and (4.121) assertion (4) follows.

This concludes the proof of Proposition 4.24. ☐

4.25. Remark. Proposition 4.24 is still valid even if the null-space of $\sigma(s, y)$ is non-trivial, and so the range of $\sigma(s, y)^*$ is not the whole space $\mathbb{R}^d$, provided we replace the (local) martingale $t \mapsto M'(t)$ with the local martingale:

$$t \mapsto \exp\left(-\int_0^t c\left(s, Y(s)\right) P\left(s, Y(s)\right)\, dB(s) - \frac{1}{2}\int_0^t \left|c\left(s, Y(s)\right) P\left(s, Y(s)\right)\right|^2\, ds\right), \tag{4.122}$$

where $P(s, y)$ is as in (4.115).

Let the process $X(t)$ solve the equation in (4.110), and put

$$M(t) = \exp\left(\int_0^t c\left(s, X(s)\right) dB(s) - \frac{1}{2}\int_0^t \left|c\left(s, X(s)\right)\right|^2\, ds\right), \tag{4.123}$$

and assume that the process $M(t)$ is not merely a local martingale, but a genuine $\mathbb{P}$-martingale.

4.26. Theorem. *Fix $T > 0$, and let the functions*

$$b(s, y), \quad \sigma(s, y), \quad c(s, y), \quad \textit{and} \quad c_1(s, y), \quad 0 \leqslant s \leqslant T,$$

be locally bounded Borel measurable vector or matrix functions such that $c(s,y) = c_1(s,y)\sigma(s,y)$, $0 \leqslant s \leqslant T$, $y \in \mathbb{R}^d$. *Suppose that the equation in (4.110) possesses unique weak solutions on the interval* $[0,T]$.

Uniqueness. *If weak solutions to the stochastic differential equation in (4.111) exist, then they are unique in the sense as explained next. In fact, let the couple* $(Y(s), B(s))$, $0 \leqslant s \leqslant t$, *be a solution to the equation in (4.111) with the property that the local martingale* $M'(t)$ *given by*

$$M'(t) = \exp\left(-\int_0^t c\left(s, Y(s)\right) dB(s) - \frac{1}{2}\int_0^t \left|c\left(s, Y(s)\right)\right|^2 ds\right). \tag{4.124}$$

satisfies $\mathbb{E}\left[M'(t)\right] = 1$. *Then the finite-dimensional distributions of the process* $Y(s)$, $0 \leqslant s \leqslant t$, *are given by the Girsanov or Cameron-Martin transform:*

$$\mathbb{E}\left[f\left(Y\left(t_1\right), \dots, Y\left(t_n\right)\right)\right] = \mathbb{E}\left[M(t) f\left(X\left(t_1\right), \dots, X\left(t_n\right)\right)\right], \tag{4.125}$$

$t \geqslant t_n > \cdots > t_1 \geqslant 0$, *where* $f : \mathbb{R}^d \times \cdots \times \mathbb{R}^d \to \mathbb{R}$ *is an arbitrary bounded Borel measurable function.*

Existence. *Conversely, let the process* $s \mapsto (X(s), B(s))$ *be a solution to the equation in (4.110). Suppose that the local martingale* $s \mapsto M(s)$, *defined by*

$$M(s) = \exp\left(\int_0^s c\left(\tau, X(\tau)\right) dB(\tau) - \frac{1}{2}\int_0^s \left|c\left(\tau, X(\tau)\right)\right|^2 d\tau\right), \quad 0 \leqslant s \leqslant t, \tag{4.126}$$

is a martingale, i.e. $\mathbb{E}\left[M(t)\right] = 1$. *Then there exists a couple* $\left(\widetilde{Y}(s), \widetilde{B}(s)\right)$, $0 \leqslant s \leqslant t$, *where* $s \mapsto \widetilde{B}(s)$, $0 \leqslant s \leqslant t$, *is a Brownian motion on a probability space* $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ *such that*

$$\begin{aligned}\widetilde{Y}(s) = x &+ \int_0^s \sigma\left(\tau, \widetilde{Y}(\tau)\right) d\widetilde{B}(\tau) + \int_0^s \sigma\left(\tau, \widetilde{Y}(\tau)\right) c\left(\tau, \widetilde{Y}(\tau)\right) d\tau \\ &+ \int_0^s b\left(\tau, \widetilde{Y}(\tau)\right) d\tau,\end{aligned} \tag{4.127}$$

and such that

$$\widetilde{\mathbb{E}}\left[\exp\left(-\int_0^t c\left(s, \widetilde{Y}(s)\right) d\widetilde{B}(s) - \frac{1}{2}\int_0^t \left|c\left(s, \widetilde{Y}(s)\right)\right|^2 ds\right)\right] = 1. \tag{4.128}$$

4.27. Remark. The formula in (4.125) is known as the Girsanov transform or Cameron-Martin transform of the measure $\mathbb{P}$. It is a martingale measure. Suppose that the process $t \mapsto M'(t)$, as defined in (4.112) is a $\mathbb{P}$-martingale. Then the proof of Theorem 4.26 shows that the process $t \mapsto M(t)$, as defined in (4.123) is a $\mathbb{P}$-martingale. By assertion (1) in Proposition 4.24 the process $t \mapsto M'(t)$ is a $\mathbb{P}$-martingale if and only if $\mathbb{E}\left[M'(t)\right] = 1$ for all $T \geqslant t \geqslant 0$, and a similar statement holds for the process $t \mapsto M(t)$. If the process $t \mapsto M'(t)$ is a martingale, then taking $G \equiv \mathbf{1}$ in (4.142) shows that $\mathbb{E}\left[M(t)\right] = 1$, and hence by 1 in Proposition 4.24 the process $t \mapsto M(t)$ is a $\mathbb{P}$-martingale. Conversely, if the process $t \mapsto M(t)$ is a $\mathbb{P}$-martingale, then we reverse the implications in the proof of Theorem 4.26 and take $F \equiv \mathbf{1}$ in (4.146) to conclude that $\mathbb{E}\left[M'(t)\right] = 1$ for all $\geqslant 0$. But then the process $t \mapsto M'(t)$ is a $\mathbb{P}$-martingale.

4.28. Remark. Let $P(s,y)$ be as in (4.115). Put $c_P(s,y) = c(s,y)\,P(s,y)$. Like Proposition 4.24 Theorem 4.26 remains valid even if the null-space of $\sigma(s,y)$ is non-trivial, provided we replace the (local) martingale $t \mapsto M'(t)$ with the local martingale in (4.122) and we replace the local martingale $t \mapsto M(t)$ with

$$t \mapsto \exp\left(\int_0^t c_P\left(s, X(s)\right)\, dB(s) - \frac{1}{2}\int_0^t \left|c_P\left(s, X(s)\right)\right|^2 ds\right). \tag{4.129}$$

In addition the equality in (4.128) should be replaced by

$$\widetilde{\mathbb{E}}\left[\exp\left(-\int_0^t c_P\left(s, \widetilde{Y}(s)\right)\, d\widetilde{B}(s) - \frac{1}{2}\int_0^t \left|c_P\left(s, \widetilde{Y}(s)\right)\right|^2 ds\right)\right] = 1. \tag{4.130}$$

Observe that $c_P(s,y)$ is automatically of the form $c_P(s,y) = c_1(s,y)\sigma\left(s,y\right)$: see the equality in (4.116) and see Remark 4.25 as well.

Notice that the process $t \mapsto M(t)$ is a $\mathbb{P}$-martingale provided Novikov's condition is satisfied, i.e. if $\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t \left|c\left(s, X(s)\right)\right|^2 ds\right)\right] < \infty$. For a precise formulation see Corollary 4.29 below. Define

$$\mathcal{E}(M)(t) = e^{M(t) - \frac{1}{2}\langle M, M\rangle(t)}. \tag{4.131}$$

4.29. Corollary. *If* $\sup_{t\geqslant 0}\mathbb{E}\left[\exp\left(\frac{1}{2}\left\langle M,M\right\rangle(t)\right)\right]<\infty$, *then*

$$\mathbb{E}\left[\exp\left(M(\infty)-\frac{1}{2}\left\langle M,M\right\rangle(\infty)\right)\right]=1,$$

and consequently the process $t\mapsto\mathcal{E}(M)(t)$ *is a* $\mathbb{P}$*-martingale relative to the filtration* $(\mathcal{F}_t)_{t\geqslant 0}$, *where* $\mathcal{F}_t=\sigma\left(M(s):0\leqslant s\leqslant t\right)$, *the* σ*-field generated by the variables* $M(s)$, $0\leqslant s\leqslant t$.

Novikov's result is a consequence of results in [**100**]; see Chapter 1 of [**184**]. Observe that $M(\infty)=\lim_{t\to\infty}M(t)$ exists $\mathbb{P}$-almost surely.

4.30. Remark. Let $s\mapsto c(s)$ be a process which is adapted to Brownian motion in $\mathbb{R}^d$, and let $\rho>0$ be such that Novikov's condition is satisfied: $\mathbb{E}\left[\exp\left(\frac{1}{2}\rho^2\int_0^t\left|c(s)\right|^2ds\right)\right]<\infty$. From assertion 4 in Proposition 4.24 and Theorem 4.26 we see that the following identity holds for all bounded Borel measurable functions F defined on $\left(\mathbb{R}^d\right)^n$:

$$\begin{aligned}&\mathbb{E}\left[F\left(Y_\rho\left(t_1\right),\dots,Y_\rho\left(t_n\right)\right)\right]\\&=\mathbb{E}\left[\exp\left(\rho\int_0^t c(s)dB(s)-\frac{1}{2}\rho^2\int_0^t\left|c(s)\right|^2ds\right)F\left(B\left(t_1\right),\dots,B\left(t_n\right)\right)\right]\end{aligned}\tag{4.132}$$

where $0\leqslant t_1<\dots<t_n\leqslant t$, and $Y_\rho(\tau)=B(\tau)+\rho\int_0^\tau c(s)\,ds$, $0\leqslant\tau\leqslant t$. In particular, if $n=1$ we get

$$\begin{aligned}&\mathbb{E}\left[F\left(B(t)+\rho\int_0^t c(s)\,ds\right)\right]\\&=\mathbb{E}\left[\exp\left(\rho\int_0^t c(s)dB(s)-\frac{1}{2}\rho^2\int_0^t\left|c(s)\right|^2ds\right)F\left(B\left(t\right)\right)\right]\end{aligned}\tag{4.133}$$

Assume that the gradient DF of the function F exists and is bounded. The equality in (4.133) can be differentiated with respect to ρ to obtain:

$$\begin{aligned}&\mathbb{E}\left[\left\langle DF\left(B(t)+\rho\int_0^t c(s)\,ds\right),\int_0^t c(s)\,ds\right\rangle\right]\\&=\mathbb{E}\left[\exp\left(\rho\int_0^t c(s)\,dB(s)-\frac{1}{2}\rho^2\int_0^t\left|c(s)\right|^2ds\right)\right.\\&\qquad\left.\times\left(\int_0^t c(s)\,dB(s)-\rho\int_0^t\left|c(s)\right|^2ds\right)F\left(B(t)\right)\right].\end{aligned}\tag{4.134}$$

The bracket in the left-hand side of (4.134) indicates the inner-product in $\mathbb{R}^d$. In (4.134) we put $\rho=0$ and we obtain the first order version of the famous integration by parts formula:

$$\mathbb{E}\left[\left\langle DF\left(B(t)\right),\int_0^t c(s)\,ds\right\rangle\right]=\mathbb{E}\left[\int_0^t c(s)\,dB(s)\,F\left(B(t)\right)\right].\tag{4.135}$$

We mention that the Cameron-Martin-Girsanov transformation is a cornerstone for the integration by parts formula, which is a central issue in Malliavin calculus. For details on this subject see e.g. Nualart [**135, 133**], Malliavin [**119**], Sanz-Solé [**154**], Kusuoka and Stroock [**103, 104, 105**], Stroock [**164**], and Norris [**129**].

For a proof of Theorem 4.26 we will need the Skorohod-Dudley-Wichura representation theorem: see Theorem 11.7.2 of Dudley [**55**]. It will be applied with $S = C\left([0,t],\mathbb{R}^d\right)$ and can be formulated as follows.

4.31. THEOREM. *Let (S,d) be a complete separable metric space (i.e. a Polish space), and let $\mathbb{P}_k$, $k \in \mathbb{N}$, and $\mathbb{P}$ be probability measures on the Borel field $\mathcal{B}_S$ of S such that the weak limit $\text{w}-\lim_{k\to\infty}\mathbb{P}_k = \mathbb{P}$, i.e. $\lim_{k\to\infty}\int F d\mathbb{P}_k = \int F d\mathbb{P}$ for all bounded continuous functions of $F \in C_b(S)$. Then there exist a probability space $\left(\widetilde{\Omega},\widetilde{\mathfrak{F}},\widetilde{\mathbb{P}}\right)$ and S-valued stochastic variables $\widetilde{Y}_k$, $k \in \mathbb{N}$, and $\widetilde{Y}$, defined on $\widetilde{\Omega}$ with the following properties:*

(1) $\mathbb{P}_k[B] = \widetilde{\mathbb{P}}\left[\widetilde{Y}_k \in B\right]$, *$k \in \mathbb{N}$, and* $\mathbb{P}[B] = \widetilde{\mathbb{P}}\left[\widetilde{Y} \in B\right]$, $B \in \mathcal{B}_S$.

(2) *The sequence $\widetilde{Y}_k$, $k \in \mathbb{N}$, converges to $\widetilde{Y}$ $\widetilde{\mathbb{P}}$-almost surely.*

4.32. REMARK. An analysis of the existence part of the proof of Theorem 4.26 shows that the invertibility of the matrix $\sigma(s,y)$ is not needed. Let $\widetilde{N}(s)$, $0 \leqslant s \leqslant t$, be a local martingale on a filtered probability space $\left(\widetilde{\Omega},\widetilde{\mathfrak{F}}_s,\widetilde{\mathbb{P}}\right)$, where the σ-field $\widetilde{\mathfrak{F}}_s$ is generated by $\left(\widetilde{Y}(\tau) : 0 \leqslant \tau \leqslant s\right)$. Suppose that the covariation process of $\widetilde{N}(s)$ is given by

$$\left\langle \widetilde{N}_{j_1}, \widetilde{N}_{j_2}\right\rangle(s) = \int_0^s \left(\sigma\left(\tau,\widetilde{Y}(\tau)\right)\sigma^*\left(\tau,\widetilde{Y}(\tau)\right)\right)_{j_1,j_2} d\tau, \quad 1 \leqslant j_1,\ j_2 \leqslant d.$$

Here $\widetilde{Y}$ is a local martingale on $\left(\widetilde{\Omega},\widetilde{\mathfrak{F}},\widetilde{\mathbb{P}}\right)$. Then by assertion (iii) in Theorem 4.17 there exists a Brownian motion $\widetilde{B}(s)$, $0 \leqslant s \leqslant t$, on this space such that

$$\begin{aligned}\int_0^s c_1\left(\tau,\widetilde{Y}(\tau)\right) d\widetilde{N}(\tau) &= \int_0^s c_1\left(\tau,\widetilde{Y}(\tau)\right)\sigma\left(\tau,\widetilde{Y}(\tau)\right) d\widetilde{B}(\tau) \\ &= \int_0^s c\left(\tau,\widetilde{Y}(\tau)\right) d\widetilde{B}(\tau). \end{aligned} \tag{4.136}$$

PROOF OF THEOREM 4.26. *Uniqueness.* Let the process $Y(s)$, $0 \leqslant s \leqslant t$, be a solution to equation (4.111). So that

$$\begin{aligned} Y(s) &= x + \int_0^s \sigma(\tau,Y(\tau))\, dB(\tau) + \int_0^s \left(b(\tau,Y(\tau)) + \sigma(\tau,Y(\tau))\, c(\tau,Y(\tau))\right) d\tau \\ &= x + \int_0^s \sigma(\tau,Y(\tau))\, dB'(\tau) + \int_0^s b(\tau,Y(\tau))\, d\tau. \end{aligned} \tag{4.137}$$

Let $F\left((Y(s))_{0\leqslant s\leqslant t}\right)$ be a bounded stochastic variable which depends on the path $Y(s)$, $0 \leqslant s \leqslant t$. As observed in 4 of Proposition 4.24 the process $B'(t)$

is a $\mathbb{P}'$-Brownian motion, provided $\mathbb{E}\left[M'(t)\right]=1$. Uniqueness of weak solutions to equation (4.110) implies that the P'-distribution of the process $s\mapsto Y(s)$, $0\leqslant s\leqslant t$, coincides with the $\mathbb{P}$-distribution of the process $s\mapsto X(s)$, $0\leqslant s\leqslant t$. In other words we have

$$\begin{aligned}
&\mathbb{E}'\left[F\left((Y(s))_{0\leqslant s\leqslant t}\right)\right]\\
&=\mathbb{E}\left[\exp\left(-\int_0^t c_1\left(s,Y(s)\right)\,dN^Y(s)-\frac{1}{2}\int_0^t\left|c\left(s,Y(s)\right)\right|^2ds\right)F\left((Y(s))_{0\leqslant s\leqslant t}\right)\right]\\
&=\mathbb{E}\left[F\left((X(s))_{0\leqslant s\leqslant t}\right)\right], \qquad (4.138)
\end{aligned}$$

where

$$\begin{aligned}
N^Y(s)&=Y(s)-\int_0^s\sigma\left(\tau,Y(\tau)\right)c\left(\tau,Y\left(\tau\right)\right)\,d\tau-\int_0^s b\left(\tau,Y(\tau)\right)\,d\tau\\
&=\int_0^s\sigma\left(\tau,Y(\tau)\right)\,dB(\tau). \qquad (4.139)
\end{aligned}$$

With

$$\begin{aligned}
&G\left((Y(s))_{0\leqslant s\leqslant t}\right)\\
&=\exp\left(-\int_0^t c_1\left(s,Y(s)\right)\,dN^Y(s)-\frac{1}{2}\int_0^t\left|c\left(s,Y(s)\right)\right|^2ds\right)F\left((Y(s))_{0\leqslant s\leqslant t}\right)
\end{aligned}$$

we have

$$\begin{aligned}
&F\left((Y(s))_{0\leqslant s\leqslant t}\right)\\
&=\exp\left(\int_0^t c_1\left(s,Y(s)\right)\,dN^Y(s)+\int_0^t\frac{1}{2}\left|c\left(s,Y(s)\right)\right|^2ds\right)G\left((Y(s))_{0\leqslant s\leqslant t}\right)
\end{aligned}$$

So, since

$$\begin{aligned}
dN^X(s)&=dX(s)-\sigma\left(s,X(s)\right)c\left(s,X(s)\right)\,ds-b\left(s,X(s)\right)\,ds\\
&=\sigma\left(s,X(s)\right)\,\left(dB(s)-c\left(s,X(s)\right)\,ds\right) \qquad (4.140)
\end{aligned}$$

it follows that

$$\begin{aligned}
&F\left((X(s))_{0\leqslant s\leqslant t}\right)\\
&=\exp\left(\int_0^t c_1\left(s,X(s)\right)\,dN^X(s)+\frac{1}{2}\int_0^t\left|c\left(s,X(s)\right)\right|^2ds\right)G\left((X(s))_{0\leqslant s\leqslant t}\right)\\
&=\exp\left(\int_0^t c\left(s,X(s)\right)\,dB(s)-\frac{1}{2}\int_0^t\left|c\left(s,X(s)\right)\right|^2ds\right)G\left((X(s))_{0\leqslant s\leqslant t}\right). \qquad (4.141)
\end{aligned}$$

From (4.138) and (4.141) we infer:

$$\begin{aligned}
&\mathbb{E}\left[G\left((Y(s))_{0\leqslant s\leqslant t}\right)\right] \qquad (4.142)\\
&=\mathbb{E}\left[\exp\left(\int_0^t c\left(s,X(s)\right)\,ds-\frac{1}{2}\int_0^t\left|c\left(s,X(s)\right)\right|^2ds\right)G\left((X(s))_{0\leqslant s\leqslant t}\right)\right].
\end{aligned}$$

By inserting $G\equiv\mathbf{1}$ in (4.142) we see that

$$\mathbb{E}\left[\exp\left(\int_0^t c\left(s,X(s)\right)\,ds-\frac{1}{2}\int_0^t\left|c\left(s,X(s)\right)\right|^2ds\right)\right]=1$$

in case there is a unique solution to the equation in (4.127). This proves the uniqueness part of Theorem 4.26.

Existence. Therefore we will approximate the solution Y by a sequence Y_k, $k \in \mathbb{N}$, which are solutions to equations of the form:

$$\begin{aligned} Y_k(s) = x &+ \int_0^s \sigma\left(\tau, Y_k(\tau)\right) \, dB(\tau) \\ &+ \int_0^s \left(b\left(\tau, Y_k(\tau)\right) + \sigma\left(\tau, Y_k(\tau)\right) c_k\left(\tau, Y_k(\tau)\right)\right) \, d\tau \\ = x &+ \int_0^s \sigma\left(\tau, Y_k(\tau)\right) dB_k'(\tau) + \int_0^s b\left(\tau, Y_k(\tau)\right) d\tau. \end{aligned} \tag{4.143}$$

Here

$$B_k'(s) = B_k(s) + \int_0^s c_k\left(\tau, Y_k(\tau)\right) \, d\tau,$$

and the coefficients $c_k(s, y) = c_{1,k}(s, y)\sigma(s, y)$ are chosen in such a way that they are bounded and that $c(s, y) = \lim_{k\to\infty} c_k(s, y)$ for all $s \in [0, t]$ and $y \in \mathbb{R}^d$. By Novikov's theorem the corresponding local martingales M_k', given by

$$M_k'(s) = \exp\left(-\int_0^s c_k\left(\tau, Y_k(\tau)\right) dB(\tau) - \frac{1}{2}\int_0^s \left|c_k\left(\tau, Y_k(\tau)\right)\right|^2 d\tau\right), \quad k \in \mathbb{N},$$

are then automatically genuine martingales: see Corollary 4.29. From the uniqueness of weak solutions to equations in $X(t)$ of the form (4.110) (and thus to equations in $Y_k(s)$ of the form (4.143) we infer

$$\mathbb{E}_k'\left[F\left(\left(Y_k(s)\right)_{0\leqslant s\leqslant t}\right)\right] = \mathbb{E}\left[F\left(\left(X(s)\right)_{0\leqslant s\leqslant t}\right)\right]. \tag{4.144}$$

In equality (4.144) the process $Y_k(s)$, $0 \leqslant s \leqslant t$, solves the equation in (4.143). The equality in (4.144) can be rewritten as

$$\mathbb{E}\left[M_k'(t)F\left(\left(Y_k(s)\right)_{0\leqslant s\leqslant t}\right)\right] = \mathbb{E}\left[F\left(\left(X(s)\right)_{0\leqslant s\leqslant t}\right)\right]. \tag{4.145}$$

By (4.119) the equality in (4.145) can be rewritten as

$$\begin{aligned} &\mathbb{E}\left[\exp\left(-\int_0^t c_k\left(s, Y_k(s)\right) \, dB(s) - \frac{1}{2}\int_0^t \left|c_k\left(s, Y_k(s)\right)\right|^2 \, ds\right) F\left(\left(Y_k(s)\right)_{0\leqslant s\leqslant t}\right)\right] \\ &= \mathbb{E}\left[\exp\left(-\int_0^t c_{1,k}\left(s, Y_k(s)\right) \, d\left(Y_k(s) - \int_0^s b\left(\tau, Y_k(\tau)\right) \, d\tau\right)\right.\right. \\ &\qquad \left.\left. + \frac{1}{2}\int_0^t \left|c_k\left(s, Y_k(s)\right)\right|^2 \, ds\right) F\left(\left(Y_k(s)\right)_{0\leqslant s\leqslant t}\right)\right] \\ &= \mathbb{E}\left[F\left(\left(X(s)\right)_{0\leqslant s\leqslant t}\right)\right]. \end{aligned} \tag{4.146}$$

Let $G\left(\left(Y_k(s)\right)_{0\leqslant s\leqslant t}\right)$ be a (bounded) stochastic variable which depends on the path $Y_k(s)$, $0 \leqslant s \leqslant t$. From the equality in (4.146) we infer

$$\begin{aligned} &\mathbb{E}\left[G\left(\left(Y_k(s)\right)_{0\leqslant s\leqslant t}\right)\right] \\ &\quad = \mathbb{E}\left[\exp\left(\left(\int_0^t c_{1,k}\left(s, X(s)\right) \, d\left(X(s) - \int_0^s b\left(\tau, X(\tau)\right) \, d\tau\right)\right.\right.\right. \end{aligned}$$

$$
\begin{aligned}
&\quad -\frac{1}{2}\int_0^t \left|c_k\left(s, X(s)\right)\right|^2\, ds\Big) G\left(\left(X(s)\right)_{0\leqslant s\leqslant t}\right)\Big] \\
&= \mathbb{E}\left[\exp\left(\int_0^t c_{1,k}\left(s, X(s)\right)\sigma\left(s, X(s)\right)\, dB(s) - \frac{1}{2}\int_0^t \left|c_k\left(s, X(s)\right)\right|^2\, ds\right)\right. \\
&\qquad\qquad \left.\times\, G\left(\left(X(s)\right)_{0\leqslant s\leqslant t}\right)\right] \\
&= \mathbb{E}\left[M_k(t)G\left(\left(X(s)\right)_{0\leqslant s\leqslant t}\right)\right]. \qquad (4.147)
\end{aligned}
$$

Here the martingales $M_k(s)$ are given by

$$
M_k(s) = \exp\left(\int_0^s c_k\left(\tau, X(\tau)\right) dB(\tau) - \frac{1}{2}\int_0^s \left|c_k\left(\tau, X(\tau)\right)\right|^2 d\tau\right), \quad k \in \mathbb{N},
$$

This fact together with the pointwise convergence of $M_k(s)$ to $M(s)$, as $k \to \infty$, and invoking the hypothesis that $\mathbb{E}\left[M(t)\right] = 1$, shows that the right-hand side of (4.147) converges to $\mathbb{E}\left[M(t)G\left(\left(X(s)\right)_{0\leqslant s\leqslant t}\right)\right]$. In other words the distribution $\mathbb{P}^{Y_k}$ of Y_k converges weakly to the measure $\mathbb{P}^{M,X}$ defined by $\mathbb{P}^{M,X}(A) = \mathbb{E}\left[M(t),\, X \in A\right]$, where A is a Borel subset of the space $C\left([0,t], \mathbb{R}^d\right)$. By the Skorohod-Dudley-Wichura representation theorem (Theorem 4.31) there exist a probability space $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ and $C\left([0,t], \mathbb{R}^d\right)$-valued stochastic variables $\widetilde{Y}_k$, $k \in \mathbb{N}$, and $\widetilde{Y}$, defined on $\widetilde{\Omega}$ with the following properties:

(1) $\mathbb{P}^{Y_k}[B] = \widetilde{\mathbb{P}}\left[\widetilde{Y}_k \in B\right]$, $k \in \mathbb{N}$, and $\mathbb{P}^{M,X}\left[B\right] = \widetilde{\mathbb{P}}\left[\widetilde{Y} \in B\right]$, for all $B \in \mathcal{B}_{C([0,t],\mathbb{R}^d)}$.

(2) The sequence $\left(\widetilde{Y}_k\right)_{k\in\mathbb{N}}$ converges to $\widetilde{Y}$ $\widetilde{\mathbb{P}}$-almost surely.

By taking the limit in (4.147) for $k \to \infty$ and using the theorem of Skorohod-Dudley-Wichura we obtain

$$\mathbb{E}\left[G\left(\left(\widetilde{Y}(s)\right)_{0\leqslant s\leqslant t}\right)\right] = \mathbb{E}\left[M(t)G\left((X(s))_{0\leqslant s\leqslant t}\right)\right] \tag{4.148}$$

where G is a bounded continuous function on $C\left([0,t],\mathbb{R}^d\right)$. Then we consider the process $\widetilde{N}(s)$, $0 \leqslant s \leqslant t$, defined by

$$\widetilde{N}(s) = \widetilde{Y}(s) - \int_0^s \sigma\left(\tau,\widetilde{Y}(\tau)\right) c\left(\tau,\widetilde{Y}(\tau)\right) d\tau - \int_0^s b\left(\tau,\widetilde{Y}(\tau)\right) d\tau. \tag{4.149}$$

If $\widetilde{Y}(s)$ were $Y(s)$, then by (4.137) $\widetilde{N}(s)$ would be $N^Y(s)$, given by the formula in (4.139). Hence the process $s \mapsto N^Y(s)$, $s \in [0,t]$, is a stochastic integral relative to Brownian motion on the space $(\Omega,\mathcal{F}_t,\mathbb{P})$. We want to do the same for the process $s \mapsto \widetilde{N}(s)$, $0 \leqslant s \leqslant t$, on the probability space $\left(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}}\right)$. Let $\mathbb{P}^{M(t)}$ be the probability measure on $(\Omega,\mathcal{F}_t)$ defined by $\mathbb{P}^{M(t)}[A] = \mathbb{E}[M(t), A]$, $A \in \mathcal{F}_t$. Then like in item (4) of Proposition 4.24 we see that the process $s \mapsto B(s) - \int_0^s \sigma(\tau,X(\tau))\,d\tau$ is a $\mathbb{P}^{M(t)}$-Brownian motion. In addition, from (4.148) and (4.149) we infer that the $\widetilde{\mathbb{P}}$-distribution of the process $\widetilde{N}(s)$, $0 \leqslant s \leqslant t$, is given by the $\mathbb{P}^{M(t)}$-distribution of the process

$$\begin{aligned} s \mapsto & X(s) - \int_0^s \sigma(\tau,X(\tau))\,c(\tau,X(\tau))\,d\tau - \int_0^s b(\tau,X(\tau))\,d\tau \\ &= \int_0^s \sigma(\tau,X(\tau))\,(dB(\tau) - c(\tau,Y(\tau))\,d\tau) \\ &= \int_0^s \sigma(\tau,X(\tau))\,dB^{M(t)}(\tau), \end{aligned} \tag{4.150}$$

where $B^{M(t)}(s)$ is a $\mathbb{P}^{M(t)}$-Brownian motion: see Proposition 4.24 item (4). It also follows that the process in (4.150) has covariation process given by the square matrix process

$$s \mapsto \int_0^s \sigma(\tau,X(\tau))\,\sigma^*(\tau,X(\tau))\,d\tau, \quad 0 \leqslant s \leqslant t.$$

Consequently, the process $s \mapsto \widetilde{N}(s)$, $0 \leqslant s \leqslant t$, is a local $\widetilde{P}$-martingale with covariation process given by

$$s \mapsto \int_0^s \sigma\left(\tau,\widetilde{Y}(\tau)\right)\sigma^*\left(\tau,\widetilde{Y}(\tau)\right) d\tau, \quad 0 \leqslant s \leqslant t. \tag{4.151}$$

In order to prove (4.151) we must show that the process

$$s \mapsto \widetilde{N}_{j_1}(s)\widetilde{N}_{j_2}(s) - \sum_{k=1}^d \int_0^s \sigma_{j_1,k}\left(\tau,\widetilde{Y}(\tau)\right)\sigma_{j_2,k}\left(\tau,\widetilde{Y}(\tau)\right) d\tau$$

is a local $\widetilde{\mathbb{P}}$-martingale. The latter can be achieved by appealing to the fact the $\widetilde{\mathbb{P}}$-distribution of the process $s \mapsto \widetilde{Y}(s)$, $0 \leqslant s \leqslant t$, coincides with the $\mathbb{P}^{M(t)}$-distribution of the process $s \mapsto X(s)$, $0 \leqslant s \leqslant t$. Then we choose a Brownian motion $\widetilde{B}(s)$, possibly on an extension of the probability space $\left(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}}\right)$, which we call again $\left(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}}\right)$ such that $\widetilde{N}(s)=\int_0^s\sigma\left(\tau,\widetilde{Y}(\tau)\right)\,d\widetilde{B}(\tau)$. For details see the proof of the implication (ii) $\Longrightarrow$ (iii) of Theorem 4.17. With such a Brownian motion we obtain:

$$\begin{aligned}\widetilde{Y}(s)=x+&\int_0^s\sigma\left(\tau,\widetilde{Y}(\tau)\right)\,d\widetilde{B}(\tau)\\&+\int_0^s\sigma\left(\tau,\widetilde{Y}(\tau)\right)\,c\left(\tau,\widetilde{Y}(\tau)\right)\,d\tau+\int_0^s b\left(\tau,\widetilde{Y}(\tau)\right)\,d\tau.\end{aligned}\tag{4.152}$$

Since

$$\widetilde{\mathbb{E}}\left[\exp\left(-\int_0^t c\left(s,\widetilde{Y}(s)\right)\,d\widetilde{B}(s)-\frac{1}{2}\int_0^t\left|c\left(s,\widetilde{Y}(s)\right)\right|^2\,ds\right)\right]=1\tag{4.153}$$

it follows that the process $s\mapsto\widetilde{B}(s)+\int_0^s c\left(\tau,\widetilde{Y}(\tau)\right)\,d\tau$ is a Brownian motion relative to the measure

$$A\mapsto\widetilde{\mathbb{E}}\left[\exp\left(-\int_0^t c\left(s,\widetilde{Y}(s)\right)\,d\widetilde{B}(s)-\frac{1}{2}\int_0^t\left|c\left(s,\widetilde{Y}(s)\right)\right|^2\,ds\right),A\right],\quad A\in\widetilde{\mathcal{F}}.$$

The equalities in (4.152) and (4.153) complete the proof of Theorem 4.26. ☐

3.1. Equations with unique strong solutions possess unique weak solutions. The following theorem shows that stochastic differential equations with unique pathwise solutions also have unique weak solutions. Its proofs puts the Lévy's characterization of Brownian motion at work: see Theorem 4.5.

4.33. THEOREM. *Let the vector and matrix functions $b(s,x)$ and $\sigma(s,x)$ be as in Theorem 4.26. Fix $x\in\mathbb{R}^d$. Suppose that the stochastic (integral) equation*

$$X(t)=x+\int_0^t\sigma\left(s,X(s)\right)dB(s)+\int_0^t b\left(s,X(s)\right)\,ds\tag{4.154}$$

possesses unique pathwise solutions. Then this equation has unique weak solutions.

In the proof we employ a certain coupling argument. In fact weak solutions to the equations in (4.3) and (4.4) are recast as two pathwise solutions of the same form as (4.154) on the same probability space.

PROOF. Let $\{(B(t):t\geqslant 0),(\Omega,\mathcal{F},\mathbb{P})\}$ and $\{(B'(t):t\geqslant 0),(\Omega',\mathcal{F}',\mathbb{P}')\}$ be two independent Brownian motions. Without loss of generality it is assumed that, for $0\leqslant t<\infty$,

$$\begin{aligned}\mathcal{F}_t&=\sigma\left(\{B(s):0\leqslant s\leqslant t\}\bigcup\left\{A\in\mathcal{F}^0:\mathbb{P}[A]=0\right\}\right),\quad\text{and}\\\mathcal{F}'_t&=\sigma\left(\{B'(s):0\leqslant s\leqslant t\}\bigcup\{A'\in\mathcal{F}':\mathbb{P}'[A']=0\}\right).\end{aligned}$$

Moreover, $\mathcal{F}=\sigma\left(\bigcup_{t\geqslant 0}\mathcal{F}_t\right)$, and a a similar assumption is made for $\mathcal{F}'$. Let $\{X(t):t\geqslant 0\}$ be an adapted process which satisfies (4.3), and let $\{X'(t):t\geqslant 0\}$

be an adapted process which satisfies (4.4) with $B'(t)$ instead of $B(t)$: see Definition 4.19. Suppose that $0 \leqslant t_1 < t_2 < \cdots < t_n < \infty$, and let $C_1, \dots, C_n$ be Borel subsets of $\mathbb{R}^d$. We have to prove the equality:

$$\mathbb{P}'\left[X'(t_1) \in C_1, \dots, X'(t_n) \in C_n\right] = \mathbb{P}\left[X(t_1) \in C_1, \dots, X(t_n) \in C_n\right]. \tag{4.155}$$

Let $(\Omega_0, \mathcal{F}^0, \mathbb{P}_0)$ be a probability space with a Brownian motion $\{B_0(t) : t \geqslant 0\}$ such that $\mathcal{F}^0 = \sigma\left(\{B_0(t) : t \geqslant 0\} \bigcup \{A_0 \in \mathcal{F}^0 : \mathbb{P}_0\left[A_0\right] = 0\}\right)$. Define the $\mathbb{R}^d$-valued processes $Y(t)$, $Y'(t)$, and $\widetilde{B}_0(t)$ on $\Omega \times \Omega' \times \Omega_0$ as follows:

$$\begin{cases} Y(t)\left(\omega, \omega', \omega_0\right) = X(t)(\omega), & (\omega, \omega', \omega_0) \in \Omega \times \Omega' \times \Omega_0; \\ Y'(t)\left(\omega, \omega', \omega_0\right) = X'(t)\left(\omega'\right), & (\omega, \omega', \omega_0) \in \Omega \times \Omega' \times \Omega_0; \\ \widetilde{B}_0(t)\left(\omega, \omega', \omega_0\right) = B_0(t)\left(\omega_0\right), & (\omega, \omega', \omega_0) \in \Omega \times \Omega' \times \Omega_0. \end{cases} \tag{4.156}$$

In fact we use the notation Ω_0 instead of Ω to distinguish the third component of the space $\Omega \times \Omega' \times \Omega_0$ from the first. The role of the first two components are very similar; the third component is related to the driving Brownian motion $\{B_0(t) : t \geqslant 0\}$. The processes $Y(t)$ and $Y'(t)$ are going to be the pathwise solutions on the same probability space $\left(\Omega \times \Omega' \times \Omega_0, \mathcal{F} \otimes \mathcal{F}' \otimes \mathcal{F}^0, \widetilde{\mathbb{Q}}_x\right)$: see (4.166) and (4.167) below. On Ω_0 the probability measure $\mathbb{P}_0$ is determined by prescribing its finite-dimensional distributions via the equality:

$$\begin{aligned} &\mathbb{P}_0\left[\left(B_0(t_1), \dots, B_0(t_n)\right) \in D\right] = \mathbb{P}\left[\left(B(t_1), \dots, B(t_n)\right) \in D\right] \\ &\quad = \mathbb{P}'\left[\left(B'(t_1), \dots, B'(t_n)\right) \in D\right]. \end{aligned} \tag{4.157}$$

In (4.157) we have $0 \leqslant t_1 < \cdots < t_n < \infty$, and D is a Borel subset of $\left(\mathbb{R}^d\right)^n$. Let C be another Borel subset of $\left(\mathbb{R}^d\right)^n$. On $\Omega \times \Omega_0$ and $\Omega' \times \Omega_0$ the probability measures $\mathbb{Q}_x$ and $\mathbb{Q}'_x$ are determined by, respectively, the equalities:

$$\begin{aligned}&\mathbb{Q}_x\left[\left(X\left(t_1\right), \ldots, X\left(t_n\right)\right) \in C,\ \left(B_0\left(t_1\right), \ldots, B_0\left(t_n\right)\right) \in D\right]\\ &= \mathbb{P}\left[\left(X\left(t_1\right), \ldots, X\left(t_n\right)\right) \in C,\ \left(B\left(t_1\right), \ldots, B\left(t_n\right)\right) \in D\right], \quad \text{and}\\ &\mathbb{Q}'_x\left[\left(X'\left(t_1\right), \ldots, X'\left(t_n\right)\right) \in C,\ \left(B_0\left(t_1\right), \ldots, B_0\left(t_n\right)\right) \in D\right]\\ &= \mathbb{P}'\left[\left(X'\left(t_1\right), \ldots, X'\left(t_n\right)\right) \in C,\ \left(B'\left(t_1\right), \ldots, B'\left(t_n\right)\right) \in D\right]. \end{aligned} \tag{4.158}$$

Notice that $\mathbb{P}_0\left[A_0\right] = 0$ implies $\mathbb{Q}_x\left[\Omega \times A_0\right] = \mathbb{Q}'_x\left[\Omega' \times A_0\right] = 0$. Consequently, by the Radon-Nikodym's theorem there are (measurable) functions

$$Q_x : \mathcal{F} \times \Omega_0 \to [0,1], \text{ and } Q'_x : \mathcal{F}' \times \Omega_0 \to [0,1]$$

such that, respectively,

$$\begin{aligned}\mathbb{Q}_x\left[A \times A_0\right] &= \int_{A_0} Q_x\left(A, \omega_0\right) d\mathbb{P}_0\left(\omega_0\right), \quad A \in \mathcal{F},\ A_0 \in \mathcal{F}^0, \quad \text{and}\\ \mathbb{Q}'_x\left[A' \times A_0\right] &= \int_{A_0} Q'_x\left(A, \omega_0\right) d\mathbb{P}_0\left(\omega_0\right), \quad A' \in \mathcal{F}',\ A_0 \in \mathcal{F}^0.\end{aligned} \tag{4.159}$$

Here $Q_x\left(\Omega, \omega_0\right) = Q'_x\left(\Omega', \omega_0\right) = 1$ for $\mathbb{P}_0$-almost all $\omega_0 \in \Omega_0$. Moreover, the functions

$$\omega_0 \mapsto Q_x\left(A, \omega_0\right), \quad \text{and} \quad \omega_0 \mapsto Q'_x\left(A, \omega_0\right) \tag{4.160}$$

are measurable relative to the $\mathbb{P}_0$-completion of $\mathcal{F}^0$. In addition, the set functions $A \mapsto Q_x\left(A, \omega_0\right)$, $A \in \mathcal{F}$, and $A' \mapsto Q'_x\left(A', \omega_0\right)$, $A' \in \mathcal{F}'$ are $\mathbb{P}_0$-almost surely probability measures. Here we use the fact that, except for negligible sets, the σ-fields $\mathcal{F}$ and $\mathcal{F}'$ are countably determined. Finally, we define the measure

$$\widetilde{\mathbb{Q}}_x : \mathcal{F} \otimes \mathcal{F}' \otimes \mathcal{F}^0 \to [0,1] \quad \text{via the equality}$$

$$\begin{aligned}\widetilde{\mathbb{Q}}_x\left[A \times A' \times A_0\right] &= \int Q_x\left(A, \omega_0\right) Q'_x\left(A', \omega_0\right) \mathbf{1}_{A_0}\left(\omega_0\right) d\mathbb{P}_0\left(\omega_0\right)\\ &= \mathbb{E}_0\left[\omega_0 \mapsto Q_x\left(A, \omega_0\right) Q'_x\left(A', \omega_0\right) \mathbf{1}_{A_0}\left(\omega_0\right)\right].\end{aligned} \tag{4.161}$$

Here A, A', and A_0 belong to $\mathcal{F}$, $\mathcal{F}'$, and $\mathcal{F}^0$ respectively. First we prove that the process $\left\{\widetilde{B}_0(t) : t \geqslant 0\right\}$ is a Brownian motion with respect to the measure $\widetilde{\mathbb{Q}}_x$. The corresponding expectation is written as $\widetilde{\mathbb{E}}_x$. From the proof of Theorem 4.5 (*i.e.*, Lévy's characterization of Brownian motion) it follows that it suffices to show that the following equality holds:

$$\begin{aligned}&\widetilde{\mathbb{E}}_x\left[\exp\left(-i\left\langle \xi, \widetilde{B}_0(t) - \widetilde{B}_0(s)\right\rangle\right) \mid \mathcal{F}_s \otimes \mathcal{F}'_s \otimes \mathcal{F}^0_s\right]\\ &= \exp\left(-\frac{1}{2}\left|\xi\right|^2(t-s)\right), \quad t > s \geqslant 0,\ \xi \in \mathbb{R}^d.\end{aligned} \tag{4.162}$$

By definition $\mathcal{F}_s = \sigma\left(B(\rho) : 0 \leqslant \rho \leqslant s\right)$. Similar definitions are employed for $\mathcal{F}'_s$ and for the σ-field $\mathcal{F}^0_s$. In order to prove (4.162) we pick $A \in \mathcal{F}_s$, $A' \in \mathcal{F}'_s$, and $A_0 \in \mathcal{F}^0_s$. Then by (4.161) we get

$$\widetilde{\mathbb{E}}_x\left[\exp\left(-i\left\langle \xi, \widetilde{B}_0(t) - \widetilde{B}_0(s)\right\rangle\right) \mathbf{1}_{A \times A' \times A_0}\right]$$

$$\begin{aligned}&= \int_{A\times A'\times A_0} \exp\left(-i\left\langle \xi, \widetilde{B}_0(t)-\widetilde{B}_0(s)\right\rangle\right) d\widetilde{\mathbb{Q}}_x \\ &= \mathbb{E}_0\left[\omega_0 \mapsto \exp\left(-i\left\langle \xi, B_0(t)\left(\omega_0\right)-B_0(s)\left(\omega_0\right)\right\rangle\right)\right. \\ &\qquad \left.\times Q_x\left(A,\omega_0\right) Q_x'\left(A',\omega_0\right) \mathbf{1}_{A_0}\left(\omega_0\right)\right]. \end{aligned} \tag{4.163}$$

The process $(\omega_0, t) \mapsto B_0(t)(\omega_0)$ is a Brownian motion relative to $\mathbb{P}_0$, and the events A, A', and A_0 belong to $\mathcal{F}_s$, $\mathcal{F}_s'$, and $\mathcal{F}_s^0$ respectively, and hence the variable $B_0(t)-B_0(s)$ is $\mathbb{P}_0$-independent of the variable

$$\omega_0 \mapsto Q_x\left(A,\omega_0\right) Q_x'\left(A',\omega_0\right) \mathbf{1}_{A_0}\left(\omega_0\right).$$

Therefore (4.163) implies

$$\begin{aligned}&\widetilde{\mathbb{E}}_x\left[\exp\left(-i\left\langle \xi, \widetilde{B}_0(t)-\widetilde{B}_0(s)\right\rangle\right) \mathbf{1}_{A\times A'\times A_0}\right] \\ &= \int_{A_0} Q_x\left(A,\omega_0\right) Q_x'\left(A',\omega_0\right) d\mathbb{P}_0\left(\omega_0\right) \times \int \exp\left(-i\left\langle \xi, B_0(t)-B_0(s)\right\rangle\right) d\mathbb{P}_0 \\ &= \widetilde{\mathbb{Q}}_x\left[A\times A'\times A_0\right] \exp\left(-\frac{1}{2}\left|\xi\right|^2(t-s)\right). \end{aligned} \tag{4.164}$$

The equality in (4.162) is a consequence of (4.164). Since, by definition (see (4.157))

$$\mathbb{P}_0\left[\left(B_0\left(t_1\right),\dots,B_0\left(t_n\right)\right)\in C\right] = \mathbb{P}\left[\left(B\left(t_1\right),\dots,B\left(t_n\right)\right)\in C\right] \tag{4.165}$$

for $0 \leqslant t_1 < \cdots < t_n < \infty$, C Borel subset of $\left(\mathbb{R}^d\right)^n$, and since the process $\{B(t) : t \geqslant 0\}$ is a Brownian motion relative to $\mathbb{P}$, the same is true for the process $\{B_0(t) : t \geqslant 0\}$ relative to $\mathbb{P}_0$. Next we compute the quantity:

$$\begin{aligned}&\widetilde{\mathbb{E}}_x\left[\left|Y(t)-x-\int_0^t \sigma\left(s,Y(s)\right) d\widetilde{B}_0(s)-\int_0^t b\left(s,Y(s)\right) ds\right|\right] \\ &= \int\left|X(t)-x-\int_0^t \sigma\left(s,X(s)\right) dB(s)-\int_0^t b\left(s,X(s)\right) ds\right| d\mathbb{P} = 0. \end{aligned} \tag{4.166}$$

Similarly we have

$$\begin{aligned}&\widetilde{\mathbb{E}}_x\left[\left|Y'(t)-x-\int_0^t \sigma\left(s,Y'(s)\right) d\widetilde{B}_0(s)-\int_0^t b\left(s,Y'(s)\right) ds\right|\right] \\ &= \int\left|X'(t)-x-\int_0^t \sigma\left(s,X'(s)\right) dB'(s)-\int_0^t b\left(s,X'(s)\right) ds\right| d\mathbb{P}' = 0. \end{aligned} \tag{4.167}$$

From (4.166) and (4.167) we infer that the following equalities hold $\widetilde{\mathbb{Q}}_x$-almost surely:

$$Y(t) = x + \int_0^t \sigma\left(s,Y(s)\right) d\widetilde{B}_0(s) + \int_0^t b\left(s,Y(s)\right) ds \quad \text{and} \tag{4.168}$$

$$Y'(t) = x + \int_0^t \sigma\left(s,Y'(s)\right) d\widetilde{B}_0(s) + \int_0^t b\left(s,Y'(s)\right) ds. \tag{4.169}$$

Moreover, the process $\left\{\widetilde{B}_0(t) : t \geqslant 0\right\}$ is a Brownian motion relative to $\widetilde{\mathbb{Q}}_x$. From the pathwise uniqueness and the equalities (4.168) and (4.169) we see that, $\widetilde{\mathbb{Q}}_x$-almost surely,

$$Y(t) = Y'(t), \quad t \geqslant 0. \tag{4.170}$$

Let $0 \leqslant 0 < t_1 < \cdots < t_n < \infty$, and let C be a Borel subset of $\left(\mathbb{R}^d\right)^n$. From (4.170) it follows that

$$\widetilde{\mathbb{Q}}_x \left[\left(Y\left(t_1\right), \ldots, Y\left(t_n\right)\right) \in C\right] = \widetilde{\mathbb{Q}}_x \left[\left(Y'\left(t_1\right), \ldots, Y'\left(t_n\right)\right) \in C\right]. \tag{4.171}$$

Using (4.171) and the definition of the measure $\widetilde{\mathbb{Q}}_x$ shows that the following identities are self-explanatory:

$$\begin{aligned} &\widetilde{\mathbb{Q}}_x \left[\left(Y\left(t_1\right), \ldots, Y\left(t_n\right)\right) \in C\right] = \mathbb{Q}_x \left[\left(X\left(t_1\right), \ldots, X\left(t_n\right)\right) \in C, \Omega_0\right] \\ &\quad = \mathbb{P}\left[\left(X\left(t_1\right), \ldots, X\left(t_n\right)\right) \in C, \Omega\right] = \mathbb{P}\left[\left(X\left(t_1\right), \ldots, X\left(t_n\right)\right) \in C\right]. \end{aligned} \tag{4.172}$$

The definition of the measure $\widetilde{\mathbb{Q}}_x$ is given in (4.161). Similarly we conclude

$$\widetilde{\mathbb{Q}}_x \left[\left(Y'\left(t_1\right), \ldots, Y'\left(t_n\right)\right) \in C\right] = \mathbb{P}' \left[\left(X'\left(t_1\right), \ldots, X'\left(t_n\right)\right) \in C\right]. \tag{4.173}$$

From (4.172), (4.173), and (4.171) we obtain

$$\mathbb{P}\left[\left(X\left(t_1\right), \ldots, X\left(t_n\right)\right) \in C\right] = \mathbb{P}\left[\left(X'\left(t_1\right), \ldots, X'\left(t_n\right)\right) \in C\right]. \tag{4.174}$$

The equality in (4.174) implies that the finite-dimensional distributions of the solution in equation in (4.3) are the same as those of the solution of equation (4.4). So that stochastic differential equations with unique pathwise solutions also possess unique weak (or distributional) solutions.

This concludes the proof of Theorem 4.33. □

4.34. EXAMPLE (Tanaka's example). Let the process $t \mapsto B(t)$, $t \geqslant 0$, be one-dimensional Browmian motion on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let the continuous process $t \mapsto X(t)$ be such that $X(t) = \int_0^t \operatorname{sgn}\left(X(s)\right) dB(s)$. Here $\operatorname{sgn}(y) = \dfrac{y}{|y|}$ for $y \neq 0$, and $\operatorname{sgn}(y) = 0$, when $y = 0$. It can be proved that such a process exists. If $t \mapsto X(t)$ solves this equation, then the process $t \mapsto -X(t)$ is a solution as well. So we see that the equation $dX(t) = \operatorname{sgn}\left(X(t)\right) dB(t)$, $X(0) = 0$, does not have pathwise unique solutions. On the other hand the process $t \mapsto X(t)$ is (local) martingale, and, since $B(t) = \int_0^t \operatorname{sgn}\left(X(s)\right) dX(s)$, we get

$$t = \left\langle B, B\right\rangle(t) = \int_0^t \left|\operatorname{sgn}\left(X(s)\right)\right|^2 d\left\langle X, X\right\rangle(s) = \left\langle X, X\right\rangle(t).$$

Hence, $\left\langle X, X\right\rangle(t) = t$. Lévy's martingale characterization of Brownian motion (see Corollary 4.7 and Theorem 4.5) then implies that the process $t \mapsto X(t)$ is a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. So that the distribution of $X(t)$ is that of Brownian motion. Consequently, the equation $dX(t) = \operatorname{sgn}\left(X(t)\right) dB(t)$ has unique weak solutions. For more details on Tanaka's example and its connection with local time see, *e.g.*, Øksendal [**138**].

Conclusion. In this chapter we treated several aspects of the theory of stochastic differential equations: strong and weak solutions, Lévy's characterization of Brownian motion, exponential martingales, Hermite polynomials with applications to exponential martingales, a version of the martingale representation theorem, and the Girsanov or the Cameron-Martin-Girsanov transformation. Stochastic differential equations for processes with jumps, that is for Lévy processes, are treated in, e.g., Applebaum [**6**].

CHAPTER 5

Some related results

In this section we will discuss, among other things, Fourier transforms of distributions of random variables, positive-definite functions, Bochner's theorem, Lévy's continuity theorem, weak convergence of measures, ergodic theorems, projective limits of distributions, Markov processes with one initial probability measure, Doob-Meyer decomposition theorem based on Komlos' theorem.

1. Fourier transforms

Since we will also need signed measures, we will discuss them first.

1.1. Signed measures. Let $\mathcal{M} = \mathcal{M}(\mathbb{R}^\nu, \mathbb{C})$ be the vector space of all complex Borel measures on $\mathbb{R}^\nu$, and let $\mathcal{M}^+$ be the convex cone of all positive finite Borel measures op $\mathbb{R}^\nu$. Then we have $\mathcal{M} = \mathcal{M}_+ - \mathcal{M}_+ + i\left(\mathcal{M}_+ - \mathcal{M}_+\right)$. Thus, every complex Borel measure μ on $\mathbb{R}^\nu$ can be written as $\mu = \mu_1 - \mu_2 + i\left(\mu_3 - \mu_4\right)$, where μ_1, μ_2, μ_3 and μ_4 are finite positive Borel measures. In fact the measures μ_j, $1 \leqslant j \leqslant 4$, can be chosen in the following manner:

$$\begin{aligned}
\mu_1(B) &= \sup\left\{\text{Re } \mu(C) : C \subseteq B,\ C \text{ Borel }\right\};\\
\mu_2(B) &= \sup\left\{-\text{Re } \mu(C) : C \subseteq B,\ C \text{ Borel }\right\};\\
\mu_3(B) &= \sup\left\{\text{Im } \mu(C) : C \subseteq B,\ C \text{ Borel }\right\};\\
\mu_4(B) &= \sup\left\{-\text{Im } \mu(C) : C \subseteq B,\ C \text{ Borel }\right\}.
\end{aligned}$$

For this choice of the measures μ_1, μ_2, μ_3 and μ_4, the measures μ_1 and μ_2 and also the measures μ_3 and μ_4 are mutually singular in the sense that for certain Borel subsets B_1 and B_3 the following equalities hold:

$$\begin{aligned}
\mu_1(B) &= \text{Re } \mu\left(B \cap B_1\right), \quad \mu_2(B) = \text{Re } \mu\left(B \cap B_1^c\right);\\
\mu_3(B) &= \text{Re } \mu\left(B \cap B_3\right), \quad \mu_4(B) = \text{Re } \mu\left(B \cap B_3^c\right).
\end{aligned}$$

This decomposition is known under the name *Hahn decomposition*. In addition, we introduce the *variation* of a complex measure μ. This measure is denoted as $|\mu|$. It is the bounded positive measure defined by

$$|\mu|(A) = \sup\left\{\sum_j |\mu(A_j)| : A \supseteq A_j \text{ and } A_j \cap A_k = \varnothing \text{ for } k \neq j\right\}. \tag{5.1}$$

Here A is a Borel subset of $\mathbb{R}^\nu$ and the same is true for the elements of the partition A_j, $j \in \mathbb{N}$. The norm $\|\mu\|$ of the complex Borel measure μ is then defined by the equality: $\|\mu\| = |\mu|(\mathbb{R}^\nu)$. Supplied with this norm $\mathcal{M}$ is turned into a Banach space. By the *Riesz representation theorem* the space $\mathcal{M}$ can be taken as the topological dual of the space $C_0(\mathbb{R}^\nu)$, being the Banach space consisting of those complex continuous functions $f : \mathbb{R}^\nu \to \mathbb{C}$ with the property that $\lim_{x\to\infty} f(x) = 0$. Then $C_0(\mathbb{R}^\nu)$ is a closed subspace of the space $C_b(\mathbb{R}^\nu)$, the space of all bounded continuous functions on $\mathbb{R}^\nu$, which is a Banach space relative to the supremum-norm $\|\cdot\|_\infty$, given by $\|f\|_\infty = \sup_{x\in\mathbb{R}^\nu} |f(x)|$, $f \in C_b(\mathbb{R}^\nu)$.

5.1. Definition. A complex Radon measure on a locally compact space E is a complex Borel measure with the property that for every $\epsilon > 0$ and every Borel subset B there exists a compact subset $K \subset B$ with the property that $|\mu(B\backslash K)| < \epsilon$.

5.2. Theorem (Riesz). *Let E be a locally compact Hausdorff space, which is σ-compact, and let $\Lambda : C_0(E) \to \mathbb{C}$ be a continuous linear functional. Then there exists a unique complex Radon measure μ on the Borel field of E such that $\Lambda(f) = \int f d\mu$, $f \in C_0(E)$. In addition, $\|\Lambda\| = \|\mu\| = |\mu|(E)$. If Λ is positive in the sense that $f \geqslant 0$ implies $\Lambda(f) \geqslant 0$, then the corresponding measure μ is positive as well and $\|\Lambda\| = \mu(E)$.*

Proof. For a proof the reader is referred to the literature. In fact the following construction can be used. Let the measures μ_j, $1 \leqslant j \leqslant 4$ be determined by

$$\begin{aligned}
\mu_1(O) &= \sup\{\mathrm{Re}\ \Lambda(f) : 0 \leqslant f \leqslant 1_O,\ f \in C_0(E)\};\\
\mu_2(O) &= \sup\{-\mathrm{Re}\ \Lambda(f) : 0 \leqslant f \leqslant 1_O,\ f \in C_0(B)\};\\
\mu_3(O) &= \sup\{\mathrm{Im}\ \Lambda(f) : 0 \leqslant f \leqslant 1_O,\ f \in C_0(E)\};\\
\mu_4(O) &= \sup\{-\mathrm{Im}\ \Lambda(f) : 0 \leqslant f \leqslant 1_O,\ f \in C_0(E)\},
\end{aligned}$$

where O is any open subset of E. Then it can be shown that, for each $1 \leqslant j \leqslant 4$, the set function μ_j extends to a genuine positive Borel measure on E. This extension is again called μ_j. Moreover,

$$\Lambda(f) = \int f\, d\mu_1 - \int f\, d\mu_2 + i\left(\int f\, d\mu_3 - \int f\, d\mu_4\right), \quad f \in C_0(E).$$

For details the reader is referred to, *e.g.*, [**174**]. This completes the proof of Theorem 5.2. □

5.3. Definition. Let μ be a complex Borel measure on $\mathbb{R}^\nu$. Then the equality

$$\widehat{\mu}(x) = \int \exp\left(-i\langle x, y\rangle\right) d\mu(y), \quad x \in \mathbb{R}^\nu,$$

defines the Fourier transform of the measure μ.

5.4. Proposition. *Let μ be a complex measure on $\mathbb{R}^\nu$ with the property that its Fourier transform is identically zero. Then the measure $\mu = 0$.*

Proof. Let ν be an arbitrary other complex Borel measure on $\mathbb{R}^\nu$ with the property that $|\widehat{\nu}(x)| \leqslant 1$, $x \in \mathbb{R}^\nu$. Then the following equality holds:

$$\int \widehat{\mu}(x) d\nu(x) = \int \widehat{\nu}(y) d\mu(y).$$

Hence,

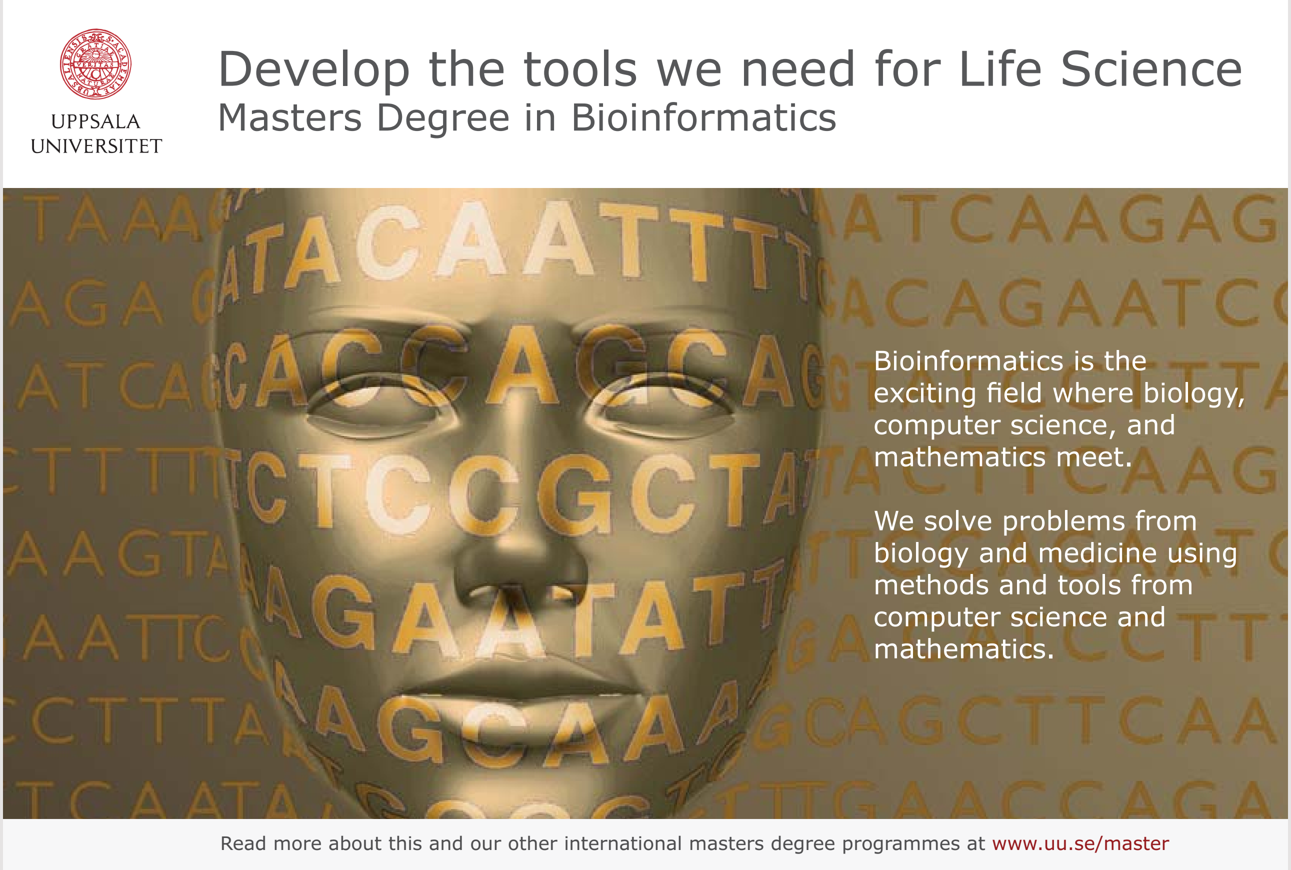
$$\begin{aligned}\|\mu\| &= \sup\left\{\left|\int f(y)d\mu(y)\right| : f \in C_b\left(\mathbb{R}^\nu\right), \|f\|_\infty \leqslant 1\right\} \\ &= \sup\left\{\left|\int \widehat{\nu}(y)d\mu(y)\right| : |\widehat{\nu}(y)| \leqslant 1, y \in \mathbb{R}^\nu\right\} \\ &= \sup\left\{\left|\int \widehat{\mu}(x)d\nu(x)\right| : |\widehat{\nu}(y)| \leqslant 1, y \in \mathbb{R}^\nu\right\} = 0.\end{aligned}$$

This completes the proof of Proposition 5.4. □

5.5. DEFINITION. Let $\varphi : \mathbb{R}^\nu \to \mathbb{C}$ be a complex valued function. This function is called *positive-definite* if for every n-tuple of complex numbers $\lambda_1, \dots, \lambda_n$ together with every choice of n vectors $\xi^{(1)}, \dots, \xi^{(n)}$ in $\mathbb{R}^\nu$, the following inequality holds:

$$\sum_{j,k=1}^{n} \lambda_j \overline{\lambda}_k \varphi\left(\xi^{(j)} - \xi^{(k)}\right) \geqslant 0,$$

and this for all $n \in \mathbb{N}$.



5.6. Proposition. *Let* μ *be a complex Borel measure on* $\mathbb{R}^\nu$ *with Fourier transform* $\widehat{\mu}$. *Then the following assertions are true:*

(a) *the following inequality holds:* $|\widehat{\mu}(x)| \leqslant \|\mu\|$.
(b) *If* μ *is positive, then the equalities* $\mu(\mathbb{R}^\nu) = \widehat{\mu}(0) = \|\mu\|$ *are valid.*
(c) *If* μ *is positive, then the function* $\widehat{\mu}$ *is positive-definite.*
(d) *The function* $\widehat{\mu}$ *is uniformly continuous.*

Proof. The proof is left as an exercise to the reader. ☐

5.7. Definition. Define for μ and ν measures in $\mathcal{M}$, the *convolution-product* $\mu * \nu$ via the equalities:

$$\begin{aligned} \mu * \nu(B) &= \int\int 1_B(x+y)d\mu(x)d\nu(y) \\ &= \mu \otimes \nu\left(S^{-1}B\right) = \mu \otimes \nu\left\{(x,y) \in \mathbb{R}^\nu : x+y \in B\right\}. \end{aligned}$$

Here B is a Borel subset of $\mathbb{R}^\nu$ and S is the (sum) mapping $S : (x,y) \mapsto x+y$. Let $x \in \mathbb{R}^\nu$. Define thee *Dirac-measure* δ_x by $\delta_x(B) = 1_B(x)$, B Borel subset of $\mathbb{R}^\nu$. Instead of δ_0 it is more customarily to write δ. Let $\mu \in \mathcal{M}$. Then $\check{\mu}$ is defined by $\check{\mu}(B) = \mu(-B)$, where B is a Borel subset of $\mathbb{R}^\nu$. Let $f : \mathbb{R}^\nu \to \mathbb{C}$ be a complex function, which is defined on all of $\mathbb{R}^\nu$. The function $\check{f}$ is given by $\check{f}(x) = f(-x)$, $x \in \mathbb{R}^\nu$.

5.8. Definition. A complex Banach algebra $(A, \|\cdot\|)$ is a complex Banach space, endowed with a *product* which is compatible with the norm. The latter means that the product $(a,b) \mapsto ab$, a, b in A, which is a bilinear operation, is continuous in both variables simultaneously. In fact it is assumed that $\|ab\| \leqslant \|a\|\,\|b\|$ for all a and b in A.

Examples of Banach algebras are the vector spaces $C_0(\mathbb{R}^\nu)$ and $C_b(\mathbb{R}^\nu)$, equipped with the supremum-norm and the pointwise multiplication. Let $\mathcal{L}(X)$ be the vector space of all continuous linear operators on the Banach space X, supplied with the operator norm and the composition as product. Then $\mathcal{L}(X)$ is a non-commutative Banach algebra. The following theorem says that $\mathcal{M}$, supplied with the convolution product, constitutes a (complex) commutative Banach algebra with identity δ. Recall that $\mathcal{M}$ stands for the space of all complex Borel measures on $\mathbb{R}^\nu$.

5.9. Theorem. *The normed vector space* $(\mathcal{M}, \|\cdot\|)$ *supplied with the convolution product* $*$ *is a commutative complex Banach algebra with identity* δ. *If* μ *and* ν *belong to* $\mathcal{M}$, *then the following equalities hold:*

$$\widehat{\mu+\nu} = \widehat{\mu} + \widehat{\nu}, \quad \widehat{a\mu} = a\widehat{\mu}, \quad \widehat{\mu * \nu} = \widehat{\mu}\widehat{\nu}, \quad \widehat{\check{\mu}} = \check{\widehat{\mu}}.$$

Here a *is a complex number.*

Proof. The proof is left as an exercise for the reader. ☐

The Banach space $L^1(\mathbb{R}^\nu)$ can be considered as a closed subspace of $\mathcal{M}(\mathbb{R}^\nu)$. This can be done via the following inclusion-mapping: $f \mapsto \mu_f$, $f \in L^1(\mathbb{R}^\nu)$.

Here μ_f is the complex measure $B \mapsto \int_B f(x)dx$, $B \in \mathcal{B} = \mathcal{B}(\mathbb{R}^\nu)$, where $\mathcal{B}$ is the Borel field of $\mathbb{R}^\nu$. Let $\mu_f = \mu_{f,1} - \mu_{f,2} + i\left(\mu_{f,3} - \mu_{f,4}\right)$ be the Hahn-Jordan decomposition of the measure μ_f. Then the following equalities hold:

$$\begin{aligned}
&\left|\mu_f\right|(B) = \int_B |f(x)|\,dx; \quad \mu_{f,1}(B) = \int_B \max\left(\operatorname{Re} f(x), 0\right)dx;\\
&\mu_{f,2}(B) = \int_B \max\left(-\operatorname{Re} f(x), 0\right)dx; \quad \mu_{f,3}(B) = \int_B \max\left(\operatorname{Im} f(x), 0\right)dx;\\
&\mu_{f,4}(B) = \int_B \max\left(-\operatorname{Im} f(x), 0\right)dx.
\end{aligned}$$

5.10. Theorem. *Let $C_{00}(\mathbb{R}^\nu)$ be the space of all complex continuous functions with compact support. Then $C_{00}(\mathbb{R}^\nu)$ is a dense subspace of $L^1(\mathbb{R}^\nu)$ for the topology of convergence in mean. This means that $C_{00}(\mathbb{R}^\nu)$ is dense in $L^1(\mathbb{R}^\nu)$ relative to the topology generated by the L^1-norm: $\|f\|_1 = \int |f(x)|\,dx$, $f \in L^1(\mathbb{R}^\nu)$.*

Proof. Let $\epsilon > 0$ and let $f \geqslant 0$ belong to $L^1(\mathbb{R}^\nu)$. It suffices that there exists a function $g \in C_{00}(\mathbb{R}^\nu)$ such that $\int |f(x) - g(x)|\,dx \leqslant \epsilon$. Since

$$f = \sup_{n\in\mathbb{N}} 2^{-n}\lfloor 2^n f\rfloor = \sup_{n\in\mathbb{N}} 2^{-n}\sum_{j=1}^{n2^n} 1_{\{f\geqslant j2^{-n}\}}$$

we only need to show that, for every pair of positive integers j and n, with $1 \leqslant j \leqslant n2^n$, there exists a function $u_{j,n} \in C_{00}(\mathbb{R}^\nu)$ such that

$$\int \left|1_{\{f\geqslant j2^{-n}\}}(x) - u_{j,n}(x)\right| dx \leqslant \frac{\epsilon}{2n}. \tag{5.2}$$

Because assume that the functions $u_{j,n}$, $1 \leqslant j \leqslant n2^n$, satisfy (5.2). Then we write $f_n = 2^{-n}\lfloor \min(n,f)2^n\rfloor$ and choose $n \in \mathbb{N}$ so large that

$$0 \leqslant \int \left(f(x) - f_n(x)\right) dx \leqslant \frac{1}{2}\epsilon.$$

Then we have

$$\begin{aligned}
&\int \left| f(x) - 2^{-n}\sum_{j=1}^{n2^n} u_{j,n}(x)\right|\\
&\leqslant \int |f(x) - f_n(x)|\,dx + 2^{-n}\sum_{j=1}^{n2^n}\int \left|1_{\{f\geqslant j2^{-n}\}}(x) - u_{j,n}(x)\right| dx\\
&\leqslant \frac{1}{2}\epsilon + 2^{-n}\sum_{j=1}^{n2^n}\frac{1}{2}\frac{\epsilon}{n} = \epsilon.
\end{aligned} \tag{5.3}$$

Let λ be the ν-dimensional Lebesgue measure. The inequality in (5.2) can be proved by employing the following identities:

$$\lambda(B) = \inf\left\{\lambda(U) : U \supseteq B, U \text{ open }\right\} = \sup\left\{\lambda(K) : K \subseteq B, K \text{ compact }\right\} \tag{5.4}$$

together with Tietsche's theorem, which, among other things, says that with a given open subset U and given compact subset K, with $K \subset U$, there exists a function $u \in C_{00}(\mathbb{R}^\nu)$ with the property that $1_K \leqslant u \leqslant 1_U$. The equalities in (5.4) follow via an argument about Dynkin systems.

This completes the proof of Theorem 5.10. ☐

5.11. PROPOSITION. *Let f belong to $L^1(\mathbb{R}^\nu)$. Then*

$$\lim_{y\to 0}\int |f(x+y)-f(x)|\,dx = 0. \tag{5.5}$$

PROOF. By theorem 5.10 it suffices to prove (5.5) for $f \in C_{00}(\mathbb{R}^\nu)$. Such a function f is uniformly continuous. Let K be the support of the function $f \in C_{00}(\mathbb{R}^\nu)$. Fix $\epsilon > 0$ and choose $\delta > 0$ in such a way that

$$\lambda\left(K + B(\delta)\right) \sup_{x\in K, y\in B(\delta)} |f(x+y)-f(x)| < \epsilon.$$

Here the symbol $B(\delta)$ stands for $B(\delta) = \delta B(1) = \{x \in \mathbb{R}^\nu : |x| \leqslant \delta\}$. Then we have

$$\int |f(x+y)-f(x)|\,dx \leqslant \epsilon$$

for $|y| \leqslant \delta$. So the proof of Proposition 5.11 is complete now. ☐

5.12. THEOREM (Riemann-Lebesgue). *Let $f \in L^1(\mathbb{R}^\nu)$. Then* $\lim_{x\to\infty} \widehat{f}(x) = 0$.

Of course here we write $\widehat{f}(x) = \int \exp\left(-i\langle x, y\rangle\right) f(y)dy$.

Proof of Theorem 5.12. By translation invariance of the Lebesgue measure we get the equality:

$$\widehat{f}(x) = \frac{1}{2} \int \exp\left(-i \langle x, y\rangle\right) \left(f(y) - f\left(y + \pi \frac{x}{|x|^2} \right) \right) dy. \tag{5.6}$$

From (5.6) the inequality:

$$\left| \widehat{f}(x) \right| \leqslant \frac{1}{2} \int \left| f(y) - f\left(y + \pi \frac{x}{|x|^2} \right) \right| dy. \tag{5.7}$$

A combination of (5.7) and Proposition 5.11 yields the desired result, and completes the proof of Theorem 5.12. □

5.13. Theorem (Stone-Weierstrass). *Let E be a locally compact Hausdorff space and let A be a subalgebra of $C_0(E)$, which separates points of E and which is closed under complex conjugation. That is, if f belongs to A, then $\overline{f}$ also belongs to A. Then A is dense in $C_0(E)$.*

Proof. Let $E^{\triangle}$ be the one-point compactification (Alexandroff compactification) and $A_1 = A \oplus \mathbb{C}1 = \{f + \lambda 1 : f \in A, \lambda \in \mathbb{C}\}$. Here 1 is the constant function with value 1 and functions $f \in A$ vanish in $\triangle$. The theorem of Stone-Weierstrass, applied to the compact Hausdorff space $E^{\triangle}$ results in the desired result, and completes the proof of Theorem 5.13. □

5.14. Theorem. *The set $\left\{\widehat{f} : f \in C_{00}\left(\mathbb{R}^{\nu}\right)\right\}$ is a subalgebra of $C_0(\mathbb{R}^{\nu})$ that is closed under taking complex conjugates. This algebra is dense in $C_0(\mathbb{R}^{\nu})$ with the supremum-norm.*

Proof. The fact that the set $A := \left\{\widehat{f} : f \in L^1(\mathbb{R}^{\nu})\right\}$ is a subalgebra of $C_0(\mathbb{R}^{\nu})$ follows from the standard properties of the Fourier transform in combination with Theorem 5.12. Since $\overline{\widehat{f}} = \widehat{\overline{\check{f}}}$ it also follows that this algebra is closed under complex conjugation. In order to apply the Theorem of Stone-Weierstrass we still have to show that A separates the points of $\mathbb{R}^{\nu}$. To this end take x_0 and $y_0 \neq x_0 \in \mathbb{R}^{\nu}$. Then there exists a bounded open neighborhood V in $\mathbb{R}^{\nu}$ such that $\exp\left(-i \langle x_0, y\rangle\right) - \exp\left(-i \langle y_0, y\rangle\right) \neq 0$ for $y \in V$. Next consider the function $f : y \mapsto \left(\exp\left(i \langle x_0, y\rangle\right) - \exp\left(i \langle y_0, y\rangle\right)\right) v(y)$, where v is a function in $C_{00}\left(\mathbb{R}^{\nu}\right)$ with $v \geqslant 1_V$. Then we see

$$\widehat{f}(x_0) - \widehat{f}(y_0) = \int \left|\exp\left(-i \langle x_0, y\rangle\right) - \exp\left(-i \langle y_0, y\rangle\right)\right|^2 v(y) dy > 0. \tag{5.8}$$

From (5.8) it immediately follows that A separates the points of $\mathbb{R}^{\nu}$. The assertion in Theorem 5.14 now follows from Theorem 5.13. □

In the following theorem we collect some properties of positive-definite functions.

5.15. Theorem. *Let $\varphi : \mathbb{R}^{\nu} \to \mathbb{C}$ be a positive-definite function. Then φ possesses the following properties:*

(a) $\varphi(-x) = \overline{\varphi(x)}$, $x \in \mathbb{R}^{\nu}$;

(b) $|\varphi(x)| \leqslant \varphi(0)$, $\in \mathbb{R}^\nu$;
(c) $|\varphi(x)-\varphi(y)|^2 \leqslant 2\varphi(0)\left(\varphi(0) - \mathit{Re}\ \varphi(x-y)\right)$, x, $y \in \mathbb{R}^\nu$;
(d) $\varphi(0)^2 \left|\varphi(x+y)\varphi(0) - \varphi(x)\varphi(y)\right|^2 \leqslant \left(\varphi(0)^2 - |\varphi(x)|^2\right)\left(\varphi(0)^2 - |\varphi(y)|^2\right)$.

Proof. Fix x and y in $\mathbb{R}^\nu$ and consider the matrices

$$\begin{pmatrix} \varphi(0) & \varphi(-x) \\ \varphi(x) & \varphi(0) \end{pmatrix} \quad \text{en} \quad \begin{pmatrix} \varphi(0) & \overline{\varphi(x)} & \overline{\varphi(y)} \\ \varphi(x) & \varphi(0) & \varphi(x-y) \\ \varphi(y) & \overline{\varphi(x-y)} & \varphi(0) \end{pmatrix}.$$

(a) and (b) Since the first one of these two matrices is positive-hermitian it follows that:

$$\varphi(-x) = \overline{\varphi(x)} \quad \text{en} \quad |\varphi(x)| \leqslant \varphi(0).$$

(c) Since the second matrix is positive-hermitian, we obtain by the choice of the constants a_1, a_2 and a_3:

$$a_1 = 1, \quad a_2 = \frac{\lambda\,|\varphi(x)-\varphi(y)|}{\varphi(x)-\varphi(y)}, \quad a_3 = -a_2$$

the following inequality for all $\lambda \in \mathbb{R}$:

$$\varphi(0)\left(1+2\lambda^2\right) + 2\lambda\,|\varphi(x)-\varphi(y)| - 2\lambda^2 \mathrm{Re}\ \varphi(x-y) \geqslant 0. \tag{5.9}$$

The inequality in (c) is a consequence of (5.9).

(d) The determinant of a positive hermitian matrix is non-negative. So that, if the 3×3 matrix

$$\begin{pmatrix} 1 & \lambda & \mu \\ \overline{\lambda} & 1 & \xi \\ \overline{\mu} & \overline{\xi} & 1 \end{pmatrix} \tag{5.10}$$

is positive-hermitian, then we get the inequality

$$1 + \lambda\overline{\mu}\xi + \overline{\lambda}\mu\overline{\xi} \geqslant |\lambda|^2 + |\mu|^2 + |\xi|^2,$$

which is equivalent with

$$\left|\xi - \overline{\lambda}\mu\right|^2 \leqslant \left(1-|\lambda|^2\right)\left(1-|\mu|^2\right). \tag{5.11}$$

The inequality in (d) then follows from (5.11) by associating the second matrix with the matrix in (5.10) and by employing (5.11).

The proof of Theorem 5.15 is complete now. ☐

5.16. Proposition. *Let g be a function in $L^1(\mathbb{R}^\nu)$. Then the following equalities hold:*

$$\textit{spectral radius of } (g) = \lim_{n\to\infty} \left\|g^{*n}\right\|_1^{1/n} = \|\widehat{g}\|_\infty.$$

In the theory of Banach algebras the Beurling-Gelfand formula gives a relationship between the spectral radius and the norm of an element. More precisely, let $(A, \|\cdot\|)$ be a Banach algebra with unit e. A Banach algebra is a Banach space with a multiplication $(x,y) \mapsto xy$ which satisfies the usual axioms of distributivity and scalar multiplication. The norm satisfies $\|xy\| \leqslant \|x\| \cdot \|y\|$, x, $y \in A$, $\|e\| = 1$. By definition, the spectrum $\sigma(x)$ of an element $x \in A$ is given by $\sigma(x) = \{\lambda \in \mathbb{C} : \lambda e - x \notin G(A)\}$. Here $G(A)$ is the group of invertible elements of A: $x \in G(A)$ if and only if there exists a (unique) element $y \in A$ such that $xy = yx = e$. Then $\sigma(x)$ is a non-empty compact subset of $\mathbb{C}$ contained in the disc of radius $\|x\|$: $\sigma(x) \subset \{\lambda \in \mathbb{C} : |\lambda| \leqslant \|x\|\}$. In fact we have the Beurling-Gelfand formula for the spectral radius:

$$\sup_{\lambda \in \sigma(x)} |\lambda| = \limsup_{n \to \infty} \|x^n\|^{1/n} = \inf_{n \in \mathbb{N}} \|x^n\|^{1/n}, \quad x \in A. \tag{5.12}$$

Let $A = L^1(\mathbb{R}^\nu) \bigoplus \mathbb{C}\delta$, where δ is the Dirac measure at zero, with a multiplication given by the convolution product:

$$(f + \alpha\delta) * (g + \beta\delta) = f * g + \alpha g + \alpha\beta, \quad f,\, g \in L^1(\mathbb{R}^\nu),\ \alpha,\, \beta \in \mathbb{C},$$

and with the norm given by $\|f + \alpha\delta\| = \|f\|_{L^1} + |\alpha|$, $f \in L^1(\mathbb{R}^\nu)$, $\alpha \in \mathbb{C}$. Here $f * g(x) = \int f(y)g(x-y)\,dy$. Then A is a commutative Banach algebra with unit δ. The spectral radius $\rho(f)$ of $f \in L^1(\mathbb{R}^\nu)$ is given by the supremum norm of its Fourier transform:

$$\rho(f) = \limsup_{n \to \infty} \|f^{*n}\|_{L^1}^{1/n} = \inf_{n \in \mathbb{N}} \|f^{*n}\|_{L^1}^{1/n} = \sup_{x \in \mathbb{R}^\nu} \left|\widehat{f}(x)\right|,$$

where $\widehat{f}(x) = \int e^{-ix\cdot y} f(y)\,dy$. The interested reader can find more information in Bonsall and Duncan [**28**], in Yosida [**197**], and in several other places like Lax [**106**].

Proof of Proposition 5.16. For a proof we refer the reader to a book on functional analysis with Banach algebras as a topic. Good references are Rudin [**153**], Theorem 11.9 together with Example (e), and Folland [**71**], Theorem 1.30 combined with Theorem 4.2. ☐

The following theorem is a very important representation theorem. It will be used in Theorem 5.25 and in the continuity theorem of Lévy: Theorem 5.42.

5.17. Theorem (Bochner). *Let* $\varphi : \mathbb{R}^\nu \to \mathbb{C}$ *be a function. The following assertions are equivalent:*

(i) *The function* φ *is continuous and positive-definite;*
(ii) *There exists a positive Borel measure* μ *op* $\mathbb{R}^\nu$ *such that* $\varphi = \widehat{\mu}$.

The Borel measure μ *in* (ii) *is unique.*

Proof. (i) ⇒ (ii). Define the linear functional $\Lambda : \mathcal{M} \to \mathbb{C}$ by means of the equality: $\Lambda(\nu) = \int \varphi(x) d\nu(x)$, $\nu \in \mathcal{M}$. Define the involution $\nu \mapsto \widetilde{\nu}$ via the equality: $\widetilde{\nu}(A) = \overline{\nu(-A)}$. Because, by hypothesis, the function φ is positive-definite we see that the functional Λ is positive in the sense that $\Lambda(\nu * \widetilde{\nu}) \geqslant 0$ for all $\nu \in \mathcal{M}$: see inequality (5.26) in Proposition 5.23 further on. By Cauchy-Schwartz inequality we then obtain

$$\begin{aligned}|\Lambda(\nu)| &= |\Lambda(\nu * \delta)| \leqslant \left(\Lambda\left(\nu * \widetilde{\nu}\right)\right)^{1/2} \left(\Lambda\left(\delta * \widetilde{\delta}\right)\right)^{1/2} \\ &\leqslant \left(\Lambda\left(\nu * \widetilde{\nu}\right)\right)^{1/2} \varphi(0)^{1/2}\end{aligned}$$

(by inductionwith respect to n)

$$\begin{aligned}&\leqslant \left(\Lambda\left(\left(\nu * \widetilde{\nu}\right)^{*2^n}\right)\right)^{1/2^{n+1}} \varphi(0)^{\sum_{j=1}^{n+1} 2^{-j}} \\ &\leqslant \|\varphi\|_\infty^{1/2^{n+1}} \left\|\left(\nu * \widetilde{\nu}\right)^{*2^n}\right\|^{1/2^{n+1}} \varphi(0)^{\sum_{j=1}^{n+1} 2^{-j}}. \qquad (5.13)\end{aligned}$$

By letting n tend to ∞ in (5.13) we deduce

$$\begin{aligned}|\Lambda(\nu)| &\leqslant \liminf_{n\to\infty} \left\|\left(\nu * \widetilde{\nu}\right)^{*2^n}\right\|^{1/2^{n+1}} \varphi(0) \\ &= \sqrt{\text{spectral radius of } \nu * \widetilde{\nu}}\,\varphi(0). \qquad (5.14)\end{aligned}$$

By applying (5.13) and (5.14) to a measure ν of the form $\nu(B) = \int_B f(x) dx$, where f belongs to $L^1(\mathbb{R}^\nu)$ we obtain

$$\left|\int \varphi(x) f(x) dx\right| \leqslant \sqrt{\text{spectral radius of } f * \widetilde{f}}\,\varphi(0). \qquad (5.15)$$

In (5.15) we wrote $\widetilde{f}(x) = \overline{f(-x)}$ and $f * g(x) = \int f(y) g(x-y) dy$, for f and g belonging to $L^1(\mathbb{R}^\nu)$. Next we realize that $L^1(\mathbb{R}^\nu)$, equipped with the L^1-norm and the convolution product $*$, is a Banach-algebra and that the spectral radius of an L^1-function f is given by the supremum-norm the Fourier transform of f: see Proposition 5.16. From (5.15) we infer

$$\left|\int \varphi(x) f(x) dx\right| \leqslant \sqrt{\left\|\widehat{f * \widetilde{f}}\right\|_\infty}\,\varphi(0) \leqslant \left\|\widehat{f}\right\|_\infty \varphi(0). \qquad (5.16)$$

Next define $\Lambda_0 : \left\{\widehat{f} : f \in L^1(\mathbb{R}^\nu)\right\} \to \mathbb{C}$ via the equality $\Lambda_0(\widehat{f}) = \int \varphi(x) f(x) dx$, $f \in L^1(\mathbb{R}^\nu)$. From (5.16) it follows that the functional Λ_0 has a unique extension as a continuous linear functional, which we call again Λ_0, on the uniform closure of the subalgebra $\left\{\widehat{f} : f \in L^1(\mathbb{R}^\nu)\right\}$. By the Stone Weierstrass theorem (Theorem 5.14) this closure coincides with $C_0(\mathbb{R}^\nu)$. The Riesz representation theorem applies to the effect that there exists a bounded Borel measure μ such that $\Lambda_0(\widehat{f}) = \int \widehat{f}(x) d\mu(x)$, $f \in L^1(\mathbb{R}^\nu)$. From this it follows that

$$\int \varphi(x) f(x) dx = \Lambda_0(\widehat{f}) = \int \widehat{f}(y) d\mu(y) = \int \widehat{\mu}(x) f(x)\, dx.$$

Consequently, $\varphi = \widehat{\mu}$. The function φ being positive-definite it follows that the measure μ is positive. This proves the implication (i) $\Longrightarrow$ (ii).

(ii) $\Rightarrow$ (i). Let μ be a finite positive Borel measure. Then its Fourier transform $\widehat{\mu}$ is a uniformly continuous positive-definite function. The proof of these assertions is left to the reader.

The proof of Theorem 5.17 is complete now. ☐

An alternative proof runs as follows: the idea is taken from Theorem 5.10 in Lőrinczi et al [**115**]. We need the following lemmas.

5.18. LEMMA. *Let* $\varphi : \mathbb{R}^\nu \to \mathbb{C}$ *be a (uniformly) continuous positive-definite function, and fix* $t > 0$. *Then the function* $\xi \mapsto e^{-\frac{1}{2}t|\xi|^2}\varphi(\xi)$ *is also (uniformly) continuous and positive-definite.*

PROOF. Let ξ_j, $1 \leqslant j \leqslant n$, belong to $\mathbb{R}^\nu$, and let λ_j, $1 \leqslant j \leqslant n$, be complex numbers. Then

$$\begin{aligned} &\sum_{j,k=1}^{n} \lambda_j \overline{\lambda_k} e^{-\frac{1}{2}t|\xi_j-\xi_k|^2} \varphi\left(\xi_j - \xi_k\right) \\ &= \frac{1}{\left(\sqrt{2\pi t}\right)^\nu} \int_{\mathbb{R}^\nu} \sum_{j,k=1}^{n} \lambda_j e^{i\xi_j\cdot y} \overline{\lambda_k e^{i\xi_k\cdot y}} \varphi\left(\xi_j - \xi_k\right) e^{-|y|^2/(2t)}\, dy \geqslant 0. \end{aligned} \tag{5.17}$$

The claim in Lemma 5.18 follows from (5.17). ☐

5.19. LEMMA. *Let* $\psi : \mathbb{R}^\nu \to \mathbb{C}$ *be a function which belongs to* $L^1(\mathbb{R}^\nu)$, *and let* V_1 *be a bounded open neighborhood of the origin in* $\mathbb{R}^\nu$. *Put* $V_n = nV_1$, $n \in \mathbb{N}$. *Let* $m(V_n) = \int \mathbf{1}_{V_n}(\xi)\, d\xi = n^\nu m(V_1)$ *be the Lebesgue measure of* V_n. *Then, uniformly in* $x \in \mathbb{R}^\nu$,

$$\int_{\mathbb{R}^\nu} e^{i\xi\cdot x}\psi(\xi)\, d\xi = \lim_{n\to\infty} \frac{\int_{V_n}\int_{V_n} e^{i(\xi-\eta)\cdot x}\psi\left(\xi-\eta\right)\, d\xi\, d\eta}{m\left(V_n\right)}. \tag{5.18}$$

Proof. By employing standard properties, like translation invariance and the homothety property of the Lebesgue measure, we deduce the following equalities:

$$\begin{aligned}
&\int_{\mathbb{R}^\nu} e^{i\xi\cdot x}\psi(\xi)\,d\xi - \frac{\int_{V_n}\int_{V_n} e^{i(\xi-\eta)\cdot x}\psi\left(\xi-\eta\right)\,d\xi\,d\eta}{m\left(V_n\right)}\\
&= \int_{\mathbb{R}^\nu} e^{i\xi\cdot x}\psi(\xi)\,d\xi - \frac{\int_{V_n}\int_{V_n-\eta} e^{i\xi\cdot x}\psi\left(\xi\right)\,d\xi\,d\eta}{m\left(V_n\right)}\\
&= \frac{\int_{V_n}\int_{\mathbb{R}^\nu\setminus(V_n-\eta)} e^{i\xi\cdot x}\psi\left(\xi\right)\,d\xi\,d\eta}{m\left(V_n\right)} = \frac{\int_{V_1}\int_{\mathbb{R}^\nu\setminus(nV_1-n\eta)} e^{i\xi\cdot x}\psi\left(\xi\right)\,d\xi\,d\eta}{m\left(V_1\right)}.
\end{aligned} \tag{5.19}$$

From (5.19) we infer

$$\left|\int_{\mathbb{R}^\nu} e^{i\xi\cdot x}\psi(\xi)\,d\xi - \frac{\int_{V_n}\int_{V_n} e^{i(\xi-\eta)\cdot x}\psi\left(\xi-\eta\right)\,d\xi\,d\eta}{m\left(V_n\right)}\right| \leqslant \frac{\int_{V_1}\int_{\mathbb{R}^\nu\setminus(nV_1-n\eta)} \left|\psi\left(\xi\right)\right|\,d\xi\,d\eta}{m\left(V_1\right)}. \tag{5.20}$$

Hence, by using the Lebesgue's dominated convergence theorem the equality in (5.18) is readily established. Moreover, this limit is uniform in $x \in \mathbb{R}^\nu$. This completes the proof of Lemma 5.19. ☐

5.20. Lemma. *Let $\psi : \mathbb{R}^\nu \to \mathbb{C}$ be a continuous positive-definite function which belongs to $L^1\left(\mathbb{R}^\nu\right)$. Then, for all $x \in \mathbb{R}^\nu$ the inequality $\int_{\mathbb{R}^\nu} e^{i\xi\cdot x}\psi(\xi)\,d\xi \geqslant 0$ holds.*

Proof of Lemma 5.20. Since the function ψ is positive-definite and continuous the right-hand side of (5.18) is non-negative. So the assertion in Lemma 5.20 follows from Lemma 5.19. ☐

5.21. Lemma. *Let* $\psi : \mathbb{R}^\nu \to \mathbb{C}$ *be a continuous positive-definite function which belongs to* $L^1(\mathbb{R}^\nu)$*, and let* μ *be a bounded complex-valued Borel measure on* $\mathbb{R}^\nu$ *with Fourier transform* $\widehat{\mu}(x) = \int_{\mathbb{R}^\nu} e^{-ix\cdot y}\, d\mu(y)$*. The the following equality holds:*

$$\int_{\mathbb{R}^\nu} \psi(\xi)\, d\mu(\xi) = \frac{1}{(2\pi)^\nu} \int_{\mathbb{R}^\nu} \int_{\mathbb{R}^\nu} e^{i\xi\cdot x} \psi(\xi)\, d\xi\, \widehat{\mu}(x)\, dx. \tag{5.21}$$

If $\varphi : \mathbb{R}^\nu \to \mathbb{C}$ *is an arbitrary continuous positive-definite function, and if* μ *is a bounded complex-valued Borel measure on* $\mathbb{R}^\nu$*, then*

$$\int_{\mathbb{R}^\nu} \varphi(\xi)\, d\mu(\xi) = \lim_{t\downarrow 0} \frac{1}{(2\pi)^\nu} \int_{\mathbb{R}^\nu} \int_{\mathbb{R}^\nu} e^{i\xi\cdot x} e^{-\frac{1}{2}t|\xi|^2} \varphi(\xi)\, d\xi\, \widehat{\mu}(x)\, dx, \tag{5.22}$$

and

$$\left|\int_{\mathbb{R}^\nu} \varphi(\xi)\, d\mu(\xi)\right| \leqslant \varphi(0) \sup_{x\in\mathbb{R}^\nu} \left|\widehat{\mu}(x)\right|. \tag{5.23}$$

Proof. From Fubini's theorem we get

$$\begin{aligned}
&\frac{1}{(2\pi)^\nu} \int_{\mathbb{R}^\nu} \int_{\mathbb{R}^\nu} e^{i\xi\cdot x} \psi(\xi)\, d\xi\, \widehat{\mu}(x)\, dx \\
&= \frac{1}{(2\pi)^\nu} \int_{\mathbb{R}^\nu} \int_{\mathbb{R}^\nu} e^{i\xi\cdot x} \psi(\xi)\, d\xi \int_{\mathbb{R}^\nu} e^{-ix\cdot y}\, d\mu(y)\, dx \\
&= \int_{\mathbb{R}^\nu} \int_{\mathbb{R}^\nu} \left(\frac{1}{(2\pi)^\nu} \int_{\mathbb{R}^\nu} e^{i\xi\cdot x} \psi(\xi)\, d\xi \right) e^{-ix\cdot y}\, dx\, d\mu(y) \\
&= \int_{\mathbb{R}^\nu} \mathcal{F}\mathcal{F}^{-1}(\psi)(y)\, d\mu(y) = \int_{\mathbb{R}^\nu} \psi(y)\, d\mu(y),
\end{aligned} \tag{5.24}$$

where $\mathcal{F}$ denotes the Fourier transform with inverse $\mathcal{F}^{-1}$. The equalities in (5.24) imply the equality in (5.21). In order to prove he equality in (5.22) we first observe that by Lemma 5.18 the functions of the form $\xi \mapsto \varphi_t(\xi) := e^{-\frac{1}{2}t|\xi|^2} \varphi(\xi)$, $t > 0$, are positive-definite and continuous, because φ is so. Applying the equality in (5.21) to the function φ_t shows

$$\begin{aligned}
\int_{\mathbb{R}^\nu} \varphi(\xi)\, d\mu(\xi) &= \lim_{t\downarrow 0} \int_{\mathbb{R}^\nu} e^{-\frac{1}{2}t|\xi|^2} \varphi(\xi)\, d\mu(\xi) \\
&= \lim_{t\downarrow 0} \frac{1}{(2\pi)^\nu} \int_{\mathbb{R}^\nu} \int_{\mathbb{R}^\nu} e^{i\xi\cdot x} e^{-\frac{1}{2}t|\xi|^2} \varphi(\xi)\, d\xi\, \widehat{\mu}(x)\, dx.
\end{aligned} \tag{5.25}$$

The equality in (5.22) follows from (5.25). Finally, the inequality in (5.23) follows from (5.22) and Lemma 5.20. So the proof of Lemma 5.21 is complete now. ☐

Second proof of Theorem 5.17. Let $\mathcal{M} = \mathcal{M}(\mathbb{R}^\nu)$ be the collection of bounded complex Borel measures on $\mathbb{R}^{|nu}$, and consider the functional

$$\Lambda_\varphi : \widehat{\mu} \mapsto \int_{\mathbb{R}^\nu} \varphi(\xi)\, d\mu(\xi), \quad \mu \in \mathcal{M}.$$

Then Λ_φ can be extended to the uniform closure of the collection $\{\widehat{\mu} : \mu \in \mathcal{M}\}$ such that $|\Lambda_\varphi(f)| \leqslant \varphi(0)\,\|f\|_\infty$ for all f in this closure. This closure contains all constant functions and all continuous functions on $\mathbb{R}^\nu$ which tend to 0 at ∞. By the Riesz representation theorem there exists a positive measure μ_φ on the Borel field of $\mathbb{R}^\nu$ such that

$$\int_{\mathbb{R}^\nu} \varphi(\xi)\,d\mu(\xi) = \int_{\mathbb{R}^\nu} \widehat{\mu}(x)\,d\mu_\varphi(x) = \int_{\mathbb{R}^\nu}\int_{\mathbb{R}^\nu} e^{-i\xi\cdot x}\,d\mu(\xi)\,d\mu_\varphi(x)$$
$$= \int_{\mathbb{R}^\nu}\int_{\mathbb{R}^\nu} e^{-i\xi\cdot x}\,d\mu_\varphi(x)\,d\mu(\xi) = \int_{\mathbb{R}^\nu} \widehat{\mu}_\varphi(\xi)\,d\mu(\xi),$$

for all $\mu \in \mathcal{M}$. It follows that $\varphi(\xi) = \widehat{\mu}_\varphi(\xi)$. This completes the proof of the theorem of Bochner: Theorem 5.17. ☐

5.22. Lemma. *Let $\varphi : \mathbb{R}^\nu \to \mathbb{C}$ be a continuous function, and let μ be a complex Borel measure on $\mathbb{R}^\nu$ with compact support. So $|\mu|\,(\mathbb{R}^\nu \backslash K) = 0$ for some compact subset K of $\mathbb{R}^\nu$. Then*

$$\inf\left\{\left|\int\int \varphi(x-y)d\mu(x)d\overline{\mu}(y) - \sum_{j,k=1}^{n} a_j\overline{a}_k\varphi\left(x_j - x_k\right)\right|\right\} = 0,$$

where the infimum is taken over all $a_j \in \mathbb{C}$, $x_j \in K_0$, $1 \leqslant j \leqslant n$, $n \in \mathbb{N}$, and where K_0 is the smallest compact set K with the property that $|\mu|\,(\mathbb{R}^\nu \backslash K) = 0$.

Proof. Fix $\epsilon > 0$, and choose a partition $(U_j : 1 \leqslant j \leqslant n)$ of K_0 with the property that

$$\left|\varphi\left(x-y\right) - \varphi\left(x'-y'\right)\right| \leqslant \frac{\epsilon}{|\mu|\left(K_0\right)^2},$$

x, $x' \in U_j$ and y, $y' \in U_k$, and write $a_j = \mu(U_j)$. Then for $x_j \in U_j$, $1 \leqslant j \leqslant n$, we have

$$\left|\iint \varphi(x-y)d\mu(x)d\overline{\mu}(y) - \sum_{j,k=1}^{n} a_j\overline{a}_k\varphi\left(x_j - x_k\right)\right|$$
$$= \left|\sum_{j,k=1}^{n}\int_{U_j}\int_{U_k}\left(\varphi(x-y) - \varphi\left(x_j - x_k\right)\right)d\mu(x)d\overline{\mu}(y)\right|$$
$$\leqslant \sum_{j,k=1}^{n}\int_{U_j}\int_{U_k}\left|\varphi(x-y) - \varphi\left(x_j - x_k\right)\right| d\,|\mu|\,(x)d\,|\overline{\mu}|\,(y)$$
$$\leqslant \frac{\epsilon}{|\mu|\left(K_0\right)^2}\sum_{j,k=1}^{n}\int_{U_j}\int_{U_k} d\,|\mu|\,(x)d\,|\overline{\mu}|\,(y) = \epsilon.$$

This proves Lemma 5.22. ☐

5.23. Proposition. *(a) Let $\varphi : \mathbb{R}^\nu \to \mathbb{C}$ be a continuous function. The following assertions are equivalent.*

(i) *The function φ is positive-definite;*
(ii) *For every function $f \in C_{00}(\mathbb{R}^\nu)$ the inequality $\int \varphi(x) f * \widetilde{f}(x)dx \geqslant 0$ holds;*

(iii) *Every Borel measure μ with compact support satisfies the inequality:*

$$\int \varphi(x) d\left(\mu * \widetilde{\mu}\right)(x) \geqslant 0. \tag{5.26}$$

(b) If φ is positive-definite and if μ is a bounded complex Borel measure on $\mathbb{R}^\nu$, then inequality (5.26) in (iii) also holds.

Proof. (a) (iii) $\Rightarrow$ (ii). Choose μ of the form $\mu(B) = \int_B f(x)dx$, with $f \in C_{00}(\mathbb{R}^\nu)$ fixed.

(ii) $\Rightarrow$ (i). Let μ be of the form $\mu = \sum_{j=1}^n a_j \delta_{x_j}$. Approximate the de Dirac measures δ_{x_j} by measures of the form $B \mapsto \int_B \int f_{j,N}(x)dx$ in the sense that

$$\lim_{N\to\infty} \int \varphi(x) d\left(\mu_N * \widetilde{\mu_N}\right)(x) = \int \varphi(x) d\left(\mu * \widetilde{\mu}\right)(x) = \sum_{j,k=1}^n a_j \overline{a}_k \varphi\left(x_j - x_k\right).$$

Here the measure μ_N is defined by $\mu_N(B) = \sum_{j=1}^n a_j \int_B f_{j,N}(x)dx$, $B \in \mathcal{B}(\mathbb{R}^\nu)$.

(i) $\Rightarrow$ (iii). Let μ be a Borel measure of compact support. Then there exists a sequence of measures $(\mu_N : N \in \mathbb{N})$,where every μ_N is of the form $\mu_N = \sum_{j=1}^N a_{j,N} \delta_{x_{j,N}}$ and where

$$\begin{aligned}\int \varphi(x) d\left(\mu * \widetilde{\mu}\right)(x) &= \lim_{N\to\infty} \int \varphi(x) d\left(\mu_N * \widetilde{\mu_N}\right)(x) \\ &= \lim_{N\to\infty} \sum_{j,k=1}^N a_{j,N} \overline{a}_{k,N} \varphi\left(x_{j,n} - x_{k,N}\right) \geqslant 0.\end{aligned}$$

That such a sequence of measures exists $(\mu_N : N \in \mathbb{N})$ follows from Lemma 5.22.

(b) Let $(K_m : m \in \mathbb{N})$ be an increasing sequence of compact subsets of $\mathbb{R}^\nu$ such that $\mathbb{R}^\nu = \bigcup_{m=1}^\infty K_m$ and such that $K_m \subset \text{interior}\left(K_{m+1}\right)$ for all $m \in \mathbb{N}$. Since, in addition,

$$\int \varphi(x) d\left(\mu * \widetilde{\mu}\right)(x) = \lim_{m\to\infty} \int \varphi(x) d\left(\left(1_{K_m}\mu\right) * \left(\widetilde{\left(1_{K_m}\mu\right)}\right)\right)(x)$$

assertion (b) follows from the results in (a).

This completes the proof of Proposition 5.23. □

5.24. Definition. The *weak topology* (or *Bernoulli topology*) on $\mathcal{M}$ is the locally convex topology $\sigma\left(\mathcal{M}, C_b(\mathbb{R}^\nu)\right)$. Let $\mu_0 \in \mathcal{M}$. So that every $\sigma\left(\mathcal{M}, C_b(\mathbb{R}^\nu)\right)$-neighborhood of μ_0 contains a neighborhood of the form

$$\bigcap_{j=1}^n \left\{\mu \in \mathcal{M} : \left|\int f_j d\left(\mu - \mu_0\right)\right| < \epsilon_j\right\}. \tag{5.27}$$

Here, the functions $f_1, \ldots, f_n$ are bounded and continuous, and the numbers $\epsilon_1, \ldots, \epsilon_n$ are strictly positive. A net $(\mu_\alpha : \alpha \in \mathcal{A})$ $\mathcal{M}$ converges to the measure μ for the topology $\sigma\left(\mathcal{M}, C_b(\mathbb{R}^\nu)\right)$ if $\lim_\alpha \int f d\mu_\alpha = \int f d\mu$ for all $f \in C_b(\mathbb{R}^\nu)$.

We write $\mu = \text{weak-}\lim_\alpha \mu_\alpha$. The space $\mathcal{M}$ can also be supplied with the *vague topology*. This is the locally convex topology $\sigma\left(\mathcal{M}, C_{00}(\mathbb{R}^\nu)\right)$. For the vague topology the functions $f_1, \dots, f_n$ in (5.27) are required to belong to $C_{00}(\mathbb{R}^\nu)$ and the net $(\mu_\alpha : \alpha \in \mathcal{A})$ converges to $\mu \in \mathcal{M}$ provided $\lim_\alpha \int f d\mu_\alpha = \int f d\mu$ for all $f \in C_{00}(\mathbb{R}^\nu)$. We write $\mu = \text{vague-}\lim_\alpha \mu_\alpha$.

Let $\mathcal{M}^+ := \{\mu \in \mathcal{M} : \mu \geqslant 0\}$ and let

$$CP := CP(\mathbb{R}^\nu) = \{\varphi \in C_b(\mathbb{R}^\nu) : \varphi \text{ positive-definite}\} .$$

The following theorem expresses the fact that the set $\mathcal{M}^+$, endowed with the weak topology and CP, endowed with the compact-open topology $\mathcal{T}$, are *homeomorphic*. The compact-open topology is also called the topology of uniform convergence on compact subsets of $\mathbb{R}^\nu$. So that a net $(\varphi_\alpha : \alpha \in \mathcal{A})$ converges to φ, if $\lim_\alpha \sup_{x \in K} |\varphi_\alpha(x) - \varphi(x)| = 0$ for every compact subset K of $\mathbb{R}^\nu$.

5.25. Theorem. *The Fourier transform $\mu \mapsto \widehat{\mu}$, $\mu \in \mathcal{M}^+$, is a homeomorphism from*

$$\left(\mathcal{M}^+, \sigma\left(\mathcal{M}^+, C_b(\mathbb{R}^\nu)\right)\right) \text{ onto } (CP, \mathcal{T}).$$

Proof. Let $(\mu_\alpha : \alpha \in \mathcal{A})$ be a net in $\mathcal{M}^+$ that weakly converges to $\mu \in \mathcal{M}^+$ relative to the weak topology. We will prove that the net $(\widehat{\mu}_\alpha : \alpha \in \mathcal{A})$ converges uniformly on compact subsets to $\widehat{\mu}$. Fix $\epsilon > 0$. Then choose $\delta > 0$ in such a way that $\delta\left(3 + \mu(\mathbb{R}^\nu)\right) < \epsilon$ and choose a function $f \in C_{00}(\mathbb{R}^\nu)$ such that

$$0 \leqslant f \leqslant 1 \quad \text{and} \quad \int (1-f)d\mu < \delta.$$

Since weak-$\lim \mu_\alpha = \mu$ there exists $\alpha_0 \in \mathcal{A}$ such that

$$\mu_\alpha(\mathbb{R}^\nu) = \int 1 d\mu_\alpha < \int 1 d\mu + 1 = \mu(\mathbb{R}^\nu) + 1 \quad \text{en} \quad \int (1-f)\, d\mu_\alpha < \delta$$

for all $\alpha \geqslant \alpha_0$. Define the zero-neighborhood V by

$$V = \left\{x \in \mathbb{R}^\nu : \left|1 - \exp\left(-i\langle x, y\rangle\right)\right| \leqslant \delta : \text{ for all } y \in \operatorname{supp}(f)\right\}.$$

Then for those $\alpha \in \mathcal{A}$ and those x_1 and $x_2 \in \mathbb{R}^\nu$ which satisfy $\alpha \geqslant \alpha_0$ and $x_1 - x_2 \in V$ the following inequalities hold:

$$\begin{aligned}
&\left|\widehat{\mu}_\alpha(x_1) - \widehat{\mu}_\alpha(x_2)\right| \leqslant \int \left|\exp\left(-i\langle x_1, y\rangle\right) - \exp\left(-i\langle x_2, y\rangle\right)\right| d\mu_\alpha(y) \\
&\leqslant \int \left|1 - \exp\left(-i\langle x_1 - x_2, y\rangle\right)\right| f(y) d\mu_\alpha(y) \\
&\qquad + \int \left|1 - \exp\left(-i\langle x_1 - x_2, y\rangle\right)\right| \left(1 - f(y)\right) d\mu_\alpha(y) \\
&\leqslant \delta \int f(y) d\mu_\alpha(y) + 2\int \left(1 - f(y)\right) d\mu_\alpha(y) \\
&\leqslant \delta\left(\mu(\mathbb{R}^\nu) + 1\right) + 2\delta \leqslant \epsilon. \qquad (5.28)
\end{aligned}$$

By (5.28) it follows that $\left|\widehat{\mu}(x_1) - \widehat{\mu}(x_2)\right| \leqslant \epsilon$ for x_1 and $x_2 \in \mathbb{R}^\nu$ for which $x_1 - x_2 \in V$. Next choose a compact subset K in $\mathbb{R}^\nu$. Then there exist $y_1, \dots, y_n$ in $\mathbb{R}^\nu$ such that $K \subseteq \bigcup_{j=1}^n (y_j + V)$ and thee exist $\alpha_j \in \mathcal{A}$, $1 \leqslant j \leqslant n$, such that

$$\left|\widehat{\mu}_\alpha(y_j) - \widehat{\mu}(y_j)\right| \leqslant \epsilon \quad \text{for} \quad \alpha \geqslant \alpha_j, \quad j = 1, \dots, n.$$

Then choose $\alpha' \in \mathcal{A}$ in such a way that $\alpha' \geqslant \alpha_j$ for $j = 1, \dots, n$. For $x \in y_j + V$ and $\alpha \geqslant \alpha'$ we get

$$\left|\widehat{\mu}_\alpha(x) - \widehat{\mu}(x)\right| \leqslant \left|\widehat{\mu}_\alpha(x) - \widehat{\mu}_\alpha(y_j)\right| + \left|\widehat{\mu}_\alpha(y_j) - \widehat{\mu}(y_j)\right| + \left|\widehat{\mu}(y_j) - \widehat{\mu}(x)\right| \leqslant \epsilon$$

and hence

$$\sup_{x \in K} \left|\widehat{\mu}_\alpha(x) - \widehat{\mu}(x)\right| \leqslant 3\epsilon.$$

This proves that the Fourier transform is continuous for the indicated topologies. Conversely, suppose that the net $(\widehat{\mu}_\alpha : \alpha \in \mathcal{A})$ converges uniformly on compact subsets to $\widehat{\mu}$. Then we will show the following two equalities:

(a) $\lim \mu_\alpha(\mathbb{R}^\nu) = \mu(\mathbb{R}^\nu)$;
(b) $\lim \int \varphi(x) d\mu_\alpha(x) = \int \varphi(x) d\mu(x)$ for all functions $\varphi \in C_{00}(\mathbb{R}^\nu)$.

From Theorem 5.26 below it then follows that weak-$\lim \mu_\alpha = \mu$. The equality in (a) follows from:

$$\lim \mu_\alpha\left(\mathbb{R}^\nu\right) = \lim \widehat{\mu}_\alpha(0) = \widehat{\mu}(0) = \mu\left(\mathbb{R}^\nu\right).$$

Let $\epsilon > 0$ be arbitrary and let $\varphi \in C_{00}(\mathbb{R}^\nu)$. Choose a function $f \in C_{00}(\mathbb{R}^\nu)$ with the property that

$$\left\|\varphi - \widehat{f}\right\|_\infty \leqslant \frac{\epsilon}{2\mu\left(\mathbb{R}^\nu\right) + 1}.$$

Then we infer

$$\begin{aligned}
&\left|\int \varphi(x) d\mu_\alpha(x) - \int \varphi(x) d\mu(x)\right| \\
&\leqslant \left|\int \left(\varphi(x) - \widehat{f}(x)\right) d\mu_\alpha(x)\right| + \left|\int \widehat{f}(x) d\left(\mu_\alpha - \mu\right)(x)\right| + \left|\int \left(\widehat{f}(x) - \varphi(x)\right) d\mu(x)\right| \\
&\leqslant \left\|\varphi - \widehat{f}\right\|_\infty \left(\mu_\alpha\left(\mathbb{R}^\nu\right) + \mu\left(\mathbb{R}^\nu\right)\right) + \int \left|\widehat{\mu}_\alpha(x) - \widehat{\mu}(x)\right| \left|f(x)\right| dx \\
&\leqslant \frac{\epsilon\left(\mu_\alpha\left(\mathbb{R}^\nu\right) + \mu\left(\mathbb{R}^\nu\right)\right)}{2\mu(\mathbb{R}^\nu) + 1} + \sup_{x \in \operatorname{supp}(f)} \left|\mu_\alpha(x) - \widehat{\mu}(x)\right| \int \left|f(x)\right| dx. \qquad (5.29)
\end{aligned}$$

The inequality

$$\limsup_\alpha \left|\int \varphi(x) d\left(\mu_\alpha - \mu\right)(x)\right| \leqslant \epsilon.$$

follows from (5.29). As a consequence we see that (b) is proved now. Together with Theorem 5.26 which follows next this completes the proof of Theorem 5.25. ☐

5.26. Theorem. *A net* $(\mu_\alpha : \alpha \in \mathcal{A})$ *in* $\mathcal{M}^+$ *converges weakly to* $\mu \in \mathcal{M}^+$ *if and only if the net* $(\mu_\alpha : \alpha \in \mathcal{A})$ *converges vaguely to* μ *and if*

$$\lim_\alpha \mu_\alpha\left(\mathbb{R}^\nu\right) = \mu\left(\mathbb{R}^\nu\right). \qquad (5.30)$$

Proof. The weak topology is stronger than the vague topology and from weak convergence the equality in (5.30) also follows. Hence, the indicated conditions are necessary. Conversely, let a net $(\mu_\alpha : \alpha \in \mathcal{A})$ converge vaguely $\mathcal{M}^+$ to μ and assume that (5.30) is satisfied. We will prove that μ is the weak limit of the net $(\mu_\alpha : \alpha \in \mathcal{A})$. Therefore pick $f \in C_b(\mathbb{R}^\nu)$ and $\epsilon > 0$ arbitrary but fixed. Choose a compact subset K such that $\mu\left(\mathbb{R}^\nu \backslash K\right) < \epsilon$. In addition, choose a function $h \in C_{00}(\mathbb{R}^\nu)$ in such a way that $1_K \leqslant h \leqslant 1$. By these hypotheses the following (in-)equalities hold:

$$\lim \int (1-h) d\mu_\alpha = \int (1-h) d\mu \leqslant \mu\left(\mathbb{R}^\nu \backslash K\right) < \epsilon$$

and also

$$\lim \int f h d\mu_\alpha = \int f h d\mu.$$

Hence, there exists an $\alpha_0 \in \mathcal{A}$ such that (for $\alpha \geqslant \alpha_0$)

$$\int (1-h)\,d\mu_\alpha < \epsilon \quad \text{en} \quad \left|\int fh\,d\left(\mu_\alpha - \mu\right)\right| < \epsilon.$$

But then for $\alpha \geqslant \alpha_0$ we get

$$\begin{aligned}
&\left|\int f\,d\left(\mu_\alpha - \mu\right)\right| \\
&\leqslant \left|\int fh\,d\left(\mu_\alpha - \mu\right)\right| + \left|\int f(1-h)\,d\mu\right| + \left|\int f(1-h)\,d\mu_\alpha\right| \\
&\leqslant \epsilon\left(1 + 2\left\|f\right\|_\infty\right),
\end{aligned}$$

which shows that $\lim \int f\,d\mu_\alpha = \int f\,d\mu$.

This completes the proof of Theorem 5.26. ☐

5.27. COROLLARY. *The following assertions are true:*

(a) *The set CP is a convex cone, which is closed for the topology of uniform convergence on compact subsets.*
(b) *With φ the functions $\overline{\varphi}$ and Re φ also belong to CP.*
(c) *If φ_1 and φ_2 belong to CP, then the same is true for the product $\varphi_1\varphi_2$.*
(d) *For every $y \in \mathbb{R}^\nu$ the function $x \mapsto \exp\left(-i\left\langle x, y\right\rangle\right)$ belongs to CP. Convex combinations of such functions belong to CP.*

PROOF. The proof is left as an exercise for the reader. ☐

5.28. Definition. A function $\psi : \mathbb{R}^\nu \to \mathbb{C}$ is called *negative-definite* if for all $n \in \mathbb{N}$ and for all complex numbers $a_1, \dots, a_n$ and for all vectors $x^{(1)}, \dots, x^{(n)}$ in $\mathbb{R}^\nu$ the inequality

$$\sum_{j,k=1}^{n} a_j \overline{a}_k \left(\psi\left(x^{(j)}\right) + \overline{\psi\left(x^{(k)}\right)} - \psi\left(x^{(j)} - x^{(k)}\right)\right) \geqslant 0 \tag{5.31}$$

holds. The symbol CN denotes the collection of all continuous negative-definite functions on $\mathbb{R}^\nu$. If ψ belongs to CN, then the same is true for $\overline{\psi}$ andRe ψ. The collection CN is a convex cone. If ψ belongs to CN, then $\psi(0) \geqslant 0$ and $\psi(x) = \overline{\psi}(-x)$ for all $x \in \mathbb{R}^\nu$. A function ψ is negative-definite if and only if ψ has the following properties:

1. $\psi(0) \geqslant 0$;
2. For every $x \in \mathbb{R}^\nu$ the equality $\psi(x) = \overline{\psi}(-x)$ holds;
3. For every $n \in \mathbb{N}$ and for every n-tuple of complex numbers $a_1, \dots, a_n$, for which $\sum_{j=1}^{n} a_j = 0$, and for all vectors $x^{(1)}, \dots, x^{(n)}$ in $\mathbb{R}^\nu$ the following inequality holds:
$$\sum_{j,k=1}^{n} a_j \overline{a}_k \psi\left(x^{(j)} - x^{(k)}\right) \leqslant 0.$$

If the function ψ is negative-definite, then so is the function $\psi - \psi(0)$. If φ is positive-definite, then the function $\varphi(0) - \varphi$ is negative-definite.

The following theorem establishes an important connection between negative- and positive-definite functions.

5.29. Theorem (Schoenberg). *A function ψ belongs to CN if and only the following two conditions are satisfied:*

(i) $\psi(0) \geqslant 0$;
(ii) *For every $t > 0$ the function $\exp(-t\psi)$ is continuous and positive-definite.*

Let ψ be a negative-definite function. Then, by Bochner's theorem together with the theorem of Schoenberg, there exists for every $t > 0$ a sub-probability measure μ_t on the Borel field of $\mathbb{R}^\nu$ such that $\widehat{\mu}_t = \exp(-t\psi)$. We return to this aspect when we discuss the notion *convolution semigroup of measures*.

Proof. First suppose that ψ belongs to CN. Let $x^{(1)}, \dots, x^{(n)}$ belong to $\mathbb{R}^\nu$. Write $a_{j,k} = \psi\left(x^{(j)}\right) + \overline{\psi}\left(x^{(k)}\right) - \psi\left(x^{(j)} - x^{(k)}\right)$. Then the matrix with entries $a_{j,k}$ is positive hermitian. But then the matrix with entries $\exp(a_{j,k})$ is also positive hermitian. Let $a_1, \dots, a_n$ belong to $\mathbb{C}$ and write $a'_j = \exp\left(-\psi\left(x^{(j)}\right)\right) a_j$. Then we see

$$\sum_{j,k=1}^{n} \exp\left(-\psi\left(x^{(j)} - x^{(k)}\right)\right) a_j \overline{a}_k = \sum_{j,k=1}^{n} \exp(a_{j,k}) a'_j \overline{a'_k} \geqslant 0. \tag{5.32}$$

From (5.32) it follows that the function $\exp(-\psi)$ is then positive-definite. The same procedure can be repeated for the function $t\psi$. Conversely, if (i) and (ii) are satisfied, then, for every $t > 0$, the function $\psi_t := 1 - \exp(-t\psi) = 1 - \exp(-t\psi(0)) + \exp(-t\psi(0)) - \exp(-t\psi)$ is negative-definite. But then the function ψ is negative-definite as well, because $\psi = \lim_{t\downarrow 0} \dfrac{\psi_t}{t}$. Since

$$\psi(x) = \frac{1 - \exp(-t\psi(x))}{\int_0^t ds \exp(-s\psi(x))},$$

for $t > 0$ but small enough, we see that the function ψ is continuous at x.

So the proof of Theorem 5.29 is now complete. ☐

5.30. Definition. A family of Borel measures $(\mu_t : t \geqslant 0)$ with the following properties:

(a) $\mu_t(\mathbb{R}^\nu) \leqslant 1$ for $t > 0$;
(b) $\mu_s * \mu_t = \mu_{s+t}$ for all s and $t \geqslant 0$;
(c) $\lim_{t\downarrow 0} \int f d\mu_t = \int f d\mu_0 = f(0) = \delta_0(f)$ for all $f \in C_{00}(\mathbb{R}^\nu)$;

is called a (vaguely continuous) *convolution semigroup* of measures on $\mathbb{R}^\nu$.

The following theorem says that a vaguely continuous convolution semigroups is in fact everywhere weakly continuous.

5.31. Theorem. *There exists a one-to-one correspondence between vaguely continuous semigroups of measures and negative-definite functions.*

(a) *If* $(\mu_t : t \geqslant 0)$ *is a vaguely continuous convolution semigroup of measures, then there exists a unique continuous negative-definite function* ψ *such that* $\widehat{\mu}_t = \exp(-t\psi)$*, for all* $t \geqslant 0$.
(b) *Conversely, if* ψ *is a negative-definite function, then there exists a vaguely continuous convolution semigroup of measures* $(\mu_t : t \geqslant 0)$ *such that* $\widehat{\mu}_t = \exp(-t\psi)$ *for all* $t \geqslant 0$*. Of course, this semigroup is unique.*

Proof. (a) Define, for $t > 0$, the function ψ via the equality

$$\psi = \frac{1 - \widehat{\mu}_t}{\displaystyle\int_0^t \widehat{\mu}_s ds}. \tag{5.33}$$

Since $\widehat{\mu}_s \widehat{\mu}_t = \widehat{\mu}_{s+t}$ we see that ψ does not depend on the choice of t. Put $g(t) = \int_0^t \widehat{\mu}_s ds$. Then we see that $g(0) = 0$ and $g(t)\psi + g'(t) = 1$, and hence $g(t) = \dfrac{1 - \exp(-t\psi)}{\psi}$. From the latter it follows that $\widehat{\mu}_t = \exp(-t\psi)$. The Theorem of Schoenberg (Theorem 5.29) implies then that the function ψ is negative-definite. The functions $\widehat{\mu}_s$, $s \geqslant 0$, are continuous. So the same is true for ψ.

(b) Since ψ is a negative-definite function, the functions $\exp(-t\psi)$ are positive-definite by the theorem of Schoenberg. The theorem of Bochner (Theorem

5.17) yields the existence of sub-probability measures $(\mu_t : t \geqslant 0)$ such that $\widehat{\mu}_t = \exp(-t\psi)$. Since

$$\lim_{t \downarrow 0} \widehat{\mu}_t(\xi) = \lim_{t \downarrow 0} \exp(-t\psi(\xi)) = 1 = \widehat{\mu}_0(\xi)$$

Theorem 5.43 in the next section implies that $\lim_{t \downarrow 0} \int f d\mu_t = f(0)$ for functions $f \in C_{00}(\mathbb{R}^\nu)$.

The proof of Theorem 5.31 is now complete. $\square$

5.32. Remark. In the proof of Theorem 5.31 part (a) there is a problem if the integral $\int_0^t \widehat{\mu}_s ds$ vanishes somewhere. However, notice that $\lim_{t \downarrow 0} \dfrac{1}{t} \displaystyle\int_0^t \widehat{\mu}_s ds = \widehat{\mu}_0$ pointwise. It follows that, certainly, for $t = t(\xi) > 0$ small enough, the expression $\int_0^t \widehat{\mu}_s(\xi) ds \neq 0$. This fact can be used to circumvent this problem.

5.33. Remark. In the proof of Theorem 5.31 part (b) Theorem 5.43 of the next section was employed. This can be averted as well. Therefore consider $\widehat{f}$, with $f \in L^1(\mathbb{R}^\nu)$. Then $\lim_{t \downarrow 0} \int \widehat{f}(x) d\mu_t(x) = \lim_{t \downarrow 0} \int f(x) \widehat{\mu}_t(x) dx = \int f(x) dx = \widehat{f}(0) = \int \widehat{f} d\mu_0$. By the theorem of Stone-Weierstrass from this we obtain $\lim_{t \downarrow 0} \int f(x) d\mu_t(x) = f(0) = \int f(x) d\mu_0(x)$.

FOSS

Sharp Minds - Bright Ideas!

Employees at FOSS Analytical A/S are living proof of the company value - First - using new inventions to make dedicated solutions for our customers. With sharp minds and cross functional teamwork, we constantly strive to develop new unique products - Would you like to join our team?

FOSS works diligently with innovation and development as basis for its growth. It is reflected in the fact that more than 200 of the 1200 employees in FOSS work with Research & Development in Scandinavia and USA. Engineers at FOSS work in production, development and marketing, within a wide range of different fields, i.e. Chemistry, Electronics, Mechanics, Software, Optics, Microbiology, Chemometrics.

We offer
A challenging job in an international and innovative company that is leading in its field. You will get the opportunity to work with the most advanced technology together with highly skilled colleagues.

Read more about FOSS at www.foss.dk - or go directly to our student site www.foss.dk/sharpminds where you can learn more about your possibilities of working together with us on projects, your thesis etc.

The Family owned FOSS group is the world leader as supplier of dedicated, high-tech analytical solutions which measure and control the quality and production of agricultural, food, pharmaceutical and chemical products. Main activities are initiated from Denmark, Sweden and USA with headquarters domiciled in Hillerød, DK. The products are marketed globally by 23 sales companies and an extensive net of distributors. In line with the corevalue to be 'First', the company intends to expand its market position.

Dedicated Analytical Solutions
FOSS
Slangerupgade 69
3400 Hillerød
Tel. +45 70103370
www.foss.dk

Click on the ad to read more

5.34. PROPOSITION. *Let* $(\mu_t : t \geqslant 0)$ *be a vaguely continuous semigroup of Borel measures on* $\mathbb{R}^\nu$. *Suppose that all these measures are probability measures. Then the following assertions hold:*

(a) *weak-*$\lim_{t\to t_0, t>0} \mu_t = \mu_{t_0}$ *for all* $t_0 \geqslant 0$;
(b) $\lim_{t\to t_0} \sup_{x\in\mathbb{R}^\nu} \left|\int f(x-y)d\mu_t(y) - \int f(x-y)d\mu_{t_0}(y)\right| = 0$ *for all* $t_0 \in [0,\infty)$ *and for all functions* $f \in C_0(\mathbb{R}^\nu)$.

PROOF. (a) First we look at

$$\mu_t\left(\mathbb{R}^\nu\right) - \mu_{t_0}\left(\mathbb{R}^\nu\right) = \exp\left(-t\psi(0)\right) - \exp\left(-t_0\psi(0)\right).$$

It follows that

$$\lim_{t\to t_0} \mu_t\left(\mathbb{R}^\nu\right) - \mu_{t_0}\left(\mathbb{R}^\nu\right) = 0.$$

For the same reason we see that

$$\lim_{t\to t_0} \widehat{\mu}_t(\xi) = \lim_{t\to t_0} \exp\left(-t\psi(\xi)\right) = \exp\left(-t_0\psi(\xi)\right) = \widehat{\mu}_{t_0}(\xi).$$

By using theorem 5.43 in the next section we see that

$$\text{weak-}\lim_{t\to t_0, t>0} \mu_t = \mu_{t_0}.$$

Of course, in this proof the function ψ denotes the negative-definite function from Theorem 5.31.

(b) Let $g \in C_0\left(\mathbb{R}^\nu\right)$ be of the form $g = \widehat{f}$ with $f \in L^1(\mathbb{R}^\nu)$. Then we see

$$\begin{aligned}
&\left|\int \widehat{f}(x-y)d\mu_t(y) - \int \widehat{f}(x-y)d\mu_{t_0}(y)\right| \\
&= \left|\iint \left(\widehat{\mu}_t(-z) - \widehat{\mu}_{t_0}(-z)\right)\exp\left(-i\langle x,z\rangle\right) f(z)dz\right| \\
&\leqslant \int \left|\exp\left(-t\psi(-z)\right) - \exp\left(-t_0\psi(-z)\right)\right| |f(z)|\,dz \\
&\leqslant \int \left|\exp\left(-|t-t_0|\,\psi(-z)\right) - 1\right| |f(z)|\,dz. \qquad (5.34)
\end{aligned}$$

The assertion in (b) now follows from (5.34) together with the theorem of Stone-Weierstrass, and completes the proof of Proposition 5.34. ☐

5.35. PROPOSITION. *Let* $(\mu_t : t \geqslant 0)$ *be a vaguely continuous semigroup of probability measures on the Borel field of* $\mathbb{R}^\nu$. *Define for every n-tuple* $t_1, \dots, t_n$ *with* $0 \leqslant t_1 < \dots < t_n$, *the probability measure* $\mathbb{P}_{t_1,\dots,t_n}$ *on the Borel field of* $\left(\mathbb{R}^\nu\right)^n$ *via de formula*

$$\begin{aligned}
&\mathbb{P}_{t_1,\dots,t_n}(B) \\
&= \mu_{t_1} \otimes \mu_{t_2-t_1} \otimes \dots \otimes \mu_{t_n-t_{n-1}}\left((x_1,\dots,x_n) \in \left(\mathbb{R}^\nu\right)^n : V_n\left(x_1,\dots,x_n\right) \in B\right) \\
&= \int d\mu_{t_1}(x_1) \dots \int d\mu_{t_n-t_{n-1}}(x_n) 1_B\left(V_n\left(x_1,\dots,x_n\right)\right), \qquad (5.35)
\end{aligned}$$

where B *is a Borel subset of* $(\mathbb{R}^\nu)^n$ *and where* $V_n : (\mathbb{R}^\nu)^n \to (\mathbb{R}^\nu)^n$ *is the linear mapping given by:* $V_n : (x_1, x_2, \dots, x_n) \mapsto (x_1, x_1 + x_2, \dots, x_1 + \cdots + x_n)$*. Then the family*

$$\left\{ \left((\mathbb{R}^\nu)^n, \mathcal{B}\left(\mathbb{R}^\nu\right)^n, \mathbb{P}_{t_1,\dots,t_n} \right) : (t_1, \dots, t_n) \in [0,\infty)^n, \quad n \in \mathbb{N} \right\}$$

forms a projective system of probability measures.

Proof. Let $B \in \mathcal{B}\left(\mathbb{R}^\nu\right)^n$ and let $B' \in \mathcal{B}\left((\mathbb{R}^\nu)^{n+1}\right)$ be defined by

$$B' = \left\{ (z_1, \dots, z_{n+1}) \in (\mathbb{R}^\nu)^{n+1} : \ (z_1, \dots, z_k, z_{k+2}, \dots, z_{n+1}) \in B \right\}.$$

Let $t_1 < \cdots < t_k < s < t_{k+1} < \cdots < t_n$ be an $(n+1)$-tuple of increasing times. We have to prove the following equality:

$$\mathbb{P}_{t_1,\dots,t_k,s,t_{k+1},\dots,t_n}(B') = \mathbb{P}_{t_1,\dots,t_n}(B).$$

Since the vector

$$V_{n+1}\left(y_1, y_2, \dots, y_{n+1}\right) := (y_1, y_1 + y_2, \dots, y_1 + \cdots + y_{n+1})$$

belongs to B' if and only if the vector

$$(y_1, y_1 + y_2, \dots, y_1 + \cdots + y_k, y_1 + \cdots + y_{k+2}, \dots, y_1 + \cdots + y_{n+1})$$

belongs to B, we get what follows:

$$\begin{aligned}
&\mathbb{P}_{t_1,\dots,t_k,s,t_{k+1},\dots,t_n}\left(B'\right) \\
&= \mu_{t_1} \otimes \cdots \otimes \mu_{t_k - t_{k-1}} \otimes \mu_{s-t_k} \otimes \mu_{t_{k+1}-s} \otimes \cdots \\
&\qquad \otimes \mu_{t_n - t_{n-1}} \left\{ (y_1, \dots, y_{n+1}) : V_{n+1}\left(y_1, \dots, y_{n+1}\right) \in B' \right\} \\
&= \int d\mu_{t_1}(y_1) \dots \int d\mu_{t_k - t_{k-1}}(y_k) \int d\mu_{s-t_k}(y) \int d\mu_{t_{k+1}-s}(z) \int d\mu_{t_{k+2}-t_{k+1}}(z_{k+2}) \dots \\
&\qquad \int d\mu_{t_n - t_{n-1}}(z_n) 1_{B'}\left(V_{n+1}\left(y_1, \dots, y_k, y, z, z_{k+2}, \dots, z_n\right)\right) \\
&= \int d\mu_{t_1}(y_1) \dots \int d\mu_{t_k - t_{k-1}}(y_k) \int d\mu_{s-t_k}(y) \int d\mu_{t_{k+1}-s}(z) \int d\mu_{t_{k+2}-t_{k+1}}(z_{k+2}) \dots \\
&\qquad \int d\mu_{t_n - t_{n-1}}(z_n) 1_{B}\left(V_{n}\left(y_1, \dots, y_k, y + z, z_{k+2}, \dots, z_n\right)\right)
\end{aligned}$$

(apply Fubini's theorem, integrate relative to $\mu_{s-t_k} \otimes \mu_{t_{k+1}-s}$ and use the equality $\int g(y+z) d\mu_u(y) d\mu_v(z) = \int g\left(z_{k+1}\right) d\mu_{u+v}(z_{k+1})$)

$$\begin{aligned}
&= \int d\mu_{t_1}(y_1) \dots \int d\mu_{t_k - t_{k-1}}(y_k) \int d\mu_{t_{k+1}-t_k}(z_{k+1}) \int d\mu_{t_{k+2}-t_{k+1}}(z_{k+2}) \\
&\qquad \dots \int d\mu_{t_n - t_{n-1}}(z_n) \\
&\quad 1_B\left(V_n\left(y_1, \dots, y_k, z_{k+1}, \dots, z_n\right)\right) \\
&= \int d\mu_{t_1}(y_1) \dots \int d\mu_{t_n - t_{n-1}}(y_n) 1_B\left(y_1, \dots, y_1 + \ldots + y_n\right) \\
&= \mathbb{P}_{t_1,\dots,t_n}(B).
\end{aligned}$$

This proves the required equality in case $1 \leqslant k \leqslant n-2$. The other cases, which are $t_{n-2} < s < t_{n-1}$, $t_{n-1} < s < t_n$, $t_n < s$ and $t_1 > s$, are left as an exercise for the reader.

So the proof of Proposition 5.35 is complete now. ☐

5.36. PROPOSITION. *Let* $(\mu_t : t \geqslant 0)$ *be a vaguely continuous semigroup of probability measures on the Borel field of* $\mathbb{R}^\nu$. *Define, for every* n*-tuple* $t_1, \dots, t_n$ *the probability measure* $\mathbb{P}_{t_1,\dots,t_n}$, *where* $t_1 < \dots < t_n$, *as in Proposition 5.35. Then there exists a unique probability measure* $\mathbb{P}$ *on the product field of* $(\mathbb{R}^\nu)^{[0,\infty)}$ *such that*

$$\mathbb{P}\left((X(t_1), \dots, X(t_n)) \in B\right) = \mathbb{P}_{t_1,\dots,t_n}(B),$$

for all Borel subsets B *of* $(\mathbb{R}^\nu)^n$. *Likewise there exists, for every* $x \in \mathbb{R}^\nu$, *a unique probability measure* $\mathbb{P}_x$ *on the product field of* $(\mathbb{R}^\nu)^{[0,\infty)}$ *such that*

$$\begin{aligned}\mathbb{P}_x\left((X(t_1), \dots, X(t_n)) \in B\right) &= \mathbb{P}\left((x + X(t_1), \dots, x + X(t_n)) \in B\right) \\ &= \int d\mu_{t_1}(x_1) \otimes \dots \otimes d\mu_{t_n - t_{n-1}}(x_n) 1_B\left(x + x_1, \dots, x + x_1 + \dots + x_n\right),\end{aligned}$$

for all Borel subsets B *of* $(\mathbb{R}^\nu)^n$.

Here the state variable $X(t) : (\mathbb{R}^\nu)^{[0,\infty)} \to \mathbb{R}^\nu$ is defined by $X(t)(\omega) = \omega(t)$, where ω belongs to the product $(\mathbb{R}^\nu)^{[0,\infty)}$.

PROOF. Apply Kolmogorov's extension theorem to get the result in 5.36. ☐

5.37. Theorem. *Let* $(\mu_t : t \geqslant 0)$ *be a vaguely continuous semigroup of probability measures on the Borel field of* $\mathbb{R}^\nu$. *Define for every* n*-tuple* $t_1, \dots, t_n$ *the probability measure* $\mathbb{P}_{t_1,\dots,t_n}$, *where* $t_1 < \dots < t_n$, *as in Proposition 5.35 and let* $\mathbb{P}_x$, $x \in \mathbb{R}^\nu$, *be the unique probability measure on the product field of* $\Omega = (\mathbb{R}^\nu)^{[0,\infty)}$ *such that*

$$\begin{aligned}
&\mathbb{P}_x\left((X(t_1), \dots, X(t_n)) \in B\right)\\
&= \int d\mu_{t_1}(x_1) \dots \int d\mu_{t_n - t_{n-1}}(x_n) 1_B\left(x + x_1, \dots, x + x_1 + \dots + x_n\right). \qquad (5.36)
\end{aligned}$$

Let $\mathcal{F}_s$ *be the* σ*-field on* Ω *generated by* $X(u)$, $0 \leqslant u \leqslant s$. *For* $t > s$ *the variable* $X(t) - X(s)$ *is independent of* $\mathcal{F}_s$ *and* $X(t) - X(s)$ *possesses the same* $\mathbb{P}_x$*-distribution as* $X(t-s) - x$, *which is* μ_{t-s}.

Proof. Fix $t > s$, let $f : (\mathbb{R}^\nu)^n \to \mathbb{R}$ be a Borel measurable function, and suppose that $0 \leqslant s_1 < \dots < s_n = s$. Let $g : \mathbb{R}^\nu \to \mathbb{R}$ be another bounded Borel measurable function. Then the following equalities hold true:

$$\begin{aligned}
&\mathbb{E}\left(f\left(X(s_1), \dots, X(s_n)\right) g\left(X(t) - X(s)\right)\right)\\
&= \int d\mu_{s_1}(x_1) \dots \int d\mu_{s_n - s_{n-1}}(x_n) \int d\mu_{t-s}(x) f\left(x_1, \dots, x_1 + \dots + x_n\right) g(x)\\
&= \mathbb{E}\left(f\left(X(s_1), \dots, X(s_n)\right)\right) \mathbb{E}\left(g\left(X(t) - X(s)\right)\right).
\end{aligned}$$

Now let H be the vector space of $\mathcal{F}_s$-measurable bounded random variables Y with the property that $\mathbb{E}\left(Y g(X(t) - X(s))\right) = \mathbb{E}(Y)\, \mathbb{E}\left(g(X(t) - X(s))\right)$. Then H satisfies the hypotheses of Lemma 5.100. Whence, H contains all bounded $\mathcal{F}_s$-measurable random variables. Since, in addition, the function g is an arbitrary bounded continuous function, it follows that the state variable $X(t) - X(s)$ is independent of $\mathcal{F}_s$. This completes the proof of Theorem 5.37. ☐

5.38. Theorem. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and let* $(X(t) : t \geqslant 0)$ *be a family of state variables with state space* $\mathbb{R}^\nu$. *Assume that these state variables are measurable relative to the* σ*-fields* $\mathcal{F}$ *and* $\mathcal{B}(\mathbb{R}^\nu)$. *Suppose that*

$$\lim_{t \downarrow 0} \mathbb{E}\left[f(X(t))\right] = f(0) \text{ for all } f \in C_{00}\left(\mathbb{R}^\nu\right),$$

and also that for every $t > s$ *the variable* $X(t) - X(s)$ *is independent of the* σ*-field* $\sigma\left(X(u) : 0 \leqslant u \leqslant s\right)$ *and that* $X(t) - X(s)$ *possesses the same distribution as* $X(t-s)$. *Then the mapping* $B \mapsto \mu_t(B) := \mathbb{P}\left(X(t) \in B\right)$ *defines a vaguely continuous semigroup of probability measures on* $\mathbb{R}^\nu$.

Proof of Theorem 5.38. It is clear that every measure μ_t is a probability measure is on the Borel σ-field of $\mathbb{R}^\nu$. Since $\int f d\mu_t = \mathbb{E}\left(f(X(t))\right)$, for $f \in C_{00}\left(\mathbb{R}^\nu\right)$, the equality $\lim_{t \downarrow 0} \mathbb{E}(f(X(t))) = f(0)$, entails that the family $(\mu_t : t \geqslant t)$ is vaguely continuous at 0. The convolution property still has to be proved. It suffices to prove that $\widehat{\mu}_s(\xi) \widehat{\mu}_t(\xi) = \widehat{\mu}_{s+t}(\xi)$ for all s and $t \geqslant 0$, and for all $\xi \in \mathbb{R}^\nu$. To this end consider

$$\int \exp\left(-i \langle \xi, x \rangle\right) d\mu_s(x) \int \exp\left(-i \langle \xi, y \rangle\right) d\mu_t(y)$$

$$= \mathbb{E}\left(\exp\left(-i\left\langle\xi, X(s)\right\rangle\right)\right)\mathbb{E}\left(\exp\left(-i\left\langle\xi, X(t)\right\rangle\right)\right)$$

(the variable $X(t)$ has the same distribution as $X(s+t) - X(s)$)

$$= \mathbb{E}\left(\exp\left(-i\left\langle\xi, X(s)\right\rangle\right)\right)\mathbb{E}\left(\exp\left(-i\left\langle\xi, X(s+t) - X(s)\right\rangle\right)\right)$$

($X(s+t) - X(s)$ does not depend on $X(s)$)

$$\begin{aligned} &= \mathbb{E}\left(\exp\left(-i\left\langle\xi, X(s) + X(s+t) - X(s)\right\rangle\right)\right) \\ &= \mathbb{E}\left(\exp\left(-i\left\langle\xi, X(s+t)\right\rangle\right)\right) = \widehat{\mu}_{s+t}(\xi). \end{aligned}$$

Since $0 = X(0) - X(0)$ it follows that μ_0 has the distribution δ_0. This proves Theorem 5.38. □

5.39. Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let the mapping $X : (t,\omega) \mapsto X(t,\omega) = X(t)(\omega)$ satisfy the hypotheses mentioned in Theorem 5.38. (So that for $t > s$ the state variable $X(t) - X(s)$ does not depend on the σ-field $\sigma\left(X(u) : 0 \leqslant u \leqslant s\right)$ and $X(t) - X(s)$ possesses the same distribution as $X(t-s)$; moreover, the equality $\lim_{s\downarrow 0}\mathbb{E}\left(f(X(s))\right) = f(0)$ holds for all $f \in C_0\left(\mathbb{R}^\nu\right)$). Then the process X is called a *Lévy-process*, that begins at $X(0) = 0$.

Important Lévy-processes are the *Poisson process with jumps 1* and the *Brownian motion*. The one-dimensional distributions of a Poisson process X (with jumps 1 and of intensity λ) are given by

$$\mathbb{P}\left(X(t) = k\right) = \frac{(\lambda t)^k}{k!}\exp\left(-\lambda t\right), \quad k \in \mathbb{N}.$$

For details on Poisson processes see Subsection 5.4 in Chapter 1. The Brownian motion B (with drift 0, intensity I and which starts in 0) possesses as one-dimensional distributions:

$$\mathbb{P}\left(B(t) \in B\right) = \frac{1}{\sqrt{(2\pi t)^\nu}}\int_B \exp\left(-\frac{|y|^2}{2t}\right)dy.$$

For more details on Brownian motion see the Section 4 in Chapter 1 and Section 3 in Chapter 2. In addition, see Chapter 3. A Lévy-process with *initial distribution* μ is a family of $\mathcal{F}$-$\mathcal{B}$-measurable mappings $X(t) : \Omega \to \mathbb{R}^\nu$ such that $X(0)$ has the distribution μ, and such that the process $t \mapsto X(t) - X(0)$ is a Lévy-process that starts at 0. If the initial distribution $\mu = \delta_x$, then it said that the process X starts at x. If $X = (X(t) : t \geqslant 0)$ is a Lévy-process that starts at 0, then $(x + X(t) : t \geqslant 0)$ is a Lévy-process, which starts at x. The Poisson process X_j (with jumps 1 and intensity λ) which starts at $j \in \mathbb{N}$ possesses as marginal or one-dimensional distributions:

$$\mathbb{P}\left(X_j(t) = k\right) = \frac{(\lambda t)^{k-j}}{k!}\exp\left(-\lambda t\right)1_{[0,\infty)}(k-j), \quad k \in \mathbb{N}.$$

Thus the distributions of the processes $(X_j(t) : t \geqslant 0)$ and $(j + X(t) : t \geqslant 0)$, where X is the Poisson-process which starts at 0, are the same. The Brownian motion B_x (with drift 0, intensity I and which starts at x) possesses the following one-dimensional distributions:

$$\mathbb{P}\left(B_x(t) \in B\right) = \frac{1}{\sqrt{(2\pi t)^{\nu}}} \int_B \exp\left(-\frac{|x-y|^2}{2t}\right) dy.$$

5.40. Definition. Let E be a locally compact Hausdorff space and let

$$\{P(t) : t \geqslant 0\}$$

be a family of linear operators of $C_0(E)$ to the space $L^\infty(E, \mathcal{E})$. Here $\mathcal{E}$ is the Borel field of E. This family is called a *Feller semigroup, or Feller-Dynkin semigroup* provided it possesses the following properties:

(i) *semigroup-property*: $P(s+t) = P(s)P(t)$ and $P(0) = I$;
(ii) *positivity preserving*: $f \geqslant 0$, $f \in C_0(E)$, implies $P(t)f \geqslant 0$;
(iii) *contractive*: $0 \leqslant f \leqslant 1$, $f \in C_0(E)$, implies $0 \leqslant P(t)f \leqslant 1$;
(iv) *continuity*: $\lim_{t\downarrow 0} \left[P(t)f\right](x) = f(x)$ for all $f \in C_0(E)$ and for all $x \in E$;
(v) *invariance*: $P(t)C_0(E) \subseteq C_0(E)$ for all $t \geqslant 0$.

In the presence of (i), (v) and (iii) assertion (iv) is equivalent with

(iv′) $\lim_{t\to t_0, t>0} \left\|P(t)f - P(t_0)f\right\|_\infty = 0$ for all $f \in C_0(E)$.

5.41. THEOREM. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, and* $(X(t) : t \geqslant 0)$ *be a family of state variables with state space* $\mathbb{R}^\nu$. *Suppose that these state variables are measurable relative to the* σ*-fields* $\mathcal{F}$ *and* $\mathcal{B}(\mathbb{R}^\nu)$. *In addition, suppose that* $\lim_{t\downarrow 0} \mathbb{E}\left(f(X(t))\right) = f(0)$ *for all* $f \in C_{00}\left(\mathbb{R}^\nu\right)$ *and also that for every* $t > s$ *the variable* $X(t) - X(s)$ *does not depend on the* σ*-field* $\sigma\left(X(u) : 0 \leqslant u \leqslant s\right)$, *an d this for all* $t > s \geqslant 0$. *Moreover, by hypothesis, the variable* $X(t) - X(s)$ *has the same distribution as* $X(t-s)$. *Define the operator* $P(t)$ *from* $L^\infty\left(\mathbb{R}^\nu\right)$ *to itself by* $\left[P(t)f\right](x) = \mathbb{E}\left(f\left(x + X(t)\right)\right)$, $f \in L^\infty\left(\mathbb{R}^\nu\right)$. *The restriction of* $P(t)$ *to* $C_0\left(\mathbb{R}^\nu\right)$ *leaves the space* $C_0\left(\mathbb{R}^\nu\right)$ *invariant, and the family* $\left\{P(t)\restriction_{C_0(\mathbb{R}^\nu)} : t \geqslant 0\right\}$ *is a Feller semigroup (also called a Feller-Dynkin semigroup).*

PROOF OF THEOREM 5.41. It is clear that every operator $P(t)$ is contractive and positivity preserving. It is also clear that $\lim_{t\downarrow 0}\left[P(t)f\right](x) = f(x)$ for all $x \in \mathbb{R}^\nu$ and for all $f \in C_0\left(\mathbb{R}^\nu\right)$. We still have to prove the invariance property. Let $f = \widehat{g}$, where g belongs to $L^1\left(\mathbb{R}^\nu\right)$. Then we obtain

$$\begin{aligned}\left[P(t)f\right](x) &= \mathbb{E}\left(f\left(x + X(t)\right)\right) = \mathbb{E}\left(\widehat{g}\left(x + X(t)\right)\right) \\ &= \int \exp\left(-i\langle\xi, x\rangle\right)\mathbb{E}\left(\exp\left(-i\langle\xi, X(t)\rangle\right)\right) g(\xi)d\xi. \qquad (5.37)\end{aligned}$$

By the lemma of Riemann-Lebesgue (Theorem 5.12), the equalities in (5.37) imply the equality

$$\lim_{x\to\infty}\left[P(t)f\right](x) = 0.$$

The continuity of the function $P(t)f$ is clear as well. As a consequence, $P(t)$ maps the space $\left\{\widehat{g} :\ g \in L^1\left(\mathbb{R}^\nu\right)\right\}$ to $C_0\left(\mathbb{R}^\nu\right)$. The theorem of Stone-Weierstrass implies that the space $\left\{\widehat{g} : g \in L^1\left(\mathbb{R}^\nu\right)\right\}$ is dense in $C_0\left(\mathbb{R}^\nu\right)$ for the uniform topology. Because of the contractive character of the operator $P(t)$ it then follows that $P(t)$ leaves the space $C_0\left(\mathbb{R}^\nu\right)$ invariant. In order to finish we prove the semigroup-property. Again we take the Fourier transform $\widehat{g}$ of a function $g \in L^1\left(\mathbb{R}^\nu\right)$ and we consider

$$\begin{aligned}&\left[P(s+t)\widehat{g}\right](x) \\ &= \mathbb{E}\left(\widehat{g}(x + X(s+t))\right) \\ &= \int e^{-i\langle\xi,x\rangle}\mathbb{E}\left(\exp\left(-i\langle\xi, X(s+t)\rangle\right)\right) g(\xi)d\xi \\ &= \int e^{-i\langle\xi,x\rangle}\mathbb{E}\left(\exp\left(-i\langle\xi, X(s+t) - X(s)\rangle\right)\exp\left(-i\langle\xi, X(s)\rangle\right)\right) g(\xi)d\xi\end{aligned}$$

(the variable $X(s+t) - X(s)$ is independent of $X(s)$)

$$= \int e^{-i\langle\xi,x\rangle}\mathbb{E}\left(\exp\left(-i\langle\xi, X(s+t) - X(s)\rangle\right)\right)\mathbb{E}\left(\exp\left(-i\langle\xi, X(s)\rangle\right)\right) g(\xi)d\xi$$

(the variable $X(s+t) - X(s)$ has the same distribution as $X(t)$)

$$\begin{aligned}&= \int e^{-i\langle\xi,x\rangle}\mathbb{E}\left(\exp\left(-i\langle\xi, X(t)\rangle\right)\right)\mathbb{E}\left(\exp\left(-i\langle\xi, X(s)\rangle\right)\right) g(\xi)d\xi \\ &= \mathbb{E}\left(\omega \mapsto \mathbb{E}\left(\omega' \mapsto \widehat{g}\left(x + X(s)(\omega) + X(t)\left(\omega'\right)\right)\right)\right). \qquad (5.38)\end{aligned}$$

The semigroup-property then follows from (5.38) together with the Theorem of Stone-Weierstrass which, among other things, implies that the space

$$\left\{\widehat{g} : g \in L^1\left(\mathbb{R}^\nu\right)\right\}$$

is dense in $C_0\left(\mathbb{R}^\nu\right)$ for the uniform topology.

This completes the proof of Theorem 5.41. □

2. Convergence of positive measures

We begin with the continuity theorem of Lévy.

5.42. Theorem (Lévy). *Let* $(\mu_n : n \in \mathbb{N})$ *be a sequence of bounded positive Borel measures on* $\mathbb{R}^\nu$*. Assume that there exists a function* $\varphi : \mathbb{R}^\nu \to \mathbb{C}$*, which is continuous at 0, such that*

$$\lim_{n\to\infty} \widehat{\mu}_n(x) = \varphi(x)$$

for all $x \in \mathbb{R}^\nu$*. Then there exists a bounded positive Borel measure such that*

$$\text{weak-}\lim_{n\to\infty} \mu_n = \mu.$$

Proof. The function φ is a point-wise limit of positive-definite functions and so it is itself positive-definite as well. Since the function φ is continuous at 0, inequality (c) in Theorem 5.15 implies the continuity of φ. By Bochner's theorem (Theorem 5.17) there exists a positive bounded Borel measure μ such that

$$\lim_{n\to\infty} \widehat{\mu}_n(x) = \varphi(x) = \widehat{\mu}(x) \tag{5.39}$$

for all $x \in \mathbb{R}^\nu$. Next, let f be an arbitrary function in $C_{00}\left(\mathbb{R}^\nu\right)$. Then we see

$$\begin{aligned}\limsup_{n\to\infty} \left|\int \widehat{f} d\mu_n - \int \widehat{f} d\mu\right| &= \limsup_{n\to\infty} \left|\int f(x)\left(\widehat{\mu}_n(x) - \widehat{\mu}(x)\right)\right| dx \\ &\leqslant \limsup_{n\to\infty} \int |f(x)|\,|\widehat{\mu}_n(x) - \widehat{\mu}(x)|\, dx. \end{aligned}\tag{5.40}$$

In view of (5.39) we see that the integrand in (5.40) converges pointwise to 0. Write $c = \sup_{n\in\mathbb{N}}\left(\widehat{\mu}_n(0) + \widehat{\mu}(0)\right)$. Then c is finite and $|f(x)|\,|\widehat{\mu}_n(x) - \widehat{\mu}(x)|$ is dominated by the L^1-function $c\,|f(x)|$. By the dominated convergence theorem (Lebesgue) it follows from (5.40) that

$$\limsup_{n\to\infty} \left|\int \widehat{f} d\mu_n - \int \widehat{f} d\mu\right| = 0. \tag{5.41}$$

Since the subspace $\left\{\widehat{f} : f \in C_{00}\left(\mathbb{R}^\nu\right)\right\}$ is uniformly dense in $C_0\left(\mathbb{R}^\nu\right)$ (see Theorem 5.14), from (5.41) it follows that $\lim_{n\to\infty} \int \varphi d\mu_n = \int \varphi d\mu$ for all functions $\varphi \in C_{00}\left(\mathbb{R}^\nu\right)$. Theorem 5.26 then implies weak-$\lim_{n\to\infty} \mu_n = \mu$. This proves the continuity theorem of Lévy: Theorem 5.42. □

In the following theorem we compare several equivalent forms of weak convergence. If $a = (a_1, \dots, a_\nu)$ and $b = (b_1, \dots, b_\nu)$ belong to $\mathbb{R}^\nu$ and if $a_j < b_j$, $1 \leqslant j \leqslant \nu$, then we write $a < b$ and also $(a, b] = (a_1, b_1] \times \cdots \times (a_\nu, b_\nu]$.

5.43. THEOREM. *Let $(\mu_\alpha : \alpha \in \mathcal{A})$ be a directed system (a net) in $\mathcal{M}^+$ consisting of sub-probability measures (so that $\mu_\alpha(\mathbb{R}^\nu) \leqslant 1$, $\alpha \in \mathcal{A}$) and let $\mu \in \mathcal{M}^+$ be a sub-probability measure as well. Let $(f_k : k \in \mathbb{N})$ be a sequence in $C_0(\mathbb{R}^\nu)$ with a linear span which is dense in $C_0(\mathbb{R}^\nu)$. The following assertions are then equivalent:*

(1) *The net $(\mu_\alpha : \alpha \in \mathcal{A})$ converges weakly to μ;*
(2) *For every bounded Borel measurable function $f : \mathbb{R}^\nu \to \mathbb{C}$ which is continuous in μ-almost all points the equality $\lim_\alpha \int f d\mu_\alpha = \int f d\mu$ holds;*
(3) *The net $(\mu_\alpha : \alpha \in \mathcal{A})$ converges vaguely to μ and $\lim \mu_\alpha(\mathbb{R}^\nu) = \mu(\mathbb{R}^\nu)$;*
(4) *For every closed subset F of $\mathbb{R}^\nu$ the inequality $\limsup_\alpha \mu_\alpha(F) \leqslant \mu(F)$ holds and*

$$\lim_\alpha \mu_\alpha(\mathbb{R}^\nu) = \mu(\mathbb{R}^\nu);$$

(5) *For every open subset G of $\mathbb{R}^\nu$ the inequality $\liminf_\alpha \mu_\alpha(G) \geqslant \mu(G)$ holds and*

$$\lim_\alpha \mu_\alpha(\mathbb{R}^\nu) = \mu(\mathbb{R}^\nu);$$

(6) *For every Borel subset B of $\mathbb{R}^\nu$, for which $\mu\left(\overline{B} \backslash \overset{\circ}{B}\right) = 0$, the equality $\lim_\alpha \mu_\alpha(B) = \mu(B)$ holds;*
(7) *For every pair of points $(a, b) \in \mathbb{R}^\nu \times \mathbb{R}^\nu$ such that $a_j < b_j$, $1 \leqslant j \leqslant \nu$, where $a = (a_1, \dots, a_\nu)$, $b = (b_1, \dots, b_\nu)$, with the property that $\mu\{x \in \mathbb{R}^\nu : x_j = a_j\} = \mu\{x \in \mathbb{R}^\nu : x_j = b_j\} = 0$, $j = 1, \dots, \nu$, the equality $\lim_\alpha \mu_\alpha\big(a, b\big] = \mu\big(a, b\big]$ holds and $\lim_\alpha \mu_\alpha(\mathbb{R}^\nu) = \mu(\mathbb{R}^\nu)$.*
(8) *For every $k \in \mathbb{N}$ the equalities $\lim_\alpha \int f_k d\mu_\alpha = \int f_k d\mu$ and $\lim_\alpha \mu_\alpha(\mathbb{R}^\nu) = \mu(\mathbb{R}^\nu)$ hold;*
(9) *For every $x \in \mathbb{R}^\nu$ the equality $\lim_\alpha \widehat{\mu}_\alpha(x) = \widehat{\mu}(x)$ holds.*
(10) *For every $a \in \mathbb{R}^\nu$ for which $\mu\{x \in \mathbb{R}^\nu : x_j = a_j\} = 0$, $j = 1, \dots, \nu$, the equality*

$$\lim_\alpha \mu_\alpha\left[(-\infty, a_1] \times \cdots \times (-\infty, a_\nu]\right] = \mu\left[(-\infty, a_1] \times \cdots \times (-\infty, a_\nu]\right]$$

holds and $\lim_\alpha \mu_\alpha(\mathbb{R}^\nu) = \mu(\mathbb{R}^\nu)$.

PROOF. The equivalence of the assertions (1) and (9) is a consequence of Theorem 5.25. The equivalence of (1) and (3) is a consequence of Theorem 5.26. The implication (1) ⇒ (8) is trivial. The implication (8) ⇒ (3) can be proved as follows. From (8) it follows that $\lim_\alpha \int \varphi d\mu_\alpha = \int \varphi d\mu$ for all φ in the linear span of $(f_k : k \in \mathbb{N}) \cup \{1\}$. So that for $f \in C_0(\mathbb{R}^\nu) + \mathbb{C}1$ and φ in the span of $(f_k : k \in \mathbb{N}) \cup \{1\}$ we see that

$$\limsup_\alpha \left| \int f \, d(\mu_\alpha - \mu) \right|$$

$$\begin{aligned}
&\leqslant \limsup_{\alpha} \left| \int (f-\varphi)\, d\left(\mu_{\alpha}-\mu\right) \right| + \limsup_{\alpha} \left| \int \varphi d\left(\mu_{\alpha}-\mu\right) \right| \\
&\leqslant \left\| f-\varphi \right\|_{\infty} \limsup_{\alpha} \left(\mu_{\alpha}\left(\mathbb{R}^{\nu}\right)+\mu\left(\mathbb{R}^{\nu}\right)\right) \\
&= 2\left\| f-\varphi \right\|_{\infty} \mu\left(\mathbb{R}^{\nu}\right). \qquad (5.42)
\end{aligned}$$

Assertion (3) follows because the linear span of $(f_k : k \in \mathbb{N}) \cup \{1\}$ is uniformly dense in $C_0\left(\mathbb{R}^{\nu}\right)+\mathbb{C}1$. From the previous arguments it follows that the assertions (1), (3), (8) and (9) are equivalent.

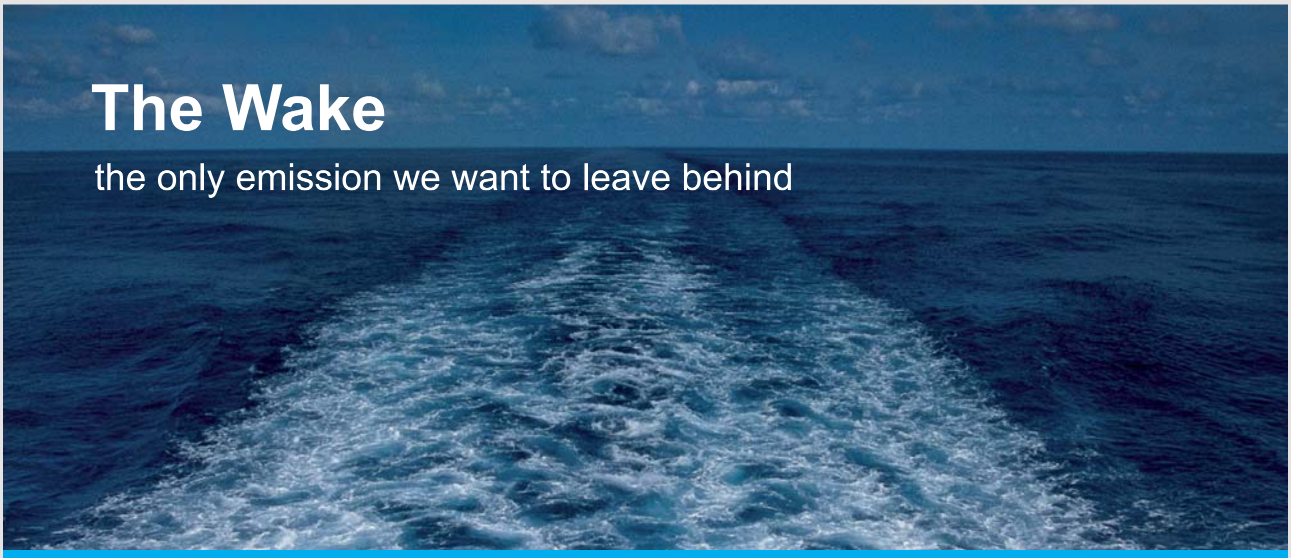
(2) $\Rightarrow$ (1). This implication is trivial.

(1) $\Rightarrow$ (4). Let F be a closed subset of $\mathbb{R}^{\nu}$. Choose a sequence of functions $(u_j : j \in \mathbb{N})$ in $C_b(\mathbb{R}^{\nu})$ in such a way that $1_F \leqslant u_{j+1} \leqslant u_j \leqslant 1$, $j \in \mathbb{N}$, and such that $1_F(x) = \lim_{j\to\infty} u_j(x)$ for all $x \in \mathbb{R}^{\nu}$. Then the equality

$$\limsup_{\alpha} \mu_{\alpha}(F) \leqslant \inf_{j\in\mathbb{N}} \limsup_{\alpha} \int u_j d\mu_{\alpha} = \inf_{j\in\mathbb{N}} \int u_j d\mu = \mu(F)$$

holds. This proves assertion (4) starting from (1).

(4) $\Leftrightarrow$ (5). These implications are easy to verify.



(5) ⇒ (6). Let B be a Borel subset of $\mathbb{R}^{\nu}$ such that $\mu\left(\overline{B}\backslash\overset{\circ}{B}\right) = 0$. Then, from (5), what is equivalent to (4), it follows that

$$\begin{aligned}\limsup \mu_{\alpha}(B) &\leqslant \limsup_{\alpha} \mu_{\alpha}\left(\overline{B}\right) \leqslant \mu\left(\overline{B}\right) = \mu\left(\overset{\circ}{B}\right)\\ &\leqslant \liminf_{\alpha} \mu_{\alpha}\left(\overset{\circ}{B}\right) \leqslant \liminf_{\alpha} \mu_{\alpha}\left(B\right). \end{aligned}\tag{5.43}$$

Hence, $\lim_{\alpha} \mu_{\alpha}(B) = \mu(B)$.

(6) ⇒ (1). Let $0 \leqslant f \leqslant 1$ be a continuous function. Because

$$\int f d\mu = \int_0^1 \mu\left\{f \geqslant \xi\right\} d\xi = \int_0^1 \mu\left\{f > \xi\right\} d\xi$$

we see that $\int_0^1 \mu\left\{f = \xi\right\} d\xi = 0$. Thus for almost all ξ the equality $\mu\left\{f = \xi\right\} = 0$ follows. For a certain sequence $(\alpha_{\ell} : \ell \in \mathbb{N})$ in $\mathcal{A}$, we then obtain by (6) the following (in-)equalities

$$\begin{aligned}\int f d\mu &= \int_0^1 \left\{f > \xi\right\} d\mu = \int_0^1 \mu\left\{f \geqslant \xi\right\} d\xi\\ &= \int_0^1 \lim_{\alpha} \mu_{\alpha}\left\{f \geqslant \xi\right\} d\xi = \int_0^1 \lim_{\alpha} \mu_{\alpha}\left\{f > \xi\right\} d\xi\\ &\leqslant \frac{1}{2^n} \sum_{k=1}^{2^n} \lim_{\alpha} \mu_{\alpha}\left\{f > k2^{-n}\right\} + \frac{1}{2^n} \leqslant \frac{1}{2^n} \sum_{k=1}^{2^n} \lim_{\ell\to\infty} \mu_{\alpha_{\ell}}\left\{f > k2^{-n}\right\} + \frac{1}{2^n}\end{aligned}$$

(Fatou's lemma)

$$\begin{aligned}&\leqslant \lim_{\ell\to\infty} \int \frac{1}{2^n} \sum_{k=1}^{2^n} 1_{\{f\geqslant k2^{-n}\}} d\mu_{\alpha_{\ell}} + \frac{1}{2^n}\\ &\leqslant \lim_{\ell\to\infty} \int f d\mu_{\alpha_{\ell}} + \frac{1}{2^n} = \liminf_{\alpha} \int f d\mu_{\alpha} + \frac{1}{2^n}.\end{aligned}\tag{5.44}$$

From (5.44) it then follows that, always for $0 \leqslant f \leqslant 1$,

$$\int f d\mu \leqslant \liminf_{\alpha} \int f d\mu_{\alpha},\tag{5.45}$$

and also

$$\int (1-f) d\mu \leqslant \liminf_{\alpha} \int (1-f) d\mu_{\alpha}.\tag{5.46}$$

Since, in addition, $\lim_{\alpha} \mu_{\alpha}\left(\mathbb{R}^{\nu}\right) = \mu\left(\mathbb{R}^{\nu}\right)$ we see by (5.45) and (5.46) that $\lim_{\alpha} f d\mu_{\alpha} = \int f d\mu$ for every function $f \in C_b\left(\mathbb{R}^{\nu}\right)$ for which $0 \leqslant f \leqslant 1$. Since the linear span of such functions coincides with $C_b\left(\mathbb{R}^{\nu}\right)$ assertion (1) follows from (6). (5) ⇒ (2). Let f be a real-valued bounded function which μ-almost everywhere continuous. Without loss of generality we assume that $0 \leqslant f \leqslant 1$

(otherwise replace f with $af + b$, with a and b appropriately chosen constants). Then define the functions $f^{\frown}$ and $f^{\smile}$ respectively by

$$f^{\frown}(x) = \inf_{U\in\mathfrak{U}(x)} \sup_{y\in U} f(y) \quad \text{en} \quad f^{\smile}(x) = \sup_{U\in\mathfrak{U}(x)} \inf_{y\in U} f(y). \tag{5.47}$$

It follows that $\int (f^{\frown} - f^{\smile})\, d\mu = 0$ and also $f^{\smile} \leqslant f \leqslant f^{\frown}$. Hence, for an appropriately chosen sequence $(\alpha_\ell : \ell \in \mathbb{N})$,

$$\begin{aligned}\int f d\mu &= \int f^{\smile} d\mu \leqslant \frac{1}{2^n} + \frac{1}{2^n}\sum_{k=1}^{2^n} \mu\left\{f^{\smile} > k2^{-n}\right\} \\ &\leqslant \frac{1}{2^n} + \frac{1}{2^n}\sum_{k=1}^{2^n} \liminf_{\ell\to\infty} \mu_{\alpha_\ell}\left\{f^{\smile} > k2^{-n}\right\} \\ &\leqslant \frac{1}{2^n} + \liminf_{\ell\to\infty} \frac{1}{2^n}\sum_{k=1}^{2^n} \mu_{\alpha_\ell}\left\{f^{\smile} > k2^{-n}\right\} \\ &\leqslant \frac{1}{2^n} + \liminf_{\alpha} \int f^{\smile} d\mu_\alpha \leqslant \frac{1}{2^n} + \liminf_{\alpha} \int f d\mu_\alpha. \end{aligned} \tag{5.48}$$

From (5.48) it follows that

$$\int f d\mu \leqslant \liminf_{\alpha} \int f d\mu_\alpha.$$

For the same reason the inequality $\int(1-f)d\mu \leqslant \liminf_\alpha \int(1-f)d\mu_\alpha$ holds. Because, in addition, $\lim_\alpha \mu_\alpha(\mathbb{R}^\nu) = \mu(\mathbb{R}^\nu)$ we see that $\int f d\mu = \lim_\alpha \int f d\mu_\alpha$. This proves (2) starting from (5).

(6) $\Rightarrow$ (7). This assertion is trivial.

(7) $\Rightarrow$ (8). In this part of the proof we write $(a, b]$ for the interval $(a_1, b_1] \times \cdots \times (a_\nu, b_\nu]$, if a and b are points in $\mathbb{R}^\nu$ for which $a_j < b_j$ for $1 \leqslant j \leqslant \nu$. From (7) it follows that $\lim_\alpha \mu_\alpha\big(a, b\big] = \mu\big(a, b\big]$ for all points a and b in $\mathbb{R}^\nu$ with the property that $\mu\{y \in \mathbb{R}^\nu : x_j = a_j\} = \mu\{y \in \mathbb{R}^\nu : x_j = b_j\} = 0$ for all $1 \leqslant j \leqslant \nu$. Next pick for f a function in $C_{00}(\mathbb{R}^\nu)$ with values in $\mathbb{R}$. Let g be an arbitrary function of the form $g = \sum_{j=1}^n f(x_j)1_{(a_j,b_j]}$, where a_j and b_j are points in $\mathbb{R}^\nu$ with the following properties: $\mu\{y \in \mathbb{R}^\nu : y_k = a_{j,k}\} = \mu\{y \in \mathbb{R}^\nu : y_k = b_{j,k}\} = 0$, for $1 \leqslant k \leqslant \nu$, and $a_{j,k} < b_{j,k}$ for $j = 1, \dots, n$, and $1 \leqslant k \leqslant \nu$. In addition, suppose that x_j belongs to the "interval" $\big(a_j, b_j\big]$. The we get

$$\begin{aligned}&\limsup_\alpha \int f d\mu_\alpha \leqslant \limsup_\alpha \int (f-g)d\mu_\alpha + \limsup_\alpha \int g d\mu_\alpha \\ &\leqslant \|f-g\|_\infty \mu(\mathbb{R}^\nu) + \int g d\mu \leqslant 2\|f-g\|_\infty \mu(\mathbb{R}^\nu) + \int f d\mu.\end{aligned}$$

Since f is uniformly continuous we are able to choose, for a given $\epsilon > 0$, a function g of the form as above in such a way that $\|f-g\|_\infty \leqslant \epsilon$. This proves the inequality $\limsup_\alpha \int f d\mu_\alpha \leqslant \int f d\mu$. The same argument can be applied to the function $-f$. It follows that $\lim_\alpha \int f d\mu_\alpha = \int f d\mu$ for functions $f \in C_{00}(\mathbb{R}^\nu)$ that are real valued. But then (8) follows.

(10) $\Rightarrow$ (7) Let $a < b$ be as in (7). Then $\mu_\alpha(a,b]$ can be written in the form

$$\mu_\alpha(a,b] = \sum_{\Lambda\subset\{1,\dots,\nu\}} (-1)^{\#\Lambda} \mu_\alpha\left[\prod_{j=1}^{\nu}(-\infty, c_{\Lambda,j}]\right], \tag{5.49}$$

where $c_{\Lambda,j} = a_j$, $j \in \Lambda$, $c_{\Lambda,j} = b_j$, $j \in \{1,\dots,\nu\}\backslash\Lambda$. The implication (10) $\Rightarrow$ (7) then easily follows from (5.49). The equality in (5.49) can be found in Durrett [**60**] Theorem 1.1.6 page 7.

(7) $\Rightarrow$ (10) Let a be as in assertion (10). Put $F = \prod_{k=1}^{\nu}(-\infty, a_k]$. Then the subset F is closed, and since assertion (7) is equivalent to (4) we know that $\limsup_\alpha \mu_\alpha(F) \leqslant \mu(F)$. Since assertion (7) is equivalent to (5) we know that $\liminf_\alpha \mu_\alpha(F) \geqslant \liminf_\alpha \mu_\alpha\left(\overset{\circ}{F}\right) \geqslant \mu\left(\overset{\circ}{F}\right)$. Since $\mu\left(\overset{\circ}{F}\right) = \mu(F)$ assertion (10) follows.

The proof of these implications completes the proof Theorem 5.43. $\square$

5.44. Remark. The implication (10) $\Rightarrow$ (1) in Theorem 5.43 can also be proved by employing the equality

$$\begin{aligned}\int_{\mathbb{R}^\nu} f(x)\,d\mu_\alpha(x) &= (-1)^\nu \int_{\mathbb{R}^\nu} D_1\cdots D_\nu f(x)\,\mu_\alpha\left[\prod_{j=1}^{\nu}(-\infty,x_j]\right]dx \qquad (5.50)\\ &= (-1)^\nu \int_{\mathbb{R}^\nu}\int_{y_1}^{\infty}\cdots\int_{y_\nu}^{\infty} D_1\cdots D_\nu f(x)\,dx_\nu\ldots dx_1\,d\mu_\alpha(y),\end{aligned}$$

where the function f is ν times continuously differentiable, and where D_j denotes differentiation with respect to the j-th coordinate, $1 \leqslant j \leqslant \nu$. The equality in (5.50) can be proved by successive integration. The second equality is a consequence of Fubini's theorem.

5.45. Definition. A topological space E is called a *Polish* space if E possesses the following properties:

(i) E is separable;
(ii) E is metrizable;
(iii) There exists a metric d on E that determines the topology and relative to which E is complete.

Since E is metrizable property (i) is equivalent with the existence of a countable basis for the topology.

5.46. Lemma. *Let E be a Polish space.*

(a) *A closed subset F of E is, with the induced metric, again Polish.*
(b) *An open subset G of E is again Polish.*

Proof. (a) The proof of assertion (a) is not difficult. If $(U_j : j \in \mathbb{N})$ is a countable basis for the topology of E, then $(U_j \cap F : j \in \mathbb{N})$ is a countable basis for the topology on F. Moreover, F is closed and hence it is complete with respect to the induced metric.

(b). Let d be a metric on E, which turns E into a complete metric topological space. Then the open subset G is Polish for the metric d_G defined by

$$d_G(x,y) = d(x,y) + \left|\frac{1}{d(x,G^c)} - \frac{1}{d(y,G^c)}\right|, \qquad (5.51)$$

where x and y belong to G, where $G^c = E\backslash G$ and where $d(x,G^c) = \inf_{z\in G^c} d(x,z)$. The separability of G is also clear. The proof of Lemma 5.46 is now complete.
□

5.47. Theorem. *A subset A of a Polish space E is again a Polish space if and only if A is the countable intersection of open subsets of E.*

Proof. Let $A = \bigcap_{j\in\mathbb{N}} G_j$, where every G_j is an open subset of E. Let d be a metric on E, which makes E into a Polish space. Define then the metrics d_j

on G_j, $j \in \mathbb{N}$, as in (5.51) and define the metric d_A on A via

$$d_A(x,y) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{d_j(x,y)}{1+d_j(x,y)}.$$

Then, endowed with the relative topology, A is a Polish space with respect to the d_A. Conversely, let $d \leqslant 1$ be a metric on A which is compatible with the topology that A inherits from E, and which turns A into a complete metric space. Let $\overline{A}$ be the closure of A. Then there exists a decreasing sequence of open subsets $(G_n)_{n\in\mathbb{N}}$ such that $\overline{A} = \bigcap_n G_n$; i.e. closed subsets of E are G_δ-subsets. For $n \in \mathbb{N}$, $n \neq 0$, we define the open subset A_n of $\overline{A}$ as follows:

$$\begin{aligned} A_n = \Big\{ x \in \overline{A} :\ & \text{there exists an open neighborhood } U(x) \text{ in } E \text{ of } x \\ & \text{for which } d(y,z) < \frac{1}{n} \text{ for all } z,\, y \in U(x) \cap A \Big\}.. \end{aligned} \tag{5.52}$$

Then the following assertions about the sets A_n will be proved:

1. For all $n \in \mathbb{N}$ we have $A \subset A_n$.
2. The sets A_n, $n \in \mathbb{N}$, are open in $\overline{A}$.
3. The inclusion $\bigcap_n A_n \subset A$ holds, and so by (1) $A = \bigcap_n A_n$.

Since, by (2), the subsets A_n, $n \in \mathbb{N}$, are open in $\overline{A}$ there exist open subsets O_n, $n \in \mathbb{N}$, of E such that $A_n = O_n \cap \overline{A}$, $n \in \mathbb{N}$. It follows that

$$A = \bigcap_n A_n = \bigcap_n O_n \cap \overline{A} = \bigcap_n O_n \cap \bigcap_m G_m = \bigcap_n O_n \cap G_n,$$

and hence, A is a countable intersection of open subsets of E. Next we prove the assertions (1), (2) and (3).

(1) Pick $x \in A$, and consider the ball

$$B_{1/(2n)}(x) = \left\{ w \in A :\ d(w,x) < \frac{1}{2n} \right\}.$$

There exists an open neighborhood $U(x)$ of x in E such that $B_{1/(2n)}(x) = A \cap U(x)$. If y, z belong to $A \cap U(x)$ we have $d(z,y) \leqslant d(z,x) + d(x,y) < 1/n$. It follows that $x \in A_n$. This is true for all $n \in \mathbb{N}$.

(2) That the subset A_n is open in $\overline{A}$ can be seen as follows. Pick $x \in A$. There exists an open neighborhood $U(x)$ in E of x such that $d(z,y) < 1/n$ for all z, $y \in A \cap U(x)$. The set $U(x) \cap \overline{A}$ is an open neighborhood of x in $\overline{A}$. It suffices to show that $U(x) \cap \overline{A} \subset A_n$. To this end choose $x' \in U(x) \cap \overline{A}$. Then $U(x)$ is an open neighborhood in E of x' as well, and since x' belongs to $\overline{A}$ it follows from the definition of A_n that x' is a member of A_n.

(3) Let x belong to A_n for all $n \in \mathbb{N}$. Then x belongs to $\overline{A}$. We will prove that $x \in A$. Let $D \leqslant 1$ be a metric on E which is compatible with its topology, and which turns E into complete metric space. For the moment fix $n \in \mathbb{N}$. Since $x \in A_n$ there exists an open ball B_n in E relative to the metric D centered at x

and with radius $< 1/n$ such that

$$d(z, y) < \frac{1}{n} \quad \text{whenever} \quad y,\, z \in A \cap B_n. \tag{5.53}$$

Since $x \in \overline{A}$ there exists $x_n \in A \cap B_n$. In this way we obtain a sequence of balls $\{B_n : n \in \mathbb{N}\}$ relative to the metric D centered at x, which we take decreasing, and which are such that the D-radius of B_n is strictly less than $1/n$. In addition, we obtain a sequence of points $\{x_n : n \in \mathbb{N}\}$ in A such that $x_n \in B_n$ for all $n \in \mathbb{N}$. Since the balls B_n are decreasing we have $x_m \in B_n$ for $m \geqslant n$. From this fact and (5.53) it follows that $d(x_m, x_n) < 1/n$ for $m \geqslant n$. Consequently, it follows that the sequence $\{x_n : n \in \mathbb{N}\}$ is a d-Cauchy sequence in A. Since A is d-complete there exists a point $x' \in A$ such $d(x_n, x') \leqslant 1/n$, $n \in \mathbb{N}$. Since D and d are topologically compatible on A it also follows that $\lim_{n\to\infty} D(x_n, x') = 0$. We also have $\lim_{n\to\infty} D(x_n, x) = 0$, and consequently $x = x' \in A$.

This completes the proof of Theorem 5.47. □

For a proof of the following theorem the reader is referred to the literature. We will give an outline of a proof. A $\mathcal{G}_\delta$-set in a topological space is a countable intersection of open subsets.

5.48. Theorem. *A Polish space E is homeomorphic with a $\mathcal{G}_\delta$-subset of the Hilbert-cube $[0,1]^{\mathbb{N}}$, endowed with the product topology.*

Technical training on
WHAT you need, WHEN you need it

At IDC Technologies we can tailor our technical and engineering training workshops to suit your needs. We have extensive experience in training technical and engineering staff and have trained people in organisations such as General Motors, Shell, Siemens, BHP and Honeywell to name a few.

Our onsite training is cost effective, convenient and completely customisable to the technical and engineering areas you want covered. Our workshops are all comprehensive hands-on learning experiences with ample time given to practical sessions and demonstrations. We communicate well to ensure that workshop content and timing match the knowledge, skills, and abilities of the participants.

We run onsite training all year round and hold the workshops on your premises or a venue of your choice for your convenience.

For a no obligation proposal, contact us today at training@idc-online.com or visit our website for more information: www.idc-online.com/onsite/

OIL & GAS ENGINEERING
ELECTRONICS
AUTOMATION & PROCESS CONTROL
MECHANICAL ENGINEERING
INDUSTRIAL DATA COMMS
ELECTRICAL POWER

Phone: +61 8 9321 1702
Email: training@idc-online.com
Website: www.idc-online.com

IDC TECHNOLOGIES

Click on the ad to read more

Proof of Theorem 5.48. Let $d : E \times E \to [0,1]$ be a metric op E which is compatible with its topology, and which turns E into a Polish space, and let $(x_\ell : \ell \in \mathbb{N})$ be a countable dense subset of E. Define the mapping $\Psi : E \to [0,1]^{\mathbb{N}}$ by $\Psi(x) = \left(d\left(x, x_\ell\right)_{\ell\in\mathbb{N}}\right)$. Then, as can be checked, the mapping Ψ is a homeomorphism from E onto a subset of $[0,1]^{\mathbb{N}}$. So far we have not yet used the fact that E, equipped with the metric d is complete; we did use the fact that E is a metrizable separable space. Since E is complete with respect to d and Ψ is a homeomorphism, it follows that the image of E under Ψ, that is $A = \Psi(E)$, is complete subspace of the Hilbert cube $[0,1]^{\mathbb{N}}$. Let $D : [0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}} \to [0,1]$ be the metric defined by

$$D\left((\xi_\ell)_{\ell\in\mathbb{N}}, (\eta_\ell)_{\ell\in\mathbb{N}}\right) = \sum_{\ell=1}^{\infty} 2^{-\ell} \left|\xi_\ell - \eta_\ell\right|, \quad (\xi_\ell)_{\ell\in\mathbb{N}},\ (\eta_\ell)_{\ell\in\mathbb{N}} \in [0,1]^{\mathbb{N}}.$$

Then D is a metric on $[0,1]^{\mathbb{N}}$ which turns this space into a Polish space. It follows that A is a subset of $[0,1]^{\mathbb{N}}$ which is homeomorphic to a Polish space, and so it itself is Polish. Since it is a Polish subspace of the Polish space $[0,1]^{\mathbb{N}}$, A is a countable intersection of open subsets of $[0,1]^{\mathbb{N}}$: see Theorem 5.47. This completes the proof of Theorem 5.48. ☐

The space $\mathbb{N}$ of positive integers with the usual metric inherited from the real numbers $\mathbb{R}$ is Polish. Then the countable product $\mathbb{N}^{\mathbb{N}}$ with metric

$$d\left(\{m_i\}, \{n_i\}\right) = \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{|m_i - n_i|}{1 + |m_i - n_i|} \tag{5.54}$$

is Polish. The proofs of Propositions 5.49 and 5.50 are taken from Garrett [**75**].

5.49. Proposition. *Totally order $\mathbb{N}^{\mathbb{N}}$ lexicographically. Then every closed subset C of $\mathbb{N}^{\mathbb{N}}$ has a least element.*

The lexicographic ordering of $\mathbb{N}^{\mathbb{N}}$ can be recursively defined. An element $a = (a_1, a_2, \dots)$ precedes an element $b = (b_1, b_2, \dots)$ if $a_1 \leqslant b_1$; however, if $a_1 = b_1$, then $a_2 \leqslant b_2$; however, if $a_1 = b_1$ and $a_2 = b_2$, then $a_3 \leqslant b_3$, and so on.

Proof. Let n_1 be the least element in $\mathbb{N}$ such that there is $x = (n_1, \dots)$ belonging to C. Let n_2 be the least element in $\mathbb{N}$ such that there is $x = (n_1, n_2, \dots)$ belonging to C, and so on. Choosing the n_i inductively, let $x_0 = (n_1, n_2, n+3, \dots)$. This x_0 satisfies $x_0 \leqslant x$ in the lexicographic ordering for every $x \in C$, and x_0 belongs to the closure of C in the metric topology introduced in (5.54). This completes the proof of Proposition 5.49. ☐

5.50. Proposition. *Let E be a Polish space. Then there exists a continuous surjective mapping $F_0 : \mathbb{N}^{\mathbb{N}} \to E$. Moreover, there exists a measurable function $G_0 : E \to \mathbb{N}^{\mathbb{N}}$ such that $F_0 \circ G_0(y) = y$ for all $y \in E$.*

Proof. The mapping F_0 can be constructed as follows. For a given $\varepsilon > 0$ there is a countable covering of E by closed sets of diameter less than ε. From this one may contrive a map F from finite sequences $\{n_1, \dots, n_k\}$ in $\mathbb{N}$ to closed sets $F(n_1, \dots, n_k)$ in E such that

1. $F(\varnothing) = E$;
2. $F(n_1, \ldots, n_k) = \bigcup_{\ell=1}^{\infty} F(n_1, \ldots, n_k; \ell)$;
3. The diameter of $F(n_1, \ldots, n_k)$ is less than 2^{-k}.

Then for $x = \{n_i\} \in \mathbb{N}^{\mathbb{N}}$ the sequence $E_k = F(n_1, \ldots, n_k)$ is a nested sequence of closed subsets of E with diameters less than 2^{-k}, respectively. Thus, the subset $\bigcap_k E_k$ consists of a single point $F_0(y)$ of E. On the other hand, every $x \in E$ lies inside some $\bigcap_k E_k$. Continuity is easy to verify. The mapping G_0 can be constructed as follows. The space $\mathbb{N}^{\mathbb{N}}$, endowed with the lexicographical ordering is totally ordered, and by Proposition 5.49 every closed subset contains a least element for this order. For $y \in E$ the subset $F_0^{-1}(y)$ is closed in $\mathbb{N}^{\mathbb{N}}$, and therefore it contains a least element $G_0(y)$. This assignment is a measurable choice (because it can be performed in countably many steps). Then $F_0 \circ G_0(y) = y$ for $y \in E$. The proof of Proposition 5.50 is complete now. □

The proof of the following theorem is based on the fact that a Polish space is homeomorphic with a $\mathcal{G}_\delta$-subset of the Hilbert cube, which, being a countable product of closed intervals, is a compact metrizable space.

5.51. Theorem. *Let μ be a finite positive measure on the Borel field of a Polish space. Then μ is regular in the sense that*

$$\mu(B) = \inf\{\mu(O) : O \text{ open, } B \subset O\} = \sup\{\mu(K) : K \text{ compact, } K \subset B\}. \tag{5.55}$$

Proof. Let Ψ and A be as in the proof of Theorem 5.48. Then $\Psi : E \to A$ is a homeomorphism. Let $\mu \geqslant 0$ be a finite measure on the Borel field of E. Define the measure ν on the Borel field of $[0,1]^{\mathbb{N}}$ by

$$\nu(B) = \mu\left[\Psi^{-1}(B \cap A)\right] = \mu\left[\Psi \in B \cap A\right], \quad B \text{ Borel subset of } [0,1]^{\mathbb{N}}. \tag{5.56}$$

Then, since the Hilbert cube is compact and complete metrizable, and A is a $\mathcal{G}_\delta$-subset of the Hilbert cube, we see that the measure ν is regular on the Borel field of $[0,1]^{\mathbb{N}}$. It also follows that the restriction of ν to the Borel field of A is regular. However, under the homeomorphism $\Psi : E \to A$ the Borel subsets of E are in a one-to-one correspondence with those of A. It easily follows that the measure μ is regular in the sense of (5.55), which completes the proof of Theorem 5.51. □

5.52. Theorem. *The following assertions hold for Banach spaces.*

(a) *Let E be a separable Banach space. Then its dual unit ball B', endowed with the weak*-topology, is a Polish space.*
(b) *(Helly) The set $\mathcal{M}^+_{\leqslant 1}$ is compact-metrizable, and thus Polish for the vague topology.*

Let E' be the topological dual space of E. The weak*-topology is denoted by $\sigma := \sigma(E', E)$.

Proof. (a) Let $(x_n : n \in \mathbb{N})$ be a sequence in the unit ball B of E of which the linear span is dense in E. Define the mapping $\Phi : (B', \sigma) \to [0,1]^{\mathbb{N}}$ via the map $x' \mapsto (< x_n, x' >)_{n=1}^{\infty}$. The mapping Φ is continuous and, by the Theorem of Banach-Alaoglu, the dual unit ball B' is compact relative the topology $\sigma(E', E)$. It follows that $\Phi(B')$ is a compact subset of $[0,1]^{\mathbb{N}}$. This image is Polish, and because the inverse of Φ is continuous, B' itself is Polish as well.

Let r be a positive real number. Let the set $\mathcal{M}^+_{\leqslant r}$ be defined by

$$\mathcal{M}^+_{\leqslant r} = \left\{\mu \in \mathcal{M}^+ : \mu\left(\mathbb{R}^\nu\right) \leqslant r\right\},$$

and let $\mathcal{M}^+_r$ be given by

$$\mathcal{M}^+_r = \left\{\mu \in \mathcal{M}^+ : \mu\left(\mathbb{R}^\nu\right) = r\right\}.$$

(b) Since $\mathcal{M}^+_{\leqslant 1}$ is a vaguely closed subset of $\mathcal{M} = C_0\left(\mathbb{R}^\nu\right)^*$, by the Theorem of Banach-Alaoglu it follows that $\mathcal{M}^+_{\leqslant 1}$ is compact for the vague topology. The fact that the set $\mathcal{M}^+_{\leqslant 1}$ is Polish will be proved in Theorem 5.54. This completes the proof of Theorem 5.52. ☐

5.53. Definition. A subset $\mathcal{A}$ of $\mathcal{M}$ is a *Prohorov* subset, if it satisfies the following two conditions:

(a) $\sup_{\mu\in\mathcal{A}} |\mu|\left(\mathbb{R}^\nu\right)$ is finite;
(b) For every $\epsilon > 0$ there exists a compact subset K of $\mathbb{R}^\nu$ such that

$$\sup_{\mu\in\mathcal{A}} |\mu|\left(\mathbb{R}^\nu \backslash K\right) \leqslant \epsilon.$$

5.54. Theorem. *The following assertions are true:*

(a) *The spaces $\mathcal{M}^+_{\leqslant 1}$ and $\mathcal{M}^+_1$ are Polish with respect to the vague topology.*
(b) *The spaces $\mathcal{M}^+_{\leqslant 1}$ and $\mathcal{M}^+_1$ are Polish with respect to the weak topology.*

Proof of Theorem 5.54. The countable collection

$$\bigcup_{n=1}^{\infty}\left\{\sum_{j=1}^{n}\alpha_j\delta_{x_j} : \alpha_j \in \mathbb{Q}, \alpha_j \geqslant 0, \sum_{j=1}^{n}\alpha_j = 1\right\}$$

is dense in $\mathcal{M}_1^+$ for the vague as well as for the weak topology. The countable collection

$$\bigcup_{n=1}^{\infty}\left\{\sum_{j=1}^{n}\alpha_j\delta_{x_j} : \alpha_j \in \mathbb{Q}, \alpha \geqslant 0, 0 \leqslant \sum_{j=1}^{n}\alpha_j \leqslant 1\right\}$$

is dense in $\mathcal{M}_{\leqslant 1}^+$ for de vague as well as the weak topology. This can be seen as follows. Let f be a bounded continuous function defined op $\mathbb{R}^\nu$. Fix $\epsilon > 0$, and choose a partition of $\mathbb{R}^\nu$ in Borel subsets $(A_j : j \in \mathbb{N})$ in such a way that $|f(x) - f(y)| \leqslant \dfrac{\epsilon}{\mu\left(\mathbb{R}^\nu\right)}$ for all x, $y \in A_j$, and this for all $j \in \mathbb{N}$. In addition, choose $N \in \mathbb{N}$ so large that

$$\left(\mu\left(\mathbb{R}^\nu\right) - \sum_{j=1}^{N}\mu\left(A_j\right)\right)\|f\|_\infty \leqslant \epsilon.$$

Put $a_j = \mu\left(A_j\right)$ and choose $x_j \in A_j$. Then we have

$$\begin{aligned}&\left|\int f d\mu - \frac{\sum_{j=1}^{N} a_j f(x_j)}{\sum_{j=1}^{N} a_j}\mu\left(\mathbb{R}^\nu\right)\right| \\ &\leqslant \left|\sum_{j=1}^{\infty}\int_{A_j}\left(f(x) - f(x_j)\right)d\mu(x)\right| + \left|\sum_{j=1}^{N} a_j f(x_j)\left\{1 - \frac{\mu\left(\mathbb{R}^\nu\right)}{\sum_{j=1}^{N} a_j}\right\}\right| \\ &\qquad + \left|\sum_{j=N+1}^{\infty}\int_{A_j} f(x)dx\right| \\ &\leqslant \frac{\epsilon}{\mu\left(\mathbb{R}^\nu\right)}\sum_{j=1}^{\infty}\mu\left(A_j\right) + 2\|f\|_\infty\left\{\mu\left(\mathbb{R}^\nu\right) - \sum_{j=1}^{N}\mu\left(A_j\right)\right\} \leqslant 3\epsilon.\end{aligned}$$

Appealing another time to the continuity of the function f, and using the fact that the rational numbers are dense in $\mathbb{R}$ we obtain the separability of the sets $\mathcal{M}_r^+$ and $\mathcal{M}_{\leqslant r}^+$ relative tot the vague as well as the weak topology. We indicate metrics which turn these spaces into Polish spaces. Therefore we choose a sequence of functions $(f_k : k \in \mathbb{N})$ in $\{f \in C_{00}\left(\mathbb{R}^\nu\right) : 0 \leqslant f \leqslant 1\}$ whose linear span is uniformly dense in $C_0\left(\mathbb{R}^\nu\right)$. We also choose a sequence $(u_\ell : \ell \in \mathbb{N})$ in $C_{00}\left(\mathbb{R}^\nu\right)$ such that $1 \geqslant u_{\ell+1} \geqslant u_\ell \geqslant 0$ and such that $1 = \lim_{\ell\to\infty} u_\ell(x)$ for all $x \in \mathbb{R}^\nu$.

(a) Define the distance d_v on $\mathcal{M}_{\leqslant 1}^+$ via the formula

$$d_v(\mu, \nu) = \sum_{j=1}^{\infty}\frac{1}{2^j}\left|\int f_j d\mu - \int f_j d\nu\right|, \tag{5.57}$$

where μ and ν are members of $\mathcal{M}_{\leqslant r}^+$. Supplied with this metric the space $\mathcal{M}_{\leqslant r}^+$ is Polish for the vague topology. The spaces $\mathcal{M}_{\leqslant r}^+$, $0 < r \leqslant 1$, are also closed for the vague topology. Since

$$\mathcal{M}_1^+ = \bigcap_{n=1}^{\infty} \mathcal{M}_{\leqslant 1}^+ \backslash \mathcal{M}_{\leqslant 1-1/n}^+ \tag{5.58}$$

we see that $\mathcal{M}_1^+$ is a Polish space: see the Theorems 5.47 and 5.52.

(b) Let $f_0 \equiv 1$ and let the sequences $(f_k : k \in \mathbb{N})$ and $(u_\ell : \ell \in \mathbb{N})$ be as above. We define the metric d_w by the equality

$$d_w(\mu,\nu) = \sup_{\ell\in\mathbb{N}} \left|\int u_\ell d\mu - \int u_\ell d\nu\right| + \sum_{j=1}^{\infty} \frac{1}{2^j} \left|\int f_j d\mu - \int f_j d\nu\right|, \tag{5.59}$$

where μ and ν belong to $\mathcal{M}_{\leqslant r}^+$. Supplied with this metric the space $\mathcal{M}_{\leqslant r}^+$ is Polish for the weak topology. If we are also able to prove that the space $\mathcal{M}_{\leqslant r}^+$ is complete relative to the metric d_w, then it follows that the spaces $\mathcal{M}_{\leqslant r}^+$, $r \geqslant 0$, are Polish. These spaces are also weakly closed. By the equality in (5.58) in Theorem 5.47 then implies that $\mathcal{M}_1^+$ is a Polish space as well. Now let $(\mu_m : m \in \mathbb{N})$ be a d_w-Cauchy sequence in $\mathcal{M}_{\leqslant r}^+$. We will prove that this sequence is a Prohorov set in $\mathcal{M}_{\leqslant r}^+$. Choose $\epsilon > 0$ arbitrary. Then there exists $M_\epsilon \in \mathbb{N}$ such that for m and $m' \geqslant M_\epsilon$ the inequality

$$\sup_{\ell\in\mathbb{N}} \left|\int u_\ell d\mu_m - \int u_\ell d\mu_{m'}\right| \leqslant \frac{\epsilon}{4} \tag{5.60}$$

holds. So it follows that

$$\left|\mu_m\left(\mathbb{R}^\nu\right) - \mu_{m'}\left(\mathbb{R}^\nu\right)\right| \leqslant \frac{\epsilon}{4}, \quad m,\ m' \geqslant M_\epsilon. \tag{5.61}$$

From (5.60) and (5.61) it follows that

$$\sup_{\ell\in\mathbb{N}} \left|\int \left(1-u_\ell\right) d\mu_m - \int \left(1-u_\ell\right) d\mu_{m'}\right| \leqslant \frac{\epsilon}{2}, \tag{5.62}$$

for m and $m' \geqslant M_\epsilon$. Then from (5.62) it follows that

$$\int \left(1-u_\ell\right) d\mu_m \leqslant \int \left(1-u_\ell\right) d\mu_{M_\epsilon} + \frac{\epsilon}{2}. \tag{5.63}$$

From (5.63) it then follows that for $\ell \geqslant \ell_\epsilon$ and $m \geqslant M_\epsilon$ the following inequality holds:

$$\int \left(1-u_\ell\right) d\mu_m \leqslant \frac{\epsilon}{4} + \frac{\epsilon}{2} = \frac{3\epsilon}{4}. \tag{5.64}$$

From (5.64) it then follows that for all $\ell \geqslant L_\epsilon$ and for all $m \in \mathbb{N}$ the inequality

$$\int \left(1-u_\ell\right) d\mu_m \leqslant \epsilon. \tag{5.65}$$

holds. Then choose the compact subset K_ϵ equal the support of the function u_{L_ℓ}. It follows that $\mu_m\left(\mathbb{R}^\nu \backslash K_\epsilon\right) \leqslant \epsilon$ for all $m \in \mathbb{N}$. Let μ be the vague limit of the sequence $(\mu_m : m \in \mathbb{N})$. This limit exists, because $\mathcal{M}_{\leqslant r}^+$ is compact for the metric d_v. Then pick a function $u \in C_{00}(\mathbb{R}^\nu)$ in such a way that $1 \geqslant u \geqslant 1_{K_\epsilon}$. By the equality

$$\int f d\mu_m - \int f d\mu = \int (f - fu)\, d\mu_m + \int fu d\mu_m - \int fu d\mu - \int (f - fu)\, d\mu$$

we infer the inequality

$$\begin{aligned}
&\left|\int f d\mu_m - \int f d\mu\right| \\
&\leqslant \|f\|_\infty \left(\int (1-u)\, d\mu_m + \int (1-u)\, d\mu\right) + \left|\int fu d\mu_m - \int fu d\mu\right| \\
&\leqslant \|f\|_\infty \left(\mu_m (\mathbb{R}^\nu \backslash K_\epsilon) + \mu (\mathbb{R}^\nu \backslash K_\epsilon)\right) + \left|\int fu d\mu_m - \int fu d\mu\right| \\
&\leqslant \|f\|_\infty \left(\mu_m (\mathbb{R}^\nu \backslash K_\epsilon) + \sup_m \mu_m (\mathbb{R}^\nu \backslash K_\epsilon)\right) + \left|\int fu d\mu_m - \int fu d\mu\right| \\
&\leqslant 2\epsilon \|f\|_\infty \left|\int fu d\mu_m - \int fu d\mu\right|. \qquad (5.66)
\end{aligned}$$

Since $\mu = \text{vague-}\lim_m \mu_m$ from (5.66) it also follows that μ is the weak limit of the sequence $(\mu_m : m \in \mathbb{N})$.

The proof Theorem 5.54 is now complete. ☐

Part of the proof of Theorem 5.54 comes back in the proof of Theorem 5.55.

5.55. THEOREM. *A subset S of $\mathcal{M}^+_{\leqslant r}$ is relatively weakly compact if and only if S is a Prohorov subset.*

PROOF OF THEOREM 5.55. First, suppose that S is a Prohorov subset of $\mathcal{M}^+_{\leqslant r}$. Let $(\mu_m : m \in \mathbb{N})$ be a sequence in S. We will prove that there exists a subsequence that converges weakly. We may assume that, possibly by passing to a subsequence, that this sequence converges vaguely. By employing the Prohorov property we will show that, in fact, this sequence converges weakly. Let $\epsilon > 0$ be arbitrary. Then there exists a compact subset K such that $\mu_m(K^c) \leqslant \epsilon$ and also that $\mu(K^c) \leqslant \epsilon$. Then choose $u \in C_{00}(\mathbb{R}^\nu)$ in such a way that $1 \geqslant u \geqslant 1_K$. Then we have (see the final part of the proof of Theorem 5.54):

$$\begin{aligned}
&\left|\int f d\mu_m - \int f d\mu\right| \\
&\leqslant \|f\|_\infty \left(\int (1-u)\, d\mu_m + \int (1-u)\, d\mu\right) + \left|\int fu d\mu_m - \int fu d\mu\right| \\
&\leqslant \|f\|_\infty \left(\mu_m(\mathbb{R}^\nu \backslash K) + \mu(\mathbb{R}^\nu \backslash K)\right) + \left|\int fu d\mu_m - \int fu d\mu\right| \\
&\leqslant \|f\|_\infty \left(\mu_m(\mathbb{R}^\nu \backslash K) + \sup_m \mu_m(\mathbb{R}^\nu \backslash K)\right) + \left|\int fu d\mu_m - \int fu d\mu\right| \\
&\leqslant 2\epsilon \|f\|_\infty \left|\int fu d\mu_m - \int fu d\mu\right|. \qquad (5.67)
\end{aligned}$$

Since $\mu = \text{vague-}\lim_m \mu_m$ from (5.67) it also follows that μ is the weak limit of the sequence $(\mu_m : m \in \mathbb{N})$.

Conversely, suppose that the set S is weakly compact. We will prove that S possesses the Prohorov property. Assume the contrary. Then there exists an $\eta > 0$ and there exists an increasing sequence of compact subsets $(K_m : m \in \mathbb{N})$ with the following properties:

(a) $K_m \subset \overset{\circ}{K}_{m+1}$ and $\mathbb{R}^\nu = \bigcup_{m=1}^\infty K_m$;
(b) For every $m \in \mathbb{N}$ there exists a measure $\mu_m \in S$ such that $\mu_m(K_m^c) \geqslant \eta$.

Then choose a sequence of functions $(u_m \in C_{00}(\mathbb{R}^\nu))$ such that $1_{K_m} \leqslant u_{m+1} \leqslant 1_{K_{m+1}}$. From (b) it follows then that, for $m \geqslant \ell$,

$$\eta \leqslant \int (1 - 1_{K_m})\, d\mu_m \leqslant \int (1 - 1_{K_\ell})\, d\mu_m \leqslant \int (1 - u_\ell)\, d\mu_m. \qquad (5.68)$$

Since the subset S is relatively weakly compact, there exists a subsequence $(\mu_{m_k} : k \in \mathbb{N})$ which converges weakly to a measure μ. From (5.68) we then see that $\eta \leqslant \int (1 - u_\ell)\, d\mu$ for all $\ell \in \mathbb{N}$. From this we obtain a contradiction, because the sequence $(1 - u_\ell : \ell \in \mathbb{N})$ decreases to 0.

This completes the proof of Theorem 5.55. □

5.56. COROLLARY. *Let* $(\mu_m : m \in \mathbb{N})$ *be a sequence of measures in* $\mathcal{M}^+$ *and let* μ *also belong to* $\mathcal{M}^+$. *Choose a sequence of functions* $(f_k : k \in \mathbb{N})$ *in*

$$\{f \in C_{00}(\mathbb{R}^\nu) : 0 \leqslant f \leqslant 1\}$$

with a linear span that is uniformly dense in $C_0(\mathbb{R}^\nu)$ *and choose another sequence* $(u_\ell : \ell \in \mathbb{N})$ *in* $C_{00}(\mathbb{R}^\nu)$ *such that* $1 \geqslant u_{\ell+1} \geqslant u_\ell \geqslant 0$ *and such that*

$1 = \lim_{\ell\to\infty} u_\ell(x)$ *for all* $x \in \mathbb{R}^\nu$. *Define the metric* d_w *by the equality*

$$d_w(\nu_1, \nu_2) = \sup_{\ell\in\mathbb{N}} \left| \int u_\ell d\nu_2 - \int u_\ell \, d\nu_1 \right| + \sum_{j=1}^{\infty} \frac{1}{2^j} \left| \int f_j d\nu_2 - \int f_j d\nu_1 \right|,$$

where ν_1 *and* ν_2 *belong to* $\mathcal{M}^+$. *Then the following assertions are equivalent:*

(i) *The sequence* $(\mu_m : m \in \mathbb{N})$ *converges weakly to* μ*;*
(ii) $\lim_{m\to\infty} d_w(\mu_m, \mu) = 0$.

PROOF. The proof is left as an exercise for the reader. ☐

3. A taste of ergodic theory

In this section we will formulate and prove the pointwise ergodic theorem of Birkhoff. We also indicate its relation with the strong law of large numbers. We will also show that the strong law of large numbers (SLLN) implies the weak law of large numbers (WLLN). However, we begin with von Neumann'a ergodic theorem in a Hilbert space. In what follows the symbol H stands for a (complex) Hilbert space with inner-product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$. An operator $U : H \to H$ is called unitary if it satisfies $U^*U = UU^* = I$. An operator $P : H \to H$ is called an orthogonal projection if $P^* = P = P^2$. Let L be closed subspace of H. Then H can be written as

$$H = L \oplus L^\perp = PH + (I - P)\, H,$$

where $P : H \to H$ is an orthogonal projection with range L. The following theorem is the same as Theorem 7.1 in Romik [**151**].

5.57. THEOREM. *Let H be a Hilbert space, and let U be a unitary operator on H. Let P be the orthogonal projection operator onto the subspace $N(U-I)$ (the subspace of H consisting of U-invariant vectors). For any vector $v \in H$ the equality*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} U^k v = Pv.$$

holds. (Equivalently, the sequence of operators $\left\{\frac{1}{n}\sum_{k=0}^{n-1} U^k : n \in \mathbb{N}\right\}$ converges to P in the strong operator topology.)

PROOF. Define the subspace $V \subset H$ by $V = N(U-I) = \{v \in H : Uv = v\} = N(U^* - I)$. Then

$$V = (R(U-I))^{\perp} := \{v \in H : \langle (U-I)u, v\rangle = 0 \text{ for all } u \in H\}, \tag{5.69}$$

where $R(U-I)$ is the range of the operator $U-I$, *i.e.*

$$R(U-I) = \{Uu - u : u \in H\}.$$

Let L be any linear subspace of H. From Hilbert space techniques it is known that $\left(L^{\perp}\right)^{\perp}$ coincides with the closure of the linear subspace L. From these observations it follows that the subspace $N(U-I) + R(U-I)$ is dense in H, and that the subspaces $N(U-I)$ and $R(U-I)$ are orthogonal. Moreover, we have

$$\left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k v\right\| \leqslant \frac{1}{n}\sum_{k=0}^{n-1}\left\|U^k v\right\| \leqslant \|v\|, \quad v \in H. \tag{5.70}$$

Define the subspace $L \subset H$ by

$$L = \left\{v \in H : \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} U^k v = Pv\right\}, \tag{5.71}$$

where P is the orthogonal projection onto the space $V = N(U-I)$. From (5.70) it follows that L is a closed subspace of H. If $v \in V$, *i.e.* if $Uv = v$, then v belongs to L. If $v = (U-I)u$ belongs to the range of $U-I$, then

$$\frac{1}{n}\sum_{k=0}^{n-1} U^k v = \frac{1}{n}\sum_{k=0}^{n-1} U^k (U-I)u = \frac{1}{n}(U^n - I)u, \tag{5.72}$$

and so by (5.70) and (5.71) it follows that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} U^k v = 0 = Pv, \tag{5.73}$$

whenever v belongs to $R(U-I)$. Again appealing to (5.70) shows that (5.73) also holds for v belonging to the closure of $R(U-I)$. Altogether it shows that the subspace L coincides with H. This completes the proof of Theorem 5.57. ☐

Let $(\Omega, \mathcal{F}, \mathbb{P}; T)$ be a measure preserving system. We associate with the measure preserving mapping $T : \Omega \to \Omega$ an operator U_T on the Hilbert space $L^2(\Omega, \mathcal{F}, \mathbb{P})$ defined by $U_T(f) = f \circ T$, $f \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. The fact that T is measure preserving implies that $U_T^* U_T = I$:

$$\langle U_T f, U_T g\rangle = \mathbb{E}\left[U_T f \overline{U_T g}\right] = \mathbb{E}\left[f \circ T \overline{g \circ T}\right] = \mathbb{E}\left[(f\overline{g}) \circ T\right] = \mathbb{E}\left[f\overline{g}\right] = \langle f, g\rangle . \tag{5.74}$$

From (5.74) it follows that $U_T^* U_T = I$. In order that U_T is a unitary operator it should also be surjective. Since the range of U_T is closed, this is the case provided that the set of functions $f \circ T$, $f \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, constitutes a dense subspace of $L^2(\Omega, \mathcal{F}, \mathbb{P})$. The latter is true if the mapping T has the property that there exists a measurable mapping $\widetilde{T} : \Omega \to \Omega$ such that $\widetilde{T} \circ T(\omega) = \omega$ for $\mathbb{P}$-almost all ω. Then the operator $\widetilde{U}$ defined by $\widetilde{U} f = f \circ \widetilde{T}$, $f \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ is the adjoint of U_T. This can be seen as follows. Now we do not only have $U_T^* U_T = I$, but we also have $U_T \widetilde{U} f = \widetilde{U} f \circ T = f \circ \widetilde{T} \circ T = f$, $f \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Hence, we see

$$\widetilde{U} = \left(U_T^* U_T\right) \widetilde{U} = U_T^* \left(U_T \widetilde{U}\right) = U_T^*.$$

Note also that the subspace $N(U - I)$ consists exactly of the invariant (square-integrable) random variables, or equivalently those random variables which are measurable with respect to the σ-algebra $\mathcal{I}$ of invariant events. Recalling the discussion of conditional expectations in Theorem 1.4, item (11), in Chapter 1, we also see that the orthogonal projection operator P is exactly the conditional expectation operator $\mathcal{E}\left[\cdot \mid \mathcal{I}\right]$ with respect to the σ-algebra of invariant events $\mathcal{I}$. Thus, Theorem 5.57 applied to this setting gives the following result.

5.58. Theorem (The L^2 ergodic theorem). *Let $(\Omega, \mathcal{F}, \mathbb{P}; T)$ be a measure preserving system. For any random variable $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ the equality*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} X \circ T^k = \mathbb{E}\left[X \mid \mathcal{I}\right] \tag{5.75}$$

holds in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. In particular, if the system is ergodic then

$$L^2\text{-}\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} X \circ T^k = \mathbb{E}\left[X\right]. \tag{5.76}$$

Since the operator $S : L^2(\Omega, \mathcal{F}, \mathbb{P}) \to L^2(\Omega, \mathcal{F}, \mathbb{P})$, defined by $Sf = f \circ T$, $f \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, is not necessarily unitary, Theorem 5.58 requires a proof. It only satisfies $S^* S = I$. The proof is based on the proof of Theorem 5.66 below.

Proof. Theorem 5.58 is a consequence of Theorem 5.59 which includes the L^1-version of Theorem 5.58. More precisely, we have to prove that

$$\lim_{n\to\infty} \mathbb{E}\left[\left|\frac{1}{n} \sum_{k=0}^{n-1} X \circ T^k - \mathbb{E}\left[X \mid \mathcal{I}\right]\right|^2\right] = 0. \tag{5.77}$$

Let $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ be bounded. Then we can use Theorem 5.59 together Lebesgue's theorem of dominated convergence that the equality in (5.77) holds for X. A general function $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ can be approximated by bounded functions in L^2-sense. Since $\int |f \circ T|^2 \, d\mu = \int |f|^2 \, d\mu$, and so the convergence in (5.77) also holds for all L^2-functions. The precise argument follows as in (5.122) below with $Sf = f \circ T$, $f \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, and relative to the L^2-norm instead of the L^1-norm. So let f belong to $L^2(\Omega, \mathcal{F}, \mathbb{P})$, and let $M > 0$ be an arbitrary real number. Then we have:

$$\begin{aligned}
&\left\| \frac{1}{n} \sum_{k=0}^{n-1} S^k f - P_\mu f \right\|_{L^2} \\
&\leqslant \left\| \left(\frac{1}{n} \sum_{k=0}^{n-1} S^k - P_\mu \right) \left(f 1_{\{|f|<M\}} \right) \right\|_{L^2} + \left\| \left(\frac{1}{n} \sum_{k=0}^{n-1} S^k - P_\mu \right) \left(f 1_{\{|f|\geqslant M\}} \right) \right\|_{L^2} \\
&\leqslant \left(\int \left| \left(\frac{1}{n} \sum_{k=0}^{n-1} S^k - P_\mu \right) \left(f 1_{\{|f|<M\}} \right) \right|^2 d\mathbb{P} \right)^{1/2} + 2 \left(\int_{\{|f|\geqslant M\}} |f|^2 \, d\mathbb{P} \right)^{1/2}.
\end{aligned} \tag{5.78}$$

As in the proof of Theorem 5.66 in (5.78) we first let $n \to \infty$, and then $M \to \infty$ to obtain the L^2-convergence in Theorem 5.58. ☐

The pointwise ergodic theorem of Birkhoff requires some more work. In what follows $(\Omega, \mathcal{F}, \mu)$ is positive measure space, and $T : \Omega \to \Omega$ is a measure preserving mapping, *i.e.* $\mu\{T \in A\} = \mu\{T^{-1}A\} = \mu\{A\}$ for all $A \in \mathcal{F}$ with $\mu\{A\} < \infty$. An equivalent formulation reads as follows. For all $f \in L^1(\Omega, \mathcal{F}, \mu)$ the equality $\int f \circ T\, d\mu = \int f\, d\mu$ holds. In other words the quadruple $(\Omega, \mathcal{F}, \mathbb{P}; T)$ is a measure preserving system, or dynamical system. The operator P_μ in (5.80) is a projection mapping from $L^1(\Omega, \mathcal{F}, \mu)$ onto a space consisting of T-invariant functions. Hence a function of the form $g = P_\mu f$ satisfies $g \circ T = g$ μ-almost everywhere, and so g is measurable with respect to the invariant σ-field $\mathcal{I}$. In addition, if h is a bounded, T-invariant function in $L^1(\Omega, \mathcal{F}, \mu)$, then we have

$$\int (P_\mu f)\, h\, d\mu = \int P_\mu (fh)\, d\mu = \int fh\, d\mu. \tag{5.79}$$

In other words the function $P_\mu f$ is the μ-conditional expectation of the function f on the σ-field of invariant subsets. A measure preserving system $(\Omega, \mathcal{F}, \mu; T)$ is called ergodic if a T-invariant function is constant μ-almost every, and so the σ-field $\mathcal{I}$ is trivial, *i.e.* $\mathcal{I}$ consists, up to sets of μ-measure zero, of the void set and of Ω.

5.59. THEOREM (The pointwise ergodic theorem in L^1). *Let* $(\Omega, \mathcal{F}, \mu; T)$ *be a measure preserving system. For any function* $f \in L^1(\Omega, \mathcal{F}, \mu)$ *the equality*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k = P_\mu f \tag{5.80}$$

holds μ*-almost everywhere. In particular, if the system is ergodic then the equality*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k = \int f\, d\mu. \tag{5.81}$$

holds μ*-almost everywhere. If* μ *is a probability measure, then the limits in (5.80) and (5.81) also hold in* L^1*-sense, and* $P_\mu f = \mathbb{E}_\mu\left[f \mid \mathcal{I}\right]$.

PROOF. The proof of Theorem 5.59 follows from Theorem 5.66 and its Corollary 5.67 with $\mu(\Omega) = 1$, and $Sf = f \circ T$, $f \in L^1(\Omega, \mathcal{F}, \mu)$. ☐

Let $(\Omega_0, \mathcal{F}_0, \mathbb{P})$ be a probability space, and let $X_j : \Omega_0 \to \mathbb{R}$, $j = 0, 1, \ldots$, be a sequence of independent and identically distributed variables (i.i.d.). Let us show that the SLLN is a consequence: see Theorem 2.54. Put $S_n = \sum_{k=0}^{n-1} X_k$.

5.60. THEOREM (Strong law of large numbers). *The equality*

$$\lim_{n\to\infty} \frac{S_n}{n} = \alpha, \quad \textit{holds } \mathbb{P}\textit{-almost surely}$$

for some finite constant α*, if and only if* $\mathbb{E}\left[|X_k|\right] < \infty$*, and then* $\alpha = \mathbb{E}\left[X_1\right]$.

PROOF. Let $\Omega = \mathbb{R}^{\mathbb{N}}$, endowed with the product σ-field $\mathcal{F} = \otimes_{j=0}^{\infty} \mathcal{B}_j$ where $\mathcal{B}_j$ is the Borel field on $\mathbb{R}$. Define the probability measure μ on $\mathcal{F}$ by

$$\mu(A) = \mathbb{E}\left[1_A(X_0, X_1, \ldots)\right] = \mathbb{P}\left[(X_0, X_1, \ldots) \in A\right], \quad A \in \mathcal{F}.$$

Put $Sf(x_0, x_1, \ldots) = f(x_1, x_2, \ldots)$, $f \in L^1(\Omega, \mathcal{F}, \mu)$, $(x_0, x_1, \ldots) \in \Omega$. Then $\int Sf\,d\mu = \int f\,d\mu$, $f \in L^1(\Omega, \mathcal{F}, \mu)$. The assertion in Theorem 5.60 follows from Theorem 5.66 and its Corollary 5.67 by applying them to the function $f_0 : \Omega \to \mathbb{R}$ defined by $f_0(x_0, x_1, \ldots) = x_0$, $(x_0, x_1, \ldots) \in \Omega$. Then

$$\sum_{k=0}^{n-1} \left(S^k f_0\right)(X_0, X_1,\, X_2,\, \ldots) = \sum_{k=0}^{n-1} f_0(X_k,\, X_{k+1}, \ldots) = \sum_{k=0}^{n-1} X_k, \quad k = 0,\, 1,\, \ldots,$$

and hence Theorem 5.60 is a consequence of Theorem 5.66 and its Corollary 5.67. Theorem 5.60 also follows from Theorem 5.59 by applying it to the mapping $T : \Omega \to \Omega$ given by $T(x_0, x_1, \ldots) = (x_1, x_2, \ldots)$, $(x_0, x_1, \ldots) \in \Omega$. ☐

We will formulate some of the results in terms of positivity preserving operators $S : L^1(\Omega, \mathcal{F}, \mu) \to L^1(\Omega, \mathcal{F}, \mu)$.

5.61. Lemma. *Let* $S : L^1(\Omega, \mathcal{F}, \mu) \to L^1(\Omega, \mathcal{F}, \mu)$ *be a linear map, and let* $f \geqslant 0$ *belong to* $L^1(\Omega, \mathcal{F}, \mu)$. *Then the following assertions are equivalent:*

(i) $\min(Sf, 1) = S(\min(f, 1))$;
(ii) $\max(Sf - 1, 0) = S(\max(f - 1, 0))$;
(iii) $\min(Sf, 1) \leqslant S(\min(f, 1))$, *and* $\max(Sf - 1, 0) = S(\max(f - 1, 0))$.

Suppose that for every $f \geqslant 0$, $f \in L^1(\Omega, \mathcal{F}, \mu)$ *the operator* S *satisfies one, and hence all of the conditions (i), (ii) and (iii). In addition, assume that*

$$\int |Sf|\, d\mu \leqslant \int f\, d\mu, \text{ for all } f \in L^1(\Omega, \mathcal{F}, \mu),\ f \geqslant 0. \tag{5.82}$$

Then S *is positivity preserving in the sense that* $f \geqslant 0$, $f \in L^1(\Omega, \mathcal{F}, \mu)$, *implies* $Sf \geqslant 0$, *and contractive in the sense that* $\int |Sf|\, d\mu \leqslant \int |f|\, d\mu$ *for all* $f \in L^1(\Omega, \mathcal{F}, \mu)$. *Then the equivalent conditions (iv), (v), (vi). and (vii) given by*

(iv) *The equality* $S(fg) = (Sf)(Sg)$ *holds for all* $f \in L^1(\Omega, \mathcal{F}, \mu)$, *and for all* $g \in L^1(\Omega, \mathcal{F}, \mu) \bigcap L^\infty(\Omega, \mathcal{F}, \mu)$,
(v) $S(\min(f, g)) = \min(Sf, Sg)$ *is true for all* f, $g \in L^1(\Omega, \mathcal{F}, \mu)$,
(vi) *The equality* $S(\max(f - 1, 0)) = \max(Sf - 1, 0)$ *holds for every* $f \in L^1(\Omega, \mathcal{F}, \mu)$.
(vii) *The* $S1_{\{f>1\}} = 1_{\{Sf>1\}}$ *holds for every* $f \in L^1(\Omega, \mathcal{F}, \mu)$.

are also true. Moreover, (vi) implies that if the assertions (i), (ii) and (iii) are true for all positive functions f *in* $L^1(\Omega, \mathcal{F}, \mu)$, *then they are true for all functions* f *in* $L^1(\Omega, \mathcal{F}, \mu)$. *If the measure* μ *is finite, then all assertions (i), (ii), (iii) (for all* $f \geqslant 0$, $f \in L^1(\Omega, \mathcal{F}, \mu)$,*) and (iv), (v), (vi) and (vii) are equivalent. Finally, the operator* $S : L^1(\Omega, \mathcal{F}, \mu) \to L^1(\Omega, \mathcal{F}, \mu)$ *is continuous, more precisely,*

$$\int |Sf|\, d\mu = \int S|f|\, d\mu \leqslant \int |f|\, d\mu, \quad f \in L^1(\Omega, \mathcal{F}, \mu). \tag{5.83}$$

5.62. Remark. Assertion (iv) also holds for all $f,\ g \in L^1(\Omega, \mathcal{F}, \mu) \bigcap L^2(\Omega, \mathcal{F}, \mu)$. The equality in (v) can be replaced with

$$S(\max(f,g)) = \max(Sf, Sg) \quad \text{for all } f,\ g \in L^1(\Omega, \mathcal{F}, \mu). \tag{5.84}$$

The latter is true because $\min(f,g) + \max(f,g) = f + g$.

Proof of Lemma 5.61. The equivalence of the assertions (i), (ii) and (iii) follows from the following identities:

$$\min(Sf, 1) + S(\max(f-1, 0)) = Sf = \min(Sf, 1) + \max(Sf - 1, 0).$$

Now assume that for all $f \geqslant 0$, the operator S satisfies (i), (ii) or (iii), and assume that S is contractive in the sense of (5.82). Let $f \geqslant 0$ belong to $L^1(\Omega, \mathcal{F}, \mu)$, and put $f_n = \max(f - n^{-1}, 0)$. Then the sequence $(f - f_n)_n$ decreases to 0, and hence, by (5.82), $\lim_{n\to\infty} \int |Sf - Sf_n|\, d\mu = 0$. Then there exists a subsequence $(f_{n_k})_k$ such that the sequence $(Sf_{n_k})_k$ converges to Sf μ-almost everywhere. Hence $Sf \geqslant 0$ μ-almost everywhere. In fact it follows that the sequence $(Sf_n)_n$ increases to Sf μ-almost everywhere. Let $f \in L^1(\Omega, \mathcal{F}, \mu)$, and write $f = f_+ - f_-$ where $f_+ = \max(f, 0)$, and $f_- = \max(-f, 0)$. Then $|f| = f_+ - f_-$, and $|Sf| \leqslant Sf_+ + Sf_-$. From (5.82) it the follows that

$$\int |Sf|\, d\mu \leqslant \int (Sf_+ + Sf_-)\, d\mu = \int S(f_+ + f_-)\, d\mu = \int S|f|\, d\mu \leqslant \int |f|\, d\mu. \tag{5.85}$$

The inequalities in (5.85) prove the contraction property of the operator S. Next we prove the assertions (iv), (v) and (vi) starting from (i), (ii) or (iii) for all $f \in L^1(\Omega, \mathcal{F}, \mu)$, $f \geqslant 0$. Let the function $f \geqslant 0$ belong to $L^1(\Omega, \mathcal{F}, \mu) \bigcap L^2(\Omega, \mathcal{F}, \mu)$. Then we write

$$f^2 = 2\int_0^\infty \max(f-\alpha, 0)\, d\alpha = \sup_n 2^{-n+1} \sum_{j=1}^{n2^n} \max\left(f - j2^{-n}, 0\right),$$

ad so

$$\begin{aligned} S\left(f^2\right) &= 2\int_0^\infty S\max(f-\alpha,0)\, d\alpha \\ &= \sup_n 2^{-n+1}\sum_{j=1}^{n2^n} S\left(\max\left(f-j2^{-n},0\right)\right) \end{aligned}$$

(apply assertion (ii))

$$\begin{aligned} &= \sup_n 2^{-n+1}\sum_{j=1}^{n2^n} \max\left(Sf - j2^{-n}, 0\right) \\ &= 2\int_0^\infty \max(Sf-\alpha,0)\, d\alpha = (Sf)^2. \end{aligned} \tag{5.86}$$

The equality in (5.86) shows that the assertion in (iv) is true provided that $f = g \geqslant 0$. For general $f = g$ we split f in its positive and negative part. For general f and g belonging to $L^1(\Omega, \mathcal{F}, \mu) \bigcap L^2(\Omega, \mathcal{F}, \mu)$ we write $2fg = (f+g)^2 - f^2 - g^2$. Altogether this shows assertion (iv). (iv) ⇒ (v) Let f belong to $L^1(\Omega, \mathcal{F}, \mu) \bigcap L^2(\Omega, \mathcal{F}, \mu)$. Then we write

$$|f| = \frac{2}{\pi}\int_0^\infty \frac{f^2}{t^2+f^2}\, dt,$$

and so by assertion (iv) we get

$$S|f| = \frac{2}{\pi}\int_0^\infty S\left\{\frac{f^2}{t^2+f^2}\right\} dt$$

(for explanation see below: equality (5.89))

$$= \frac{2}{\pi}\int_0^\infty \frac{S\left(f^2\right)}{t^2+S\left(f^2\right)}\, dt$$

((iv) implies $S\left(f^2\right) = (Sf)^2$)

$$= \frac{2}{\pi}\int_0^\infty \frac{(Sf)^2}{t^2+(Sf)^2}\, dt = |Sf|. \tag{5.87}$$

The equality in (5.87) shows that (5.84) is true for $g = -f$. For general f, $g \in L^1(\Omega, \mathcal{F}, \mu) \bigcap L^2(\Omega, \mathcal{F}, \mu)$ we write $2\max(f,g) = |f-g| + f + g$. Consequently, assertion (v) follows for f, $g \in L^1(\Omega, \mathcal{F}, \mu) \bigcap L^2(\Omega, \mathcal{F}, \mu)$. If f and g are arbitrary functions in $L^1(\Omega, \mathcal{F}, \mu)$, then we approximate them by $f_n := f1_{\{|f| \leqslant n\}}$

and by $g_n := g1_{\{|g|\leqslant n\}}$ respectively. This shows that assertion (v) is a consequence of (iv), except that the proof of the second equality in (5.87) is not provided yet. In order to prove this equality it suffices to prove that, for $a > 0$ and $g \geqslant 0$, $g \in L^1(\Omega, \mathcal{F}, \mu)$ the equality

$$S\left\{\frac{g}{a+g}\right\} = \frac{Sg}{a+Sg} \tag{5.88}$$

holds. By assertion (iv) we have

$$\begin{aligned} S\left\{\frac{g}{a+g}\right\}(a+Sg) &= S\left\{\frac{ag}{a+g}\right\} + S\left\{\frac{g}{a+g}\right\}Sg \\ &= S\left\{\frac{ag}{a+g}\right\} + S\left\{\frac{g^2}{a+g}\right\} = Sg. \end{aligned} \tag{5.89}$$

The equality in (5.89) shows the validity of (5.88). Therefore the second equality in (5.87) is proved now.

(ii) plus (v) ⇒ (vi) We apply (5.84), which is equivalent to (v), with $f \in L^1(\Omega, \mathcal{F}, \mu)$ arbitrary and $g = 0$ to obtain

$$S(\max(f-1,0)) = S(\max(\max(f,0)-1,0))$$

(employ assertion (ii))

$$= \max(S\max(f,0)-1,0)$$

(apply assertion (v))

$$\begin{aligned} &= \max(\max(Sf, S0)-1,0) \\ &= \max(\max(Sf,0)-1,0) = \max(Sf-1,0). \end{aligned}$$

Hence, assertion (vi) follows from (ii) and (v).

(vi) ⇒ (v) Let $f \in L^1(\Omega, \mathcal{F}, \mu)$. By assertion (vi) we have

$$S\max(f,0) = \lim_{\varepsilon\downarrow 0} S\max(f-\varepsilon,0) = \lim_{\varepsilon\downarrow 0}\max(Sf-\varepsilon,0) = \max(Sf,0). \tag{5.90}$$

Whence, $S\max(f,0) = \max(Sf,0)$. Since $|f| = 2\max(f,0) - f$ we easily infer $S|f| = |Sf|$, and assertion (v) follows: see the proof of the implication (iv) ⇒ (v).

(ii) plus (v) ⇒ (iv) Let f belong to $L^1(\Omega, \mathcal{F}, \mu) \bigcap L^2(\Omega, \mathcal{F}, \mu)$. Then we write

$$f^2 = 2\int_0^\infty \max(|f|-\alpha,0)\,d\alpha,$$

and so by assertion (ii) and (v) we get

$$\begin{aligned} S\left(f^2\right) &= 2\int_0^\infty \max(S|f|-\alpha,0)\,d\alpha \\ &= 2\int_0^\infty \max(|Sf|-\alpha,0)\,d\alpha \\ &= |Sf|^2 = (Sf)^2, \end{aligned}$$

and hence (iv) follows. (vi) $\Rightarrow$ (vii) Let f belong to $L^1(\Omega, \mathcal{F}, \mu)$. Then $1_{\{f>1\}}$ is μ-integrable as well. Then we have

$$\begin{aligned} S1_{\{f>1\}} &= \lim_{m\to\infty} S\left(\min\left(m\max\left(f-1,0\right)\right)\right) \\ &= \lim_{m\to\infty}\left(\min\left(m\max\left(Sf-1,0\right)\right)\right) = 1_{\{Sf>1\}}. \end{aligned} \tag{5.91}$$

The equality of the ultimate terms in (5.91) proves the implication (vi) $\Rightarrow$ (vii).

(vii) $\Rightarrow$ (vi) Let f belong to $L^1(\Omega, \mathcal{F}, \mu)$. Then the functions $1_{\{f>\alpha\}}$, $\alpha > 0$, are μ-integrable as well. We have

$$\begin{aligned} S\left(\max\left(f-1,0\right)\right) &= S\int_0^\infty 1_{\{f-1>\alpha\}}\,d\alpha = \int_0^\infty S1_{\{f-1>\alpha\}}\,d\alpha \\ &= \int_0^\infty 1_{\{Sf-1>\alpha\}}\,d\alpha = \max\left(Sf-1,0\right). \end{aligned} \tag{5.92}$$

The equality of the ultimate terms in (5.92) proves the implication (vii) $\Rightarrow$ (vi).

If the measure μ is finite, then the constant functions belong to $L^1(\Omega, \mathcal{F}, \mu)$. Since $S1 = 1$, it is easy to see that assertion (iv) implies assertion (i), and hence by what is proved above, we see that for a finite measure μ all assertion (i) through (vi) are equivalent. The equality and inequality in (5.83) follow from assertion (v) and the inequality in (5.82). This completes the proof of Lemma 5.61. $\square$

5.63. Proposition. *Let g and h be functions in $L^1(\Omega, \mathcal{F}, \mu)$. Define, for $-\infty < a < b < \infty$, the subset $C_{a,b}^{\{g,h\}}$ by*

$$C_{a,b}^{\{g,h\}} = \{g < a < b < h\}. \tag{5.93}$$

Then the following equality holds:

$$S1_{C_{a,b}^{\{g,h\}}} = 1_{C_{a,b}^{\{Sg,Sh\}}}. \tag{5.94}$$

Proof. Since $C_{-b,-a}^{\{-h,-g\}} = \{g < a < b < h\}$ we may assume that $b > 0$. If $a < 0$, then $1_{\{g<a<b<h\}} = 1_{\{-g>-a\}}1_{\{h>b\}}$. From assertions (iv) and (vii) of Lemma 5.61 it follows that

$$S1_{\{g<a<b<h\}} = 1_{\{-Sg>-a\}}1_{\{Sh>b\}} = 1_{\{Sg<a<b<Sh\}},$$

and consequently (5.94) follows for $a < 0 < b$. If $a = 0$, then we replace a with $a - \varepsilon$ and let $\varepsilon \downarrow 0$. If $a > 0$, then we consider, for $0 < \varepsilon < a$,

$$1_{\{g\leqslant a-\varepsilon<b<h\}} = 1_{\{h>b\}} - 1_{\{g>a-\varepsilon\}}1_{\{h>b\}}. \tag{5.95}$$

Another application of the assertions (iv) and (vii) of Lemma 5.61 then yields by employing (5.95) the equality:

$$S1_{\{g\leqslant a-\varepsilon<b<h\}} = 1_{\{Sh>b\}} - 1_{\{Sg>a-\varepsilon\}}1_{\{Sh>b\}} = 1_{\{Sg\leqslant a-\varepsilon<b<Sh\}}. \tag{5.96}$$

In (5.96) we let $\varepsilon \downarrow 0$ to obtain the equality in (5.93) for $0 < a < b$. This completes the proof of Proposition 5.63. □

We also need the following proposition. It will be used with $g_n = h_n$ of the form $h_n = \frac{1}{n}\sum_{k=0}^{n-1} S^k f$ where $f \in L^1(\Omega, \mathcal{F}, \mu)$.

5.64. Proposition. *Let $\{g_n\}_n$ and $\{h_n\}_n$ be sequences in $L^1(\Omega, \mathcal{F}, \mu)$ with the property that for every $c > 0$ the subsets $\{\sup_n |g_n| > c\}$ and $\{\sup_n |h_n| > c\}$ have finite μ-measure. Define, for $-\infty < a < b < \infty$, the subset $C_{a,b}^{\{g_n\}_n,\{h_n\}_n}$ by*

$$C_{a,b}^{\{g_n\}_n,\{h_n\}_n} = \left\{\liminf_{n\to\infty} g_n < a < b < \limsup_{n\to\infty} h_n\right\}. \tag{5.97}$$

Then the following equality holds:

$$S1_{C_{a,b}^{\{g_n\}_n,\{h_n\}_n}} = 1_{C_{a,b}^{\{Sg_n\},\{Sh_n\}_n}}. \tag{5.98}$$

Proof. We write the function $1_{C_{a,b}^{\{g_n\}_n,\{h_n\}_n}}$ as follows:

$$1_{C_{a,b}^{\{g_n\}_n,\{h_n\}_n}} = \sup_{N_1}\inf_{N_1'}\sup_{N_2}\inf_{N_2'}\min_{N_1\leqslant n_1\leqslant N_1'}\max_{N_2\leqslant n_2\leqslant N_2'} 1_{\{g_{n_1}<a<b<h_{n_2}\}}, \tag{5.99}$$

where the suprema and infima are monotone limit operations in $L^1(\Omega, \mathcal{F}, \mu)$. An appeal to assertion (v) in Lemma 5.61, to (5.83), and to (5.84) the equality in (5.99) implies

$$S1_{C_{a,b}^{\{g_n\}_n,\{h_n\}_n}} = \sup_{N_1}\inf_{N_1'}\sup_{N_2}\inf_{N_2'}\min_{N_1\leqslant n_1\leqslant N_1'}\max_{N_2\leqslant n_2\leqslant N_2'} S1_{\{g_{n_1}<a<b<h_{n_2}\}}. \tag{5.100}$$

The equality in (5.96) in combination with (5.100) then shows

$$\begin{aligned} S1_{C_{a,b}^{\{g_n\}_n,\{h_n\}_n}} &= \sup_{N_1}\inf_{N_1'}\sup_{N_2}\inf_{N_2'}\min_{N_1\leqslant n_1\leqslant N_1'}\max_{N_2\leqslant n_2\leqslant N_2'} 1_{\{Sg_{n_1}<a<b<Sh_{n_2}\}} \\ &= 1_{C_{a,b}^{\{Sg_n\}_n,\{Sh_n\}_n}}. \end{aligned} \tag{5.101}$$

The equality in (5.101) completes the proof of Proposition 5.64. ☐

5.65. Theorem (Maximal ergodic theorem). *Let* $S: L^1(\Omega,\mathcal{F},\mu)\to L^1(\Omega,\mathcal{F},\mu)$ *be a linear map, which is positivity preserving and contractive. So that* $f\in L^1(\Omega,\mathcal{F},\mu)$ *and* $f\geqslant 0$ *implies* $Sf\geqslant 0$ *and* $\|Sf\|_1\leqslant\|f\|_1$. *Define for* $f\in L^1(\Omega,\mathcal{F},\mu)$ *the to* S *corresponding maximal function* $\widetilde{f}$ *by*

$$\widetilde{f}=\sup_{n\in\mathbb{N}}\frac{1}{n}\sum_{k=0}^{n-1}S^kf. \tag{5.102}$$

Then the following assertions are valid:

(a) *If* f *belongs to* $L^1(\Omega,\mathcal{F},\mu)$, *then* $\displaystyle\int_{\{\widetilde{f}>0\}} f\,d\mu\geqslant 0$;

(b) *If, in addition,* $\min(Sf,1)=S(\min(f,1))$, *for all* $f\in L^1(\Omega,\mathcal{F},\mu)$, $f\geqslant 0$, *then for any* $a>0$ *and any* $f\in L^1(\Omega,\mathcal{F},\mu)$, *the following inequalities hold:*

$$\mu\{Sf>a\}\leqslant\mu\{f>a\},\quad \text{and}\quad a\mu\left\{\widetilde{f}>a\right\}\leqslant\|f\|_1. \tag{5.103}$$

Observe that the second inequality in (5.103) resembles the Doob's maximal inequality for sub-martingales: see Theorem 5.110 or Proposition 3.107.

Proof. (a) Let $f\in L^1(\Omega,\mathcal{F},\mu)$, and define, for n a positive integer, the function h_n by

$$h_n=\max_{0\leqslant k\leqslant n-1}\max\left(0,\sum_{j=0}^{k}S^jf\right). \tag{5.104}$$

Then we have $h_{n+1}\geqslant h_n\geqslant 0$, and for $\omega\in\Omega$ such that $h_{n+1}(\omega)>0$, we have $Sh_n(\omega)+f(\omega)\geqslant_{n+1}(\omega)$. The latter inequality is a consequence of the inequality

$$\begin{aligned} f+Sh_n &= f+S\left(\max_{0\leqslant k\leqslant n-1}\max\left(0,\sum_{j=0}^{k}S^jf\right)\right) \\ &\geqslant f+\max_{0\leqslant k\leqslant n-1}S\left(\max\left(0,\sum_{j=0}^{k}S^jf\right)\right) \\ &\geqslant \max_{0\leqslant k\leqslant n-1}\left(\max\left(f+S0,f+\sum_{j=0}^{k}S^{j+1}f\right)\right) \\ &= \max_{0\leqslant k\leqslant n-1}\left(\max\left(f,\sum_{j=0}^{k+1}S^jf\right)\right)=\max_{0\leqslant k\leqslant n}\left(\sum_{j=0}^{k}S^jf\right). \end{aligned} \tag{5.105}$$

From (5.105) it readily follows that

$$f + Sh_n \geqslant h_{n+1} \quad \text{on} \quad \{h_{n+1} > 0\}. \tag{5.106}$$

Notice that in the arguments leading to $h_{n+1} \geqslant h_n$, and also to (5.105), we employed the fact that $g \geqslant 0$, $g \in L^1(\Omega, \mathcal{F}, \mu)$, implies $Sg \geqslant 0$. From (5.106) we infer

$$\begin{aligned}
\int_{\{h_{n+1}>0\}} f\,d\mu &\geqslant \int_{\{h_{n+1}>0\}} \left(h_{n+1} - Sh_n\right)\,d\mu \\
&= \int_{\{h_{n+1}>0\}} h_{n+1}\,d\mu - \int_{\{h_{n+1}>0\}} Sh_n\,d\mu \\
&\geqslant \int h_{n+1}\,d\mu - \int Sh_n\,d\mu \geqslant \int h_{n+1}\,d\mu - \int h_n\,d\mu \\
&= \int \left(h_{n+1} - h_n\right)\,d\mu \geqslant 0.
\end{aligned} \tag{5.107}$$

From (5.107) we obtain

$$\int_{\{\tilde{f}>0\}} f\,d\mu = \int_{\bigcup_{n=0}^{\infty}\{h_{n+1}>0\}} f\,d\mu = \lim_{n\to\infty} \int_{\{h_{n+1}>0\}} f\,d\mu \geqslant 0. \tag{5.108}$$

The inequality in (5.108) entails assertion (a).

(b) Let $f \geqslant 0$ belong to $L^1(\Omega, \mathcal{F}, \mu)$, and fix $a > 0$. We first prove that $\mu\{Sf > a\} \leqslant \mu\{f > a\}$, *i.e.* the first inequality in (5.103). Let m be a positive integer. By the extra hypothesis in (b), together with assertion (ii) in Lemma 5.61, we have

$$\min(m\max(Sf - a, 0), 1) = S(\min(m\max(f - a, 0), 1)). \tag{5.109}$$

From (5.109) we deduce

$$\begin{aligned}\int \min(m\max(Sf - a, 0), 1)\, d\mu &= \int S(\min(m\max(f - a, 0), 1))\, d\mu \\ &\leqslant \int \min(m\max(f - a, 0), 1)\, d\mu. \end{aligned}\tag{5.110}$$

In (5.110) we let m tend to ∞ to obtain:

$$\begin{aligned}\mu\{Sf > a\} &= \lim_{m\to\infty} \int \min(m\max(Sf - a, 0), 1)\, d\mu \\ &\leqslant \lim_{m\to\infty} \int \min(m\max(f - a, 0), 1)\, d\mu = \mu\{f > a\}. \end{aligned}\tag{5.111}$$

This proves the first inequality in (5.103). In order to show the second inequality in (5.103) we proceed as follows. Let f be a member of $L^1(\Omega, \mathcal{F}, \mu)$, and define, always for $a > 0$ fixed, the subset D by $D = \left\{\widetilde{f} > a\right\}$. Here $\widetilde{f}$ is as in (5.106). In addition, define for n a positive integer, the subset D_n by

$$D_n = \bigcup_{k=0}^{n} \left\{S^k f > a\right\} \cap D.$$

Then we have

$$\mu\{D_n\} \leqslant \sum_{k=0}^{n} \mu\left\{S^k f > a\right\} \leqslant \sum_{k=0}^{n} \mu\{f > a\} = (n+1)\mu\{f > a\} \leqslant \frac{n+1}{a}\|f\|_1,$$

and so $\mu\{D_n\}$ is finite. We also have $D \supset D_{n+1} \supset D_n$, and $D = \bigcup_{n=1}^{\infty} D_n$. Hence, because for $f \in L^1(\Omega, \mathcal{F}, \mu)$ we have

$$\widetilde{f} - a \leqslant \widetilde{f} - a\widetilde{1_{D_n}} \leqslant (\widetilde{f - a1_{D_n}}),$$

it follows that

$$\begin{aligned}a\mu\{D_n\} &= \int_{\{\widetilde{f}>a\}} a1_{D_n}\, d\mu = \int_{\{\widetilde{f}-a>0\}} a1_{D_n}\, d\mu \leqslant \int_{\{(\widetilde{f-a1_{D_n}})>0\}} a1_{D_n}\, d\mu \\ &= \int_{\{(\widetilde{f-a1_{D_n}})>0\}} (a1_{D_n} - f)\, d\mu + \int_{\{(\widetilde{f-a1_{D_n}})>0\}} f\, d\mu \end{aligned}\tag{5.112}$$

(apply assertion (a) to the first terem in (5.112))

$$\leqslant 0 + \int_{\{(\widetilde{f-a1_{D_n}})>0\}} f\, d\mu \leqslant \|f\|_1. \tag{5.113}$$

In (5.113) we let n tend to ∞ and infer $a\mu\left\{\widetilde{f} > a\right\} \leqslant \|f\|_1$. This is the second inequality in (5.103), and completes the proof of Theorem 5.65. $\square$

5.66. THEOREM (Theorem of Birkhoff). *Let $S : L^1(\Omega, \mathcal{F}, \mu) \to L^1(\Omega, \mathcal{F}, \mu)$ be a linear operator such that for every $f \geqslant 0$, $f \in L^1(\Omega, \mathcal{F}, \mu)$, the following two conditions are satisfied:*

(i) $S(\min(f, 1)) = \min(Sf, 1)$;
(ii) $\|Sf\|_1 := \int |Sf|\, d\mu \leqslant \int f\, d\mu = \|f\|_1$.

(It follows that all properties mentioned in Lemma 5.61 are available as well as the Propositions 5.63 and 5.64, and Theorem 5.65.) Then for every $f \in L^1(\Omega, \mathcal{F}, \mu)$ the pointwise limit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} S^k f =: P_\mu f \tag{5.114}$$

exists μ-almost everywhere. In addition, $P_\mu f$ belongs to $L^1(\Omega, \mathcal{F}, \mu)$, and the operator P_μ is a projection operator, i.e. *$P_\mu^2 = P_\mu$, with the following properties:*

(a) $\int |P_\mu f|\, d\mu \leqslant \int |f|\, d\mu$, *and*
(b) $SP_\mu f = P_\mu S f = P_\mu f$, *where f belongs to $L^1(\Omega, \mathcal{F}, \mu)$.*

If the measure μ is a probability measure, then the limit in (5.114) is also an L^1-limit, and $P_\mu f = \mathbb{E}_\mu\left[f \mid \mathcal{I}\right]$, $f \in L^1(\Omega, \mathcal{F}, \mu)$, where $\mathbb{E}_\mu\left[f \mid \mathcal{I}\right]$ denotes the conditional expectation on the σ-field of invariant events: $\mathcal{I} = \{A \in \mathcal{F} : S1_A = 1_A\}$.

Before we prove this theorem we insert a corollary.

5.67. COROLLARY. *Let the notation and hypotheses be as in Theorem 5.66. Suppose that the operator S is ergodic in the sense that $S1 = 1$, and $Sf = f$, $f \in L^1(\Omega, \mathcal{F}, \mu)$, implies $f =$ constant μ-almost everywhere. If $\mu(\Omega) = \infty$, then $P_\mu f = 0$, $f \in L^1(\Omega, \mathcal{F}, \mu)$. If $\mu(\Omega) = 1$, then $P_\mu f = \int f\, d\mu$, $f \in L^1(\Omega, \mathcal{F}, \mu)$.*

PROOF. Let $f \in L^1(\Omega, \mathcal{F}, \mu)$. Then $SP_\mu f = P_\mu f$, and so by ergodicity $P_\mu f$ is a constant μ-almost everywhere. If $\mu(\Omega) = \infty$, then this constant must be zero, because $P_\mu f$ belongs to $L^1(\Omega, \mathcal{F}, \mu)$. If $\mu(\Omega) = 1$, then, by the L^1-version of Theorem 5.66, we have

$$\int P_\mu f\, d\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int S^k f\, d\mu = \int f\, d\mu, \tag{5.115}$$

and the inequality in (5.115) completes the proof of Corollary 5.67. $\square$

PROOF OF THEOREM 5.66. Define for $f \in L^1(\Omega, \mathcal{F}, \mu)$, and $-\infty < a < b < \infty$ the subset $C^f_{a,b}$ by

$$C^f_{a,b} = \left\{ \liminf_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} S^k f < a < b < \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} S^k f \right\}. \tag{5.116}$$

Then $C^f_{a.b}$ belongs to $\mathcal{F}$, and by Theorem 5.65 it follows that $\mu\left[C^f_{a,b}\right] < \infty$. By Proposition 5.64 we see that

$$S1_{C^f_{a,b}} = 1_{C^{Sf}_{a,b}} = 1_{C^f_{a,b}}. \tag{5.117}$$

As in equality (5.102) in Theorem 5.65 we write $\widetilde{g}$ for the maximal function corresponding to $g \in L^1(\Omega, \mathcal{F}, \mu)$. Then, with $C = C^f_{a,b}$, we see

$$\begin{aligned}(\widetilde{(a-f)\,1_C}) &= \sup_n \frac{1}{n}\sum_{k=0}^{n-1} S^k\left\{(a-f)\,1_C\right\} = \sup_n \frac{1}{n}\sum_{k=0}^{n-1}\left\{a - S^{k-1}f\right\}1_C \\ &= \left(a - \inf_n \frac{1}{n}\sum_{k=0}^{n-1} S^k f\right)1_C, \end{aligned}\tag{5.118}$$

and so from (5.118) and the definition of $C = C^f_{a,b}$ we see that

$$C \subset \left\{(\widetilde{(a-f)\,1_C}) > 0\right\}. \tag{5.119}$$

The inclusion in (5.119) together with Theorem 5.65 yields

$$\int (a-f)\,1_C\,d\mu = \int_C (a-f)\,1_C\,d\mu = \int_{\left\{(\widetilde{(a-f)1_C})>0\right\}} (a-f)\,1_C\,d\mu \geqslant 0. \tag{5.120}$$

As a consequence (5.120) implies $\int_C f\,d\mu \leqslant a\mu(C)$. A similar reasoning shows that $C \subset \left\{(\widetilde{(f-b)\,1_C}) > 0\right\}$, and therefore, like in (5.120),

$$\int (f-b)\,1_C\,d\mu = \int_{\left\{(\widetilde{(f-b)1_C})>0\right\}} (f-b)\,1_C\,d\mu \geqslant 0,$$

and hence $b\mu(C) \leqslant \int_C f\,d\mu$. Since (5.120) entails $\int_C f\,d\mu \leqslant a\mu(C)$, we obtain $b\mu(C) \leqslant a\mu(C)$. Since $b > a$ and $\mu(C)$ we get $\mu\left(C^f_{a,b}\right) = 0$. The subset C_0 defined by

$$C_0 = \left\{\liminf_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} S^k f < \limsup_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} S^k f\right\}$$

can be written in the form

$$C_0 = \bigcup_{-\infty<a<b<\infty,\, a,b\in\mathbb{Q}} C^f_{a,b},$$

where the symbol $\mathbb{Q}$ denotes the set of rational numbers. So, by what is proved above the set C_0 can be covered by a countable collection of subsets of the form $C^f_{a,b}$, $-\infty < a < b < \infty$, all of which have μ-measure 0. Whence, $\mu(C_0) = 0$. So the pointwise limit in (5.114) exists μ-almost everywhere.

(a) Next we prove assertion (a), and therefore $P_\mu f$ belongs to $L^1(\Omega, \mathcal{F}, \mu)$. By Fatou's lemma we have

$$\int |P_\mu f|\,d\mu = \int \left|\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} S^k f\right| d\mu \leqslant \int \liminf_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \left|S^k f\right|\,d\mu$$

$$\leqslant \liminf_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1}\int \left|S^k f\right| \, d\mu \leqslant \liminf_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1}\int |f| \, d\mu = \int |f| \, d\mu. \tag{5.121}$$

The inequality in (5.121) shows property (a).

The fact that $P_\mu^2 = P_\mu$ follows from (b). First let $f \geqslant 0$ belong to $L^1(\Omega, \mathcal{F}, \mu)$. Then

$$P_\mu f = \liminf_{n\to} \frac{1}{n}\sum_{k=0}^{n-1} S^k f = \sup_N \inf_{N'\geqslant N} \min_{N\leqslant n\leqslant N'} \frac{1}{n}\sum_{k=0}^{n-1} S^k f,$$

and hence

$$\begin{aligned} SP_\mu f &= \liminf_{n\to} \frac{1}{n}\sum_{k=0}^{n1} S^k f = \sup_N \inf_{N'\geqslant N} \min_{N\leqslant n\leqslant N'} \frac{1}{n}\sum_{k=0}^{n-1} S^{k+1} f \\ &= \liminf_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} S^{k+1} f = P_\mu S f = P_\mu f, \end{aligned}$$

which implies property (b) for non-negative functions in $L^1(\Omega, \mathcal{F}, \mu)$. A general function can be written as a difference of non-negative functions in $L^1(\Omega, \mathcal{F}, \mu)$. This proves property (b).

Next we assume that $\mu(\Omega) = 1$. First we will show that the pointwise limit in (5.114) is in fact also an L^1-limit. For this purpose we fix $f \in L^1(\Omega, \mathcal{F}, \mu)$ and a real number $M > 0$. Then we have

$$\begin{aligned} &\int \left|\frac{1}{n}\sum_{k=0}^{n-1} S^k f - P_\mu f\right| d\mu \\ &\leqslant \int \left|\left(\frac{1}{n}\sum_{k=0}^{n-1} S^k - P_\mu\right)\left(f 1_{\{|f|<M\}}\right)\right| d\mu + \int \left|\left(\frac{1}{n}\sum_{k=0}^{n-1} S^k - P_\mu\right)\left(f 1_{\{|f|\geqslant M\}}\right)\right| d\mu \\ &\leqslant \int \left|\left(\frac{1}{n}\sum_{k=0}^{n-1} S^k - P_\mu\right)\left(f 1_{\{|f|<M\}}\right)\right| d\mu + 2\int_{\{|f|\geqslant M\}} |f| \, d\mu. \end{aligned} \tag{5.122}$$

Since $|g| \leqslant M$, $g \in L^1(\Omega, \mathcal{F}, \mu)$ implies $|Sg| \leqslant M$ and $|P_\mu g| \leqslant M$, the integrand in the first term of the right-hand side of (5.122) is dominated by the constant L^1-function $2M$. So from Lebesgue's dominated convergence theorem, (5.114) and (5.122) it follows that

$$\limsup_{n\to\infty} \int \left|\frac{1}{n}\sum_{k=0}^{n-1} S^k f - P_\mu f\right| d\mu \leqslant 2\int_{\{|f|\geqslant M\}} |f| \, d\mu. \tag{5.123}$$

Since $M > 0$ is arbitrary and $f \in L^1(\Omega, \mathcal{F}, \mu)$ the inequality in (5.123) implies

$$\limsup_{n\to\infty} \int \left|\frac{1}{n}\sum_{k=0}^{n-1} S^k f - P_\mu f\right| d\mu = 0,$$

which is the same as saying that the limit in (5.114) also holds in L^1-sense. The equality $P_\mu f = \mathbb{E}_\mu\left[f \mid \mathcal{I}\right]$, $f \in L^1(\Omega, \mathcal{F}, \mu)$, follows from the following two facts:

(a) the collection $\{A \in \mathcal{F} : S1_A = 1_A\}$ is a σ-field, which is readily established.
(b) Moreover, if $f \geqslant 0$ is such that $Sf = f$, and if $\alpha > 0$, then $S1_{\{f>\alpha\}} = 1_{\{Sf>\alpha\}} = 1_{\{f>\alpha\}}$.

This completes the proof of Theorem 5.66. ☐

4. Projective limits of probability distributions

This section is dedicated to a proof of Kolmogorov's extension theorem. We will also present Carathédory's extension theorem. Let I be an arbitrary set of indices. Denote by $\mathcal{H}(I)$ the class of all finite subsets of I, by $\mathcal{H}'(I)$ the collection of all countable subsets of I, and by $\mathcal{H}''(I) = 2^I$ the class of all subsets of I. Consider a collection of measurable spaces $(\Omega_i, \mathcal{A}_i)$ indexed by $i \in I$. For $J \in \mathcal{H}''(I)$ we write $\Omega_J = \prod_{j \in J} \Omega_j$, and for $J \subset K \subset I$ we denote by p^K_J the canonical projection of Ω_K onto Ω_J. If $J = \{j\} \subset K$ we write p^K_j instead of $p^K_{\{j\}}$, and if $K = I$ we write p_J instead of P^I_J. Hence p_j denotes the (one-dimensional-)projection of Ω_I on its j-th coordinate Ω_j. Often these coordinate functions $\{p_j : j \in I\}$ serve as a canonical stochastic process. On each Ω_J, $J \in \mathcal{H}''(I)$, we consider the σ-field $\mathcal{A}_J = \otimes_{j \in J} \mathcal{A}_j$, generated by the set of projections $\{p^J_j : j \in J\}$, *i.e.* the smallest σ-field containing the sets

$$\left\{\left(p^J_j\right)^{-1}(A) : j \in J,\, A \in \mathcal{A}_J\right\} = \left\{\left\{p^J_j \in A\right\} : j \in J,\, A \in \mathcal{A}_J\right\}.$$

The collection $\mathcal{A}_J$ is called the product σ-algebra on Ω_J. One easily sees that, if $J \in \mathcal{H}(I)$, then $\mathcal{A}_J$ is generated by the set of rectangles, *i.e.* by

$$\prod_{j\in J}\mathcal{A}_j = \left\{\prod_{j\in J} A_j : A_j \in \mathcal{A}_j,\, j \in J\right\}.$$

If $J \subset K \subset L \subset I$, then we clearly have

$$p_J^L = p_J^K \circ p_K^L. \tag{5.124}$$

It is easily seen that the projection p_J^K, where K, $J \in \mathcal{H}''(I)$, $J \subset K$, from Ω_K onto Ω_J is measurable for the σ-fields $\mathcal{A}_K$ and $\mathcal{A}_J$. The latter is also written as: p_J^K is $\mathcal{A}_K$-$\mathcal{A}_J$-measurable. On Ω_I we consider two classes of subsets:

$$\mathcal{B} = \left\{\{p_J \in A\} = p_J^{-1}(A) : J \in \mathcal{H}(I),\, A \in \mathcal{A}_J\right\}, \quad \text{and} \tag{5.125}$$

$$\mathcal{B}' = \left\{\{p_J \in A\} = p_J^{-1}(A) : J \in \mathcal{H}'(I),\, A \in \mathcal{A}_J\right\}. \tag{5.126}$$

The subsets belonging to $\mathcal{B}$ are called *cylinders* or *cylinder sets*. If $Z \in \mathcal{B}$, respectively $Z \in \mathcal{B}'$, then there exists $J \in \mathcal{H}(I)$, respectively $J' \in \mathcal{H}'(I)$, such that

$$Z = A \times \Omega_{I\backslash J}. \tag{5.127}$$

The inclusions $\mathcal{B} \subset \mathcal{B}' \subset \mathcal{A}_I$ are obvious.

5.68. DEFINITION. Let $\mathcal{B}$ be a subset of the powerset of Ω_I. Then $\mathcal{B}$ is called a *Boolean algebra*, if it is closed under finite union, and under taking complements.

5.69. LEMMA. *The set $\mathcal{B}$ is a Boolean algebra, $\mathcal{B}'$ is a σ-field, and*

$$\sigma\{\mathcal{B}\} = \mathcal{B}' = \mathcal{A}_I. \tag{5.128}$$

PROOF. First we show that $\mathcal{B}$ is a Boolean algebra. Let $Z = (p_J)^{-1}(A) = \{p_J \in A\}$, $J \in \mathcal{H}(I)$, $A \in \mathcal{A}_J$, be a cylinder. Then

$$Z^c := \Omega_I\backslash Z = \Omega_I\backslash (p_J)^{-1}(A) = \{p_J \in \Omega_J\backslash A\} = p_J^{-1}(A^c),$$

which shows that Z^c belongs to $\mathcal{B}$ whenever $Z \in \mathcal{B}$. Furthermore, let $Z_i = p_{J_i}^{-1}(A_i)$, $J_i \in \mathcal{H}(I)$, $A_i \in \mathcal{A}_{J_i}$, $i = 1, \dots, n$, be n cylinders. Then for $J = J_1 \cup \dots \cup J_n$ we have

$$\begin{aligned} Z_1 \cup \dots \cup Z_n &= p_{J_1}^{-1}(A_1) \cup \dots \cup p_{J_n}^{-1}(A_n) \\ &= p_J^{-1}\left(p_{J_1}^J\right)^{-1}(A_1) \cup \dots \cup p_J^{-1}\left(p_{J_n}^J\right)^{-1}(A_n) \\ &= p_J^{-1}\left(\left(p_{J_1}^J\right)^{-1}(A_1) \cup \cdots \left(p_{J_n}^J\right)^{-1}(A_n)\right). \end{aligned} \tag{5.129}$$

Since the sets $p_{J_i}^{-1}(A_i)$ belong to A_J for $i = 1, \dots, n$ the set $Z_1 \cup \cdots Z_n$ is a cylinder.

In the same manner one proves that $\mathcal{B}'$ is a σ-field.

In order to get the equalities in (5.128) it remains to show that $\mathcal{A}_I \subset \sigma\{\mathcal{B}\}$. Considering the definition of $\mathcal{A}_I$ it is sufficient to prove that p_i, $i \in I$, is measurable for $\sigma\{\mathcal{B}\}$ and $\mathcal{A}_I$. However, this follows from (5.125). So the proof of Lemma 5.69 is complete now. ☐

5.70. Remark. The fact that $\mathcal{B}' = \mathcal{A}_J$ is important. It shows that each $B \in \mathcal{A}_I$ only depends on at most a countable number of indices, in the sense that B can be written as $B = A \times \Omega_{I\backslash J}$ where J is countable or finite and where $A \in \mathcal{A}_J$.

The observation in this remark shows that the product σ-field is relatively "poor" when the index set I is uncountable. The following two examples will clarify this.

5.71. Example. Take I uncountable, let each Ω_i, $i \in I$, be an arbitrary topological Hausdorff space with at least two points, and let $\mathcal{A}_i$ be the Borel field of Ω_i. For every $i \in I$ we select $\omega_i \in \Omega_i$. Since the singleton $\{(\omega_i)_{i\in I}\}$ is a closed subset of Ω_I with respect to the product topology, it belongs to the Borel σ-field of Ω_I. But it does not belong to $\mathcal{A}_I$ because it cannot be written as the set B in Remark 5.70.

5.72. Example. Take $I = [0, \infty)$ and suppose that $\Omega_i = \Omega$, where Ω is a topological Hausdorff space consisting of at least two points. Hence $\Omega_{[0,\infty)} = \Omega^{[0,\infty)}$ is the set of all mappings from $[0, \infty)$ to Ω. Let B be the subset of $\Omega^{[0,\infty)}$ consisting of all right-continuous (or all continuous) mappings from $[0, \infty)$ to Ω. Assuming that B belongs to $\mathcal{A}_{[0,\infty)} = \mathcal{A}^{\otimes[0,\infty)}$, where $\mathcal{A}$ is the Borel σ-field of Ω will lead to a contradiction. Because, if B belongs to $\mathcal{A}_{[0,\infty)}$, then by Remark 5.70 B is of the form $B = A \times \Omega_{[0,\infty)\backslash J}$ where $J \subset [0, \infty)$ is countable, and where $A \in \mathcal{A}_J$. We may suppose that J contains all rational numbers. Pick $f \in B$ and $t \in [0, \infty)\backslash J$. We define the function $g : [0, \infty) \to \Omega$ as follows: $g(s) = f(s)$ if $s \neq t$, and $g(s) \neq f(t)$ if $s = t$. Then $g \in B$, but it is not right-continuous, which can be seen as follows. In J there exists a sequence $(t_n)_n$ which decreases to t. Then

$$\lim_{n\to\infty} g(t_n) = \lim_{n\to\infty} f(t_n) = f(t) \neq g(t).$$

It follows that $B \notin \mathcal{A}_{[0,\infty)}$.

5.73. Definition. Consider a family of measurable spaces $(\Omega_i, \mathcal{A}_i)$, $i \in I$. Suppose that for every $J \in \mathcal{H}(I)$ $\mathbb{P}_J$ is a probability measure on $(\Omega_J, \mathcal{A}_J)$ such that

$$\mathbb{P}_K\left[p_J^K \in A\right] = \mathbb{P}_K\left[\left(p_J^K\right)^{-1}(A)\right] = \mathbb{P}_J[A], \tag{5.130}$$

whenever J, $K \in \mathcal{H}(I)$, $J \subset K$, and $A \in \mathcal{A}_J$. Then the family $\{\mathbb{P}_J : J \in \mathcal{H}(I)\}$, or the family $\{(\Omega_J, \mathcal{A}_J, \mathbb{P}_J) : J \in \mathcal{H}(I)\}$, is called a projective system of probability measures, or spaces. Such a system is also called a consistent system, or a cylindrical measure.

The following theorem says that a cylinder measure is a genuine measure provided that the spaces Ω_i are topological Hausdorff spaces which are Polish, endowed with their Borel σ-fields $\mathcal{B}_i$. From Theorem 5.51 it follows that all probability measures μ on a Polish space S are inner and outer regular in the sense that

$$\mu(B) = \sup\{\mu(K) : K \subset B,\ K \text{ compact}\} = \inf\{\mu(O) : O \supset B,\ O \text{ open}\} \tag{5.131}$$

whenever B belongs to the Borel σ-field of S. The following theorem is a slight reformulation of Theorem 3.1. We also make the following observations. A second-countable locally-compact Hausdorff space is Polish. See Theorem 1.16, and see the formula in (1.18) which gives the metric. As mentioned earlier this construction can be found in Garrett [**75**].

A countable disjoint union of Polish spaces (E_j, d_j) is Polish, with metric

$$d(x,y) = \begin{cases} 1, & \text{(for } x,\ y \text{ in distinct spaces in the union)}, \\ d_n(x,y) & \text{(for } x,\ y \text{ in the } n\text{th space in the union)}. \end{cases} \tag{5.132}$$

Here we assume that $d_j(x,y) \leqslant 1$, x, $y \in E_j$. From this result it follows that a σ-compact metrizable Hausdorff space $E = \cup_{j=1}^{\infty} K_j$, $K_j \subset K_{j+1}$, K_j compact, is a Souslin space, *i.e.* a continuous image of a Polish space. This is so because every subset K_j is compact metrizable, and therefore separable. Therefore the complements $K_{j+1} \backslash K_j$, $j \in \mathbb{N}$, are Polish, and since E is the disjoint union of such spaces E itself is Polish. It is known that probability measures on the class of Borel subsets of Souslin spaces are regular. For details the reader is referred to Bogachev [**27**]. Before we formulate and prove the Kolmogorov's extension theorem we will discuss the Carathéodory's extension theorem. We need the notion of *semi-ring*, *ring*, and *(Boolean) algebra* of subsets of a given set Ω.

5.74. Definition (Definitions). Let Ω be a given set. A semi-ring is a subset $\mathcal{S}$ of $\mathcal{P}(\Omega)$, the power set of Ω, which has the following properties:

(i) $\varnothing \in \mathcal{S}$;
(ii) For all A, $B \in \mathcal{S}$, the intersection $A \cap B$ belongs to $\mathcal{S}$ ($\mathcal{S}$ is closed under pairwise intersections);
(iii) For all A, $B \in \mathcal{S}$, there exist disjoint sets $K_i \in \mathcal{S}$, with $i = 1, 2, \dots, n$, such that $A \backslash B = \bigcup_{i=1}^{n} K_i$ (relative complements can be written as finite disjoint unions).

A ring $\mathcal{R}$ is a subset of the power set of Ω which has the following properties:

(i) $\varnothing \in \mathcal{R}$;
(ii) For all A, $B \in \mathcal{R}$, the union $A \cup B$ belongs to $\mathcal{R}$ ($\mathcal{R}$ is closed under pairwise unions);
(iii) For all A, $B \in \mathcal{R}$, the relative complement $A \backslash B$ belongs to $\mathcal{R}$ ($\mathcal{R}$ is closed under relative complements).

Thus any ring on Ω is also a semi-ring.

A Boolean algebra $\mathcal{B}$ is defined as a subset of the power set of Ω with the following properties:

(i) $\varnothing \in \mathcal{B}$;
(ii) For all $A \in \mathcal{B}$ and $B \in \mathcal{B}$ the union $A \cup B$ belongs to $\mathcal{B}$;
(iii) If A belongs to $\mathcal{B}$, then its complement $A^c = \Omega \backslash A$ belongs to $\mathcal{B}$.

Sometimes, the following constraint is added in the measure theory context: Ω is the disjoint union of a countable family of sets in $\mathcal{S}$.

Without proof we mention some properties. Arbitrary (possibly uncountable) intersections of rings on Ω are still rings on Ω. If $\mathcal{A}$ is a non-empty subset of $\mathcal{P}(\Omega)$, then we define the ring generated by $\mathcal{A}$ (noted $\mathcal{R}(\mathcal{A})$) as the smallest ring containing $\mathcal{A}$. It is straightforward to see that the ring generated by $\mathcal{A}$ is equivalent to the intersection of all rings containing $\mathcal{A}$.

For a semi-ring $\mathcal{S}$, the set containing all finite disjoint union of sets of $\mathcal{S}$ is the ring generated by $\mathcal{S}$:

$$\mathcal{R}(\mathcal{S}) = \left\{ A : \, A = \bigcup_{i=1}^{n} A_i, \, A_i \in \mathcal{S} \right\}.$$

This means that $\mathcal{R}(\mathcal{S})$ is simply the set containing all finite unions of sets in $\mathcal{S}$.

A *content* μ defined on a semi-ring $\mathcal{S}$ can be extended on the ring generated by $\mathcal{S}$. Such an extension is unique. The *extended content* is necessarily given by:

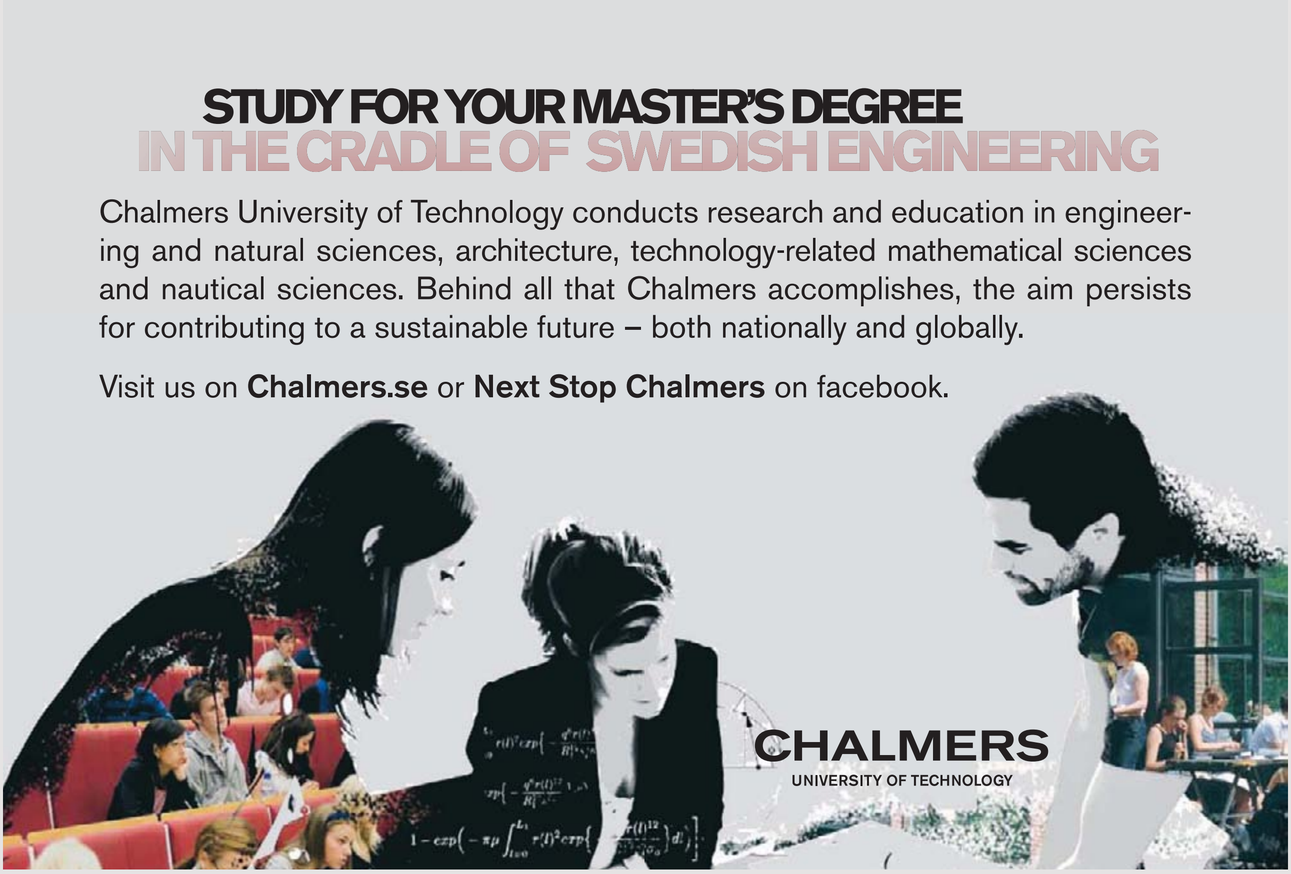
$$\mu(A) = \sum_{i=1}^{n} \mu(A_i) \quad \text{for } A = \bigcup_{i=1}^{n} A_i, \text{ with the } A_i \in \mathcal{S}\text{'s mutually disjoint.}$$

In addition, it can be proved that μ is a pre-measure if and only if the extended content is also a pre-measure, and that any pre-measure on $\mathcal{R}(\mathcal{S})$ that extends the pre-measure on $\mathcal{S}$ is necessarily of this form.

Some motivation is at place here. In measure theory, one is usually not interested in semi-rings and rings themselves, but rather in σ-algebras (or σ-fields) generated by them. The idea is that it is possible to build a pre-measure on a semi-ring $\mathcal{S}$ (for example Stieltjes measures), which can then be extended to a pre-measure on $\mathcal{R}(\mathcal{S})$, which can finally be extended to a genuine measure on a σ-algebra through Carathéeodory's extension theorem. As σ-algebras generated by semi-rings and rings are the same, the difference does not really matter (in the measure theory context at least). Actually, the Carathéodory's extension theorem can be slightly generalized by replacing ring with semi-ring.

5.75. DEFINITION. Let $\mathcal{S}$ be a semi-ring in $\mathcal{P}(\Omega)$. A *pre-measure* on $\mathcal{S}$ is a map $\mu : \mathcal{S} \to [0, \infty]$ such that

(i) $\mu(\varnothing) = 0$.
(ii) If $(A_n)_n$ is a mutually disjoint sequence in $\mathcal{S}$, and if $A := \bigcup_n A_n$ belongs to $\mathcal{S}$, then $\mu(A) = \sum_n \mu(A_n) = \lim_{N\to\infty} \sum_{n=1}^{N} \mu(A_n)$.



We also need the concept of outer or exterior measure.

5.76. Definition. An *outer measure* on $\mathcal{P}(\Omega)$ is a map $\lambda : \mathcal{P}(\Omega) \to [0, \infty]$ with the following properties:

(i) $\lambda(\varnothing) = 0$.
(ii) $A \subset B$ implies $\lambda(A) \leqslant \lambda(B)$.
(iii) If $(A_n)_n$ is a sequence in $\mathcal{P}(\Omega)$, then $\lambda\left(\bigcup_n A_n\right) \leqslant \sum_n \lambda(A_n)$.

By taking all but finitely many A_n to be the empty set one sees that an outer measure is sub-additive: $\lambda(A \cup B) \leqslant \lambda(A) + \lambda(B)$, A, $B \in \mathcal{P}(\Omega)$. Let λ be an outer measure on $\mathcal{P}(\Omega)$. We define Σ_λ to be the set of all subsets $A \subset \Omega$ such that for any $D \subset \Omega$ we have

$$\lambda(D) = \lambda(A \cap D) + \lambda(A^c \cap D). \tag{5.133}$$

Since an outer measure λ is sub-additive we may replace the equality in (5.133) be an inequality of the form

$$\lambda(D) \geqslant \lambda(A \cap D) + \lambda(A^c \cap D). \tag{5.134}$$

In other words, Σ_λ consists of all subsets $A \subset \Omega$ that split Ω in two in a good way. Clearly, $\Omega \in \Sigma_\lambda$ and by the very form of the definition of Σ_λ, we have a subset A belongs to Σ_λ if and only if its complement A^c belongs to Σ_λ. We now present the following proposition, whose proof is a bit tedious. For details the reader is referred to, *e.g.*, [**7**] or [**15**]. The reader may also want to consult the Probability Tutorials by Noel Vaillant: `http://www.probability.net/`. The sets in Σ_λ are called *Carathéodory measurable* relative to the outer measure λ.

5.77. Proposition. *Let λ be an outer measure on Ω, and let Σ_λ be as defined above. Then Σ_λ is a σ-algebra on Ω.*

The Lebesgue-Stieltjes integral $\int_a^b f(x)\,dg(x)$ is defined when $f : [a,b] \to \mathbb{R}$ is Borel-measurable and bounded and $g : [a,b] \to \mathbb{R}$ is of bounded variation in $[a,b]$ and right-continuous, or when f is Borel-measurable and non-negative and g is non-decreasing, and right-continuous. Define $w((s,t]) := g(t) - g(s)$ and $w(\{a\}) := 0$ (Alternatively, the construction works for g left-continuous, $w([s,t)) := g(t) - g(s)$ and $w(\{b\}) := 0$). By Carathéodory's extension theorem (Theorem 5.79), there is a unique Borel measure μ_g on $[a,b]$ which agrees with w on every interval $I \subset [a,b]$. The measure μ_g arises from the outer measure

$$\mu_g(E) = \inf\left\{\sum_i \mu_g(I_i) : E \subset \bigcup_i I_i\right\},$$

where the infimum is taken over all coverings of E by countably many semi-open intervals I_i. This measure is sometimes called the Lebesgue-Stieltjes, or Stieltjes measure associated with g. The Lebesgue-Stieltjes integral $\int_a^b f(x)\,dg(x)$ is defined as the Lebesgue integral of f with respect to the measure μ_g in the usual way. If g is non-increasing, then define $\int_a^b f(x)\,dg(x) := -\int_a^b f(x)\,d(-g)(x)$. If the function $g : [a,b] \to \mathbb{R}$ is right-continuous, and of bounded variation on $[a,b]$, then g may be written in the form $g = g_1 - g_2$, where the functions g_1

and g_2 are monotone non-decreasing and right-continuous. So that $\mu_g\,(s,t] = g(t) - g(s)$, $a \leqslant s < t \leqslant b$, extends to a real-valued measure on the Borel field of $[a,b]$. Of course, if g were right-continuous, complex-valued and of bounded variation, then g can be split as follows $g = \text{Re } g + i\text{Im } g = g_1 - g_2 + i\,(g_3 - g_4)$ where the functions g_j, $1 \leqslant j \leqslant 4$, are right-continuous, and non-decreasing. To the function g we can associate a complex-valued measure μ_g such that $\mu_g\,(s,t] = g(t) - g(s)$, $a \leqslant s < t \leqslant b$. For more details on Riemann-Stieltjes integrals the reader is referred to [**167**]. The book by Tao [**175**] contains a discussion on Stieltjes measures. The definition of semi-ring may seem a bit convoluted, but the following simple example shows why it is useful.

5.78. Example. Think about the subset of $\mathcal{P}\,(\mathbb{R})$ defined by the set of all half-open intervals $(a,b]$ for a and b reals. This is a semi-ring, but not a ring. Stieltjes measures are defined on intervals; the countable additivity on the semi-ring is not too difficult to prove because we only consider countable unions of intervals which are intervals themselves. Proving it for arbitrary countably union of intervals is proved using Carathéodory's extension theorem.

Now we are ready to formulate the Carathéodory's extension theorem.

5.79. Theorem (Carathéodory's extension theorem). *Let $\mathcal{R}$ be a ring on Ω and $\mu : \mathcal{R} \to [0,\infty]$ be a pre-measure on a $\mathcal{R}$. Then there exists a measure $\mu' : \sigma\,(\mathcal{R}) \to [0,\infty]$ such that μ' is an extension of μ. (That is, $\mu'\,\big|_{\mathcal{R}} = \mu$). Here $\sigma\,(\mathcal{R})$ is the σ-algebra generated by $\mathcal{R}$.*

If μ is σ-finite then the extension μ' is unique (and also σ-finite).

If $\mathcal{R}$ is a Boolean algebra, then Theorem 5.79 is also called the Hahn-Kolmogorov extension theorem. A complete proof can also be found in [**27**] Theorem 1.5.6. We will present just an outline. Another interesting book is Tao [**175**]; in particular see Theorems 1.7.3 (Carathéodory's extension theorem) and 1.7.8 together with Exercise 1.7.7 (Hahn-Kolmogorov's extension theorem). An (older) paper, which treats Carathéodory's extension theorem thoroughly, is Maharam [**117**].

Proof. The proof is based on the σ-field corresponding to the outer (or exterior) measure associated to pre-measure μ. This exterior measure μ^* is defined by

$$\mu^*\,(A) = \inf\left\{\sum_{k=1}^{\infty} \mu\,(A_k) : \ A_k \in \mathcal{R},\ A \subset \bigcup_{k=1}^{\infty} A_k\right\}, \quad A \subset \Omega. \tag{5.135}$$

(If A can not be covered by a countable union of sets in $\mathcal{R}$, then we put $\mu^*\,(A) = \infty$.) Then it is not too difficult to prove that μ^* is an outer measure. Like in Proposition 5.77 let Σ_{μ^*} be the σ-field consisting of those subsets A of Ω for which

$$\mu^*\,(D) \geqslant \mu^*\,(A \cap D) + \mu^*\,(A^c \cap D)$$

for all $D \subset \Omega$. Then it follows that the σ-field Σ_{μ^*} contains the ring $\mathcal{R}$. Put $\mu'\,(B) = \mu^*\,(B)$, $B \in \Sigma_{\mu^*}$. Then μ' is a measure on Σ_{μ^*} which extends μ, and which unique provided that μ is σ-finite. For details see [**27**] Theorem 1.5.6. This concludes an outline of the proof of Theorem 5.79. □

5.80. Example. Let E be a σ-compact topological Hausdorff space, and assume that each compact subset K_j is metrizable, and hence separable. Define the sequence of open subsets $(O_j)_j$ of E as follows: $O_0 = \emptyset$, $O_1 = E \backslash K_1$, $O_{j+1} = (E \backslash K_1) \cap \cdots \cap (E \backslash K_j)$, $j \geqslant 1$. Then, for an appropriate metric $(x, y) \mapsto d_j(x, y)$, x, $y \in K_j \cap O_j$, $0 \leqslant d_j(x, y) \leqslant 1$, the spaces $K_j \cap O_j$ is complete metrizable and separable, and so a Polish space. Moreover, by construction the spaces $K_j \cap O_j$, $j = 0, 1, \ldots$, are mutually disjoint, and so the E can be supplied with the metric $d(x, y)$ defined by $d(x, y) = 1$, if x, y belong to different spaces $K_j \cap O_j$, and $d(x, y) = d_j(x, y)$, if x and y belong to $K_j \cap O_j$, $j = 0, 1, \ldots$. Then this metric turns E written as a disjoint union of $K_j \cap O_j$ into a Polish space. Its topology is stronger than the original one, and hence E itself is continuous image of a Polish space (via the identity map). It follows that E is a Souslin space.

Example 5.80 should be compared with the notion of disjoint unions of Polish spaces are again Polish: see (5.132).

5.81. Theorem (Kolmogorov's extension theorem). *Let*

$$\{(\Omega_J, \mathcal{B}_J, \mathbb{P}_J) : J \in \mathcal{H}(I)\}$$

be a projective system of probability spaces. Suppose that for every $i \in I$, Ω_i is a Polish space (or Souslin space) endowed with its Borel σ-field $\mathcal{B}_i$. Then there exists a unique probability measure $\mathbb{P}_I$ on $(\Omega_I, \mathcal{B}_I)$ such that

$$\mathbb{P}_I [p_J \in A] = \mathbb{P}_I \left[p_J^{-1}(A)\right] = \mathbb{P}_J [A], \quad A \in \mathcal{B}_J, \tag{5.136}$$

for every $J \in \mathcal{H}(I)$.

Proof. If $Z = \{p_J \in A\} = p_J^{-1}(A)$, $J \in \mathcal{H}(I)$, $A \in \mathcal{B}_J$, is a cylinder in Ω_I, then we define $\mathbb{P}_I[Z]$ by

$$\mathbb{P}_I [Z] = \mathbb{P}_J [A]. \tag{5.137}$$

This definition is unambiguous. Indeed, let

$$Z = \{p_J \in A\} = p_J^{-1}(A) = p_K^{-1}(B) = \{p_K \in B\}, \quad \text{with } A \in \mathcal{B}_J \text{ and } B \in \mathcal{B}_K,$$

with $J, K \in \mathcal{H}(I)$. We have to show that $\mathbb{P}_J[A] = \mathbb{P}_K[B]$. Indeed, with $L = J \cup K$, we get

$$\begin{aligned} Z &= p_L^{-1}\left(p_J^L\right)^{-1}(A) = \left\{p_J^L \circ p_L \in A\right\} \\ &= \left\{p_J^L \in A\right\} = \left(p_J^L\right)^{-1}(A) = \left(p_K^L\right)^{-1}(B) = \left\{p_K^L \in B\right\} \\ &= \left\{p_K^L \circ p_L \in B\right\} = p_L^{-1}\left(p_K^L\right)^{-1}(B). \end{aligned} \tag{5.138}$$

From (5.130) together with (5.138) we infer

$$\mathbb{P}_J [A] = \mathbb{P}_L \left[\left(p_J^L\right)^{-1}(A)\right] = \mathbb{P}_L \left[\left(p_K^L\right)^{-1}(B)\right] = \mathbb{P}_K [B]. \tag{5.139}$$

The equality in (5.139) shows that $\mathbb{P}_I$ is well defined. We also have

$$\mathbb{P}_I [\Omega_I] = \mathbb{P}_I \left[p_i^{-1}(\Omega_i)\right] = \mathbb{P}_{\{i\}} [\Omega_i] = 1.$$

Next we show that $\mathbb{P}_I$ is finitely additive on $\mathcal{B}$, the collection of cylinders. Let $Z = p_J^{-1}(A)$, with $J \in \mathcal{H}(I)$ and $A \in \mathcal{B}_J$, and $Z' = p_K^{-1}(B)$, with $K \in \mathcal{H}(I)$ and $B \in \mathcal{B}_K$ be two disjoint cylinders. Put $L = J \cup K$. Then we have

$$\begin{aligned} \emptyset = Z \cap Z' &= p_L^{-1}\left(p_J^L\right)^{-1}(A) \cap p_L^{-1}\left(p_K^L\right)^{-1}(B) \\ &= p_L^{-1}\left(\left(p_J^L\right)^{-1}(A) \cap \left(p_K^L\right)^{-1}(B)\right). \end{aligned} \tag{5.140}$$

From (5.140) we infer $\left(p_J^L\right)^{-1}(A) \cap \left(p_K^L\right)^{-1}(B) = \emptyset$. Consequently, we obtain

$$\begin{aligned} \mathbb{P}_I [Z \cup Z'] &= \mathbb{P}_I \left[p_L^{-1}\left(p_J^L\right)^{-1}(A) \cup p_L^{-1}\left(p_K^L\right)^{-1}(B)\right] \\ &= \mathbb{P}_I \left[p_L^{-1}\left(\left(p_J^L\right)^{-1}(A) \cup \left(p_K^L\right)^{-1}(B)\right)\right] \\ &= \mathbb{P}_L \left[\left(p_J^L\right)^{-1}(A) \cup \left(p_K^L\right)^{-1}(B)\right] \\ &= \mathbb{P}_L \left[\left(p_J^L\right)^{-1}(A)\right] + \mathbb{P}_L \left[\left(p_K^L\right)^{-1}(B)\right] \end{aligned}$$

(apply the equality in (5.130))

$$= \mathbb{P}_J [A] + \mathbb{P}_K [B] = \mathbb{P}_I [A] + \mathbb{P}_I [B]. \tag{5.141}$$

The equality in (5.141) proves the finite additivity of the mapping $\mathbb{P}_I$ on the collection of cylinder sets $\mathcal{B}$.

Finally we prove that the mapping $\mathbb{P}_I$ is σ-additive on $\mathcal{B}$. For that purpose we consider a decreasing sequence $(Z_n)_n$ of cylinder sets such that $\mathbb{P}_I[Z_n] \geqslant a > 0$ for all $n \in \mathbb{N}$. We will show that $\bigcap_n Z_n \neq \emptyset$. By contraposition it then follows that $\bigcap_n Z_n = \emptyset$ implies $\lim_{n\to\infty} \mathbb{P}_I[Z_n] = 0$. For each n we have $Z_n = p_J^{-1}(A_n)$ with $J_n \in \mathcal{H}(I)$ and $A_n \in \mathcal{B}_{J_n}$. Of course we may suppose that $J_1 \subset J_2 \subset \cdots \subset J_n \subset \cdots$. Put $J = \bigcup_n J_n$. Then

$$Z_n = p_J^{-1}\left(p_{J_n}^J\right)^{-1}(A_n) = \left(p_{J_n}^J\right)^{-1}(A_n) \times \Omega_{I\setminus J}.$$

Since

$$\bigcap_n Z_n = \left(\bigcap_n \left(p_{J_n}^J\right)^{-1}(A_n)\right) \times \Omega_{I\setminus J} \tag{5.142}$$

we see that $\bigcap_n Z_n \neq \emptyset$ if and only if $\bigcap_n \left(p_{J_n}^J\right)^{-1}(A_n) \neq \emptyset$. This means that our problem is reduced to the problem with $I = J$, *i.e.* to a countable problem. For every $m \in \mathbb{N}$ there exists a compact subset L_{j_m} of Ω_{j_m} with

$$\mathbb{P}_{j_m}\left[\Omega_{j_m}\setminus L_{j_m}\right] \leqslant \frac{a}{4\times 2^m}.$$

Then $L := \prod_m L_{j_m}$ is a compact subset of Ω_J. Furthermore, for every $n \in \mathbb{N}$ we have

$$\mathbb{P}_{J_n}\left[\left(\prod_{j\in J_n} L_j\right)^c\right] = \mathbb{P}_{J_n}\left[\bigcup_{j\in J_n}\left(p_j^{J_n}\right)^{-1}\left(L_j^c\right)\right] \leqslant \sum_{j\in J_n} \mathbb{P}_j\left[\Omega_j\setminus L_j\right] < \frac{a}{4}. \tag{5.143}$$

On the other hand for every $n \in \mathbb{N}$ we choose a compact subset K_n of A_n (in $\mathcal{B}_{J_n}$) such that

$$\mathbb{P}_{J_n}\left[A_n\setminus K_n\right] \leqslant \frac{a}{4\times 2^n}. \tag{5.144}$$

For every $n \in \mathbb{N}$ the set Y_n defined by

$$Y_n = \left(p_{J_1}^{J_n}\right)^{-1}(K_1) \cap \cdots \cap \left(p_{J_{n-1}}^{J_n}\right)^{-1}(K_{n-1}) \cap K_n$$

is a closed subset Ω_{J_n}, and so $Z_n' := p_{J_n}^{-1}(Y_n)$ is a closed cylinder in Ω_J, and $Z_n' \subset Z_n$. In addition, we have

$$Z_n' = p_{J_1}^{-1}(K_1) \cap \cdots \cap p_{J_{n-1}}^{-1}(K_{n-1}) \cap p_{J_n}^{-1}(K_n)$$

the sequence $(Z_n')_n$ is decreasing. We also have

$$\begin{aligned}
\mathbb{P}_I\left[Z_n\setminus Z_n'\right] &= \mathbb{P}_I\left[Z_n\setminus\left(\bigcap_{1\leqslant k\leqslant n} p_{J_k}^{-1}(K_k)\right)\right] \\
&\leqslant \sum_{k=1}^n \mathbb{P}_I\left[Z_n\setminus p_{J_k}^{-1}(K_k)\right] \leqslant \sum_{k=1}^n \mathbb{P}_I\left[Z_k\setminus p_{J_k}^{-1}(K_k)\right] \\
&= \sum_{k=1}^n \mathbb{P}_I\left[p_{J_k}^{-1}(A_k)\setminus p_{J_k}^{-1}(K_k)\right] = \sum_{k=1}^n \mathbb{P}_I\left[p_{J_k}^{-1}(A_k\setminus K_k)\right]
\end{aligned}$$

$$= \sum_{k=1}^{n} \mathbb{P}_{J_k}\left[A_k \backslash K_k\right] < \frac{a}{4}. \tag{5.145}$$

Since, by assumption, $\mathbb{P}_I\left[Z_n\right] \geqslant a$, (5.145) implies

$$\mathbb{P}_{J_n}\left[Y_n\right] = \mathbb{P}_I\left[Z_n'\right] = \mathbb{P}_I\left[Z_n\right] - \mathbb{P}_I\left[Z_n \backslash Z_n'\right] > a - \frac{a}{4} = \frac{3a}{4}. \tag{5.146}$$

Since, by (5.146) and (5.143) we have

$$\mathbb{P}_{J_n}\left[Y_n \bigcap \prod_{j \in J_n} L_j\right] \geqslant 1 - \mathbb{P}_{J_n}\left[Y_n^c\right] - \mathbb{P}_{J_n}\left[\left(\prod_{j \in J_n} L_j\right)^c\right] \geqslant \frac{3a}{4} - \frac{a}{4} = \frac{a}{2}, \tag{5.147}$$

it follows that $Y_n \bigcap \prod_{j \in J_n} L_j \neq \varnothing$. Moreover, observe that

$$Z_n' \cap L = \left(Y_n \times \Omega_{J \backslash J_n}\right) \bigcap \left(\prod_m L_{j_m}\right) = \left(Y_n \bigcap \prod_{j \in J_n} L_j\right) \times \prod_{j \in J \backslash J_n} L_j, \tag{5.148}$$

and consequently, $Z_n' \cap L \neq \varnothing$. Hence, the decreasing sequence $\left(Z_n' \cap L\right)_n$ consists of non-empty compact subsets of Ω_J. By compactness we get that $\bigcap_n Z_n' \cap L \neq \varnothing$. So we infer

$$\bigcap_n p_{J_n}^{-1}\left(A_n\right) \supset \bigcap_n Z_n' \supset \bigcap_n \left(Z_n' \cap L\right) \neq \varnothing. \tag{5.149}$$

As a consequence of the previous arguments, we see that $\mathbb{P}_I$ is a σ-additive on the Boolean algebra $\mathcal{B}$ which consists of cylinders in Ω_I. This measure $\mathbb{P}_I$ satisfies (5.136). By the classical Carathéodory theorem the mapping $\mathbb{P}_I$ extends in a unique fashion as a probability measure on the σ-field $\sigma\{\mathcal{B}\} = \mathcal{B}_I$. Then, technically speaking, the mapping $\mathbb{P}_I$, defined on the Boolean algebra $\mathcal{B}$, is a pre-measure. This corresponding exterior measure $\mathbb{P}_I^*$ is defined by

$$\mathbb{P}_I^*(A) = \inf\left\{\sum_{k=1}^{\infty} \mu\left(Z_k\right) : Z_k \in \mathcal{B},\ A \subset \bigcup_{k=1}^{\infty} Z_k\right\}. \tag{5.150}$$

Then it is not so difficult to prove that the set function defined by (5.150) is an outer measure indeed. Define the associated σ-field $\mathcal{D}$ by

$$\mathcal{D} = \left\{A \subset \Omega_I : \mathbb{P}_I^*(D) \geqslant \mathbb{P}_I^*\left(A \cap D\right) + \mathbb{P}_I^*\left(A^c \cap D\right) : \text{for all } D \subset \Omega_I\right\}. \tag{5.151}$$

The fact that $\mathcal{D}$ is a σ-field indeed follows from Proposition 5.77: see Theorem 5.79 as well. It is fairly easy to see that $\mathcal{D}$ contains the Boolean algebra $\mathcal{B}$ which consists of the cylinder sets in Ω_I.

This completes the proof of Theorem 5.81. ☐

5. Uniform integrability

The next Theorem is often used as a replacement for the dominated convergence theorem of Lebesgue.

5.82. Theorem (Theorem of Scheffé). *Let* $(\Omega, \mathcal{F}, \mu)$ *be an arbitrary measure space and let* $(f_n : n \in \mathbb{N})$ *be a sequence of non-negative functions in* $L^1(\Omega, \mathcal{F}, \mu)$*. In addition, let the function* f *belong to* $L^1(\Omega, \mathcal{F}, \mu)$*. Suppose that* $f(x) = \lim_{n\to\infty} f_n(x)$ *for* μ*-almost all* $x \in \Omega$*. The following assertions are equivalent:*

(i) $\lim_{n\to\infty} \int |f_n - f| \, d\mu = 0$;
(ii) *The sequence* $(f_n : n \in \mathbb{N})$ *is uniformly integrable;*
(iii) $\lim_{n\to\infty} \int f_n d\mu = \int f d\mu$.

Instead of *uniformly integrable* the term *equi-integrable* is often used. A family $(f_\alpha : \alpha \in \mathcal{A})$ in $L^1(\Omega, \mathcal{F}, \mu)$ is uniformly integrable, if for every $\epsilon > 0$ there exists a function $g \geqslant 0$ in $L^1(\Omega, \mathcal{F}, \mu)$ such that $\int_{\{f_\alpha \geqslant g\}} |f_\alpha| \, d\mu \leqslant \epsilon$ for all $\alpha \in \mathcal{A}$.

5.83. Proposition. *If* μ *is a probability measure, then a family* $(f_\alpha : \alpha \in \mathcal{A})$ *in* $L^1(\Omega, \mathcal{F}, \mu)$ *is uniformly integrable, if and only if for every* $\epsilon > 0$ *there exists a constant* $M_\varepsilon \geqslant 0$ *such that* $\int_{\{|f_\alpha| \geqslant M_\varepsilon\}} |f_\alpha| \, d\mu \leqslant \epsilon$ *for all* $\alpha \in \mathcal{A}$.

Proof of Proposition 5.83. The sufficiency is clear: choose for g_ε a constant function M_ε. Next we show that, if the family $(f_\alpha : \alpha \in \mathcal{A})$ is uniformly integrable, then necessarily for every $\varepsilon > 0$ there exists a constant M_ε such that

$$\int_{\{|f_\alpha| \geqslant M_\varepsilon\}} |f_\alpha| \, d\mu \leqslant \varepsilon, \quad \alpha \in \mathcal{A}. \tag{5.152}$$

Fix $\varepsilon > 0$. By hypothesis we know that there exists a function $g_\varepsilon \in L^1(\Omega, \mathcal{F}, \mu)$, $g_\varepsilon > 0$, such that

$$\int_{\{|f_\alpha| \geqslant g_\varepsilon\}} |f_\alpha| \, d\mu \leqslant \frac{\varepsilon}{2}, \quad \alpha \in \mathcal{A}. \tag{5.153}$$

Then we choose M_ε so large that

$$\int_{\{g_\varepsilon \geqslant M_\varepsilon\}} g_\varepsilon \, d\mu \leqslant \frac{\varepsilon}{2}. \tag{5.154}$$

Then by (5.153) and (5.154) we have

$$\begin{aligned}\int_{\{|f_\alpha| \geqslant M_\varepsilon\}} |f_\alpha| \, d\mu &= \int_{\{M_\varepsilon \leqslant |f_\alpha| < \max(M_\varepsilon, g_\varepsilon)\}} |f_\alpha| \, d\mu + \int_{\{|f_\alpha| \geqslant \max(M_\varepsilon, g_\varepsilon)\}} |f_\alpha| \, d\mu \\ &\leqslant \int_{\{g_\varepsilon > M_\varepsilon\}} g_\varepsilon \, d\mu + \int_{\{|f_\alpha| \geqslant g_\varepsilon\}} |f_\alpha| \, d\mu \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned} \tag{5.155}$$

The inequality in (5.155 completes the proof of Proposition 5.83. ☐

Proof of Theorem 5.82. (i) $\Rightarrow$ (ii). Put $g = \sup_{n \in \mathbb{N}} f_n$. The following inequalities hold for $m \in \mathbb{N}$:

$$\begin{aligned}\int_{\{f_n \geqslant mf\}} f_n d\mu &\leqslant \int_{\{f_n \geqslant mf\}} |f_n - f| \, d\mu + \int_{\{f_n \geqslant mf\}} f d\mu \\ &\leqslant \int |f_n - f| \, d\mu + \int_{\{f_n \geqslant mf\}} f d\mu \\ &\leqslant \int |f_n - f| \, d\mu + \int_{\{g \geqslant mf\}} f d\mu. \end{aligned} \tag{5.156}$$

Let $\epsilon > 0$, but arbitrary. By (i) there exists $N(\epsilon) \in \mathbb{N}$ such that $\int |f_n - f| \, d\mu \leqslant \epsilon/2$ for $n \geqslant N(\epsilon) + 1$. The inequalities below then follow for $m \geqslant M(\epsilon)$:

$$\int_{\{g \geqslant mf\}} f d\mu \leqslant \epsilon/2, \quad \text{and} \quad \int_{\{f_n \geqslant mf\}} f_n d\mu \leqslant \epsilon, \quad 1 \leqslant n \leqslant N(\epsilon). \tag{5.157}$$

From (5.156) and (5.157) we see $\int_{\{f_n \geqslant M(\epsilon) f\}} f_n d\mu \leqslant \epsilon$. But this means that the sequence $(f_n : n \in \mathbb{N})$ is uniformly integrable.

(ii) $\Rightarrow$ (iii). Let $\epsilon > 0$ be arbitrary and choose a function $g_\epsilon \in L^1(\Omega, \mathcal{F}, \mu)$ such that

$$\int_{\{f_n \geqslant g_\epsilon\}} f_n d\mu + \int_{\{f_n \geqslant g_\epsilon\}} f d\mu \leqslant \epsilon. \tag{5.158}$$

From (5.158) we obtain

$$\left| \int f_n d\mu - \int f d\mu \right| \leqslant \int_{\{f_n \leqslant g_\epsilon\}} |f_n - f| \, d\mu + \int_{\{f_n \geqslant g_\epsilon\}} f_n d\mu + \int_{\{f_n \geqslant g_\epsilon\}} f d\mu$$

$$\leqslant \int_{\{f_n \leqslant g_\epsilon\}} |f_n - f|\, d\mu + \epsilon. \tag{5.159}$$

By the theorem of dominated convergence, it follows from (5.159) that

$$\limsup_{n\to\infty} \left| \int f_n d\mu - \int f d\mu \right| \leqslant \epsilon.$$

Since ϵ is arbitrary assertion (iii) follows. The same argumentation shows the implication (ii) $\Rightarrow$ (i).

(iii) $\Rightarrow$ (i). The equality

$$|f_n - f| = f_n - f + 2\left(f - \min\left(f, f_n\right)\right)$$

is obvious. From (iii) together with the theorem of dominated convergence it then follows that

$$\begin{aligned} &\lim_{n\to\infty} \int |f_n - f|\, d\mu \\ &= \lim_{n\to\infty} \int \left(f_n - f\right) d\mu + 2 \int \lim_{n\to\infty} \left(f - \min\left(f, f_n\right)\right) d\mu = 0. \end{aligned}$$

The proof of Theorem 5.82 is now complete. ☐

5.84. Corollary. *Let* $(\mu_m : m \in \mathbb{N})$ *be a sequence of probability measures on the Borel* σ*-field of* $\mathbb{R}^\nu$. *Let every measure* μ_m *have a probability density* g_m *relative to the Lebesgue measure* λ. *Furthermore, let* $g \geqslant 0$ *be a probability density. Suppose that for* λ*-almost all* $x \in \mathbb{R}^\nu$ *the equality* $\lim_{m\to\infty} g_m(x) = g(x)$ *is true. Let the measure* μ *have density* g. *Then the sequence* $(\mu_m : m \in \mathbb{N})$ *converges weakly to* μ.

Proof of Corollary 5.84. From the theorem of Scheffé (Theorem 5.82) we see

$$\lim_{m\to\infty}\int |g_m(x)-g(x)|\,dx = 0.$$

Let f be a bounded continuous function. Then

$$\left|\int f d\mu_m - \int f d\mu\right| = \left|\int (f(x)g_m(x) - f(x)g(x))\,dx\right| \leqslant \|f\|_\infty \int |g_m(x)-g(x)|\,dx. \tag{5.160}$$

The assertion in Corollary 5.84 follows from (5.160). ☐

5.85. Theorem. *Let $(X_m : m \in \mathbb{N})$ be a sequence of stochastic variables, which are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*

(a) *If the sequence $(X_m : m \in \mathbb{N})$ converges in probability to a stochastic variable X, then the sequence of probability measures $(\mathbb{P}_{X_m} : m \in \mathbb{N})$ converges weakly to the distribution $\mathbb{P}_X$;*
(b) *If the sequence $(\mathbb{P}_{X_m} : m \in \mathbb{N})$ converges vaguely to the Dirac-measure δ_a, then the sequence $(X_m : m \in \mathbb{N})$ converges in probability to a stochastic variable X, which is $\mathbb{P}$-almost surely equal to the constant a.*

Proof. (a) Suppose that the sequence $(X_m : m \in \mathbb{N})$ converges in probability to X. We pick $f \in C_{00}(\mathbb{R}^\nu)$ and we will prove that $\lim_{m\to\infty}\int f d\mathbb{P}_{X_m} = \int f d\mathbb{P}_X$. The latter is equivalent to $\lim_{m\to\infty}\int f(X_m)\,d\mathbb{P} = \int f(X)\,d\mathbb{P}$. The function f is uniformly continuous. So, for $\epsilon > 0$ given, there exists $\delta > 0$ such that

$$|x_2 - x_1| \leqslant \delta \quad \text{impliceert} \quad |f(x_2) - f(x_1)| \leqslant \epsilon. \tag{5.161}$$

Put $A_m = \{|X - X_m| \geqslant \delta\}$. For $\omega \notin A_m$ the inequality

$$|f(X_m(\omega)) - f(X(\omega))| \leqslant \epsilon$$

holds. From this it follows that

$$\begin{aligned}\left|\int f d\mathbb{P}_{X_m} - \int f d\mathbb{P}_X\right| &\leqslant \int_{A_m^c} |f(X) - F(X_m)|\,d\mathbb{P} + \int_{A_m} |f(X) - f(X_m)|\,d\mathbb{P} \\ &\leqslant \epsilon\mathbb{P}(A_m^c) + 2\|f\|_\infty \mathbb{P}\{|X_m - X| \geqslant \delta\} \\ &\leqslant \epsilon + 2\|f\|_\infty \mathbb{P}\{|X_m - X| \geqslant \delta\}. \end{aligned} \tag{5.162}$$

The assertion in (a) follows from (5.162) together with assertion (3) in Theorem 5.43.

(b) Suppose that the sequence $(\mathbb{P}_{X_m} : m \in \mathbb{N})$ vaguely converges to the Dirac-measure δ_a. Let $I(\epsilon)$ be the interval $I(\epsilon) = [a-\epsilon, a+\epsilon]$ and choose functions f and $g \in C_{00}(\mathbb{R}^\nu)$ such that $f \leqslant 1_I \leqslant g$ and such that $f(a) = g(a) = 1$. Then the equalities follow:

$$\begin{aligned} f(a) &= \liminf_m \int f d\mathbb{P}_{X_m} \leqslant \liminf_m \mathbb{P}_{X_m}(I) \leqslant \limsup_m \mathbb{P}_{X_m}(I) \\ &\leqslant \limsup_m \int g d\mathbb{P}_{X_m} = g(a). \end{aligned} \tag{5.163}$$

From (5.163) it follows that

$$\lim_m \mathbb{P}\left(|X_m - a| \leqslant \epsilon\right) = 1,$$

which amounts to the same as

$$\lim_m \mathbb{P}\left(|X_m - a| > \epsilon\right) = 0.$$

This proves assertion (b). So the proof of Theorem 5.85 is now complete. □

6. Stochastic processes

We begin with some definitions.

5.86. Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(E, \mathcal{E})$ be a locally compact Hausdorff space, that satisfies the second countability axiom, with Borel σ-field $\mathcal{E}$. Often E will be chosen as $\mathbb{R}$ or as $\mathbb{R}^\nu$. A *stochastic process* X with values in the *state space* E is a mapping $X : [0, \infty) \times \Omega \to E$. For every $\omega \in \Omega$ the mapping $t \mapsto X(t, \omega)$ defines a *path* of the process. A path is sometimes also called a *realization*. If we fix $n \in \mathbb{N}$, then the mappings $\mathbb{P}_{t_1,\dots,t_n} : \underbrace{\mathcal{E} \otimes \cdots \otimes \mathcal{E}}_{n\times} \to [0, 1]$, where $(t_1, \dots, t_n)$ varies over $[0, \infty)^n$, and which are defined by

$$\mathbb{P}_{t_1,\dots,t_n}(B) = \mathbb{P}\left\{(X(t_1), \dots, X(t_n)) \in B\right\}, \quad B \in \underbrace{\mathcal{E} \otimes \cdots \otimes \mathcal{E}}_{n\times}, \tag{5.164}$$

are called the *n-dimensional distributions* of the process X. Here $X(t)$ is the mapping $X(t)(\omega) = X(t, \omega)$, $\omega \in \Omega$.

Sometimes we write X_t instead of $X(t)$. If $n = 1$, then the distributions in (5.164) are also called the *marginal* distributions, or *marginals*. However, notice that a process is much more than the corresponding collection of finite-dimensional distributions. In particular the paths or realizations of a process are very important. For example, the continuity properties of the paths are relevant. Often we will suppose that the paths are continuous, or that they are continuous from the right, and possess limits from the left {càdlàg paths}, or *cadlag paths*. So that the process X is *cadlag* provided that for all $t \geqslant 0$ the equality $\lim_{s\downarrow t} X(s) = X(t)$ holds $\mathbb{P}$-almost surely (this is continuity from the right, or continue à droite in French) and if the limit $\lim_{s\uparrow t} X(s)$ exists in E (this means that the left limits exist in E, limité à gauche in French).

5.87. Definition. A family sub-σ-fields $(\mathcal{F}_t : t \geqslant 0)$ of $\mathcal{F}$ is called a *filtration* (or, sometimes, also called *history*), if $t < s$ implies $\mathcal{F}_t \subset \mathcal{F}_s$. Thus the probability $\mathbb{P}$ is defined on all σ-fields $\mathcal{F}_t$. With $\mathcal{F}_\infty$, or also $\mathcal{F}_{\infty-}$ the σ-field generated by $\bigcup_{t\geqslant 0} \mathcal{F}_t$ is meant. If for every $t \geqslant 0$ the equality $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ holds, then the filtration $(\mathcal{F}_t : t \geqslant 0)$ is called *continuous from the right*, or *right-continuous*. Let $(\mathcal{F}_t : t \geqslant 0)$ be a filtration, and put $\mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s$. Then the family $(\mathcal{F}_{t+} : t \geqslant 0)$ is a right-continuous filtration. This filtration is called the right closure of the filtration $(\mathcal{F}_t : t \geqslant 0)$. A subset A of Ω is called a $\mathbb{P}$-*null* *set* if there exists a subset $A_0 \in \mathcal{F}$ with the following properties: $A \subseteq A_0$ and $\mathbb{P}[A_0] = 0$. Usually this is expressed by saying that A is a null set instead of A is a $\mathbb{P}$-null set. Often it is assumed that $\mathcal{F}_0$ contains all null sets, and that the filtration $(\mathcal{F}_t : t \geqslant 0)$ is right-continuous. Sometimes it is said that $\mathcal{F}_0$ has the usual properties. The process X is called *adapted* to the filtration $(\mathcal{F}_t : t \geqslant 0)$ if for every $t \geqslant 0$ the *state variable* $X(t)$ is measurable with respect to σ-fields $\mathcal{F}_t$ and $\mathcal{E}$. Let $\mathcal{H}_t = \sigma(X(u) : 0 \leqslant u \leqslant t)$ be the σ-field generated by the state variables $X(u)$, $0 \leqslant u \leqslant t$. The filtration $(\mathcal{H}_t : t \geqslant 0)$ is called the *internal history* of the process X. If $t > 0$ is given, then $\mathcal{H}_t$ is called the (information from the) past, $\sigma(X(t))$ is called the (information from the) present, and $\sigma(X(u) : u \geqslant t)$ the (information from the) future. The process X is adapted if and only if $\mathcal{H}_t \subseteq \mathcal{F}_t$ for every $t \geqslant 0$.

5.88. DEFINITION. Let X and Y be two processes. The processes X and Y are said to be *non-$\mathbb{P}$-distinguishable* or *$\mathbb{P}$-indistinguishable* provided there exists a $\mathbb{P}$-null subset N with the property that for every $\omega \notin N$ and for every $t \geqslant 0$ the equality $X(t,\omega) = Y(t,\omega)$ holds. The process X is called a *modification* of the process Y (or also Y is a modification of X) if for every $t \geqslant 0$ there exists a $\mathbb{P}$-null set N_t with the property that $X(t,\omega) = Y(t,\omega)$ for $\omega \notin N_t$. Thus the null set is t-dependent. If the processes X and Y are not distinguishable, then X is a modification of Y. In general, the converse statement is not true.

5.89. THEOREM. *Suppose that the process X as well as the process Y possesses right-continuous paths. If X is a modification of Y, then X and Y are not distinguishable (also called stochastically equivalent).*

PROOF OF THEOREM 5.89. Let X be a modification of the process Y. For every $t \geqslant 0$ there then exists a null set N_t such that $X(t) = Y(t)$ on the complement of N_t. Put $N = \bigcup_{t\in\mathbb{Q}} N_t$. Then $\mathbb{P}(N) = 0$ and for every $t \in \mathbb{Q}$ the equality $X(t) = Y(t)$ holds on the complement of N. By right-continuity of the paths it then follows that

$$X(t) = \lim_{s\downarrow 0, s\in\mathbb{Q}} X(s) = \lim_{s\downarrow t, s\in\mathbb{Q}} Y(s) = Y(t)$$

on the complement of N and completes the proof of Theorem 5.89. ☐

5.90. DEFINITION. Let $(\mathcal{F}_t : t \geqslant 0)$ be a filtration and let $T : \Omega \to [0, \infty]$ be a "stochastic time". The function T is called a *stopping time* for the filtration $(\mathcal{F}_t : t \geqslant 0)$ if for every fixed time t the event $\{T \leqslant t\}$ belongs to $\mathcal{F}_t$. Since the event $\{T < \infty\} = \bigcup_{n\in\mathbb{N}} \{T \leqslant n\}$ belongs to $\mathcal{F}_\infty$, the complementary event $\{T = \infty\}$ is also an element of $\mathcal{F}_\infty$.

5.91. THEOREM. *Let $(\mathcal{F}_t : t \geqslant 0)$ be a filtration. Let $(\mathcal{F}_{t+} : t \geqslant 0)$ be the so-called right closure of the filtration $(\mathcal{F}_t : t \geqslant 0)$. Then a stochastic time $T : \Omega \to [0, \infty]$ is a stopping time for the filtration $(\mathcal{F}_{t+} : t \geqslant 0)$ if and only if, for every $t > 0$, the event $\{T < t\}$ belongs to $\mathcal{F}_t$.*

PROOF. "*Sufficiency*" Suppose that for every $t \geqslant 0$ the event $\{T < t\}$ belongs to $\mathcal{F}_t$. Then the event $\{T \leqslant t\} = \bigcap_{n\in\mathbb{N}} \left\{T < 1 + \frac{1}{n}\right\}$ belongs to the σ-field $\bigcap_{n\in\mathbb{N}} \mathcal{F}_{t+n^{-1}} = \mathcal{F}_t$.

"*Necessity*" Assume that for every $t \geqslant 0$ the event $\{T \leqslant t\}$ belongs to $\mathcal{F}_{t+}$. Then the event $\{T < t\} = \bigcup_{n\in\mathbb{N}} \left\{T \leqslant 1 - \frac{1}{n}\right\}$ belongs to $\bigcup_{n\in\mathbb{N}} \mathcal{F}_{t-n^{-1}+} \subset \mathcal{F}_t$. This completes the proof of Theorem 5.91. ☐

5.92. COROLLARY. *Let $(\mathcal{F}_t : t \geqslant 0)$ be a right-continuous filtration. Then the stochastic time T is a $(\mathcal{F}_t : t \geqslant 0)$-stopping time if and only if for every $t \geqslant 0$ the event $\{T < t\}$ belongs to $\mathcal{F}_t$ and this is the case for every $t > 0$ if and only if for every $t > 0$ the event $\{T \leqslant t\}$ belongs to $\mathcal{F}_t$.*

5.93. THEOREM. *Let $(\mathcal{F}_t : t \geqslant 0)$ be a right-continuous filtration, let X be an adapted cadlag process, let G be an open subset and let F be a closed subset of E. In addition, let $(G_n : n \in \mathbb{N})$ be a sequence of open subsets of E such that $F = \bigcap_n G_n$ and such that $G_n \supset G_{n+1}$, $n \in \mathbb{N}$. Finally, let $(F_n : n \in \mathbb{N})$ be an increasing sequence of closed subsets with the property that $G = \bigcup_n F_n$. Define the times S, S_n, T and T_n by means of the equalities:*

$$\begin{aligned} S &= \inf\{s \geqslant 0 : X(s) \in F \text{ or } X(s-) \in F\}; \\ S_n &= \inf\{s \geqslant 0 : X(s) \in F_n \text{ or } X(s-) \in F_n\}; \\ T_n &= \inf\{s \geqslant 0 : X(s) \in G_n\} \quad \text{and } T = \inf\{s \geqslant 0 : X(s) \in G\}. \end{aligned} \tag{5.165}$$

Then these times are stopping times and the following assertions hold: $S_n \downarrow T$ and $T_n \uparrow S$.

Proof. Let $t > 0$. Since the paths are continuous from the right se see

$$\{T < t\} = \bigcup_{0<r<t} \{X(r) \in G\} = \bigcup_{0<r<t, r\mathbb{Q}} \{X(r) \in G\} \in \mathcal{F}_t.$$

This proves that T is a stopping time. Since

$$\{S \leqslant t\} = \{X(t) \in F \text{ or } X(t-) \in F\} \cup \left(\bigcap_{n\in\mathbb{N}} \bigcup_{r<t, r\in\mathbb{Q}} \{X(r) \in G_n\}\right) \in \mathcal{F}_t$$

it follows that S is a stopping time as well. Since $G_n \supset G_{n+1}$ it follows that $T_{n+1} \geqslant T_n$. Put $S_0 = \sup T_n$. The ultimate equalities in

$$\begin{aligned}\{S_0 < t\} &= \bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{0\leqslant s\leqslant t-m^{-1}} \{X(s) \in G_n\} \\ &= \bigcup_{m=1}^{\infty} \bigcup_{0\leqslant s\leqslant t-m^{-1}} \{X(s) \in F \text{ of } X(s-) \in F\} \\ &= \bigcup_{0\leqslant s<t} \{X(s) \in F \text{ of } X(s-) \in F\} = \{S < t\}\end{aligned}$$

prove the equalities $\{S_0 < t\} = \{S < t\}$ for all $t > 0$ and hence, $S = S_0$. The fact that $S_n \downarrow T$ is left to the reader as an exercise. This completes the proof of Theorem 5.93. ☐

5.94. Theorem. *Let S and T be stopping times for the filtration $(\mathcal{F}_t : t \geqslant 0)$. Then $\min(S,T)$, $\max(S,T)$ and $S+T$ are also stopping times for this filtration. If $(S_n : n \in \mathbb{N})$ is a sequence of stopping times, then $\sup_n S_n$ is also a stopping time, and if, moreover, the filtration $(\mathcal{F}_t : t \geqslant 0)$ is right continuous, then $\inf_n S_n$ is stopping time as well.*

Proof. The proof is left as an exercise for the reader. ☐

5.95. Definition. Let T be a stopping time for the filtration $(\mathcal{F}_t : t \geqslant 0)$. The σ-field of events which *precedes* T is defined by

$$\mathcal{F}_T := \bigcap_{t\geqslant 0} \{A \in \mathcal{F}_\infty : A \cap \{T \leqslant t\} \in \mathcal{F}_t\}.$$

Indeed, the collection $\mathcal{F}_T$ is a σ-field and if $T = t$ is a fixed time, then $\mathcal{F}_T = \mathcal{F}_t$. If $S \leqslant T$ is also a stopping time, then $\mathcal{F}_S \subset \mathcal{F}_T$. If the filtration $(\mathcal{F}_t : t \geqslant 0)$ is continuous from the right, then an event A belongs to $\mathcal{F}_T$ if and only if A belongs to $\mathcal{F}_\infty$, and if for every $t > 0$ the event $A \cap \{T < t\}$ belongs to $\mathcal{F}_t$. If S and T are stopping times, then $\mathcal{F}_{\min(S,T)} = \mathcal{F}_S \cap \mathcal{F}_T$. If the filtration $(\mathcal{F}_t : t \geqslant 0)$ is right continuous and if $(S_n : n \in \mathbb{N})$ is a sequence of stopping times which converges downward to S, then S is a stopping time and $\bigcap_{n\in\mathbb{N}} \mathcal{F}_{S_n} = \mathcal{F}_S$.

5.96. Definition. A process $X : [0,\infty) \times \Omega \to E$ is called *progressively measurable* for the filtration $(\mathcal{F}_t : t \geqslant 0)$ if for every $t > 0$ the restriction of X to $[0,t] \times \Omega$ is measurable for the σ-fields $\mathcal{B}[0,t] \otimes \mathcal{F}_t$ and $\mathcal{E}$.

5.97. Theorem. *If X is right-continuous adapted process, then X is progressively measurable.*

PROOF. Define the sequence of processes $(X^n : n \in \mathbb{N})$ by means of the formula:

$$X^n(u,\omega) = \begin{cases} X\left(\frac{k+1}{2^n}t,\omega\right), & \text{if } k2^{-n}t < u \leqslant (k+1)2^{-n}t,\ 0 \leqslant k \leqslant 2^n-1; \\ 0, & \text{if } u = 0. \end{cases} \tag{5.166}$$

Let $B \in \mathcal{E}$. Then we have

$$\begin{aligned} &\{X^n \in B\} \\ &= \{0\} \times \{X(0) \in B\} \cup \bigcup_{0 \leqslant k \leqslant 2^n-1} \left(\left(\frac{k}{2^n}, \frac{k+1}{2^n}\right] \times \left\{X\left(\frac{k+1}{2^n}\right) \in B\right\}\right) \\ &\in \mathcal{B}[0,t] \otimes \mathcal{F}_t. \end{aligned} \tag{5.167}$$

So X^n is progressively measurable. Because the process X is $\mathbb{P}$-almost surely right-continuous it follows that $\lim_{n\to\infty} X^n = X$, and, consequently, X is progressively measurable. This completes the proof of Theorem 5.97. ☐

5.98. THEOREM. *Suppose that X is progressively measurable for the filtration $(\mathcal{F}_t : t \geqslant 0)$. Let T be a stopping time. The the state variable $X(T) : \omega \mapsto X(T(\omega),\omega)$ measurable for the σ-fields $\mathcal{E}$ and $\mathcal{F}_T$.*

Proof of Theorem 5.98. On the event $\{T \leqslant t\}$ the mapping

$$\omega \mapsto X\left(T(\omega), \omega\right)$$

is the composition of the mapping $\omega \mapsto (T(\omega), \omega)$, which goes from $\{T \leqslant t\}$ to $[0, t] \times \Omega$ and which is measurable for the σ-fields $\mathcal{F}_t$ and $\mathcal{B}[0, t] \otimes \mathcal{F}_t$, and the mapping $(u, \omega) \mapsto X(u, \omega)$, which goes from $[0, t] \times \Omega$ to E and which is measurable for the σ-fields $\mathcal{B}[0, t] \otimes \mathcal{F}_t$ and $\mathcal{E}$. In the latter argument the progressive measurability of X was used. The composition of measurable mappings is again measurable, and hence $X(T)$ is measurable for de σ-fields $\mathcal{F}_T$ and $\mathcal{E}$.

This completes the proof of Theorem 5.98. ☐

5.99. Corollary. *If T is a stopping time and if X is progressively measurable, then the process X^T defined by $X^T(u) = X(\min(T, u))$ is adapted to the stopped filtration $\left(\mathcal{F}_{\min(T,u)} : u \geqslant 0\right)$.*

Proof. The proof is left as an exercise for the reader. ☐

The next lemma is often employed instead of the monotone class theorem.

5.100. Lemma. *Let $\mathcal{F}$ be a σ-field on Ω and H a vector space consisting of $\mathcal{F}$-measurable real-valued bounded functions on Ω. Suppose that the following hypotheses are fulfilled:*

1. *H contains the constant functions;*
2. *If f and g belong to H, then the product fg belongs to H;*
3. *If f is the pointwise limit of a sequence of functions $(f_n : n \in \mathbb{N})$ in H, for which $|f_n| \leqslant 1$, then f belongs to H;*
4. *$\mathcal{F} = \sigma\left(f : f \in H\right)$.*

Then H contains all bounded $\mathcal{F}$-measurable functions.

Proof. Let $\mathcal{D}$ be the collection $\mathcal{D} = \{A \in \mathcal{F} : 1_A \in H\}$. Then $\mathcal{D}$ is a Dynkin system and by (2) $\mathcal{D}$ is closed for taking finite intersections. So $\mathcal{D}$ is a σ-field. Pick $f \in H$ and let $a \in \mathbb{R}$. We will prove that the set $\{f \geqslant a\}$ belongs to $\mathcal{D}$. By taking an appropriate combination of f and the constant function $\mathbf{1}$ we may assume that $0 \leqslant f \leqslant 1$ and that $0 \leqslant a \leqslant 1$. Let p be a polynomial. By (2) $p(f)$ belongs to H. Let $\varphi : [0, 1] \to \mathbb{R}$ be a continuous function. By the theorem of Stone-Weierstrass there exists a sequence of polynomials $(p_n : n \in \mathbb{N})$ such that $\sup_{x \in [0,1]} |\varphi(x) - p_n(x)| \leqslant n^{-1}$. Consequently, $\varphi(f)$ belongs to H. Since the function $1_{[a,\infty)}$ is a (decreasing) pointwise limit of a sequence of continuous functions, it follows that $1_{[a,\infty)}(f) = 1_{\{f \geqslant a\}}$ belongs to H. So the set $\{f \geqslant a\}$ belongs to $\mathcal{D}$. From which it follows that $\mathcal{D} = \mathcal{F}$. But then we infer $\mathcal{F} \subset \{A \in \mathcal{F} : 1_A \in H\}$. From this the assertion in Lemma 5.100 immediately follows. ☐

5.101. Definition. Let $(\mathcal{F}_t : t \geqslant 0)$ be a filtration on the probability space

$$(\Omega, \mathcal{F}, \mathbb{P}),$$

and let X be an adapted process.

(i) The process X is called a *martingale* (relative to $\mathbb{P}$ and to the filtration $(\mathcal{F}_t : t \geqslant 0)$) if for every $t \geqslant 0$ the variable $X(t)$ belongs to $L^1(\Omega, \mathcal{F}, \mathbb{P})$ and if for every pair $0 \leqslant s < t$ the equality $X(s) = \mathbb{E}\left(X(t) \mid \mathcal{F}_s\right)$ holds $\mathbb{P}$-almost surely.
(ii) The process X is called a *sub-martingale* (relative to $\mathbb{P}$ and to the filtration $(\mathcal{F}_t : t \geqslant 0)$) if for every $t \geqslant 0$ the variable $X(t)$ belongs to $L^1(\Omega, \mathcal{F}, \mathbb{P})$ and if for every $s < t$ the inequality $X(s) \leqslant \mathbb{E}\left(X(t) \mid \mathcal{F}_s\right)$ holds $\mathbb{P}$-almost surely.
(iii) The process X is called a *super-martingale* (relative to $\mathbb{P}$ and to the filtration $(\mathcal{F}_t : t \geqslant 0)$) if for every $t \geqslant 0$ the variable $X(t)$ belongs to $L^1(\Omega, \mathcal{F}, \mathbb{P})$ and if for every $s < t$ the inequality $X(s) \geqslant \mathbb{E}\left(X(t) \mid \mathcal{F}_s\right)$ holds $\mathbb{P}$-almost surely.

Instead of assuming $X(t) \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ in (ii) it is sometimes assumed that the variable $X(t)^+ = \max(X(t), 0)$ belongs to $L^1(\Omega, \mathcal{F}, \mathbb{P})$. In (iii) it is sometimes only assumed that $X(t)^- = \max(-X(t), 0)$ belongs to $L^1(\Omega, \mathcal{F}, \mathbb{P})$. If T is a (discrete) subset of $[0, \infty)$ and if $(X(t), \mathcal{F}_t)_{t \geqslant 0}$ is a martingale (sub-martingale, super-martingale), then the process $(X(t), \mathcal{F}_t)_{t \in T}$ is so as well. Then we can use "discrete results" and via a limiting procedure we then obtain results in the "continuous case".

5.102. Definition. Let $f : [0, \infty) \to \mathbb{R}$ be a function, let $T \subseteq [0, \infty)$ and let $a < b$ be real numbers. Define the number of *upcrossings* $U_T(f, a, b)$ of $f \big|_T$ between a and b by

$$\begin{aligned} &U_T(f, a, b) \\ &= \sup\left\{m : \text{ there exist } t_1 < t_2 < \ldots < t_{2m},\ t_j \in T\ f\left(t_{2k-1}\right) \leqslant a,\ f\left(t_{2k}\right) \geqslant b\right\}. \end{aligned} \tag{5.168}$$

5.103. Lemma. *Let D by the set of non-negative dyadic numbers and let $f : D \to \mathbb{R}$ be a function, which is bounded on $D \cap [0, n]$ for all $n \in \mathbb{N}$. Assume that, for all $n \in \mathbb{N}$ and for all real numbers $a < b$, with a and b (dyadic) rational, the number of upcrossings $U_{D \cap [0,n]}(f, a, b)$ of f is finite. Then the following assertions are true:*

(a) *For every $t \in \mathbb{R}$ the following left and right limits exist:*

$$\lim_{s \uparrow t, s \in D} f(s) \quad \text{and} \quad \lim_{s \downarrow t, s \in D} f(s); \tag{5.169}$$

(b) *Define the function g by $g(t) = \lim\limits_{s \downarrow t, s \in D} f(s)$. Then g is right-continuous and for every $t > 0$ the left limit $\lim_{s \uparrow t} g(s)$ exists.*

Proof. (a) We will show that the limit $\lim_{s \downarrow t, s > t} f(s)$ exists. Since the function f is bounded it suffices to prove that $\liminf_{s \downarrow t, s > t} f(s) = \limsup_{s \downarrow t, s > t} f(s)$.

Assume that this not the case. Then there exist dyadic rational numbers a and b such that $\liminf_{s\downarrow t, s>t} f(s) < a < b < \limsup_{s\downarrow t, s>t} f(s)$. This means that there exists $s_0 > t$, $s_0 \in D$, with $f(s_0) > b$. There also exists $s_1 < s_0$, $s_1 > t$, $s_1 \in D$, such that $f(s_1) < a$. In general we obtain $t < s_{2k-1} < s_{2k-2}$, $s_{2k-1} \in D$, for which $f(s_{2k-1}) < a$ and we obtain $t < s_{2k} < s_{2k-1}$, $s_{2k} \in D$, with $f(s_{2k}) > b$. For $m \in \mathbb{N}$ we write $t_{2m} = s_0$, $t_{2m-1} = s_1$, $\ldots$, $t_2 = s_{2m-1}$, $t_1 = s_{2m-1}$. Pick $n > s_0$. Then we have $U_{D\cap[0,n]}(f,a,b) \geqslant m$. Since $m \in \mathbb{N}$ is arbitrary it follows that $U_{D\cap[0,n]}(f,a,b) = \infty$. So we obtain a contradiction. The existence of the left limit can be treated similarly.

(b) Put $g(t) = \lim_{s\downarrow t, s>t, s\in D} f(s)$. By (a) this function is well defined. Since, for every $n \in \mathbb{N}$, the function f is bounded on the set $D \cap [0,n]$ the function g possesses this property as well. Let now $(t_n : n \in \mathbb{N})$ be a sequence that decreases to t and for which $t_n > t$ for all $n \in \mathbb{N}$. We will prove $\lim_{n\to\infty} g(t_n) = g(t)$. Then this shows that g is right-continuous at t. Assume $\liminf_{n\to\infty} g(t_n) < g(t)$. This will lead to a contradiction. By passing to a subsequence, which we call again $(t_n : n \in \mathbb{N})$, we may suppose that $\liminf_{n\to\infty} g(t_n) = \lim_{n\to\infty} g(t_n)$ and that there are numbers a and $b \in D$ such that for all $n \in \mathbb{N}$, $g(t_n) < a < b < g(t)$. Then pick $s_0 > t_0$ such that $f(s_0) < a$: this possible, because $g(t_0) < a$. The pick $s_0 > t_0 > s_1 > t$ in such a way that $f(s_1) > b$: this is possible, because $g(t) > b$. Then choose t_{n_2}, $s_1 > t_{n_2} > t$, with $g(t_{n_2}) < a$. Then there exists $s_1 > s_2 > t_{n_2}$, such that $f(s_2) < a$. This is so because $g(t_{n_2}) < a$. This procedure can be continued. Like in (a) we arrive at $U_{D\cap[0,n]}(f,a,b) = \infty$, for a certain $n\mathbb{N}$, $n > t$. This is a contradiction. But then it follows that $\liminf_{n\to\infty} g(t_n) \geqslant g(t)$. In the same fashion we see that $\limsup_{n\to\infty} g(t_n) \leqslant g(t)$. Consequently, $g(t) = \lim_{n\to\infty} g(t_n)$. In order to prove the existence of the left limit of the function g at t, we choose a sequence $(t_n : n \in \mathbb{N})$, that increases to t, and which has the property that $t_n < t$ for all $n \in \mathbb{N}$. Assuming that $\liminf_{n\to\infty} g(t_n) < \limsup_{n\to\infty} g(t_n)$, then, as above, we arrive at the conclusion that, for certain dyadic numbers $a < b$, for which $\liminf_{n\to\infty} g(t_n) < a < b < \limsup_{n\to\infty} g(t_n)$, the number of upcrossings of the function f on the interval $D \cap [0,n]$ with $n > t$ is infinite.

This completes the proof of Lemma 5.103. □

5.104. THEOREM (Doob's optional time theorem for sub-martingales). *Let*

$$(X(j) : j \in \mathbb{N})$$

be a sub-martingale relative to the filtration $(\mathcal{F}_n : n \in \mathbb{N})$*, and let* $T \geqslant S$ *be stopping times. Suppose that* $\mathbb{E}[|X(T)|] < \infty$ *and also* $\mathbb{E}[|X(S)|] < \infty$*. If, additionally,* $\lim_{m\to\infty} \mathbb{E}[X(m) : T \geqslant m \geqslant S] = 0$*, then* $X(S)$ *is measurable for the* σ*-field* $\mathcal{F}_S$ *and the inequality* $\mathbb{E}\left[X(T) \mid \mathcal{F}_S\right] \geqslant X(S)$ *holds* $\mathbb{P}$*-almost surely.*

PROOF. Let A be an event in $\mathcal{F}_S$. For every j, $j \geqslant 1$, and for every $\ell \in \mathbb{N}$, $\ell \geqslant 0$, the event $A \cap \{T \geqslant \ell + j\} \cap \{S = \ell\} \cap A$ then belongs to the σ-field $\mathcal{F}_{\ell+j-1}$. To see this, observe that the event $\{T \geqslant k\} = \Omega \backslash \{T \leqslant k-1\}$ belongs to $\mathcal{F}_{k-1}$. Since

$$(X(\min(T,m)) - X(\min(S,m)))\, 1_A$$

$$= \sum_{\ell=0}^{m} \sum_{j=1}^{m-\ell} \left(X(\ell+j) - X(\ell+j-1)\right) 1_{\{T \geqslant \ell+j\} \cap \{S=\ell\} \cap A},$$

it follows that

$$\begin{aligned} &\mathbb{E}\left(\left(X(\min(T,m)) - X(\min(S,m))\right) 1_A\right) \\ &= \sum_{\ell=0}^{m} \sum_{j=1}^{m-\ell} \mathbb{E}\left(\left(X(\ell+j) - X(\ell+j-1)\right) 1_{\{T \geqslant \ell+j\} \cap \{S=\ell\} \cap A}\right). \end{aligned}$$

Hence, $\mathbb{E}\left(\left(X(\min(T,m)) - X(\min(S,m))\right) 1_A\right) \geqslant 0$. Since, in addition,

$$\begin{aligned} &\mathbb{E}\left(X(T) - X(S) - X(\min(T,m)) + X(\min(S,m))\right) \\ &= \mathbb{E}\left(X(T) - X(S) : S \geqslant m\right) + \mathbb{E}\left(X(T) - X(m) : T \geqslant m > S\right), \end{aligned}$$

the claim in Theorem 5.104 follows. □

5.105. Proposition. *Let $(X(n) : n \in \mathbb{N})$ be a (sub-)martingale relative to the discrete filtration $(\mathcal{F}_n : n \in \mathbb{N})$.*

(a) *Let $H = (H(n) : n \in \mathbb{N}, n \geqslant 1)$ be a positive bounded process with the property that H_n is measurable for the σ-field $\mathcal{F}_{n-1}$. Define the process $(Y(n) : n \in \mathbb{N})$ by*

$$Y(0) = X(0), \quad Y(n) = X(0) + \sum_{k=1}^{n} H(k)\left(X(k) - X(k-1)\right), \quad n \geqslant 1.$$

Then the process Y is a (sub-)martingale. By putting $H(n) = 1_{\{n \leqslant T\}}$, where T is a stopping time we see that process

$$X^T := (X(\min(T, n)) : n \in \mathbb{N})$$

is a (sub-)martingale.

(b) *Let S and T be a pair of bounded stopping times such that $0 \leqslant S \leqslant T$. Then*

$$X(S) \leqslant \mathbb{E}\left(X(T) \mid \mathcal{F}_S\right), \quad \mathbb{P}\text{-almost surely,} \tag{5.170}$$

and if X is a martingale, then there is an equality in (5.170).

Moreover, an adapted and integrable process X is a martingale if and only if $\mathbb{E}(X(T)) = \mathbb{E}(X(S))$ for each pair of bounded stopping times S and T for which $S \leqslant T$.

Proof. (a) The first assertion in (a) is easy to see. To understand the second assertion we observe that $1_{\{T \geqslant n\}} = 1 - 1_{\{T \leqslant n-1\}}$ is measurable for the σ-field $\mathcal{F}_{n-1}$ and we notice that $X(0) + \sum_{k=1}^{n} 1_{\{T \geqslant k\}}(X(k) - X(k-1)) = X(\min(T, n))$. This proves assertion (a) in Proposition 5.105.

(b) De inequality $X(S) \leqslant \mathbb{E}\left(X(T) \mid \mathcal{F}_S\right)$, $\mathbb{P}$-almost surely was already proved in Theorem 5.104 and can be obtained from (a) by putting $H(n) = 1_{\{T \geqslant n\}} - 1_{\{S \geqslant n\}}$. If we use the equality $\mathbb{E}(X(S^B)) = \mathbb{E}(X(T^B))$ for de times $S^B = S1_B + M1_{B^c}$ and $T^B = T1_B + M1_{B^c}$, where B belongs to $\mathcal{F}_S$ and where $M \geqslant T \geqslant S$, then we get

$$\mathbb{E}\left(X(T)1_B + X(M)1_{B^c}\right) = \mathbb{E}\left(X(S)1_B + X(M)1_{B^c}\right).$$

But, then it follows that $\mathbb{E}(X(T)1_B) = \mathbb{E}(X(S)1_B)$ for all $B \in \mathcal{F}_S$ and hence $X(S) = \mathbb{E}\left(X(T) \mid \mathcal{F}_S\right)$.

The proof of Proposition 5.105 is now complete. □

5.106. Theorem (Doob-Meyer decomposition for discrete sub-martingales). *Let $(X(j) : j \in \mathbb{N})$ be a sub-martingale. Then there exists a unique martingale $M = (M(k) : k \in \mathbb{N})$ together with a unique predictable increasing process $A = (A(k) : k \in \mathbb{N})$, with $A(0) = 0$, such that $X(k) = M(k) + A(k)$, for $k \in \mathbb{N}$.*

5.107. Remark. This theorem is, in an appropriate form, also true for sub-martingales X of de form $X = (X(t) : t \geqslant 0)$ (continuous time). A process $A = (A(k) : k \in \mathbb{N})$ is called predictable, if $A(k)$ is measurable for $\mathcal{F}_{k-1}$, and this for every $k \in \mathbb{N}$.

Proof. *Existence* Define the process A by $A(0) = 0$ and

$$A(k) = \sum_{j=1}^{k} \mathbb{E}\left(X(j) - X(j-1) \mid \mathcal{F}_{j-1}\right).$$

Define the process M by $M(k) = X(k) - A(k)$. Then the process M is a martingale and the process A is increasing (i.e. non-decreasing) and predictable. Moreover, the equality $X = M + A$ holds.

Uniqueness Let the process X be such that $X = M + A$ where M is a martingale and where A is predictable and increasing. In addition, suppose that $A(0) = 0$. Then the equalities

$$\begin{aligned}
&\sum_{j=1}^{k} \mathbb{E}\left(X(j) - X(j-1) \mid \mathcal{F}_{j-1}\right) \\
&= \sum_{j=1}^{k} \mathbb{E}\left(M(j) - M(j-1) \mid \mathcal{F}_{j-1}\right) + \sum_{j=1}^{k} \mathbb{E}\left(A(j) - A(j-1) \mid \mathcal{F}_{j-1}\right) \\
&= \sum_{j=1}^{k} \left(A(j) - A(j-1)\right) = A(k),
\end{aligned}$$

hold for $k \geqslant 1$. So the proof of Theorem 5.106 is complete now. ☐

5.108. Theorem. *Let $X = (X(k) : 1 \leqslant k \leqslant N)$ be a sub-martingale. Then the following inequality holds:*

$$\mathbb{E}\left(U_{\{1,\dots,N\}}(X, a, b)\right) \leqslant \frac{\mathbb{E}\left[\max(X(N) - a, 0)\right]}{b - a}.$$

Proof. For a proof we refer the reader to Proposition 3.71 of Chapter 3. Notice that, with X the process $\max(X - a, 0)$ is also a sub-martingale. ☐

5.109. Theorem. *Let $X = (X(t) : t \geqslant 0)$ be a sub-martingale for the filtration $(\mathcal{F}_t : t \geqslant 0)$. For $a < b$ the inequality*

$$\mathbb{E}\left(U_{D\cap[0,N]}\left(X, a, b\right)\right) \leqslant \frac{\mathbb{E}\left[\max\left(X(N) - a, 0\right)\right]}{b - a}.$$

holds.

Proof. Write $D_n = \dfrac{\mathbb{Z}}{2^n}$ and define U_n by $U_n = U_{D_n\cap[0,N]}(X, a, b)$. The sequence U_n then increases to $U_{D\cap[0,N]}(X, a, b)$. So it follows that

$$\mathbb{E}\left(U_{D\cap[0,n]}\right) = \lim_{n\to\infty} \mathbb{E}\left(U_n\right) \leqslant \frac{\mathbb{E}\left[\max(X(N) - a, 0)\right]}{b - a}.$$

This completes the proof of Theorem 5.109. ☐

The following theorem contains Doob's maximal inequalities for submartingales.

5.110. THEOREM. *Let* $X = (X(0), \dots, X(n))$ *be a sub-martingale. Then the following maximal inequalities of Doob hold:*

$$\mathbb{P}\left(\max_{0 \leqslant j \leqslant n} X_j \geqslant \lambda\right) \leqslant \frac{1}{\lambda}\mathbb{E}(X(n)); \qquad (a)$$

$$\mathbb{P}\left(\max_{0 \leqslant j \leqslant n} |X_j| \geqslant \lambda\right) \leqslant \frac{2}{\lambda}\mathbb{E}\left(5\,|X(n)| - 2X(0)\right); \qquad (b)$$

and if X *is a martingale*

$$\mathbb{P}\left(\max_{0 \leqslant j \leqslant n} |X_j| \geqslant \lambda\right) \leqslant \frac{1}{\lambda}\left\{\mathbb{E}\left(|X(n)|\right)\right\}. \qquad (c)$$

PROOF. We begin with a proof of (c). Consider the mutually disjoint events

$$A_0 = \{|X_0| > \lambda\}, \text{ and } A_k := \left\{|X(k)| > \lambda, \max_{0 \leqslant j \leqslant k-1} |X(j)| \leqslant \lambda\right\},$$

$1 \leqslant k \leqslant n$. Then $\bigcup_{k=0}^{n} A_k = \{\max_{0 \leqslant j \leqslant n} |X(j)| \geqslant \lambda\}$. Therefore

$$\mathbb{P}\left[\max_{0 \leqslant j \leqslant n} |X(j)| \geqslant \lambda\right] = \sum_{j=0}^{n} \mathbb{P}\left(A_j\right),$$

and so, using the martingale property

$$\mathbb{P}(A_k) = \mathbb{E}\left(1_{A_k}\right) \leqslant \frac{1}{\lambda}\mathbb{E}\left[1_{A_k}\,|X(k)|\right] \qquad (5.171)$$

(martingale property)

$$= \frac{1}{\lambda}\mathbb{E}\left[1_{A_k}\,|X(n)| \bigm| \mathcal{F}_k\right] \leqslant \frac{1}{\lambda}\mathbb{E}\left[1_{A_k}\mathbb{E}\left(|X(n)| \bigm| \mathcal{F}_k\right)\right] = \frac{1}{\lambda}\mathbb{E}\left[1_{A_k}\,|X(n)|\right].$$

By summing over k in (5.171)we get (c).

(a) The proof of (a) follows almost the same lines, except that in the definitions of the events A_k the absolute value signs have to be omitted.

(b) For the proof of this assertion we employ the Doob-Meyer decomposition theorem (Theorem 5.106). Write $X = M + A$ with M a martingale, and A (predictable) increasing process. We let $M(0) = X(0)$. Then, by (c), we see

$$\begin{aligned}
&\mathbb{P}\left(\max_{0\leqslant j\leqslant n}|X(j)| \geqslant \lambda\right) \leqslant \mathbb{P}\left(\max_{0\leqslant j\leqslant n}|M(j)| \geqslant \frac{\lambda}{2}\right) + \mathbb{P}\left(A(n) \geqslant \frac{\lambda}{2}\right)\\
&\leqslant \frac{2}{\lambda}\mathbb{E}\,|M(n)| + \frac{2}{\lambda}\mathbb{E}(A_n) \leqslant \frac{2}{\lambda}\mathbb{E}\left(|M(n)| - M(n) + X(n)\right)\\
&\leqslant \frac{2}{\lambda}\mathbb{E}\left(2\,|X(n)| + 2A(n) + X(n)\right) \leqslant \frac{2}{\lambda}\mathbb{E}\left(5\,|X(n)| - 2X(0)\right).
\end{aligned}$$

This proves assertion (b).

The proof of Theorem 5.110 is complete now. □

5.111. Lemma. *Let $(\mathcal{A}_n : n \in \mathbb{N})$ be a sequence of σ-fields decreasing to the σ-field $\mathcal{A}_\infty$. So that $\mathcal{A}_{n+1} \subseteq \mathcal{A}_n$, $n \in \mathbb{N}$, and $\mathcal{A}_\infty = \bigcap_{n\in\mathbb{N}}\mathcal{A}_n$. Let $(f_n : n \in \mathbb{N}) \cup \{f_\infty\}$ be a sequence of stochastic variables with the following properties:*

- (i) *f_n is $\mathcal{A}_n$-measurable, $n \in \mathbb{N}$, and f_∞ is $\mathcal{A}_\infty$-measurable;*
- (ii) *$f_m \leqslant \mathbb{E}\left(f_n \,\middle|\, \mathcal{A}_m\right)$, for all $m \geqslant n$, and $f_\infty \leqslant \mathbb{E}\left(f_n \,\middle|\, \mathcal{A}_\infty\right)$, $n \in \mathbb{N}$;*
- (iii) *$\lim_{n\to\infty}\mathbb{E}(f_n) = \mathbb{E}(f_\infty)$.*

Then the sequence $(f_n : n \in \mathbb{N})$ is uniformly integrable.

Proof. For $m = 1, 2, \dots, \infty$ we have

$$f_m \leqslant \mathbb{E}\left(f_1 \,\middle|\, \mathcal{A}_m\right) \text{ and } \max(f_m, 0) \leqslant \mathbb{E}\left(\max(f_m,0) \,\middle|\, \mathcal{A}_m\right).$$

From this it follows that the sequence $(\max(f_m, 0) : 1 \leqslant m \leqslant \infty)$ is dominated by an integrable function (in fact by $\mathbb{E}(\max(f_1, 0))$). So it follows that this sequence is uniformly integrable. The fact that the sequence $(\max(-f_n, 0) : n \in \mathbb{N})$ is also uniformly integrable, is much less trivial. To this end we consider

$$\begin{aligned}
&-\lambda\mathbb{P}(f_n < -\lambda) \geqslant \mathbb{E}(f_n : f_n < -\lambda) = \mathbb{E}(f_n) - \mathbb{E}(f_n : f_n \geqslant -\lambda)\\
&\geqslant \mathbb{E}\left(\mathbb{E}\left(f_n \,\middle|\, \mathcal{A}_\infty\right)\right) - \mathbb{E}\left(\mathbb{E}\left(f_1 \,\middle|\, \mathcal{A}_n\right) : f_n \geqslant -\lambda\right)\\
&\geqslant \mathbb{E}(f_\infty) - \mathbb{E}(f_1 : f_n \geqslant -\lambda)\\
&\geqslant \mathbb{E}(f_\infty) - \mathbb{E}(\max(f_1, 0)).
\end{aligned} \tag{5.172}$$

From (5.172) it follows that

$$\lambda\mathbb{P}(f_n < -\lambda) \leqslant \mathbb{E}\left(\max(f_1, 0) + \max(-f_\infty, 0)\right) < \infty.$$

Then choose $\epsilon > 0$ and m_0 in such a way that $\mathbb{E}(f_{m_0}) \leqslant \mathbb{E}(f_\infty) + \epsilon$. For $n \geqslant m_0$ we then see $\mathbb{E}(f_{m_0}) \leqslant \mathbb{E}(f_n) + \epsilon$. Hence, $\mathbb{E}(f_n) \geqslant \mathbb{E}(f_{m_0}) - \epsilon$. Then choose $\delta > 0$ such that $\mathbb{P}(A) \leqslant \delta$ implies $\mathbb{E}(|f_k| : A) \leqslant \epsilon$ for $k = 1, \dots, m_0$. After that choose λ_0 so large that $\mathbb{P}(f_k < \lambda) \leqslant \delta$ for all $k \in \mathbb{N}$ and for all $\lambda \geqslant \lambda_0$. For $1 \leqslant k \leqslant m_0$ we then get $\mathbb{E}(|f_k| : f_k \leqslant -\lambda) \leqslant \epsilon$, $\lambda \geqslant \lambda_0$. For $k \geqslant m_0$ we see

$$\mathbb{E}(f_k : f_k < -\lambda) = \mathbb{E}(f_k) - \mathbb{E}(f_k : f_k \geqslant \lambda) \tag{5.173}$$

$$\geqslant \mathbb{E}\left(f_{m_0}\right) - \mathbb{E}\left(\mathbb{E}\left(f_{m_0} \,\middle|\, \mathcal{A}_k\right) : f_k \geqslant -\lambda\right) - \epsilon \geqslant \mathbb{E}\left(f_{m_0} :\, f_k < -\lambda\right) - \epsilon.$$

By (5.173) we obtain

$$\begin{aligned} \mathbb{E}\left(|f_k| : f_k < -\lambda\right) &= -\mathbb{E}\left(f_k : f_k < -\lambda\right) \\ &\leqslant \left|\mathbb{E}\left(f_{m_0} : f_k < -\lambda\right)\right| + \epsilon \leqslant 2\epsilon \end{aligned}$$

for a certain $\lambda > 0$. Thus we see that the sequence $(\max(-f_n, 0) : n \in \mathbb{N})$ is also uniformly integrable. This yields the desired result in Lemma 5.111. ☐

5.112. THEOREM. *Let, relative to the right-continuous filtration $(\mathcal{F}_t : t \geqslant 0)$, the process X be a sub-martingale. Suppose that $\mathcal{F}_0$ contains the zero-sets, and that the function $t \mapsto \mathbb{E}(X(t))$ is right-continuous. Then there exists a process $Y = (Y(t) : t \geqslant 0)$ which is cadlag is which cannot be distinguished from X. So for every $t \geqslant 0$ the equality $Y(t) = X(t)$ holds $\mathbb{P}$-almost surely.*

PROOF. There exists an event Ω' in Ω, with $\mathbb{P}(\Omega') = 1$, such that on Ω' the following claims hold:

$$\begin{aligned} &\sup_{t \in D \cap [0,n]} |X(t)| < \infty, \text{ for all } n \in \mathbb{N}; \\ &U_{D\cap[0,n]}(X, a, b) < \infty, \text{ or all } n \in \mathbb{N} \text{ and for all } a < b,\ a \text{ and } b \text{ rational.} \end{aligned}$$

Since $\mathbb{P}(\Omega') = 1$ we see that Ω' belongs to $\mathcal{F}_0$ and, hence Ω' belongs to $\mathcal{F}_t$ for all $t \geqslant 0$. On Ω' we define the process $Y = (Y(t); t \geqslant 0)$ as follows: $Y(t) = \lim_{s\downarrow t, s>t, s\in D} X(s)$. Then $Y(t)$ is measurable for all σ-fields $\mathcal{F}_u$ with $u > t$. By the right continuity of the filtration $(\mathcal{F}_t : t \geqslant 0)$ we then see that $Y(t)$ is measurable for the σ-field $\mathcal{F}_t$. Then take $t = \lim_{n\to\infty} s_n$, where $s_n \downarrow t$, and where, for every $n \in \mathbb{N}$, s_n belongs to D. Then $Y(t) = \lim_{n\to\infty} X(s_n)$ in probability. Then apply Lemma 5.111 to conclude that the sequence $(X(s_n) : n \in \mathbb{N})$ is uniformly integrable, and hence $Y(t) = L^1 - \lim_{n\to\infty} X(s_n)$. We may apply Lemma 5.111. for $f_n := X(s_n)$, $f_\infty = Y(t)$, $\mathcal{A}_\infty = \mathcal{F}_t$ and $\mathcal{A}_n = \mathcal{F}_{s_n}$. Then notice that $X(t) \leqslant \mathbb{E}\left(X(s_n) \,\middle|\, \mathcal{F}_t\right)$, $\mathbb{P}$-almost surely. By L^1-convergence, from the latter we see that $X(t) \leqslant \mathbb{E}\left(Y(t) \,\middle|\, \mathcal{F}_t\right)$ and thus $X(t) \leqslant Y(t)$ $\mathbb{P}$-almost surely. Since, in addition, $\mathbb{E}(Y(t)) = \lim_{n\to\infty} \mathbb{E}(X(s_n)) = \mathbb{E}(X(t))$, the equality $Y(t) = X(t)$ follows $\mathbb{P}$-almost surely.

This completes the proof of Theorem 5.112. ☐

5.113. THEOREM. *Let $X = (X(t) : t \geqslant 0)$ be a sub-martingale with property that $\sup_{t\geqslant 0} \mathbb{E}\left[X(t)^+\right] < \infty$. The following assertions hold true.*

(a) *The limit $X(\infty) := \lim_{s\to\infty, s\in D} X(s)$ exists $\mathbb{P}$-almost surely.*
(b) *If X is a cadlag process, then the limit $X(\infty) := \lim_{s\to\infty}$ exists $\mathbb{P}$-almost surely.*
(c) *If, in addition, the process $(X(t)^+ : t \geqslant 0)$ is uniformly integrable, then the inequality $X(t) \leqslant \mathbb{E}\left(X(\infty) \,\middle|\, \mathcal{F}_t\right)$ holds.*

Proof of Theorem 5.113. (a) From the maximal inequality of Doob it follows that, for $\lambda > 0$, the following inequality holds:

$$\lambda \mathbb{P}\left(\sup_{t\in D\cap[0,n]} |X(t)| > \lambda\right) \leqslant 10\mathbb{E}\left(X(n)^+ + X(0)^-\right). \tag{5.174}$$

By letting n tend to ∞ in (5.174) we obtain

$$\lambda \mathbb{P}\left(\sup_{t\in D} |X(t)| > \lambda\right) \leqslant 10 \sup_n \mathbb{E}\left(X(n)^+ + X(0)^-\right),$$

and hence, $\sup_{t\in D} |X(t)| < \infty$ $\mathbb{P}$-almost surely. In the same manner we see

$$\mathbb{E}\left(U_{D\cap[0,\infty)}(X, a, b)\right) \leqslant \sup_n \frac{\mathbb{E}\left(X(n) - a\right)^+}{b-a}. \tag{5.175}$$

From (5.175) we see that $U_{D\cap[0,\infty)}(X, a, b) < \infty$ $\mathbb{P}$-almost surely. As we proved regularity starting from (5.175) and (5.174) (in fact from their consequences), we now obtain that $X(\infty) := \lim_{s\to\infty, s\in D} X(s)$ exists.

(b) If X is cadlag, then, like in the proof of the regularity, the limit $X(\infty) = \lim_{s\to\infty} X(s)$ exists.

(c) Since the process $(X(t)^+ : t \geqslant 0)$ is uniformly integrable, it also follows that the process $t \mapsto \max(X(t), a) = (X(t) - a)^+ + a$ is uniformly integrable as well. So, for $A \in \mathcal{F}_t$ and for $u > t$, the folowing (in-)equalities hold true:

$$\begin{aligned}
&\int_A \max(X(t), a)\, d\mathbb{P} \leqslant \lim_{u\to\infty} \int_A \max(X(u), a)\, d\mathbb{P} \\
&= \int_A \lim_{u\to\infty} \max(X(u), a)\, d\mathbb{P} = \int_A \max(X(\infty), a)\, d\mathbb{P}.
\end{aligned}$$

Since

$$\int X(\infty)^{+}\,d\mathbb{P} = \lim_{u\to\infty}\int X(u)^{+}\,d\mathbb{P} < \infty$$

we see that $X(\infty)^{+}$ belongs to $L^1(\Omega, \mathcal{F}, \mathbb{P})$. But then we get

$$\begin{aligned}\int_A X(t)\,d\mathbb{P} &= \lim_{a\to-\infty}\int_A \max(X(t),a)\,d\mathbb{P} \leqslant \lim_{a\to-\infty}\int_A \max(X(\infty),a)\,d\mathbb{P}\\ &= \int_A X(\infty)\,d\mathbb{P}. \qquad (5.176)\end{aligned}$$

From (5.176) the inequality $X(t) \leqslant \mathbb{E}\left(X(\infty) \,\middle|\, \mathcal{F}_t\right)$ follows. This proves item (c). The proof of Theorem 5.113 is now complete. ☐

5.114. THEOREM. *Let* $X = (X(t) : t \geqslant 0)$ *be a sub-martingale with the property that the process* $(X(t)^{+} : t \geqslant 0)$ *is uniformly integrable. In addition, suppose that* X *is cadlag. If* S *and* T *is a pair of stopping times such that* $0 \leqslant S \leqslant T \leqslant \infty$, *then the following inequality holds:* $X(S) \leqslant \mathbb{E}\left(X(T) \,\middle|\, \mathcal{F}_S\right)$.

PROOF. Put $S_n = 2^{-n}\lceil 2^n T\rceil$ and, similarly, $T_n = 2^{-n}\lceil 2^n T\rceil$. Then the stopping times S_n and T_n attain exclusively discrete values (in fact they take their values in $2^{-n}\mathbb{N}$). It is true that $S_n \downarrow S$ (if $n \to \infty$) and the same is true for the sequence $(T_n : n \in \mathbb{N})$. Moreover, $S_n \leqslant T_m$ for $n \geqslant m$. From Doob's theorem about discrete optional stopping times it follows that

$$\begin{aligned}&X(S_n) \leqslant \mathbb{E}\left(X(T_m) \,\middle|\, \mathcal{F}_{S_n}\right), \quad X(S_n) \leqslant \mathbb{E}\left(X(\infty) \,\middle|\, \mathcal{F}_{S_n}\right),\\ &X(S_n) \leqslant \mathbb{E}\left(X(\infty) \,\middle|\, \mathcal{F}_{S_n}\right).\end{aligned}$$

From this it follows that the processes $(X(S_n)^{+} : n \in \mathbb{N})$ and $(X(T_n)^{+} : n \in \mathbb{N})$ are uniformly integrable. For all n, m in $\mathbb{N}$, $n \geqslant m$, the following inequality holds for $A \in \mathcal{F}_S$:

$$\int_A \max(X(S_n),a)\,d\mathbb{P} \leqslant \int_A \max(X(T_m),a)\,d\mathbb{P}; \qquad (5.177)$$

(let n tend to ∞ in (5.177) to obtain)

$$\int_A \max(X(S),a)\,d\mathbb{P} \leqslant \int_A \max(X(T_m),a)\,d\mathbb{P}; \qquad (5.178)$$

(in (5.178) let m tend to ∞ to obtain)

$$\int_A \max(X(S),a)\,d\mathbb{P} \leqslant \int_A \max(X(T),a)\,d\mathbb{P}; \qquad (5.179)$$

(in (5.179) let a tend to $-\infty$ to obtain)

$$\int_A X(S)\,d\mathbb{P} \leqslant \int_A X(T)\,d\mathbb{P}, \qquad (5.180)$$

and that $\lim_{n\to\infty} X(S_n) = X(S)$ and that the same is true for the stopping time T. By (5.180) we then see $X(S) \leqslant \mathbb{E}\left(X(T) \,\middle|\, \mathcal{F}_S\right)$. This completes the proof of Theorem 5.114. ☐

5.115. Corollary. *Let $X = (X(s) : 0 \leqslant s \leqslant t)$ be a cadlag martingale and let $0 \leqslant S \leqslant T \leqslant t$ be two stopping times. The following equalities are true:*

$$X(S) = \mathbb{E}\left(X(T) \mid \mathcal{F}_S\right) \text{ and } \mathbb{E}\left(X(T)\right) = \mathbb{E}\left(X(\infty)\right) = \mathbb{E}(X(0)).$$

Proof. The proof is left as an exercise for the reader. Among other things we notice that the martingale $(X(s) : 0 \leqslant s \leqslant t)$ is uniformly integrable. ☐

5.116. Corollary. *Let X be a cadlag martingale in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ which is uniformly integrable. Then the limit $X(\infty) := \lim_{t\to\infty} X(t)$ exists $\mathbb{P}$-almost surely, and if S and T are stopping times such that $0 \leqslant S \leqslant T \leqslant \infty$, then the following equalities hold:*

$$X(S) = \mathbb{E}\left(X(T) \mid \mathcal{F}_S\right) \text{ and } \mathbb{E}\left(X(T)\right) = \mathbb{E}\left(X(\infty)\right) = \mathbb{E}(X(0)).$$

Proof. The proof of this corollary is left as an exercise for the reader. Observe that for $n \in \mathbb{N}$ fixed the martingale $(X(\min(n,t)) : t \geqslant 0)$ is uniformly integrable. ☐

In what follows the process $X : [0,\infty) \times \Omega \to \mathbb{R}^\nu$ is a process with values in $\mathbb{R}^\nu$, where ν may be 1.

5.117. Definition. Let X be a stochastic process, which is adapted to the filtration $(\mathcal{F}_t : t \geqslant 0)$. The process X is said to be a *Lévy process* if X possesses the following properties:

(a) For all $s < t$ the variable $X(t) - X(s)$ is independent of $\mathcal{F}_s$;
(b) For all $s \leqslant t$ the variable $X(t) - X(s)$ has the same distribution as $X(t-s)$;
(c) For all $t \geqslant 0$ and for every sequence $(t_n : n \in \mathbb{N})$ in $[0,\infty)$ that converges to t, the limit $\lim_{n\to\infty} X(t_n) = X$ exists in $\mathbb{P}$-law (or in $\mathbb{P}$-measure). Sometimes this is denoted by $\mathbb{P}\text{-}\lim_{n\to\infty} X(t_n) = X(t)$.

5.118. Theorem. *Let X be a stochastic process, which is adapted to the filtration $(\mathcal{F}_t : t \geqslant 0)$, and which takes it values in $\mathbb{R}^\nu$. The following assertions are true:*

(a) *Let X be a Lévy-process. Define for $t \geqslant 0$ the probability measure μ_t as being the distribution of $X(t)$. So $\mu_t(B) = \mathbb{P}(X(t) \in B)$, where B is a Borel subset of $\mathbb{R}^\nu$. Then the family $\{\mu_t : t \geqslant 0\}$ is a vaguely continuous semigroup of probability measures.*
(b) *Conversely, let $\{\mu_t : t \geqslant 0\}$ be a vaguely continuous semigroup of probability measures on $\mathbb{R}^\nu$. Then there exists a Lévy-process*

$$X = \{X(t) : t \geqslant 0\}$$

with cadlag paths such that $\mu_t(B) = \mathbb{P}(X(t) \in B)$ for all Borel subsets B of $\mathbb{R}^\nu$.

Proof of Theorem 5.118. (a) Define for $t \geqslant 0$ the *characteristic function* f_t of $X(t)$ as being the Fourier transform of the $\mathbb{P}$-distribution of $X(t)$. So that

$$f_t(\xi) = \mathbb{E}\left(\exp\left(-i\left\langle \xi, X(t)\right\rangle\right)\right), \quad \xi \in \mathbb{R}^\nu.$$

Since, for $t, s \in [0, \infty)$, $X(s+t) = X(s+t) - X(s) + X(s)$, since, in addition, $X(s+t) - X(s)$ is independent of $X(t)$, and because $X(s+t) - X(s)$ possesses the same distribution as $X(t)$ we infer

$$\begin{aligned} f_{s+t}(\xi) &= \mathbb{E}\left(\exp\left(-i\left\langle \xi, X(s+t)\right\rangle\right)\right) \\ &= \mathbb{E}\left(\exp\left(-i\left\langle \xi, X(s+t) - X(s)\right\rangle\right)\exp\left(-i\left\langle \xi, X(s)\right\rangle\right)\right) \\ &= \mathbb{E}\left(\exp\left(-i\left\langle \xi, X(s+t) - X(s)\right\rangle\right)\right)\mathbb{E}\left(\exp\left(-i\left\langle \xi, X(s)\right\rangle\right)\right) \\ &= \mathbb{E}\left(\exp\left(-i\left\langle \xi, X(t)\right\rangle\right)\right)\mathbb{E}\left(\exp\left(-i\left\langle \xi, X(s)\right\rangle\right)\right) \\ &= f_t(\xi) f_s(\xi). \end{aligned} \tag{5.181}$$

Since $X(0)$ and $X(0) - X(0) = 0$ have the same distribution we see $f_0(\xi) = 1$. Since $\mathbb{P}\text{-}\lim_{u\downarrow 0} X(u) = X(0)$ we see, for example by Theorem 5.85 in combination with the implication (1) $\Rightarrow$ (9) of Theorem 5.43, that $\lim_{s\downarrow 0} f_s(\xi) = f_0(\xi) = 1$. From (5.181) it then follows that

$$\lim_{t\downarrow s} f_t(\xi) - f_s(\xi) = \lim_{t\downarrow s} f_s(\xi)\left(f_{t-s}(\xi) - f_0(\xi)\right) = 0$$

for all $s \geqslant 0$. Because, by applying equality (5.181) repeatedly, we see $f_t(\xi) = \left(f_{t2^{-n}}(\xi)\right)^{2^n}$. In addition we have $\lim_{s\downarrow 0} f_s(\xi) = 1$. So it follows that for no value of $t \in [0, \infty)$ the function $f_t(\xi)$ vanishes for any ξ. Since, for $t < s$, $f_t(\xi) - f_s(\xi) = \left(f_0(\xi) - f_{s-t}(\xi)\right) f_t(\xi)$, it also follows that $\lim_{t\uparrow s} f_t(\xi) = f_s(\xi)$, for $s > 0$. From the previous considerations it follows that the function $t \mapsto f_t(\xi)$, $t \in [0, \infty)$, is a continuous function, which satisfies the relation $f_{s+t}(\xi) = f_s(\xi) f_t(\xi)$ for all s, $t \geqslant 0$ and this for all $\xi \in \mathbb{R}^\nu$. Furthermore, we define the family of measures $\{\mu_t : t \geqslant 0\}$ as being the $\mathbb{P}$-distributions of the Lévy process X. So that $\mu_t(B) = \mathbb{P}\left(X(t) \in B\right)$, B Borel subset of $\mathbb{R}^\nu$.

From the previous arguments it then follows that

$$\widehat{\mu}_{s+t}(\xi) = f_{s+t}(\xi) = f_s(\xi) f_t(\xi) = \widehat{\mu}_s(\xi)\widehat{\mu}_t(\xi) \tag{5.182}$$

and that $\lim_{s\downarrow 0} \widehat{\mu}_s(\xi) = 1$. So that the family $\{\mu_t : t \geqslant 0\}$ is a vaguely continuous semigroup of probability measures on $\mathbb{R}^\nu$. By Theorem 5.31 there then exists a continuous negative-definite function ψ such that $f_t(\xi) = \widehat{\mu}_t(\xi) = \exp(-t\psi(\xi))$.

(b) Define $(\Omega, \mathcal{F}, \mathbb{P})$ as in Proposition 5.36. Likewise we define the state variables $X(t) : \Omega \to \mathbb{R}^\nu$ as in Proposition 5.36. Let the filtration $(\mathcal{F}_t : t \geqslant 0)$ be determined by $\mathcal{F}_t = \sigma(X(u) : 0 \leqslant u \leqslant t)$. So the filtration $(\mathcal{F}_t : t \geqslant 0)$ is the internal history of the process X. Then X is a Lévy-process, which possesses the properties as described in (b). The fact that for $t > s$ the variable $X(t) - X(s)$ is independent of $\mathcal{F}_s$ was proved in Theorem 5.37. We must show that, for $\epsilon > 0$ fixed, $\lim_{s\downarrow 0} \mathbb{P}(|X(s) - X(0)| > \epsilon) = 0$. Therefore, notice first that $\mathbb{P}(X(0) = 0) = \mu_0\{0\} = 1$. Hence, with $B(\epsilon) = \{x \in \mathbb{R}^\nu : |x| \leqslant \epsilon\}$, we have

$$\begin{aligned}\mathbb{P}(|X(s) - X(0)| > \epsilon) &= \mathbb{P}(|X(s) - X(0)| > \epsilon, X(0) = 0) \\ &= \mathbb{P}(|X(s)| > \epsilon, X(0) = 0) \\ &= \mathbb{P}(|X(s)| > \epsilon) \\ &= \mu_s\{\mathbb{R}^\nu \backslash B(\epsilon)\} = 1 - \mu_s\{B(\epsilon)\}. \end{aligned} \tag{5.183}$$

Since the convolution semigroup $\{\mu_t : t \geqslant 0\}$ is vaguely continuous it follows that $\lim\limits_{s\downarrow 0} \mu_s\{B(\epsilon)\} = 1$. From (5.183) we then see that $\lim_{s\downarrow 0} \mathbb{P}(|X(s) - X(0)| > \epsilon) = 0$. The only problem which is still left, is the fact that the process X is not necessarily cadlag. In the following propositions and lemmas we will, among other things, resolve this problem. From Theorem 5.121 it follows that the process X is also a Lévy process for the filtration $(\mathcal{G}_t : t \geqslant 0)$, where $\mathcal{G}_t = \mathcal{F}_t \cup \mathcal{N}$. By Theorem 5.123 we then see that the process X possesses a cadlag version.

The proof of Theorem 5.118 is now complete. □

5.119. Proposition. *Suppose* $0 \leqslant s_1 < \cdots < s_m$ *and choose* $t \geqslant 0$. *Let* X *be a Lévy process for the filtration* $(\mathcal{F}_t : t \geqslant 0)$, *where* $\mathcal{F}_t = \sigma\{X(u) : 0 \leqslant u \leqslant t\}$. *Let* $\{\mu_t : t \geqslant 0\}$ *be the corresponding convolution semigroup and* ψ *the corresponding negative-definite function. So* $\mu_t(B) = \mathbb{P}(X(t) \in B)$ *for all Borel subsets* B *and* $\widehat{\mu}_t(\xi) = \exp(-t\psi(\xi))$ *for all* $t \geqslant 0$. *For* $\xi^1, \dots, \xi^m$ *in* $\mathbb{R}^\nu$ *the following equalities hold:*

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-i\sum_{j=1}^m \left\langle \xi^j, X(t+s_j)\right\rangle\right) \mid \mathcal{F}_t\right] \\ &\quad = \mathbb{E}\left[\exp\left(-i\sum_{j=1}^m \left\langle \xi^j, X(t+s_j)\right\rangle\right) \mid \mathcal{F}_{t+}\right] \\ &\quad = \exp\left(-i\left\langle \sum_{j=1}^m \xi^j, X(t)\right\rangle\right) \exp\left(-\sum_{j=1}^m (s_j - s_{j-1})\,\psi\left(\sum_{k=j}^m \xi^k\right)\right). \end{aligned} \tag{5.184}$$

Here we write $\mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s$ and $s_0 = 0$.

Proof. We apply induction with respect to m. We begin with the conditioning on $\mathcal{F}_t$. For $m = 1$ we have

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-i\left\langle \xi^1, X\left(t+s_1\right)\right\rangle\right) \mid \mathcal{F}_t\right]\\&= \mathbb{E}\left[\exp\left(-i\left\langle \xi^1, X\left(t+s_1\right)-X(t)\right\rangle\right) \mid \mathcal{F}_t\right] \exp\left(-i\left\langle \xi^1, X(t)\right\rangle\right)\end{aligned}$$

($X(t+s_1) - X(t)$ does not depend on $\mathcal{F}_t$)

$$= \mathbb{E}\left[\exp\left(-i\left\langle \xi^1, X\left(t+s_1\right)-X(t)\right\rangle\right)\right] \exp\left(-i\left\langle \xi^1, X(t)\right\rangle\right)$$

($X(t+s_1) - X(t)$ has the same distribution as $X(s_1)$)

$$\begin{aligned}&= \mathbb{E}\left[\exp\left(-i\left\langle \xi^1, X\left(s_1\right)\right\rangle\right)\right] \exp\left(-i\left\langle \xi^1, X(t)\right\rangle\right)\\&= \widehat{\mu}_{s_1}\left(\xi^1\right) \exp\left(-i\left\langle \xi^1, X(t)\right\rangle\right)\\&= \exp\left(-s_1\psi\left(\xi^1\right)\right) \exp\left(-i\left\langle \xi^1, X(t)\right\rangle\right)\\&= \exp\left(-\left(s_1-s_0\right)\psi\left(\xi^1\right)\right) \exp\left(-i\left\langle \xi^1, X(t)\right\rangle\right). \qquad (5.185)\end{aligned}$$

Notice that (5.185) is the same as the equality in (5.184) for $m = 1$. Suppose now that we already know (5.184) for every $t \geqslant 0$, for every m-tuple $s_1 < \cdots < s_m$ and for every m-tuple $\xi^1, \dots, \xi^m$ in $\mathbb{R}^\nu$. We keep working with the original filtration $(\mathcal{F}_t : t \geqslant 0)$. For $s_{m+1} > s_m$ and for $\xi^{m+1} \in \mathbb{R}^\nu$ we then see

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-i\sum_{j=1}^{m+1}\left\langle \xi^j, X\left(t+s_j\right)\right\rangle\right) \mid \mathcal{F}_t\right]\\&= \mathbb{E}\left[\exp\left(-i\sum_{j=1}^{m}\left\langle \xi^j, X\left(t+s_j\right)\right\rangle\right)\right.\\&\qquad\left.\mathbb{E}\left[\exp\left(-i\left\langle \xi^{m+1}, X\left(t+s_{m+1}\right)\right\rangle\right) \mid \mathcal{F}_{t+s_m}\right] \mid \mathcal{F}_t\right]\end{aligned}$$

(employ (5.185) for $t + s_{m+1}$ instead of t)

$$\begin{aligned}&= \mathbb{E}\left[\exp\left(-i\sum_{j=1}^{m}\left\langle \xi^j, X\left(t+s_j\right)\right\rangle\right)\right.\\&\qquad\left.\exp\left(-i\left\langle \xi^{m+1}, X\left(t+s_{m+1}\right)\right\rangle\right) \exp\left(-\left(s_{m+1}-s_m\right)\psi\left(\xi^{m+1}\right)\right) \mid \mathcal{F}_t\right]\end{aligned}$$

(induction hypothesis)

$$= \exp\left(-i\left\langle \sum_{j=1}^{m+1}\xi^j, X(t)\right\rangle\right) \exp\left(-\sum_{j=1}^{m+1}\left(s_j - s_{j-1}\right)\psi\left(\sum_{k=j}^{m+1}\xi^k\right)\right). \qquad (5.186)$$

But (5.186) is the same as (5.184) with m replaced by $m + 1$. Next we look at the situation for the filtration $\{\mathcal{F}_{t+} : t \geqslant 0\}$ which is closed from the right.

Without loss of generality we may assume that $s_1 > 0$. In case $s_1 = 0$ we have indeed

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-i\sum_{j=1}^{m}\left\langle \xi^j, X\left(t+s_j\right)\right\rangle\right) \mid \mathcal{F}_{t+}\right]\\ &= \exp\left(-i\left\langle \xi^1, X(t)\right\rangle\right)\mathbb{E}\left[\exp\left(-i\sum_{j=2}^{m}\left\langle \xi^j, X\left(t+s_j\right)\right\rangle\right) \mid \mathcal{F}_{t+}\right].\end{aligned}$$

So assume that $s_1 > 0$ and choose $n \in \mathbb{N}$ such that $s_1 > n^{-1}$. Then we see, by (5.184) for $t + n^{-1}$ instead of t,

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-i\sum_{j=1}^{m}\left\langle \xi^j, X\left(t+s_j\right)\right\rangle\right) \mid \mathcal{F}_{t+}\right] \qquad (5.187)\\ &= \mathbb{E}\left[\mathbb{E}\left[\exp\left(-i\sum_{j=1}^{m}\left\langle \xi^j, X\left(t+s_j\right)\right\rangle\right) \mid \mathcal{F}_{t+n^{-1}}\right] \mid \mathcal{F}_{t+}\right]\end{aligned}$$

(write $s_0 = n^{-1}$ in what follows)

$$\begin{aligned}&= \mathbb{E}\left[\exp\left(-i\left\langle \sum_{j=1}^{m}\xi^j, X\left(t+n^{-1}\right)\right\rangle\right)\right.\\ &\qquad\left.\exp\left(-\sum_{j=1}^{m}\left(s_j - s_{j-1}\right)\psi\left(\sum_{k=j}^{m}\xi^k\right)\right) \mid \mathcal{F}_{t+}\right].\end{aligned}$$

In (5.187) we let n tend to ∞. Apparently it follows that

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-i\sum_{j=1}^{m}\left\langle \xi^j, X\left(t+s_j\right)\right\rangle\right) \mid \mathcal{F}_{t+}\right]\\ &= \mathbb{E}\left[\exp\left(-i\left\langle \sum_{j=1}^{m}\xi^j, X(t)\right\rangle\right)\exp\left(-\sum_{j=1}^{m}\left(s_j - s_{j-1}\right)\psi\left(\sum_{k=j}^{m}\xi^k\right)\right) \mid \mathcal{F}_{t+}\right]\\ &= \exp\left(-i\left\langle \sum_{j=1}^{m}\xi^j, X(t)\right\rangle\right)\exp\left(-\sum_{j=1}^{m}\left(s_j - s_{j-1}\right)\psi\left(\sum_{k=j}^{m}\xi^k\right)\right). \qquad (5.188)\end{aligned}$$

From (5.188) it then follows that (5.184) holds for the filtration $\{\mathcal{F}_{t+} : t \geqslant 0\}$ which is closed from the right.

This completes the proof of Proposition 5.119. ☐

5.120. COROLLARY. *Let the assumptions and hypotheses be as in Proposition 5.119. For every bounded complex-valued random variable Y, that is measurable for the σ-field $\sigma\{X(u) : u \geqslant 0\}$ the following equality holds $\mathbb{P}$-almost surely de equality:*

$$\mathbb{E}\left[Y \mid \mathcal{F}_t\right] = \mathbb{E}\left[Y \mid \mathcal{F}_{t+}\right]. \qquad (5.189)$$

Proof. Put $Y = \exp\left(-i\sum_{j=1}^m \langle \xi^j, X(s_j)\rangle\right)$. By Proposition 5.119 we see that for all such random variables Y the equality in (5.189) holds, provided that $s_j \geqslant t$, for $1 \leqslant j \leqslant m$. By splitting and using the standard properties of a conditional expectation we see that the restriction $s_j \geqslant t$ is superfluous. In other words the equality in (5.189) holds for all variables Y of the form $Y = \exp\left(-i\sum_{j=1}^m \langle \xi^j, X(s_j)\rangle\right)$ where all s_j belong to $[0,\infty)$ and where all ξ^j are members of $\mathbb{R}^\nu$. Let Y_0 be a bounded complex-valued random variable, which belongs to the linear span of variables of the form $\exp\left(-i\sum_{j=1}^m \langle \xi^j, X(s_j)\rangle\right)$. Then consider the vector space $\mathcal{H}(Y_0)$ defined by

$$\begin{aligned}&\mathcal{H}(Y_0)\\ &\ = \{Y : \Omega \to \mathbb{C} : Y \text{ is bounded and measurable for the } \sigma\text{-field } \sigma\{X(u) : u \geqslant 0\}\\ &\qquad \text{and } \mathbb{E}\left(YY_0 \mid \mathcal{F}_t\right) = \mathbb{E}\left(YY_0 \mid \mathcal{F}_{t+}\right)\}.\end{aligned}$$

By employing Lemma 5.100 or, even better, the monotone class theorem we see that $\mathcal{H}(Y_0)$ contains all complex-valued bounded random variables, which are measurable for the σ-field $\sigma\{X(u) : u \geqslant 0\}$. Among others we may put $Y_0 = 1$, and the claim in Corollary 5.120 follows. ☐

5.121. Theorem. *Let* $X = \{X(t) : t \geqslant 0\}$ *be a Lévy-process. Let* $\mathcal{H} = \{\mathcal{H}_t : t \geqslant 0\}$ *be the internal history of the process* X*. Let* $\mathcal{N}$ *be the null sets in* $\mathcal{H}_\infty$*. Then the filtration* $\mathcal{G}$*, with* $G_t = \sigma\{\mathcal{H}_t \cup \mathcal{N}\}$*, is continuous from the right.*

Proof of Theorem 5.121. Let $A \in \mathcal{G}_{t+}$. By Corollary 5.120 we have $1_A = \mathbb{E}\left(1_A \mid \mathcal{F}_t\right)$. Let $B \in \mathcal{F}_t$ be such that $1_B = \mathbb{E}\left(1_A \mid \mathcal{F}_t\right)$, $\mathbb{P}$-almost surely. Then $\mathbb{P}(A \triangle B) = 0$. Since $A = B \triangle (A \triangle B)$, we see that A in fact belongs to $\mathcal{G}_t$. $\square$

5.122. Lemma. *Let $(x_n : n \in \mathbb{N})$ be a sequence of vectors in $\mathbb{R}^\nu$ with the property that the sequence $(\exp(-i\langle \xi, x_n\rangle) : n \in \mathbb{N})$ converges for almost all $\xi \in \mathbb{R}^\nu$. Then the sequence $(x_n : n \in \mathbb{N})$ converges.*

Proof. Fix $1 \leqslant j \leqslant \nu$, and let U^j be a vector valued stochastic variable which is zero for the coordinates $k \neq j$ and with the property that U^j_j is uniformly distributed on the interval $[0,1]$. The following inequalities are true for $0 < \delta < 1$:

$$
\begin{aligned}
&2\mathrm{Re}\,\mathbb{E}\left(1 - \exp\left(-i\left\langle U^j, x_n - x_m\right\rangle\right)\right) = \mathbb{E}\left|1 - \exp\left(-i\left\langle U^j, x_n - x_m\right\rangle\right)\right|^2 \\
&\quad = 4\mathbb{E}\left(\sin^2 \frac{1}{2}\left\langle U^j, x_n - x_m\right\rangle\right) \geqslant 4\delta^2 \mathbb{P}\left\{\sin^2 \frac{1}{2}\left\langle U^j, x_n - x_m\right\rangle \geqslant \delta^2\right\} \\
&\quad \geqslant 4\delta^2 \mathbb{P}\left\{\left(\frac{2}{\pi}\right)^2 \min\left(\frac{1}{4}\left\langle U^j, x_n - x_m\right\rangle^2, \frac{\pi^2}{4}\right) \geqslant \delta^2\right\} \\
&\quad \geqslant 4\delta^2 \mathbb{P}\left\{\left|\left\langle U^j, x_n - x_m\right\rangle\right| \geqslant \delta\pi\right\} \\
&\quad = 4\delta^2 \max\left(1 - \frac{\delta\pi}{\left|x_{n,j} - x_{m,j}\right|}, 0\right).
\end{aligned}
$$

Since this inequality holds for every $0 < \delta < 1$, it follows that the j-th coordinate $(x_{n,j} : n \in \mathbb{N})$ of the sequence $(x_n : n \in \mathbb{N})$ converges. This holds for $1 \leqslant j \leqslant \nu$. Hence, the limit $\lim_{n\to\infty} x_n$ exists. This makes the proof of Lemma 5.122 complete. $\square$

Among other things, in Theorem 5.123 the proof of item (b) in 5.118 is completed.

5.123. Theorem. *Let $(X(t), \mathcal{F}_t)_{t\geqslant 0}$ be a Lévy-process. Suppose that the filtration $(\mathcal{F}_t : t \geqslant 0)$ is right-continuous, and that $\mathcal{F}_0$ contains the null sets. Then there exists a cadlag modification of $X = (X(t) : t \geqslant 0)$.*

The result in Theorem 5.123 can be applied in case we take the internal history, completed with null sets, as the filtration $(\mathcal{F}_t)_{t\geqslant 0}$.

Proof of Theorem 5.123. We will make use of the following product set:

$$
E = \mathbb{C}^{\mathbb{Q}_+} = \left\{(\alpha_t)_{t\in\mathbb{Q}_+} : \alpha_t \in \mathbb{C} \text{ for all } t \in \mathbb{Q}_+\right\},
$$

endowed with the product-σ-field $\mathcal{E} = \otimes_{t\in\mathbb{Q}_+}\mathcal{B}(\mathbb{C})$. Put

$$
D = \left\{(\alpha_t)_{t\in\mathbb{Q}_+} : \exists\varphi : [0,\infty) \to \mathbb{C} \text{ cadlag with } \varphi(t) = \alpha_t \text{ for all } t \in \mathbb{Q}_+\right\}.
$$

Upon writing φ as the pointwise limit $\varphi(t) = \lim_{n\to\infty} \varphi_n(t)$, where

$$\varphi_n(t) = \sum_{k=0}^{\infty} \varphi\left((k+1)2^{-n}\right) 1_{[k2^{-n},(k+1)2^{-n})}(t),$$

it can be proved that D belongs to the σ-field $\mathcal{E}$. Consider the mapping:

$$\Phi : \mathbb{R}^\nu \times \Omega \to E$$

defined by

$$\Phi(\xi,\omega) = \exp\left(-i\left\langle \xi, X(t)\right\rangle\right) =: \alpha_t. \tag{5.190}$$

The mapping Φ is measurable for the σ-fields $\mathcal{F}_\infty$ and $\mathcal{E}$. As a consequence $\Lambda := \Phi^{-1}(D)$ belongs to the σ-field $\mathcal{B}(\mathbb{R}^\nu)\otimes\mathcal{F}_\infty$. So for every pair $(\xi,\omega) \in \mathbb{R}^\nu \times \Omega$ there exists a cadlag function $f : [0,\infty) \to \mathbb{C}$ with the property that the equality $f(t) = \exp\left(-i\left\langle \xi, X(t)\right\rangle\right)$ holds for all $t \in \mathbb{Q}_+$. Now let the negative-definite function corresponding to the process X be given by ψ. Then the process $t \mapsto \exp\left(-i\left\langle \xi, X(t)\right\rangle + t\psi(\xi)\right)$ is a martingale. This is so, because, for $0 \leqslant s < t$, we have the following equalities:

$$\begin{aligned}&\mathbb{E}\left(\exp\left(-i\left\langle \xi, X(t)\right\rangle + t\psi(\xi)\right) \,\middle|\, \mathcal{F}_s\right)\\ &\quad= \mathbb{E}\left(\exp\left(-i\left\langle \xi, X(t)-X(s)\right\rangle + (t-s)\psi(\xi)\right) \,\middle|\, \mathcal{F}_s\right) \exp\left(-i\left\langle \xi, X(s)\right\rangle + s\psi(\xi)\right)\end{aligned}$$

($X(t) - X(s)$ does not depend on $\mathcal{F}_s$)

$$= \mathbb{E}\left(\exp\left(-i\left\langle \xi, X(t)-X(s)\right\rangle + (t-s)\psi(\xi)\right)\right) \exp\left(-i\left\langle \xi, X(s)\right\rangle + s\psi(\xi)\right)$$

($X(t) - X(s)$ heeft dezelfde distribution als $X(t-s)$)

$$= \mathbb{E}\left(\exp\left(-i\left\langle \xi, X(t-s)\right\rangle + (t-s)\psi(\xi)\right)\right) \exp\left(-i\left\langle \xi, X(s)\right\rangle + s\psi(\xi)\right)$$

(definition of ψ)

$$= \exp\left(-i\left\langle \xi, X(s)\right\rangle + s\psi(\xi)\right). \tag{5.191}$$

From martingale theory it follows that there exists a cadlag version $M^\xi = \left(M^\xi(t) : t \geqslant 0\right)$ of the martingale $t \mapsto \exp\left(-i\left\langle \xi, X(t)\right\rangle + t\psi(\xi)\right)$. By this we mean that for every $(\xi, t) \in \mathbb{R}^\nu \times [0,\infty)$ there exists an event $N_{t,\xi}$ with the following properties: $\mathbb{P}\left(N_{t,\xi}\right) = 0$ and for $\omega \notin N_{t,\xi}^c$ the equality $M^\xi(t)(\omega) = \exp\left(-i\left\langle \xi, X(t,\omega)\right\rangle + t\psi(\xi)\right)$ holds. Hence, for every $\xi \in \mathbb{R}^\nu$ there exists a $\mathbb{P}$-null set N_ξ such that for every $t \in [0,\infty) \cap \mathbb{Q}$ the equality

$$M^\xi(t)(\omega) = \exp\left(-i\left\langle \xi, X(t,\omega)\right\rangle + t\psi(\xi)\right)$$

holds for all $\omega \notin N_\xi$. In other words for all $\xi \in \mathbb{R}^\nu$ the equality:

$$\mathbb{P}\left\{\omega : (\xi,\omega) \in \Lambda\right\} \geqslant \mathbb{P}\left\{N_\xi^c\right\} = 1 \tag{5.192}$$

holds. From (5.192) it then follows that

$$0 = \int_{\mathbb{R}^\nu} d\xi \int_\Omega d\mathbb{P} 1_{\Lambda^c} = \int_\Omega d\mathbb{P} \int_{\mathbb{R}^\nu} 1_{\Lambda^c}\, d\xi. \tag{5.193}$$

The equality in (5.193) implies that for $\mathbb{P} \otimes \lambda$-almost all (ω, ξ) the function $t \mapsto \exp\left(-i \langle \xi, X(t,\omega)\rangle\right)$ belongs to D. By Lemma 5.122 we see that $\mathbb{P}$-almost surely the following limits exist for all $t \geqslant 0$:

$$\lim_{\substack{s\downarrow t\\ s\in\mathbb{Q}}} X(s) \text{ and } \lim_{\substack{s\uparrow t\\ s\in\mathbb{Q}}} X(s).$$

Define the process Y by $Y(t) = \lim_{s\downarrow t,\, s\in\mathbb{Q}} X(s)$. Then the process Y is cadlag: see (the proof of) Lemma 5.103 (b). Furthermore, $X(t) = \lim_{s\downarrow t, s\in\mathbb{Q}} X(s)$ (in $\mathbb{P}$-distributional sense), and thus $X(t) = Y(t)$ $\mathbb{P}$-almost surely. The proof of Theorem 5.123 is now complete. ☐

5.124. THEOREM (Dynkin-Hunt). *Let $(X(t), \mathcal{F}_t)_{t\geqslant 0}$ be a cadlag Lévy process with a right-continuous filtration $(\mathcal{F}_t : t \geqslant 0)$. Let $T : \Omega \to [0,\infty)$ be a stopping time which is not identically ∞. So that $\mathbb{P}\{T < \infty\} > 0$. On the event $\{T < \infty\}$ the process $Y = \{Y(t) : t \geqslant 0\}$ is defined by $Y(t) = X(t+T) - X(T)$.*

(a) *Under $\widetilde{\mathbb{P}}$ the process Y has the same distribution as the process X under $\mathbb{P}$.*
(b) *The σ-fields $\mathcal{F}_T$ and $\sigma\{Y(s) : s \geqslant 0\}$ are $\mathbb{P}$-independent.*

Proof of Theorem 5.124. (a) For $n \in \mathbb{N}$ we write $T_n = 2^{-n}[2^nT]$. Then $(T_n : n \in \mathbb{N})$ is a sequence of stopping times with the following properties:

(i) $\{T_n < \infty\} = \{T < \infty\}$;
(ii) $T \leqslant T_{n+1} \leqslant T_n \leqslant T + 2^{-n}$, $n \in \mathbb{N}$.

Define the sequence of processes $(Y^n : n \in \mathbb{N})$ via the formula:

$$Y^n(t) = X(t + T_n) - X(T_n) \text{ op de event } \{T_n < \infty\} = \{T < \infty\}.$$

Let now $f : (\mathbb{R}^\nu)^m \to \mathbb{C}$ be a bounded continuous function, let A be an event in $\mathcal{F}_T \subseteq \mathcal{F}_{T_n}$, and let $s_1 < \cdots < s_m$ be an increasing sequence of fixed times. Then the following equalities hold:

$$\begin{aligned}
&\mathbb{E}\left[f\left(Y^n(s_1), Y^n(s_2) - Y^n(s_1), \dots, Y^n(s_m) - Y^n(s_{m-1})\right) 1_{A\cap\{T<\infty\}}\right] \\
&= \sum_{k=0}^{\infty} \mathbb{E}\left[1_{A\cap\{T_n=k2^{-n}\}} f\left(X\left(s_1 + k2^{-n}\right) - X\left(k2^{-n}\right), \dots\right.\right. \\
&\qquad \left.\left., X\left(s_m + k2^{-m}\right) - X\left(s_{m-1} + k2^{-m}\right)\right)\right]
\end{aligned}$$

$(X(s_j + k2^{-n}) - X(s_{j-1} + k2^{-n})$ does not depend on $\mathcal{F}_{s_{j-1}+k2^{-n}})$

$$\begin{aligned}
&= \sum_{k=0}^{\infty} \mathbb{P}\left[A \cap \left\{T_n = k2^{-n}\right\}\right] \\
&\quad \mathbb{E}\left[f\left(X\left(s_1 + k2^{-n}\right) - X\left(k2^{-n}\right), \dots, X\left(s_m + k2^{-m}\right) - X\left(s_{m-1} + k2^{-m}\right)\right)\right]
\end{aligned}$$

$(X(s_j + k2^{-n}) - X(s_{j-1} + k2^{-n})$ has the same distribution as $X(s_j) - X(s_{j-1}))$

$$\begin{aligned}
&= \sum_{k=0}^{\infty} \mathbb{P}\left[A \cap \left\{T_n = k2^{-n}\right\}\right] \mathbb{E}\left[f\left(X(s_1) - X(0), \dots, X(s_m) - X(s_{m-1})\right)\right] \\
&= \mathbb{P}\left[A \cap \{T_n < \infty\}\right] \mathbb{E}\left[f\left(X(s_1) - X(0), \dots, X(s_m) - X(s_{m-1})\right)\right] \\
&= \mathbb{P}\left[A \cap \{T < \infty\}\right] \mathbb{E}\left[f\left(X(s_1) - X(0), \dots, X(s_m) - X(s_{m-1})\right)\right]. \qquad (5.194)
\end{aligned}$$

In (5.194) we let n tend to ∞. Since the process X is right-continuous it follows that $\lim_{n\to\infty} Y^n(t) = Y(t)$ $\mathbb{P}$-almost surely on the event $\{T < \infty\}$. By the continuity of the function f the equality

$$\begin{aligned}
&\mathbb{E}\left[f\left(Y(s_1), Y(s_2) - Y(s_1), \dots, Y(s_m) - Y(s_{m-1})\right) 1_{A\cap\{T<\infty\}}\right] \\
&= \mathbb{P}\left[A \cap \{T < \infty\}\right] \mathbb{E}\left[f\left(X(s_1) - X(0), \dots, X(s_m) - X(s_{m-1})\right)\right] \qquad (5.195)
\end{aligned}$$

follows. By taking, in (5.195), the function f of the form $f = f_0 \circ V_m$, where $V_m : (\mathbb{R}^\nu)^m \to (\mathbb{R}^\nu)^m$ is give by $V_m(x_1, \dots, x_m)) = (x_1, \dots, x_1 + \cdots + x_m)$ we see

$$\begin{aligned}
&\widetilde{\mathbb{E}}\left[f_0\left(Y(s_1), \dots, Y(s_m)\right) 1_{A\cap\{T<\infty\}}\right] \\
&= \mathbb{P}\left[A \cap \{T < \infty\}\right] \mathbb{E}\left[f_0\left(X(s_1), \dots, X(s_m)\right)\right]. \qquad (5.196)
\end{aligned}$$

Here $f_0 : (\mathbb{R}^\nu)^m \to \mathbb{C}$ is an arbitrary bounded continuous function. By passing to limits (5.196) follows for arbitrary bounded Borel measurable functions $f_0 : (\mathbb{R}^\nu)^m \to \mathbb{C}$. Via the monotone class theorem the assertion in (a) follows.

(b) By taking the function f of the form $f = f_0 \circ V_m$, where $V_m : (\mathbb{R}^\nu)^m \to (\mathbb{R}^\nu)^m$ is given by $V_m\left(x_1, \dots, x_m)\right) = (x_1, \dots, x_1 + \cdots x_m)$ in (5.195), we get

$$\begin{aligned} &\widetilde{\mathbb{E}}\left[f_0\left(Y(s_1), \dots, Y(s_m)\right) 1_{A \cap \{T<\infty\}}\right] \\ &\quad = \mathbb{P}\left[A \cap \{T < \infty\}\right] \mathbb{E}\left[f_0\left(X(s_1), \dots, X(s_m)\right)\right], \qquad (5.197) \end{aligned}$$

where $f_0 : (\mathbb{R}^\nu)^m \to \mathbb{C}$ is an arbitrary bounded continuous function. Then choose $A = \Omega$ and divide by $\mathbb{P}\{T < \infty\}$. We get

$$\widetilde{\mathbb{E}}\left[f_0\left(Y(s_1), \dots, Y(s_m)\right)\right] = \mathbb{E}\left[f_0\left(X(s_1), \dots, X(s_m)\right)\right]. \qquad (5.198)$$

Inserting the result in (5.198) into (5.197) entails

$$\widetilde{\mathbb{E}}\left[f_0\left(Y(s_1), \dots, Y(s_m)\right) 1_A\right) = \widetilde{\mathbb{E}}\left[f_0\left(Y(s_1), \dots, Y(s_m)\right)\right] \widetilde{\mathbb{P}}(A). \qquad (5.199)$$

From (5.199) it follows that the σ-field $\mathcal{F}_T$ is independent of the one generated by $\{Y(s) : s \geqslant 0\}$: for this employ the monotone class theorem.

This completes the proof of Theorem 5.124. ☐

7. Markov processes

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let X, Y and Z be stochastic variables on Ω with values in a topological Hausdorff space E. We assume that E is locally compact and that E is second countable, or, what is the same, that E satisfies the second countability axiom. In other words E has a countable basis for its topology. The space E is supplied with the Borel σ-field $\mathcal{E}$ and we suppose that the variables X, Y and Z are measurable for the σ-fields $\mathcal{F}$ and $\mathcal{E}$. The symbol $\mathbb{P}_X$ stands for the *image measure* on $\mathcal{E}$ of the probability $\mathbb{P}$ under the mapping X. So $\mathbb{P}_X(B) = \mathbb{P}(X \in B)$, $B \in \mathcal{E}$. The symbol $\mathbb{P}_{Y\big|X}$ is a probability kernel from Ω to E with the property that

$$\int_B \mathbb{P}_{Y\big|X}(x, C)\mathbb{P}_X(dx) = \mathbb{P}(Y \in C, X \in B)$$

for all B and C in $\mathcal{E}$. As function of the first variable the probability kernel $\mathbb{P}_{Y\big|X}$ is $\mathbb{P}_X$-almost surely determined. Putting it differently, the function $x \mapsto \mathbb{P}_{Y\big|X}(x, C)$ is the Radon-Nikodym derivative of the measure $B \mapsto \mathbb{P}(Y \in C, X \in B)$ with respect to the measure $B \mapsto \mathbb{P}_X(B) = \mathbb{P}(X \in B)$. In the following proposition we collect some useful formulas for (conditional) probability kernels.

5.125. PROPOSITION. *Let $(\Omega, \mathcal{F}, \mathbb{P})$, E, X, Y and Z be as described above. Let $g : E \times E \to \mathbb{C}$ be a bounded measurable function and let B and C belong to $\mathcal{E}$. Then the following equalities hold:*

$$\int\int g(x, y)\mathbb{P}_{Y\big|X}(x, dy)\mathbb{P}_X(dx) = \mathbb{E}(g(X, Y)); \tag{5.200}$$

$$\int_B \mathbb{P}_{Y\big|X}(x, C)\mathbb{P}_X(dx) = \mathbb{E}(1_C(Y), X \in B); \tag{5.201}$$

$$\mathbb{P}_{Y\big|X}(X, C) = \mathbb{E}\left(1_C(Y) \bigm| \sigma(X)\right); \tag{5.202}$$

$$\int \mathbb{P}_{Z\big|Y}(y, C)\mathbb{P}_{Y\big|X}(x, dy) = \mathbb{P}_{Z\big|X}(x, C), \quad \mathbb{P}_X\text{-almost surely,} \tag{5.203}$$

provided that $\mathbb{E}\left(Z \bigm| \sigma(Y)\right) = \mathbb{E}\left(Z \bigm| \sigma(X, Y)\right)$.

PROOF. The equality in (5.201) follows in fact from the definition of $\mathbb{P}_{Y\big|X}$. By choosing the function g of the form $g(x, y) = 1_B(x)1_C(y)$ in (5.200) we see that (5.200) coincides with (5.201). An arbitrary bounded measurable function g can be approximated by linear combinations of functions of the form $(x, y) \mapsto 1_B(x)1_C(y)$i, with B and C in $\mathcal{E}$. Let $g : E \to E$ be a bounded measurable function. Then the following equalities hold:

$$\mathbb{E}\left(g(X)\mathbb{P}_{Y\big|X}(X, C)\right) = \int g(x)\mathbb{P}_{Y\big|X}(x, C)\mathbb{P}_X(dx) = \mathbb{E}(g(X)1_C(Y)). \tag{5.204}$$

From (5.204) the equality in (5.202) follows. Let $g : E \to \mathbb{C}$ be a bounded measurable function. Then by, among others, (5.202) the following equalities are true:

$$
\begin{aligned}
&\int g(x)\int \mathbb{P}_{Z\big|Y}(y,C)\mathbb{P}_{Y\big|X}(x,dy)\mathbb{P}_X(dx)\\
&= \mathbb{E}\left(\mathbb{P}_{Z\big|Y}(Y,C)g(X)\right) = \mathbb{E}\left(\mathbb{E}\left(1_C(Z)\mid \sigma(Y)\right)g(X)\right)\\
&= \mathbb{E}\left(\mathbb{E}\left(1_C(Z)\mid \sigma(X,Y)\right)g(X)\right) = \mathbb{E}\left(\mathbb{E}\left(1_C(Z)g(X)\mid \sigma(X,Y)\right)\right)\\
&= \mathbb{E}\left(1_C(Z)g(X)\right) = \int g(x)\mathbb{P}_{Z\big|X}(x,C)\mathbb{P}_X(dx). \qquad (5.205)
\end{aligned}
$$

From (5.205) the equality in (5.203) follows, and completes the proof of Proposition 5.125. ☐

5.126. THEOREM. *Let* $(\Omega,\mathcal{F},\mathbb{P})$ *and* $(E,\mathcal{E})$ *be as above. Let* $X=\{X(t):t\geqslant 0\}$ *be a stochastic process with values in the state space* E *adapted to the filtration* $(\mathcal{F}_t:t\geqslant 0)$*. So every state variable* $X(t)$ *is a mapping from* Ω *to* E*, measurable for the* σ*-fields* $\mathcal{F}_t$ *and* $\mathcal{E}$*. In addition, suppose that the family of operators* $\{\vartheta_t:t\geqslant 0\}$ *from* Ω *to* Ω *satisfies the translation property* $X(s)\circ\vartheta_t = X(s+t)$ *for all* s *and* $t\geqslant 0$*. Then the following assertions are equivalent (for the implication (iii)* $\Rightarrow$ *(i) it is assumed that* $\mathcal{F}_t=\sigma\{X(u):0\leqslant u\leqslant t\}$*):*

(i) *For every* $C\in\mathcal{E}$ *and every* s *and* $t\geqslant 0$ *the following equality holds:*

$$\mathbb{E}\left[1_C(X(s+t))\mid \mathcal{F}_t\right] = \mathbb{E}\left[1_C(X(s+t))\mid \sigma\left(X(t)\right)\right] \quad \mathbb{P}\text{-almost surely;} \qquad (5.206)$$

(ii) *For every bounded random variable* $Y:\Omega\to\mathbb{C}$*, that is measurable for* $\mathcal{F}_\infty$ *and* $\mathcal{E}$*, and for every* $t\geqslant 0$ *the following equality holds:*

$$\mathbb{E}\left[Y\circ\vartheta_t\mid\mathcal{F}_t\right] = \mathbb{E}\left[Y\circ\vartheta_t\mid\sigma\left(X(t)\right)\right]\mathbb{P}\text{-almost surely;} \qquad (5.207)$$

(iii) *For every* $m\in\mathbb{N}$ *and for all* $(m+1)$*-tuple of bounded Borel measurable functions* $f_0,\dots,f_m:E\to\mathbb{C}$ *the equality:*

$$
\begin{aligned}
&\mathbb{E}\left[f_0\left(X(0)\right)f_1\left(X(s_1)\right)\dots f_m\left(X(s_m)\right)\right] \qquad (5.208)\\
&= \underbrace{\int\int\dots\int}_{m+1 \text{ times}} f_0\left(x_0\right)f_1\left(x_1\right)\dots f_m\left(x_m\right)\\
&\qquad \mathbb{P}_{X(s_m)\big|X(s_{m-1})}\left(x_{m-1},dx_m\right)\dots\mathbb{P}_{X(s_1)\big|X(0)}(x_0,dx_1)\mathbb{P}_{X(0)}(dx_0),
\end{aligned}
$$

holds for every $s_1<\dots<s_m$ *in* $[0,\infty)$*.*
If the process X *is right-continuous, then (i) and (ii) are also equivalent with the following assertions:*

(iv) *For every bounded Borel measurable function* $f:E\to\mathbb{C}$ *and for every stopping time* $T:\Omega\to[0,\infty]$ *the following equality holds* $\mathbb{P}$*-almost surely on the event* $\{T<\infty\}$*:*

$$\mathbb{E}\left[f(X(s+T))\mid\mathcal{F}_T\right] = \mathbb{E}\left[f(X(s+T))\mid\sigma\left(T,X(T)\right)\right];$$

(v) *For every bounded random variable* $Y:\Omega\to\mathbb{C}$*, which is measurable for* $\mathcal{F}_\infty$*, and for every stopping time* $T:\Omega\to[0,\infty]$ *the equality*

$$\mathbb{E}\left[Y\circ\vartheta_T\mid\mathcal{F}_T\right] = \mathbb{E}\left[Y\circ\vartheta_T\mid\sigma\left(T,X(T)\right)\right] \qquad (5.209)$$

holds $\mathbb{P}$*-almost surely on the event* $\{T < \infty\}$.

If the process X *is right-continuous and if as filtration the internal history is chosen, then all assertions (i) through (v) are equivalent.*

Proof. (i) $\Rightarrow$ (ii). Upon invoking the monotone class theorem it suffices to prove (ii) for functions $Y : \Omega \to \mathbb{C}$ of the form $Y = \prod_{j=1}^{m} f_j\left(X(s_j)\right)$, where the functions f_j, $1 \leqslant j \leqslant m$ are bounded and measurable. For $m = 1$ (i) is clearly equivalent with (ii). Next we prove (ii) for $Y = \prod_{j=1}^{m+1} f_j\left(X(s_j)\right)$ starting from (ii), but with $Y = \prod_{j=1}^{k} f_j\left(X(s_j)\right)$, with $1 \leqslant k \leqslant m$. The equalities below then show that (5.207) follows for $Y = \prod_{j=1}^{m+1} f_j\left(X(s_j)\right)$:

$$
\begin{aligned}
&\mathbb{E}\left[\prod_{j=1}^{m+1} f_j\left(X(s_j+t)\right) \mid \mathcal{F}_t\right] \\
&\quad = \mathbb{E}\left[\mathbb{E}\left(\prod_{j=1}^{m+1} f_j\left(X(s_j+t)\right) \mid \mathcal{F}_{s_m+t}\right) \mid \mathcal{F}_t\right] \\
&\quad = \mathbb{E}\left[\prod_{j=1}^{m} f_j\left(X(s_j+t)\right) \mathbb{E}\left(f_{m+1}\left(X(s_{m+1}+t)\right) \mid \mathcal{F}_{s_m+t}\right) \mid \mathcal{F}_t\right]
\end{aligned}
$$

(the equality in (5.207) for $Y = f_{m+1}\left(X(s_{m+1})\right)$)

$$
= \mathbb{E}\left[\prod_{j=1}^{m} f_j\left(X(s_j+t)\right) \mathbb{E}\left[f_{m+1}\left(X(s_{m+1}+t)\right) \mid \sigma\left(X\left(s_m+t\right)\right)\right] \mid \mathcal{F}_t\right]
$$

(the equality in (5.207) for $Y = \prod_{j=1}^{m} g_j\left(X(s_j)\right)$, where $g_j = f_j$, $1 \leqslant j \leqslant m-1$, and where $g_m(x) = f_m(x) \int f_{m+1}(y) \mathbb{P}_{X(s_{m+1}+t)\big| X(s_m+t)}(x, dy))$

$$
\begin{aligned}
&= \mathbb{E}\left[\prod_{j=1}^{m} f_j\left(X(s_j+t)\right) \mathbb{E}\left(f_{m+1}\left(X(s_{m+1}+t)\right) \mid \sigma(X\left(s_m+t\right)\right) \mid \sigma(X(t))\right] \\
&= \mathbb{E}\left[\prod_{j=1}^{m} f_j\left(X(s_j+t)\right) \mathbb{E}\left(f_{m+1}\left(X(s_{m+1}+t)\right) \mid \mathcal{F}_{s_m+t}\right) \mid \sigma(X(t))\right] \\
&= \mathbb{E}\left[\mathbb{E}\left(\prod_{j=1}^{m} f_j\left(X(s_j+t)\right) f_{m+1}\left(X(s_{m+1}+t)\right) \mid \mathcal{F}_{s_m+t}\right) \mid \sigma(X(t))\right] \\
&= \mathbb{E}\left[\prod_{j=1}^{m} f_j\left(X(s_j+t)\right) f_{m+1}\left(X(s_{m+1}+t)\right) \mid \sigma(X(t))\right]. \qquad (5.210)
\end{aligned}
$$

Then observe that (5.210) is the same as (5.207), but for $Y = \prod_{j=1}^{m+1} f_j\left(X(s_j)\right)$. This proves the implication (i) $\Rightarrow$ (ii).

(ii) $\Rightarrow$ (i). This implication follows by putting $Y = 1_C(X(s))$.

(ii) $\Rightarrow$ (iii). The equality in (5.208) is correct for $m = 0$ and for $m = 1$. This is a consequence of Proposition 5.125. Again we will apply induction with respect to m. We assume that (5.208) is correct for m and for the increasing m-tuple $s_1 < s_2 < \cdots s_m$. Then we see

$$\begin{aligned}\mathbb{E}\left[\prod_{j=0}^{m+1} f_j\left(X(s_j)\right)\right] &= \mathbb{E}\left[\mathbb{E}\left(\prod_{j=0}^{m+1} f_j\left(X(s_j)\right) \mid \mathfrak{F}_{s_m}\right)\right] \\ &= \mathbb{E}\left[\prod_{j=0}^{m} f_j\left(X(s_j)\right) \mathbb{E}\left(f_{m+1}\left(X(s_{m+1})\right) \mid \mathfrak{F}_{s_m}\right)\right] \\ &= \mathbb{E}\left[\prod_{j=0}^{m} f_j\left(X(s_j)\right) \mathbb{E}\left(f_{m+1}\left(X(s_{m+1})\right) \mid \sigma\big(X(s_m)\big)\right)\right]\end{aligned}$$

(Proposition 5.125)

$$\begin{aligned}&= \mathbb{E}\left[\prod_{j=0}^{m} f_j\left(X(s_j)\right) \int f_{m+1}\left(x_{m+1}\right) \mathbb{P}_{X(s_{m+1})\big|X(s_m)}\left(X(s_m), dx_{m+1}\right)\right] \\ &= \underbrace{\int\int \dots \int}_{m+2 \text{ times}} f_0(x_0) \dots f_{m+1}(x_{m+1}) \mathbb{P}_{X(s_{m+1})\big|X(s_m)}\left(x_m, dx_{m+1}\right) \dots \\ &\qquad\qquad \mathbb{P}_{X(s_1)\big|X(0)}\left(x_0, dx_1\right) \mathbb{P}_{X(0)}\left(dx_0\right). \qquad (5.211)\end{aligned}$$

From the equality in (5.208) for m the equality in (5.200) follows for $m + 1$ instead of m.

(iii) ⇒ (i). Let $C \in \mathcal{E}$ and let s and $t > 0$. Starting from (iii) we will prove that the following equality holds $\mathbb{P}$-almost surely:

$$\mathbb{E}\left(1_C\left(X(s+t)\right) \mid \mathcal{F}_t\right) = \mathbb{E}\left(1_C\left(X(s+t)\right) \mid \sigma\left(X(t)\right)\right). \tag{5.212}$$

Choose $0 < t_1 < \cdots < t_m = t$ and choose bounded Borel measurable functions $f_0, \dots, f_m$. The following equality is a consequence of (iii):

$$\begin{aligned}
&\mathbb{E}\left(f_0(X_0)\dots f_m\left(X(t_m)\right)1_C(X(s+t))\right) \\
&= \underbrace{\int\int\dots\int\int}_{m+2 \text{ times}} f_0(x_0)\dots f_m(x_m)1_C(x_{m+1}) \\
&\qquad \mathbb{P}_{X(s+t)\big|X(t_m)}\left(x_m, dx_{m+1}\right)\mathbb{P}_{X(t_m)\big|X(t_{m-1})}\left(x_{m-1}, dx_m\right)\dots \\
&\qquad\qquad \mathbb{P}_{X(t_1)\big|X(0)}\left(x_0, dx_1\right)\mathbb{P}_{X(0)}(dx_0) \\
&= \underbrace{\int\int\dots\int\int}_{m+1 \text{ times}} f_0(x_0)\dots f_m(x_m) \\
&\qquad \mathbb{P}_{X(s+t)\big|X(t_m)}\left(x_m, C\right)\mathbb{P}_{X(t_m)\big|X(t_{m-1})}\left(x_{m-1}, dx_m\right)\dots \\
&\qquad\qquad \mathbb{P}_{X(t_1)\big|X(0)}\left(x_0, dx_1\right)\mathbb{P}_{X(0)}(dx_0) \\
&= \mathbb{E}\left[f_0(X_0)\dots f_m\left(X(t_m)\right)\mathbb{P}_{X(s+t)\big|X(t)}\left(X(t_m), C\right)\right].
\end{aligned} \tag{5.213}$$

The monotone class theorem applies to the effect that (5.212) follows from (5.213), provided that the internal history is chosen as filtration.

(iv) ⇒ (v). By the monotone class theorem it suffices to prove (ii) for functions $Y : \Omega \to \mathbb{C}$ of the form $Y = \prod_{j=1}^{m} f_j\left(X(s_j)\right)$, where the functions f_j, $1 \leqslant j \leqslant m$ are bounded and measurable. For $m = 1$ it is clear that (iv) is equivalent to (v). We prove (v) for $Y = \prod_{j=1}^{m+1} f_j\left(X(s_j)\right)$ starting from (iv), but with $Y = \prod_{j=1}^{k} f_j\left(X(s_j)\right)$, for $1 \leqslant k \leqslant m$. The following equalities show that the equality (5.209) then follows for $Y = \prod_{j=1}^{m+1} f_j\left(X(s_j)\right)$:

$$\begin{aligned}
&\mathbb{E}\left[\prod_{j=1}^{m+1} f_j\left(X(s_j+T)\right) \mid \mathcal{F}_T\right] \\
&= \mathbb{E}\left[\mathbb{E}\left(\prod_{j=1}^{m+1} f_j\left(X(s_j+T)\right) \mid \mathcal{F}_{s_m+T}\right) \mid \mathcal{F}_T\right] \\
&= \mathbb{E}\left[\prod_{j=1}^{m} f_j\left(X(s_j+T)\right)\mathbb{E}\left[f_{m+1}\left(X(s_{m+1}+T)\right) \mid \mathcal{F}_{s_m+T}\right] \mid \mathcal{F}_T\right]
\end{aligned}$$

(apply equality (5.209) for $Y = f_{m+1}\left(X(s_{m+1})\right)$)

$$= \mathbb{E}\left[\prod_{j=1}^{m} f_j\left(X(s_j+T)\right)\mathbb{E}\left[f_{m+1}\left(X(s_{m+1}+T)\right) \mid \sigma\left(s_m+T, X\left(s_m+T\right)\right)\right] \mid \mathcal{F}_T\right]$$

(use equality (5.209) for $Y = \prod_{j=1}^{m} g_j\left(X(s_j)\right)$, where $g_j = f_j$, $1 \leqslant j \leqslant m-1$, and where $g_m(x) = f_m(x) \int f_{m+1}(y) \mathbb{P}_{(s_m+T, X(s_{m+1}+T))\big|(s_m+T, X(s_m+T))}(x, dy))$

$$
\begin{aligned}
&= \mathbb{E}\Bigg[\prod_{j=1}^{m} f_j\left(X(s_j+T)\right) \mathbb{E}\left[f_{m+1}\left(X(s_{m+1}+T)\right) \bigm| \sigma\left(T, X\left(s_m+T\right)\right)\right] \bigm| \\
&\qquad\qquad \sigma(T, X(T))\Bigg] \\
&= \mathbb{E}\left[\prod_{j=1}^{m} f_j\left(X(s_j+T)\right) \mathbb{E}\left[f_{m+1}\left(X(s_{m+1}+T)\right) \bigm| \mathcal{F}_{s_m+T}\right] \bigm| \sigma(T, X(T))\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\prod_{j=1}^{m} f_j\left(X(s_j+T)\right) f_{m+1}\left(X(s_{m+1}+T)\right) \bigm| \mathcal{F}_{s_m+T}\right] \bigm| \sigma(T, X(T))\right] \\
&= \mathbb{E}\left[\prod_{j=1}^{m} f_j\left(X(s_j+T)\right) f_{m+1}\left(X(s_{m+1}+T)\right) \bigm| \sigma(T, X(T))\right]. \qquad (5.214)
\end{aligned}
$$

Then realize that (5.214) is the same as (5.209) for $Y = \prod_{j=1}^{m+1} f_j\left(X(s_j)\right)$. This proves the implication (iv) $\Rightarrow$ (v). The implication (v) $\Rightarrow$ (iv) is again trivial.

(i) $\Rightarrow$ (iv). By the fact E satisfies the second countability axiom, and by the fact $\mathcal{E}$ is the Borel field it suffices to prove (iv) for functions $f \in C_0(E)$ instead of 1_C (verify this precisely). So we have to show the following equality:

$$
\mathbb{E}\left[f\left(X(s+T)\right) \bigm| \mathcal{F}_T\right] 1_{\{T<\infty\}} = \mathbb{E}\left[f\left(X(s+T)\right) \bigm| \sigma(T, X(T))\right] 1_{\{T<\infty\}}, \quad (5.215)
$$

for $f \in C_0(E)$ and for $s \geqslant 0$. By employing the right-continuity of paths, it suffices to prove (5.215) for the stopping times $T_n := 2^{-n}\lceil 2^n T\rceil$, $n \in \mathbb{N}$, instead of T. The equality for T then follows from those of T_n by letting n tend to ∞. For this notice that $0 \leqslant T - T_{n+1} \leqslant T - T_n \leqslant 2^{-n}$. Choose the event $A \in \mathcal{F}_{T_n}$. Then the event $A \cap \{T_n = k2^{-n}\}$ belongs to $\mathcal{F}_{k2^{-n}}$ and the following equalities hold:

$$
\begin{aligned}
&\mathbb{E}\left[f\left(X(s+T_n)\right), A \cap \left\{T_n = k2^{-n}\right\}\right] \\
&\quad = \mathbb{E}\left[\mathbb{E}\left(f\left(X(s+T_n)\right) \bigm| \mathcal{F}_{k2^{-n}}\right), A \cap \left\{T_n = k2^{-n}\right\}\right] \\
&\quad = \mathbb{E}\left[\int f(y) \mathbb{P}_{X(s+k2^{-n})\big|X(k2^{-n})}\left(X(k2^{-n}), dy\right),\ A \cap \left\{T_n = k2^{-n}\right\}\right] \\
&\quad = \mathbb{E}\left[\omega \mapsto \int f(y) \mathbb{P}_{X(s+T_n(\omega))\big|X(T_n(\omega))}\left(X(T_n)(\omega), dy\right) 1_{\{A\cap\{T_n=k2^{-n}\}\}}(\omega)\right]. \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (5.216)
\end{aligned}
$$

We also have

$$
\begin{aligned}
&\mathbb{E}\left[f\left(X(s+T_n)\right),\ A \cap \left\{T_n = k2^{-n}\right\}\right] \\
&\quad = \mathbb{E}\left[\mathbb{E}\left(f\left(X(s+k2^{-n})\right) \bigm| \mathcal{F}_{k2^{-n}}\right),\ A \cap \left\{T_n = k2^{-n}\right\}\right]
\end{aligned}
$$

(because of (i))

$$\begin{aligned}
&= \mathbb{E}\left[\mathbb{E}\left(f\left(X(s+k2^{-n})\right) \mid \sigma\left(X\left(k2^{-n}\right)\right)\right),\ A \cap \left\{T_n = k2^{-n}\right\}\right] \\
&= \mathbb{E}\left[\int f(y)\mathbb{P}_{X(s+k2^{-n})\big|X(k2^{-n})}\left(X(k2^{-n}), dy\right),\ A \cap \left\{T_n = k2^{-n}\right\}\right] \\
&= \mathbb{E}\left[\omega \mapsto \int f(y)\mathbb{P}_{X(s+T_n(\omega))\big|X(T_n(\omega))}\left(X(T_n)(\omega), dy\right) 1_{\{A\cap\{T_n=k2^{-n}\}\}}(\omega)\right].
\end{aligned} \tag{5.217}$$

We see that (5.216) and (5.217) are the same. It follows that the assertion in (iv) is proved for T_n instead of T. By letting n tend to ∞ we then obtain (iv) for T (by employing the right-continuity of paths of the process).

So the proof of Theorem 5.126 is complete now. ☐

Sharp Minds - Bright Ideas!

Employees at FOSS Analytical A/S are living proof of the company value - First - using new inventions to make dedicated solutions for our customers. With sharp minds and cross functional teamwork, we constantly strive to develop new unique products - Would you like to join our team?

FOSS works diligently with innovation and development as basis for its growth. It is reflected in the fact that more than 200 of the 1200 employees in FOSS work with Research & Development in Scandinavia and USA. Engineers at FOSS work in production, development and marketing, within a wide range of different fields, i.e. Chemistry, Electronics, Mechanics, Software, Optics, Microbiology, Chemometrics.

We offer
A challenging job in an international and innovative company that is leading in its field. You will get the opportunity to work with the most advanced technology together with highly skilled colleagues.

Read more about FOSS at www.foss.dk - or go directly to our student site www.foss.dk/sharpminds where you can learn more about your possibilities of working together with us on projects, your thesis etc.

The Family owned FOSS group is the world leader as supplier of dedicated, high-tech analytical solutions which measure and control the quality and production of agricultural, food, pharmaceutical and chemical products. Main activities are initiated from Denmark, Sweden and USA with headquarters domiciled in Hillerød, DK. The products are marketed globally by 23 sales companies and an extensive net of distributors. In line with the corevalue to be 'First', the company intends to expand its market position.

Dedicated Analytical Solutions
FOSS
Slangerupgade 69
3400 Hillerød
Tel. +45 70103370
www.foss.dk

Click on the ad to read more

We continue with some definitions.

5.127. Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let E be a locally compact Hausdorff space with a countable basis for its topology. In addition, let $(\mathcal{F}_t : t \geqslant 0)$ be a filtration on Ω. Let $X = \{X(t) : t \geqslant 0\}$ be a process attaining values in E. The state space E is equipped with the Borel filed and it is assumed that X is an adapted process. Suppose that for every $x \in E$ the (sub-)probability kernel $\mathbb{P}_{X(s+t)\big|X(t)}(x, C)$, $C \in \mathcal{E}$, is defined. Here the (sub-)probability kernel $\mathbb{P}_{Y|X}(x, C)$ possesses the following defining property:

$$\int_B \mathbb{P}_{Y\big|X}(x, C)\mathbb{P}(X \in dx) = \mathbb{P}\{Y \in C,\ X \in B\},$$

where B and C are Borel subsets of E and where X and Y are stochastic variables with values in E. In addition, it is assumed that there are so-called translation operators $\vartheta_t : \Omega \to \Omega$ with the property that $X(s) \circ \vartheta_t = X(s+t)$ for all s, $t \geqslant 0$. Moreover, by hypothesis the process X is cadlag. We say that the process X is a *Markov process* if for every $C \in \mathcal{E}$ and every $t \geqslant 0$ the equality

$$\mathbb{E}\left[1_C(X(s+t)) \mid \mathcal{F}_t\right] = \mathbb{E}\left[1_C(X(s+t)) \mid \sigma(X(t))\right] \tag{5.218}$$

is $\mathbb{P}$-almost surely true for all $s \geqslant 0$. The process X is called a *strong Markov* process if equality (5.218) also holds for stopping times. More precisely, if for every $s \geqslant 0$, for every $C \in \mathcal{E}$ and for every stopping time $T : \Omega \to [0, \infty]$ the equality

$$\mathbb{E}\left[1_C(X(s+T)) \mid \mathcal{F}_T\right] = \mathbb{E}\left[1_C(X(s+T)) \mid \sigma(T, X(T))\right]$$

holds $\mathbb{P}$-almost surely on the event $\{T < \infty\}$. If the process X is cadlag is, then a Markov process is automatically a strong Markov: see Theorem 5.126. We say that a Markov process X is *time homogeneous* if for all $C \in \mathcal{E}$ and for all s and $t \geqslant 0$ the equality

$$\mathbb{P}_{X(s+t)\big|X(t)}(x, C) = \mathbb{P}_{X(s)\big|X(0)}(x, C) \tag{5.219}$$

is true for all $x \in E$. In what follows we always suppose that X is a cadlag, time homogeneous Markov process. Furthermore we define the operators $\{P(t) : t \geqslant 0\}$ via the formula

$$\left[P(s)f\right](x) = \int f(y)\mathbb{P}_{X(s)\big|X(0)}(x, dy). \tag{5.220}$$

Here $s \geqslant 0$ and f belongs to $C_0(E)$. Since we have (see equality (5.203) in Proposition 5.125)

$$\int \mathbb{P}_{X(s+t)\big|X(t)}(y, C)\mathbb{P}_{X(t)\big|X(0)}(x, dy) = \mathbb{P}_{X(s+t)\big|X(0)}(x, C), \quad \mathbb{P}_{X(0)}\text{-almost surely,} \tag{5.221}$$

we get, for a time-homogeneous Markov process X the following equalities:

$$\begin{aligned}\left[P(s)P(t)f\right](x) &= \int \left[P(t)f\right](y)\mathbb{P}_{X(s)\big|X(0)}(x, dy)\\ &= \int\int f(z)\mathbb{P}_{X(t)\big|X(0)}(y, dz)\mathbb{P}_{X(s)\big|X(0)}(x, dy)\end{aligned}$$

(X is time homogeneous)

$$= \int\int f(z)\mathbb{P}_{X(s+t)\big|X(s)}(y,dz)\mathbb{P}_{X(s)\big|X(0)}(x,dy)$$

(employ equality (5.221))

$$= \int\int f(z)\mathbb{P}_{X(s+t)\big|X(0)}(x,dz) = \left[P(s+t)f\right](x)$$

The cadlag property of X implies $\lim_{s\downarrow 0}\left[P(s)f\right](x) = f(x)$ for all $f \in C_0(E)$ and for all $x \in E$. If $P(s)f$ belongs to $C_0(E)$ for every $f \in C_0(E)$ and for every $s \geqslant 0$, then the family $\{P(t) : t \geqslant 0\}$ apparently constitutes a *Feller semigroup*. Put $P(s,x,C) = \mathbb{P}_{X(s)\big|X(0)}(x,C)$, $s \geqslant 0$, $x \in E$, $C \in \mathcal{E}$. Let the expectation values of $\mathbb{E}_x(Y)$, $x \in E$, $Y = \prod_{j=1}^m f_j\left(X\left(s_j\right)\right)$, $s_1 < s_2 < \ldots < s_m$, be determined by the formula:

$$\begin{aligned}&\mathbb{E}_x\left(\prod_{j=1}^m f_j(X(s_j))\right)\\ &= \int\ldots\int\prod_{j=1}^m f_j\left(x_j\right)P\left(s_1,x,dx_1\right)\ldots P\left(s_m-s_{m-1},x_{m-1},dx_m\right).\end{aligned} \tag{5.222}$$

Instead of (5.222) most of the time we write $\mathbb{E}_x(Y) = \mathbb{E}\left[Y \;\middle|\; X(0)=x\right]$, for a bounded stochastic variable Y. Since X is a time homogeneous Markov process we see that the following equality also holds $\mathbb{P}_x$-almost surely:

$$\mathbb{E}_x\left(Y\circ\vartheta_t \;\middle|\; \mathcal{F}_t\right) = \mathbb{E}_{X(t)}(Y), \tag{5.223}$$

for all $t \geqslant 0$ and for all bounded random variables Y. The equality in (5.223) is first proved for random variables Y of the form $Y = \prod_{j=1}^m f_j(X(s_j))$, where the functions f_j, $1 \leqslant j \leqslant m$, are bounded Borel functions. Equality (5.223) is also true if $\mathbb{P}_x$ and $\mathbb{E}_x$ are replaced by $\mathbb{P}$ and $\mathbb{E}$ respectively.

5.128. Remark. The expectation value $\mathbb{E}_x(Y)$ is in fact the *Radon-Nikodym* derivative of de measure $B \mapsto \mathbb{E}\left[Y, X(0)\in B\right]$ with respect to the measure $B \mapsto \mathbb{P}\left[X(0)\in B\right]$. If in this definition we take for Y the variable $Y = 1_C\left(X(s)\right)$, then we obtain the probability kernel $\mathbb{P}_{X(s)\big|X(0}(x,C)$. Hence, these quantities are defined as Radon-Nikodym derivatives. So, in general, the expression $\mathbb{P}_{X(s)\big|X(0)}(x,C)$ is not defined for every $x \in E$. However, we will assume that these probability kernels exist for every $x \in E$ indeed, and that the corresponding semigroup is a Feller. Many authors define a (time homogeneous) Markov process X relative to a family of probability measures $\{\mathbb{P}_x : x \in E\}$ by means of the following equality:

$$\mathbb{E}_x\left(Y\circ\vartheta_t \;\middle|\; \mathcal{F}_t\right) = \mathbb{E}_{X(t)}\left(Y\right), \tag{5.224}$$

$\mathbb{P}_x$-almost surely for all $x \in E$, for all $t \geqslant 0$ and for all bounded random variables $Y : \Omega \to \mathbb{C}$. In fact we also do this. In the time homogeneous case the equality in (5.224) also holds for stopping times T:

$$\mathbb{E}_x\left(Y\circ\vartheta_T \;\middle|\; \mathcal{F}_T\right) = \mathbb{E}_{X(T)}\left(Y\right), \tag{5.225}$$

$\mathbb{P}_x$-almost surely on the event $\{T < \infty\}$, provided that the process X is cadlag. The equality in (5.225) can be proved in the same manner as equality (5.209) in Theorem 5.126. Therefore pick $f \in C_0(E)$ and a stopping time $T : \Omega \to [0, \infty]$. Consider the stopping times $T_n := 2^{-n}\lceil 2^n T\rceil$, $n \in \mathbb{N}$, instead of T. Then, for an event $A \in \mathcal{F}_{T_n}$, we have

$$\begin{aligned}
&\mathbb{E}_x\left[f\left(X(s+T_n)\right)1_{A\cap\{T_n=k2^{-n}\}}\right] \\
&= \mathbb{E}_x\left[f\left(X\left(s+k2^{-n}\right)\right)1_{A\cap\{T_n=k2^{-n}\}}\right] \\
&= \mathbb{E}_x\left[\mathbb{E}_x\left(f\left(X\left(s+k2^{-n}\right)\right) \bigm| \mathcal{F}_{k2^{-n}}\right)1_{A\cap\{T_n=k2^{-n}\}}\right] \\
&= \mathbb{E}_x\left[\mathbb{E}_x\left(f\left(X\left(s+k2^{-n}\right)\right) \bigm| \mathcal{F}_{k2^{-n}}\right)1_{A\cap\{T_n=k2^{-n}\}}\right] \\
&= \mathbb{E}_x\left[\mathbb{E}_{X(k2^{-n})}\left(f\left(X(s)\right)\right)1_{A\cap\{T_n=k2^{-n}\}}\right] \\
&= \mathbb{E}_x\left[\mathbb{E}_{X(T_n)}\left(f\left(X(s)\right)\right)1_{A\cap\{T_n=k2^{-n}\}}\right]. \qquad (5.226)
\end{aligned}$$

From (5.226) it follows that (5.225) for $Y = f(X(s))$ and for T_n in the place of T. By taking the limit in (5.226) for $n \to \infty$ the equality in (5.225) follows for $Y = f(X(s))$. Precisely as in the proof of the implication (iv) ⇒ (v) in Theorem 5.126 the equality in (5.225) then follows for arbitrary random variables $Y : \Omega \to \mathbb{C}$, which are bounded and measurable for the σ-field $\mathcal{F}_\infty$.

8. The Doob-Meyer decomposition via Komlos theorem

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\{\mathcal{F}_t : t \geqslant 0\}$ be a right continuous filtration in $\mathcal{F}$ and let $\{X(t) : t \geqslant 0\}$ be a real-valued $\mathcal{F}_t$-submartingale. The Doob-Meyer decomposition theorem states that there exists an $\mathcal{F}_t$-martingale $\{M(t) : t \geqslant 0\}$ together with an increasing predictable adapted process $\{A(t) : t \geqslant 0\}$, which is right continuous $\mathbb{P}$-almost surely, such that $X(t) = M(t) + A(t)$, $t \geqslant 0$, provided that the process $\{X(t) : t \geqslant 0\}$ is of class (DL). The latter means that for every $t > 0$ the family $\{X(\tau) : 0 \leqslant \tau \leqslant t, \tau \text{ stopping time}\}$ is uniformly integrable. Moreover this decomposition is unique in case we assume that $A(0) = 0$. By Doob's optional sampling theorem every martingale is automatically of class (DL) (see e.g. Ikeda and Watanabe [**81**], p.35, Ethier and Kurtz [**68**], p.74). An interesting discussion of the Doob-Meyer decomposition and (sub-)martingale theory can be found in Kopp [**98**]. For a nice account of the Doob-Meyer decomposition theorem the reader may also consult van Neerven [**188**].

We shall employ the following result of Komlos [**97**]. In fact it can be interpreted as kind of a law of large numbers.

5.129. Theorem (Komlos). *Let $\{f_k : k \in \mathbb{N}\}$ be a sequence in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ such that*

$$\sup\left\{\mathbb{E}\left(|f_k|\right) : k \in \mathbb{N}\right\} < \infty.$$

Then there exists an infinite large subset Λ_0 of $\mathbb{N}$ together with a function f in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ such that for every infinite subset Λ of Λ_0

$$\lim_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} f_j = f, \quad \mathbb{P}\text{-almost surely.} \tag{5.227}$$

Examples show that this limit need not be an L^1-limit. Set $\Omega = \mathbb{N}$ with the discrete σ-field and with $\mathbb{P}\{k\} = 2^{-k}$, $k \in \mathbb{N}$. Let $\{f_k : k \in \mathbb{N}\}$ be the sequence defined by $f_k = 2^k e_k$, $k \in \mathbb{N}$, where $\{e_k : k \in \mathbb{N}\}$ is the sequence of the unit vectors. Then $n^{-1}\sum_{j=1}^n f_j \to 0$ pointwise, but $n^{-1}\int\sum_{j=1}^n f_j d\mathbb{P} = 1$, $n \in \mathbb{N}$.

Standard results on continuity properties of submartingales yield the existence of a realization (version) which is continuous from the right and possesses left limits $\mathbb{P}$-almost surely. Henceforth we shall assume that the $\mathcal{F}_t$-submartingale $\{X(t) : t \geqslant 0\}$ is continuous from the right and has left limits $\mathbb{P}$-almost surely. We shall prove that there exists a predictable increasing process $\{A(t) : t \geqslant 0\}$ together with an infinite Λ_0 of $\mathbb{N}$ such that for every infinite subset Λ of Λ_0 and every $t \geqslant 0$ the variable $A(t)$ is given as the limit:

$$A(t) = \lim_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t), \tag{5.228}$$

where

$$A_j(t) = \sum\nolimits_{0\leqslant k<2^j t} \left\{\mathbb{E}\left(X\left(\frac{k+1}{2^j}\right) \mid \mathcal{F}_{k2^{-j}}\right) - X\left(\frac{k}{2^j}\right)\right\}. \tag{5.229}$$

Moreover the process $\{X(t) - A(t) : t \geqslant 0\}$ is an $\mathcal{F}_t$-martingale. The limit in (5.228) is a point-wise almost sure limit as well as an L^1-limit.

Again let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\{\mathcal{F}_t : t \geqslant 0\}$ be a right-continuous filtration in $\mathcal{F}$ and let $\{X(t) : t \geqslant 0\}$ be right continuous submartingale of class (DL) which possesses almost sure left limits. We want to prove the following version of the Doob-Meyer decomposition theorem.

5.130. Theorem. *There exists a unique predictable right continuous increasing process* $\{A(t) : t \geqslant 0\}$ *with* $A(0) = 0$ *such that the process* $\{X(t) - A(t) : t \geqslant 0\}$ *is an* $\mathcal{F}_t$*-martingale.*

It is perhaps useful to insert the following proposition.

5.131. Proposition. *Processes of the form* $M(t) + A(t)$*, with* M *a martingale and with* A *an increasing process in* $L^1(\Omega, \mathcal{F}, \mathbb{P})$ *are of class (DL).*

Proof of Proposition 5.131. Let $\{X(t) = M(t) + A(t) : t \geqslant 0\}$ be the decomposition of the submartingale $\{X(t) : t \geqslant 0\}$ in a martingale $\{M(t) : t \geqslant 0\}$ and an increasing process $\{A(t) : t \geqslant 0\}$ with $A(0) = 0$ and $0 \leqslant \tau \leqslant t$ be any $\mathcal{F}_t$-stopping time. Here t is some fixed time. For $N \in \mathbb{N}$ we have

$$\begin{aligned}\mathbb{E}\left(|X(\tau)| : |X(\tau)| \geqslant N\right) &\leqslant \mathbb{E}\left(|M(\tau)| : |X(\tau)| \geqslant N\right) + \mathbb{E}\left(A(\tau) : |X(\tau)| \geqslant N\right)\\ &\leqslant \mathbb{E}\left(|M(t)| : |X(\tau)| \geqslant N\right) + \mathbb{E}\left(A(\tau) : |X(\tau)| \geqslant N\right)\\ &\leqslant \mathbb{E}\left(|M(t)| + A(t) : |X(\tau)| \geqslant N\right)\\ &\leqslant \mathbb{E}\left(|M(t)| + A(t) : \sup_{0\leqslant s\leqslant t} |X(s)| \geqslant N\right). \qquad (5.230)\end{aligned}$$

Since $N \times \mathbb{P}\left\{\sup_{0\leqslant s\leqslant t} |X(s)| \geqslant N\right\} \leqslant \mathbb{E}\left(|X(t)|\right)$, it follows that

$$\lim_{N\to\infty} \sup\left\{\mathbb{E}\left(|X(\tau)| : |X(\tau)| \geqslant N\right) : 0 \leqslant \tau \leqslant t, \tau \quad \text{stopping time}\right\} = 0. \qquad (5.231)$$

This shows Proposition 5.131 □

Similarly we have the following result.

5.132. Proposition. *Let* $\{X(t) : t \geqslant 0\}$ *be an* $\mathcal{F}_t$*-submartingale. For any real number* N *the process* $\{\max(X(t), N) : t \geqslant 0\}$ *is an* $\mathcal{F}_t$*-submartingale which is of class (DL).*

Next we come to the heart of the matter. The symbol $\lceil x \rceil$, $x \in \mathbb{R}$, denotes the integer k with $k < x \leqslant k+1$.

Proof of Theorem 5.130. It will be convenient to introduce the following processes:

$$X_j(t) = \mathbb{E}\left(X\left(\frac{\lceil 2^j t\rceil}{2^j}\right) \mid \mathcal{F}_t\right), \ t \geqslant 0, \ j \in \mathbb{N}; \tag{5.232}$$

$$A_j(t) = \sum\nolimits_{0 \leqslant k < 2^j t} \left\{\mathbb{E}\left(X\left(\frac{k+1}{2^j}\right) \mid \mathcal{F}_{k2^{-j}}\right) - X\left(\frac{k}{2^j}\right)\right\}. \tag{5.233}$$

The processes $\{A_j(t) : t \geqslant 0\}$ are right continuous and have left limits. The processes $\{A_j(t) : t \geqslant 0\}$ are predictable in the sense that, for j, N in $\mathbb{N}$, the functions $(t, \omega) \mapsto A_j(t, \omega)$ are measurable with respect to the σ-field generated by the collection $\left\{1_{(a,b]} \times A : 0 \leqslant a < b, \ A \in \mathcal{F}_a\right\}$: see e.g. Durrett [**58**], p. 49. Moreover it is readily verified that the process

$$\{X_j(t) - A_j(t) : t \geqslant 0\} \tag{5.234}$$

is an $\mathcal{F}_t$-martingale and that

$$\lim_{j\to\infty} \mathbb{E}\left(A_j(t) - A_j(t-)\right) = 0. \tag{5.235}$$

Equality (5.235) is true because

$$\lim_{s\downarrow t} \mathbb{E}\left(X(s)\right) = \mathbb{E}\left(X(t)\right). \tag{5.236}$$

Equality (5.236) can be proved in the following manner. Put

$$X''(t) = \lim_{h\downarrow 0} \mathbb{E}\left(X(t+h) \mid \mathcal{F}_t\right) = \inf_{h>0} \mathbb{E}\left(X(t+h) \mid \mathcal{F}_t\right).$$

Then $X''(t) \geqslant X(t)$, $\mathbb{P}$-almost surely. The following argument shows that $X''(t) = X(t)$, $\mathbb{P}$-almost surely. Define for $m \in \mathbb{N}$ the stopping time τ_m by

$$\tau_m = \inf\{s > 0 : |X(s)| > m\}.$$

Then, $\mathbb{P}$-almost surely, $\tau_m \uparrow \infty$. Moreover, we have

$$\begin{aligned}
&\mathbb{E}\left[X''(t) - X(t) : \tau_m > t\right] \qquad (5.237)\\
&= \lim_{h\downarrow 0} \mathbb{E}\left[\mathbb{E}\left(X(t+h) \mid \mathcal{F}_t\right) - X(t) : \tau_m > t\right]\\
&= \lim_{h\downarrow 0} \mathbb{E}\left[\mathbb{E}\left(\left(X(t+h) - X(t)\right) 1_{\{\tau_m > t\}}\right) \mid \mathcal{F}_t\right]\\
&= \lim_{h\downarrow 0} \mathbb{E}\left[X(t+h) - X(t) : \tau_m > t\right]\\
&= \lim_{h\downarrow 0} \{\mathbb{E}\left[X(t+h) - X(t) : \tau_m > t+h\right]\\
&\quad + \mathbb{E}\left[X(t+h) - X(t) : t < \tau_m \leqslant t+h\right]\}
\end{aligned}$$

$$\begin{aligned}&\leqslant \lim_{h\downarrow 0}\left\{\mathbb{E}\left[X(t+h)-X(t):\tau_m>t+h\right]\right.\\&\quad\left.+\mathbb{E}\left[\mathbb{E}\left(X(t+1)\mid\mathcal{F}_{t+h}\right)-X(t):t<\tau_m\leqslant t+h\right]\right\}\\&=\lim_{h\downarrow 0}\left\{\mathbb{E}\left[X(t+h)-X(t):\tau_m>t+h\right]\right.\\&\quad\left.+\mathbb{E}\left[X(t+1)-X(t):t<\tau_m\leqslant t+h\right]\right\}=0,\end{aligned}$$

by dominated convergence (twice: on $\{\tau_m > t+h\}$ we have $|X(t+h)-X(t)| \leqslant 2m$, $\mathbb{P}$-almost surely). Consequently

$$0\leqslant \mathbb{E}\left(X''(t)-X(t)\right)=\lim_{m\to\infty}\mathbb{E}\left(X''(t)-X(t):\tau_m>t\right)=0, \tag{5.238}$$

and hence $X''(t)=X(t)$, $\mathbb{P}$-almost surely. We also infer

$$\begin{aligned}\mathbb{E}(X(t))=\mathbb{E}(X''(t))&=\mathbb{E}\left(\lim_{h\downarrow 0}\mathbb{E}\left(X(t+h)\mid\mathcal{F}_t\right)-X(t)\right)+\mathbb{E}(X(t))\\&=\mathbb{E}\left(\lim_{h\downarrow 0}\mathbb{E}\left(\left(X(t+h)-X(t)\right)\mid\mathcal{F}_t\right)\right)+\mathbb{E}(X(t))\\&=\lim_{h\downarrow 0}\mathbb{E}\left(\mathbb{E}\left(\left(X(t+h)-X(t)\right)\mid\mathcal{F}_t\right)\right)+\mathbb{E}(X(t))\\&=\lim_{h\downarrow 0}\mathbb{E}\left(X(t+h)-X(t)\right)+\mathbb{E}(X(t))=\lim_{h\downarrow 0}\mathbb{E}\left(X(t+h)\right).\end{aligned}\tag{5.239}$$

This proves (5.236). In addition we write

$$f(t)=\mathbb{E}\left(X(t)-X(0)\right) \tag{5.240}$$

and we define the countable dense subset D of $[0,\infty)$ by

$$D=\{t\geqslant 0:t\in\mathbb{Q}\}\cup\{t\geqslant 0:f(t+)>f(t-)\}. \tag{5.241}$$

(Notice that the functions f is increasing.)

Let Λ_0 be any infinite subset of $\mathbb{N}$ and let $\{A_{\Lambda_0}(t):t\in D\}$ be a process such that for every infinite subset Λ of Λ_0 and $\mathbb{P}$-almost surely,

$$A_{\Lambda_0}(t)=\lim_{n\to\infty}\frac{1}{|\Lambda\cap[1,n]|}\sum\nolimits_{j\in\Lambda\cap[1,n]}A_j(t),\quad t\in D. \tag{5.242}$$

By Komlos' theorem (Theorem 5.129) and a diagonal procedure such a subset Λ_0 exists. We shall prove that for $t\in D$ the limit in (5.241) also exists in L^1-sense. In view of a theorem of Scheffé (Corollary 2.12.5 in Bauer [**15**], p. 105, it suffices to prove that

$$\mathbb{E}\left(A_{\Lambda_0}(t)\right)=\lim_{n\to\infty}\frac{1}{|\Lambda\cap[1,n]|}\sum\nolimits_{j\in\Lambda\cap[1,n]}\mathbb{E}\left(A_j(t)\right). \tag{5.243}$$

It is readily verified that

$$\mathbb{E}\left(A_j(t)\right)=\mathbb{E}\left(X\left(\frac{[2^jt]}{2^j}\right)\right)-\mathbb{E}\left(X(0)\right), \tag{5.244}$$

so that

$$\lim_{n\to\infty}\frac{1}{|\Lambda\cap[1,n]|}\sum\nolimits_{j\in\Lambda\cap[1,n]}\mathbb{E}\left(A_j(t)\right)=f(t+),\quad t\in D.) \tag{5.245}$$

On the other hand we have, by Fatou's lemma,

$$\mathbb{E}\left(A_{\Lambda_0}(t)\right) \leqslant \liminf_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} \mathbb{E}\left(A_j(t)\right) = f(t+). \tag{5.246}$$

In addition we have for $\lambda > 0$

$$\begin{aligned}\mathbb{E}\left(A_{\Lambda_0}(t)\right) &\geqslant \mathbb{E}\left(\limsup_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t)\right) \\ &\geqslant \mathbb{E}\left(\limsup_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j\left(\min\left(t,\tau_\lambda\right)-\right)\right),\end{aligned} \tag{5.247}$$

where τ_λ is the stopping time defined by

$$\tau_\lambda = \inf\left\{s > 0 : s \in D,\ \ \sup_{n\in\mathbb{N}} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(s) \geqslant \lambda\right\}. \tag{5.248}$$

Since

$$\frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j\left(\min\left(t,\tau_\lambda\right)-\right) \leqslant \lambda, \tag{5.249}$$

we infer from (5.247) and (5.235) that, for any $\lambda > 0$,

$$\begin{aligned}\mathbb{E}\left(A_{\Lambda_0}(t)\right) &\geqslant \limsup_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} \mathbb{E}\left(A_j\left(\min\left(t,\tau_\lambda\right)-\right)\right) \\ &\geqslant \mathbb{E}\left(X(t) : \tau_\lambda > t\right) + \mathbb{E}\left(X\left(\min(\tau_\lambda, t)\right) : \tau_\lambda \leqslant t\right) - \mathbb{E}\left(X(0)\right).\end{aligned} \tag{5.250}$$

Since $\tau_\lambda \uparrow \infty$, $\mathbb{P}$-almost surely, as λ tends to infinity, we infer from (5.250) together with the fact that the collection $\{X(\tau) : \tau \leqslant t,\ \tau \text{ stopping time}\}$ is uniformly integrable,

$$\mathbb{E}\left(A_{\Lambda_0}(t)\right) \geqslant \mathbb{E}\left(X(t) - X(0)\right) = f(t). \tag{5.251}$$

(In fact in (5.250) we first take the sum, then we write $\Omega = \{\tau_\lambda > t\} \cup \{\tau_\lambda \leqslant t\}$.) The right continuity of the submartingale $\{X(t) : t \geqslant 0\}$ together with (5.236) implies the equality

$$f(t) = f(t+). \tag{5.252}$$

Hence the equality in (5.243) now follows from (5.251), (5.252) and (5.246). So the limit in (5.241) is also an L^1-limit. Since the submartingale $\{X(t) : t \geqslant 0\}$ is continuous from the right we also deduce

$$\begin{aligned}\limsup_{j\to\infty} \mathbb{E}\left(\left|X_j(t) - X(t)\right|\right) &= \limsup_{j\to\infty} \mathbb{E}\left(X\left(\frac{\lceil 2^j t\rceil}{2^j}\right)\right) - \mathbb{E}\left(X(t)\right) \\ &= \lim_{s\downarrow t} f(s) - f(t) = 0.\end{aligned} \tag{5.253}$$

Hence the L^1-convergence in the equality

$$\begin{aligned}&\frac{1}{|\Lambda \cap [1,n]|} \sum_{j\in\Lambda\cap[1,n]} X_j(t) \\ &\quad= \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} M_j(t) + \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t)\end{aligned}$$

yields

$$X(t) = M_{\Lambda_0}(t) + A_{\Lambda_0}(t), \quad t \in D, \tag{5.254}$$

where the process $\{A_{\Lambda_0}(t) : t \in D\}$ is increasing and predictable. We shall extend (5.254) to all $t \geqslant 0$ and we shall prove that the process $\{A_{\Lambda_0}(t) : t \in D\}$ has right continuous extensions to all of $[0, \infty)$. In order to achieve this fix $t_0 \notin D$, $t_0 > 0$, and let s, t be arbitrary numbers in D with $0 < s < t_0 < t < \infty$. Then

$$\begin{aligned} A_{\Lambda_0}(s) &\leqslant \liminf_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(s) \\ &\leqslant \liminf_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t_0) \\ &\leqslant \limsup_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t) \leqslant A_{\Lambda_0}(t). \end{aligned} \tag{5.255}$$

From (5.255) it follows that

$$\begin{aligned} &\mathbb{E}\left(\limsup_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t_0) - \liminf_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t_0)\right) \\ &\quad \leqslant \mathbb{E}\left(A_{\Lambda_0}(t) - A_{\Lambda_0}(s)\right) = \mathbb{E}\left(X(t) - X(s)\right) = f(t) - f(s). \end{aligned} \tag{5.256}$$

So that

$$\begin{aligned} &\mathbb{E}\left(\limsup_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t_0) - \liminf_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t_0)\right) \\ &\quad \leqslant f(t_0+) - f(t_0-) = f(t_0) - f(t_0) = 0, \end{aligned} \tag{5.257}$$

since t_0 does not belong to D. Hence, for every $t_0 \geqslant 0$,

$$\begin{aligned} A_{\Lambda_0}(t_0) &= \limsup_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t_0) \\ &= \liminf_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t_0), \end{aligned} \tag{5.258}$$

$\mathbb{P}$-almost surely. In addition, as above we also have

$$\mathbb{E}\left(A_{\Lambda_0}(t)\right) = f(t), \quad t \geqslant 0,) \tag{5.259}$$

and hence

$$\mathbb{E}\left(A_{\Lambda_0}(t) - A_{\Lambda_0}(s)\right) = f(t) - f(s). \tag{5.260}$$

So that the process $\{A_{\Lambda_0}(t) : t \geqslant 0\}$ is almost surely right continuous. Again we have decomposition (5.254) for all $t \geqslant 0$. From (5.236) and (5.258) it follows that, $\mathbb{P}$-almost surely,

$$A_{\Lambda_0}(t_0) = \lim_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t_0-)$$

and consequently the process $\{A_{\Lambda_0}(t) : t \geqslant 0\}$ is predictable. □

The uniqueness of the Doob-Meyer decomposition does not depend on the (DL)-property. So the processes $\{M_{\Lambda_0}(t) : t \geqslant 0\}$ and $\{A_{\Lambda_0}(t) : t \geqslant 0\}$ do not depend on the particular choice of Λ_0. Henceforth we write

$$X(t) = M(t) + A(t), \qquad t \geqslant 0, \tag{5.261}$$

where $\{M(t) : t \geqslant 0\}$ is an $\mathcal{F}_t$-martingale and where $\{A(t) : t \geqslant 0\}$ is an increasing right continuous process which is predictable. Proposition 5.131 shows that the process $\{X(t) : t \geqslant 0\}$ must possess the (DL)-property. Let D_0 be the countable dense subset of $[0, \infty)$ given by

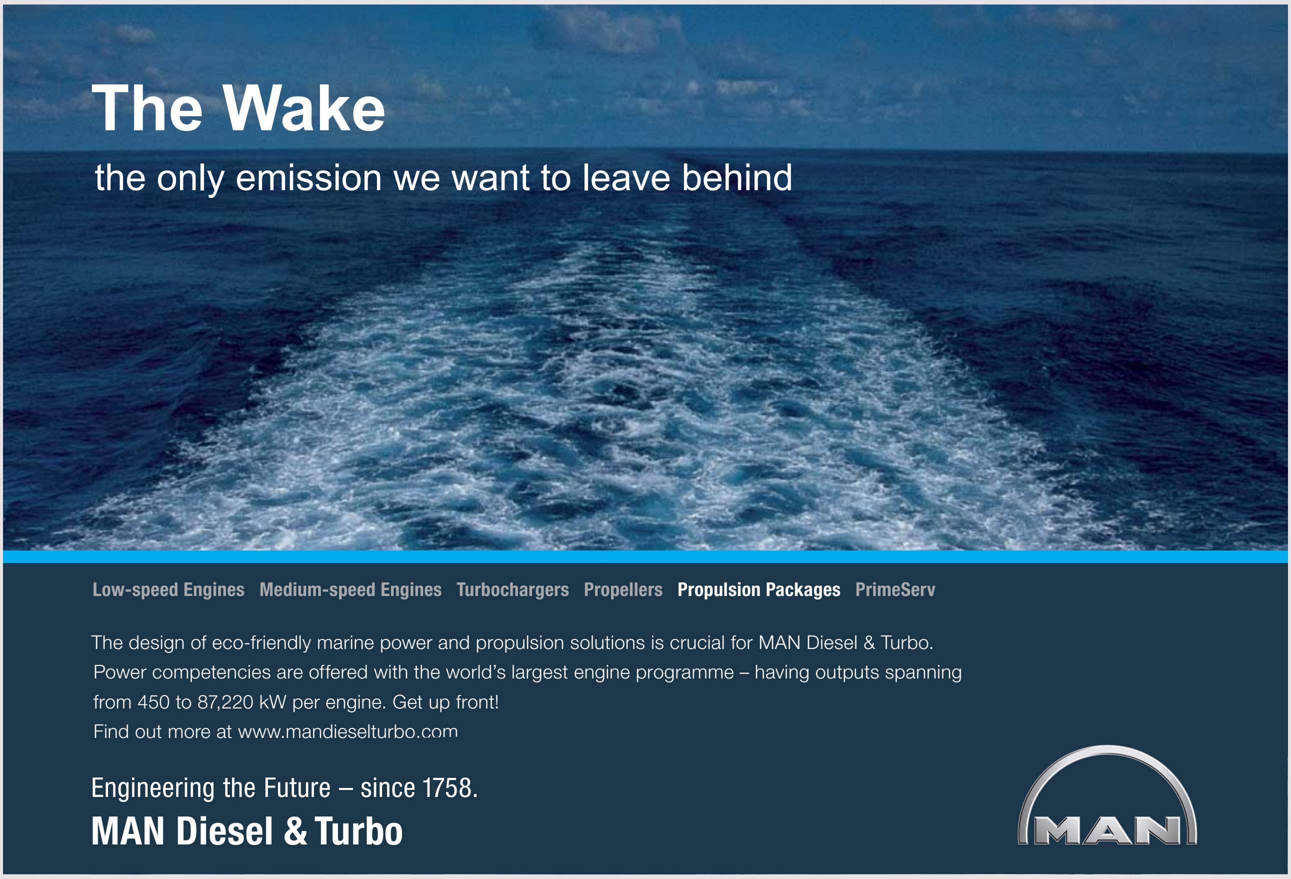
$$D_0 = \{t \in \mathbb{Q} : t \geqslant 0\} \cup \{t \geqslant 0 : f(t+) > f(t-)\} \tag{5.262}$$

and choose $\Lambda_0 \subseteq \mathbb{N}$, $|\Lambda_0| = \infty$, and the process $\{B(t) : t \in D_0\}$ in such a way that for every infinite subset Λ of Λ_0,

$$B(t) = \lim_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t), \quad t \in D_0. \tag{5.263}$$

Then, as in the case of $\{A_j(t) : j \in \mathbb{N}\}$ it follows that the convergence in (5.263) is an L^1-convergence as well. Again as above the convergence in (5.263) occurs for all $t \geqslant 0$. Consequently the process $\{X(t) - B(t) : t \geqslant 0\}$ is a martingale, because the processes $\{X_j(t) - A_j(t) : t \geqslant 0\}$, $j \in \mathbb{N}$, are martingales. Here

$$X_j(t) = \mathbb{E}\left[X\left(\frac{[2^j t]}{2^j}\right) \mid \mathcal{F}_t\right].$$



These remarks prove the following corollary.

5.133. Corollary. *Write a submartingale* $\{X(t) : t \geqslant 0\}$ *in the form* $X(t) = M(t) + A(t)$, $t \geqslant 0$, *where the process* $\{M(t) : t \geqslant 0\}$ *is a martingale and where* $\{A(t) : t \geqslant 0\}$ *is a right continuous increasing predictable process with* $A(0) = 0$. *Then there exists an infinite subset* Λ_0 *of* $\mathbb{N}$ *such that for every infinite subset* Λ *of* Λ_0 *and every* $t \geqslant 0$:

$$A(t) = \lim_{n\to\infty} \frac{1}{|\Lambda \cap [1,n]|} \sum\nolimits_{j\in\Lambda\cap[1,n]} A_j(t). \tag{5.264}$$

Here

$$A_j(t) = \sum\nolimits_{0\leqslant k<2^j t} \left(\mathbb{E}\left(X\left(\frac{k+1}{2^j}\right) \mid \mathcal{F}_{k2^{-j}} \right) - X\left(\frac{k}{2^j}\right) \right)$$

and the convergence in (5.264) is a $\mathbb{P}$-almost sure as well as an L^1-convergence. Of course the process $\{A(t) : t \geqslant 0\}$ does not depend on the particular choice of Λ_0 for which all the limits in (5.264) exist.

Next the uniqueness part of the Doob-Meyer decomposition will follow from Proposition 5.134.

5.134. Proposition. *Let* $Z = \{Z(t) : t \geqslant 0\}$ *be a bounded martingale and let* $A = \{A(t) : t \geqslant 0\}$ *and* $\{B(t) : t \geqslant 0\}$ *be adapted increasing processes such that* $B - A$ *is a martingale. Also suppose that* $\mathbb{E}(A(t)) < \infty$, *for* $t \geqslant 0$. *Then*

$$\begin{aligned} &\mathbb{E}\left[Z(t+)\left(B(t+) - A(t+)\right) - Z(0)\left(B(0) - A(0)\right)\right] \\ &= \mathbb{E}\left(\int_0^t \left(Z(s+) - Z(s-)\right) d(B-A)(s)\right). \end{aligned} \tag{5.265}$$

5.135. Remark. The integral $\int_0^t \left(Z(s+) - Z(s-)\right) d(B-A)(s)$ should be interpreted as follows:

$$\begin{aligned} &\int_0^t \left(Z(s+) - Z(s-)\right) d(B-A)(s) \\ &= \int_0^\infty \left(Z(s+) - Z(s-)\right) 1_{(0,t]}(s) dB(s) - \int_0^\infty \left(Z(s+) - Z(s-)\right) 1_{(0,t]}(s) dA(s). \end{aligned}$$

Proof of Proposition 5.134. Let $n \in \mathbb{N}$. Since Z is a martingale we have:

$$\begin{aligned} &\mathbb{E}\left[Z\left(\lceil 2^n t\rceil 2^{-n}\right)\left(B\left(\lceil 2^n t\rceil 2^{-n}\right) - A\left(\lceil 2^n t\rceil 2^{-n}\right)\right)\right] - \mathbb{E}\left[Z(0)\left(B(0) - A(0)\right)\right] \\ &= \mathbb{E}\left[\sum_{0\leqslant j<\lceil 2^n t\rceil} Z\left((j+1)2^{-n}\right)\right. \\ &\qquad \left.\left\{\left(B\left((j+1)2^{-n}\right) - B\left(j2^{-n}\right)\right) - \left(A\left((j+1)2^{-n}\right) - A\left(j2^{-n}\right)\right)\right\}\right]. \end{aligned}$$

Since $B - A$ is a martingale, it follows that:

$$\mathbb{E}\left[Z\left(\lceil 2^n t\rceil 2^{-n}\right)\left(B\left(\lceil 2^n t\rceil 2^{-n}\right) - B(0) - A\left(\lceil 2^n t\rceil 2^{-n}\right) + A(0)\right)\right]$$

$$
\begin{aligned}
&\quad - \mathbb{E}\left[Z(0)\left(B(0)-A(0)\right)\right] \\
&= \mathbb{E}\Big(\sum\nolimits_{0\leqslant j<\lceil 2^n t\rceil}\left(Z\left((j+1)2^{-n}\right)-Z\left(j2^{-n}\right)\right) \\
&\qquad \Big(\left(B\left((j+1)2^{-n}\right)-B\left(j2^{-n}\right)\right)-\left(A\left((j+1)2^{-n}\right)-A\left(j2^{-n}\right)\right)\Big)\Big).
\end{aligned}
$$

Put

$$
\begin{aligned}
Z_n^+(s) &= Z\left(\lceil 2^n s\rceil 2^{-n}\right) = \sum\nolimits_{j=0}^{\infty} Z\left((j+1)2^{-n}\right) 1_{(j2^{-n},(j+1)2^{-n}]}(s), \quad \text{and} \\
Z_n^-(s) &= Z\left(\left(\lceil 2^n s\rceil -1\right)2^{-n}\right) = \sum\nolimits_{j=0}^{\infty} Z\left(j2^{-n}\right) 1_{(j2^{-n},(j+1)2^{-n}]}(s).
\end{aligned}
$$

Then we obtain

$$
\begin{aligned}
&\mathbb{E}\left(Z\left(\lceil 2^n t\rceil 2^{-n}\right)\left(B\left(\lceil 2^n t\rceil 2^{-n}\right)-A\left(\lceil 2^n t\rceil 2^{-n}\right)\right)\right)-\mathbb{E}\left[Z(0)\left(B(0)-A(0)\right)\right] \\
&= \mathbb{E}\left(\int_0^{\infty}\left(Z_n^+(s)-Z_n^-(s)\right)1_{(0,\lceil 2^n t\rceil 2^{-n}]}(s)d\left(B(s)-A(s)\right)\right).
\end{aligned}
$$

So, upon letting n tend to infinity, Proposition 5.134 follows. □

5.136. PROPOSITION. *In addition to the hypotheses in Proposition 5.134, suppose that the martingale $B-A$ is predictable. Then $B(t+)=A(t+)$ $\mathbb{P}$-almost surely. So that, if $B-A$ is right-continuous, then $B=A$ $\mathbb{P}$-almost surely, provided $B(0)=A(0)=0$.*

PROOF. First we prove that $\mathbb{E}\left(\int_0^t\left(Z(s+)-Z(s-)\right)d(B-A)(s)\right)=0$. Here we shall employ the predictability of the process $B-A$. It suffices to prove that, for all $s>0$,

$$
\mathbb{E}\left(\left(Z(s+)-Z(s-)\right)\left(B(s+)-A(s+)-B(s-)+A(s-)\right)\right)=0. \tag{5.266}
$$

Since the predictable field on $\Omega\times[0,\infty)$ is generated by the collection $\{C\times(a,b]: C\in\mathcal{F}_a,\ 0\leqslant a<b\}$ it suffices to prove (5.266) for all $s\geqslant 0$ if $B-A$ is of the form

$$
B(s)-A(s)=1_C\times 1_{(a+\varepsilon,\infty)}(s), \quad C\in\mathcal{F}_a. \tag{5.267}
$$

So let C belong to $\mathcal{F}_a$ and let $B(s)-A(s)=1_C\times 1_{(a+\varepsilon,\infty)}(s)$. Then, for $s=a+\varepsilon$ (and $C\in\mathcal{F}_a$), we have by the martingale property of Z,

$$
\begin{aligned}
&\mathbb{E}\left(Z\left((a+\varepsilon)+\right)-Z\left((a+\varepsilon)-\right)1_C\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(Z(((a+\varepsilon)+)-Z\left((a+\varepsilon)-\right)\mid\mathcal{F}_a\right)1_C\right) \\
&= \mathbb{E}\left(\left(Z(a)-Z(a)\right)1_C\right)=0.
\end{aligned} \tag{5.268}
$$

Notice that, for $s=a+\varepsilon$,

$$
\begin{aligned}
&\mathbb{E}\left(\left(Z(s+)-Z(s-)\right)\left(B(s+)-A(s+)-B(s)+A(s)\right)\right) \\
&\qquad = \mathbb{E}\left(\left(Z\left((a+\varepsilon)+\right)-Z\left((a+\varepsilon)-\right)\right)1_C\right).
\end{aligned}
$$

From Proposition 5.134 it now follows that

$$
\mathbb{E}\left(Z(t+)\left(B(t+)-A(t+)\right)\right)=\mathbb{E}\left[Z(0)\left(B(0)-A(0)\right)\right]=0.
$$

Next, fix $t>0$ and define the martingale $Z(s)$ by

$$
Z(s)=\mathbb{E}\left(\frac{B(t+)-A(t+)}{|B(t+)-A(t+)|+1}\mid\mathcal{F}_s\right).
$$

Then

$$0 = \mathbb{E}\left(Z(t+)\left(B(t+) - A(t+)\right)\right) = \mathbb{E}\left(\frac{\left|B(t+) - A(t+)\right|^2}{\left|B(t+) - A(t+)\right| + 1}\right)$$

and hence $B(t+) = A(t+)$, $\mathbb{P}$-almost surely for all $t \geqslant 0$. It also follows that $B(t-) = A(t-)$, $\mathbb{P}$-almost surely for all $t > 0$. If the process $B - A$ is right continuous almost surely, we infer $B(t) = A(t)$, $t \geqslant 0$, $\mathbb{P}$-almost surely. This completes the proof of Proposition 5.136. ☐

As a special case the following result contains the uniqueness part of the Doob-Meyer decomposition theorem.

5.137. PROPOSITION. *Let A and B be increasing adapted processes. Suppose that $B - A$ is a predictable right continuous martingale. Then $B(t) = A(t) + B(0) - A(0)$, $\mathbb{P}$-almost surely.*

PROOF. This result is an immediate consequence of Proposition 5.134 and Proposition 5.136. ☐

5.138. COROLLARY. *There is only one way to write a semi-martingale Y in the form $Y = M + A$, where M is a (local) martingale and where A is a predictable right continuous process of finite variation locally with $A(0) = 0$.*

5.139. REMARK. An increasing, predictable right continuous real-valued process $\{A(t) : t \geqslant 0\}$, with $\mathbb{E}\left(A(t)\right) < \infty$ for $t \geqslant 0$, is called a *Meyer process.*

It is perhaps worthwhile to isolate the following result in the existence part of Doob-Meyer decomposition theorem: notation is that of the proof of Theorem 5.130. We also use $A_\Lambda(t) = \dfrac{1}{|\Lambda|}\sum_{n\in\Lambda} A_n(t)$, for a finite subset Λ of $\mathbb{N}$.

5.140. Theorem. *Let $\{X(t) : t \geqslant 0\}$ be a right continuous submartingale of class (DL). For every infinite subset Λ_0 of $\mathbb{N}$ there exists an infinite subset Λ of Λ_0 such that for every further infinite subset Λ' of Λ, the limit*

$$A_\Lambda(t) := \lim_{N\to\infty} A_{\Lambda'\cap[1,N]}(t), \quad \text{exists } \mathbb{P}\text{-almost surely},$$

and does not depend on the choice of Λ'. Moreover, since we are dealing with (DL)-submartingales, $\lim_{N\to\infty} \mathbb{E}\left[\left|A_\Lambda(t) - A_{\Lambda'\cap[1,N]}(t)\right|\right] = 0$. In addition, the process $\{A_\Lambda(t) : t \geqslant 0\}$ is predictable and right continuous.

Proof. Write

$$Q' = \{t_\ell : \ell \in \mathbb{N}\} = \{t \geqslant 0 : \mathbb{E}(X(t+)) > \mathbb{E}(X(t-))\} \bigcup (\mathbb{Q} \cap [0,\infty)).$$

Define the measure μ on Q' by

$$\mu(I) = \sum_{\ell\in I} \frac{1}{2^\ell} \frac{1}{1 + \mathbb{E}(X(t_\ell) - X(0))}.$$

Let Λ_0 be an infinite subset of $\mathbb{N}$. Komlos' theorem, applied to the sequence $\{A_n(t_\ell) : \ell \in \mathbb{N}\}_{n\in\mathbb{N}}$ on the measure space $\{\mathbb{N} \times \Omega, \mathcal{P}(\mathbb{N}) \otimes \mathcal{F}, \mu \otimes \mathbb{P}\}$ applies to the effect that there exists an infinite subset Λ of Λ_0 such that for every further infinite subset Λ' of Λ, $A_\Lambda(t) = \lim_{N\to\infty} A_{\Lambda'\cap[1,N]}(t_\ell)$ exists for $\ell = 1, 2, \dots$ and does not depend on the particular choice of Λ'. In addition,

$$\lim_{N\to\infty} \mathbb{E}\left(\left|A_\Lambda(t_\ell) - A_{\Lambda'\cap[1,N]}(t_\ell)\right|\right) = 0.$$

Next let $t \geqslant 0$ be arbitrary with $\mathbb{E}(X(t)) = \mathbb{E}(X(t-)) = \mathbb{E}(X(t+))$ and let $\Lambda' \subseteq \Lambda_0$, Λ' infinitely large. For $t' < t < t''$, t', t'' in Q', we have

$$\begin{aligned}\mathbb{E}\left(\limsup_{N\to\infty} A_{\Lambda'\cap[1,N]}(t)\right) &\leqslant \mathbb{E}\left(\limsup_{N\to\infty} A_{\Lambda'\cap[1,N]}(t'')\right)\\ &= \mathbb{E}\left(\liminf_{N\to\infty} A_{\Lambda'\cap[1,N]}(t'')\right) \leqslant \mathbb{E}(X(t'') - X(0)).\end{aligned}$$

Similarly we have

$$\begin{aligned}\mathbb{E}\left(\liminf_{N\to\infty} A_{\Lambda'\cap[1,N]}(t)\right) &\geqslant \mathbb{E}\left(\liminf_{N\to\infty} A_{\Lambda'\cap[1,N]}(t')\right)\\ &= \mathbb{E}\left(\limsup_{N\to\infty} A_{\Lambda'\cap[1,N]}(t')\right) \geqslant \mathbb{E}(X(t') - X(0)).\end{aligned}$$

Since $\mathbb{E}(X(t+)) = \mathbb{E}(X(t-))$, it follows that the limit

$$A_{\Lambda'}(t) := \lim_{N\to\infty} A_{\Lambda'\cap[1,N]}(t)$$

exists $\mathbb{P}$-almost surely. Consequently, the limits

$$A_\Lambda(t) := \lim_{N\to\infty} A_{\Lambda'\cap[1,N]}(t), \quad t \geqslant 0,$$

all exist $\mathbb{P}$-almost surely and $\lim_{N\to\infty} \mathbb{E}\left(\left|A_\Lambda(t) - A_{\Lambda'\cap[1,N]}(t)\right|\right) = 0$. Finally we shall prove that the process $\{A_\Lambda(t) : t \geqslant 0\}$ is right continuous. Fix $t_0 \geqslant 0$ and let $t > t_0$. Then

$$\mathbb{E}(A_\Lambda(t) - A_\Lambda(t_0)) = \mathbb{E}(X(t) - X(t_0)) \geqslant 0.$$

Since $t \mapsto \mathbb{E}(X(t))$ is right continuous we infer that $\lim_{t\downarrow t_0} \mathbb{E}\left(A_\Lambda(t) - A_\Lambda(t_0)\right) = 0$. It follows that, $\mathbb{P}$-almost surely, $\lim_{t\downarrow t_0} A_\Lambda(t) = A_\Lambda(t_0)$.

This completes the proof of Theorem 5.140. ☐

5.141. Corollary. *Let $X = M + A$ be the Doob-Meyer decomposition of a submartingale into a martingale and an increasing right continuous predictable process A. Then, for an appropriate sequence $(n_\ell : \ell \in \mathbb{N})$ in $\mathbb{N}$,*

$$A(t) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \sum_{j=0}^{\infty} \left(\mathbb{E}\left(A\left((j+1)2^{-n_k}\right) \mid \mathcal{F}_{j2^{-n_k}}\right) - A\left(j2^{-n_k}\right)\right) 1_{(j2^{-n_k},\infty)}(t).$$

This limit is an $\mathbb{P}$-almost sure limit as well as a limit in $L^1(\Omega, \mathcal{F}, \mathbb{P})$.

Proof. A combination of the existence and uniqueness of the Doob-Meyer decomposition yields the desired result. Notice that by Proposition 5.131 a process of the form $M+A$, where M is a martingale and where A is an increasing adapted process in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ is of class (DL): see (5.230) and (5.231). So the proof of Corollary 5.141 is complete now. ☐

Another corollary is the following one.

5.142. Corollary. *Let $\{X(t) : t \geqslant 0\}$ be a right continuous submartingale of class (DL) with left limits. Fix $t_0 > 0$ and let $\{\tau_\ell : \ell \in \mathbb{N}\}$ be sequence of stopping times which increases to the fixed time t_0. Suppose $\tau_\ell < t_0$, $\mathbb{P}$-almost surely, for all $\ell \in \mathbb{N}$. Then $\mathbb{E}\left(|X(t_0-)|\right) < \infty$ and $\lim_{\ell\to\infty} \mathbb{E}\left(\left|X\left(\tau_\ell\right) - X\left(t_0-\right)\right|\right) = 0$. In addition, $\lim_{h\downarrow 0} \mathbb{E}\left(\left|X\left(t_0+h\right) - X\left(t_0\right)\right|\right) = 0$.*

The following result also follows from our discussion.

5.143. Corollary. *Let $\{X(t) : t \geqslant 0\}$ be a submartingale. If the function $t \mapsto \mathbb{E}\left(X(t)\right)$ is $\mathbb{P}$-almost surely continuous, then the process $\{A(t) : t \geqslant 0\}$ is $\mathbb{P}$-almost surely continuous as well.*

5.144. Remark. Several people have reformulated and extended Komlos' result as a principle of subsequences, e.g. see Chatterji [**40**]. Others have treated an infinite dimensional version, e.g. see Balder [**12**]. In [**124**], Exercise 3, p. 103 the authors give an example of a submartingale which is not of class (DL). In fact Métivier and Pellaumail give the following example. Let Ω be the interval $[0,1]$ with Lebesgue measure and let $0 = t_0 < t_1 < \ldots < t_n < \ldots < 1$ be a sequence such that $\lim_{n\to\infty} t_n = 1$. Define the process X by

$$X(t,\omega) = -\sum\nolimits_{n=1}^{\infty} 2^n 1_{[t_{n-1},t_n)}(t) 1_{(1-2^{-n},1]}(\omega), \quad \omega \in [0,1], \quad t \geqslant 0.$$

Then X is a submartingale, X is not of class (DL) and X is a martingale on the interval $[0,1)$. If $t_{n-1} \leqslant t < t_n$, we write $\mathcal{F}_t$ for the σ-field generated by $\{(j-1)2^{-n}, j2^n] : 1 \leqslant j \leqslant 2^n\}$. If $t \geqslant 1$, then $\mathcal{F}_t$ is the Borel field of $[0,1]$.

5.145. Definition. Let $\{Y(t) : t \geqslant 0\}$ be a martingale in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Then $\{|Y(t)|^2 : t \geqslant 0\}$ is a submartingale of class (DL). So by Theorem 5.130 there exists a unique martingale $\{M(t) : t \geqslant 0\}$ with $M(0) = |Y(0)|^2$ and an increasing predictable right-continuous process $\{\langle Y\rangle(t) : t \geqslant 0\}$ in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$|Y(t)|^2 = M(t) + \langle Y\rangle(t), \quad \mathbb{P}\text{-almost surely}.$$

The process $\{\langle Y\rangle(t) : t \geqslant 0\}$ is called the (quadratic) *variation* or *variance* process of $\{Y(t) : t \geqslant 0\}$.

5.146. Example. Let $\{B(t) : t \geqslant 0\}$ be ν-dimensional Brownian motion. Then the process $\{t \mapsto \nu t : t \geqslant 0\}$ is the corresponding quadratic variation process.

5.147. Example. let $t \mapsto \int_0^t F_1(s)\,dB(s)$ and $t \mapsto \int_0^t F_2(s)\,dB(s)$ be two local martingales. Then the process $t \mapsto \int_0^t F_1(s)F_2(s)\,ds$ is the corresponding covariation process.

Subjects for further research and presentations

The following topics may be of interest for a presentation and/or further research:

1. Certain pseudo-differential operators of order less than or equal to 2 can be put into correspondence with space-homogeneous or non-space-homogeneous Markov processes. A detailed exposition can be found in Jacob [**83, 84, 85**].
2. Viscosity solutions to partial differential equations. The standard reference for this subject is Crandall, Ishii, and Lions [**47**]. This topic can also be treated in the context of Backward Stochastic Differential Equations (BSDEs): see, *e.g.*, Pardoux [**142**].
3. Elliptic differential operators of second order (and Markov processes); see, *e.g.*, Øksendael [**138**]. A recent interesting book on stochastic calculus is Stroock [**170**].
4. Parabolic differential operators (of second order and Markov processes). An interesting article in this context is Bossy and Champagnat [**29**]. The abstract of this paper reads: “We present the main concepts of the theory of Markov processes: transition semigroups, Feller processes, infinitesimal generator, Kolmogorov’s backward and forward equations, and Feller diffusion. We also give several classical examples including stochastic differential equations (SDEs) and backward stochastic differential equations (BSDEs) and describe the links between Markov processes and parabolic partial differential equations (PDEs). In particular, we state the Feynman-Kac formula for linear PDEs and BSDEs, and we give some examples of the correspondence between stochastic control problems and Hamilton-Jacobi-Bellman (HJB) equations and between optimal stopping problems and variational inequalities. Several examples of financial applications are given to illustrate each of these results, including European options, Asian options, and American put options.”
5. Solutions to stochastic differential equations and the corresponding second order differential equation (of parabolic type) satisfied by the one-dimensional distributions. For stochastic differential equations with singularities see, *e.g.*, [**44**]. For classical partial differential equations, see, *e.g.*, [**186**] and the references therein.
6. Backward stochastic differential equations and their viscosity solutions; see, *e.g.* Pardoux [**142**], Van Casteren [**185**], Boufoussi and Van Casteren [**30, 31**], Boufoussi, Van Casteren and Mhardy [**32**]. For more recent work on backward stochastic differential equations, the reader is referred to Pardoux and Răşcanu [**143**], and Zhang [**198**].
7. Heat equation on a Riemannian manifold. A relevant book in this context is [**78**]. For connections with stochastic differential equations on manifolds see, *e.g.*, Elworthy [**66, 67**].
8. Oscillatory integrals and related path integrals. There is a lot of literature on this subject. Nice papers on this topic are Albeverio and

Mazzucchi [**1, 2**]. Interesting books are, *e.g.*, Mazzucchi [**123**], Johnson and Lapidus [**86**], and Kleinert [**93**].

(9) Malliavin calculus, or stochastic calculus of variations, and applications to regularity properties of integral kernels. For details see *e.g.* Nualart [**135, 136**]. Other references which contain results on and applications of Malliavin calculus include: Cruzeiro and Malliavin [**49**], Stroock [**164, 165, 166**], Cruzeiro and Zambrini [**50**], [**51**]. Of course the original work by Malliavin should not be forgotten: [**119**]. The book by Bismut [**23**] combines Malliavin calculus with the theory of large deviations. For a discussion on Malliavin calculus in relation to Lévy processes see, *e.g.*, Osswald [**140**]. A rather elementary approach to Malliavin calculus can be found in Friz [**74**]. For application to stochastic differential equations see, *e.g.*, Takeuchi [**173**]. For applications of Malliavin calculus to operator semigroups see, *e.g.*, Léandre [**108, 109**]. For Malliavin calculus without probability theory see [**107**].

Technical training on
WHAT you need, *WHEN* you need it

At IDC Technologies we can tailor our technical and engineering training workshops to suit your needs. We have extensive experience in training technical and engineering staff and have trained people in organisations such as General Motors, Shell, Siemens, BHP and Honeywell to name a few.

Our onsite training is cost effective, convenient and completely customisable to the technical and engineering areas you want covered. Our workshops are all comprehensive hands-on learning experiences with ample time given to practical sessions and demonstrations. We communicate well to ensure that workshop content and timing match the knowledge, skills, and abilities of the participants.

We run onsite training all year round and hold the workshops on your premises or a venue of your choice for your convenience.

For a no obligation proposal, contact us today at training@idc-online.com or visit our website for more information: www.idc-online.com/onsite/

OIL & GAS ENGINEERING

ELECTRONICS

AUTOMATION & PROCESS CONTROL

MECHANICAL ENGINEERING

INDUSTRIAL DATA COMMS

ELECTRICAL POWER

Phone: **+61 8 9321 1702**
Email: **training@idc-online.com**
Website: **www.idc-online.com**

IDC TECHNOLOGIES

Download free eBooks at bookboon.com

Click on the ad to read more

(10) Books and papers with literature on financial mathematics include: León, Solé, Utzet, and Vives [**112**], Nualart and Schoutens [**137**], Malliavin and Thalmaier [**120**], Karatsas and Shreve [**89**], Gulisashvili [**79**], El Karoui and Mazliak [**65**], El Karoui, Pardoux and Quenez [**63**], Lim [**114**]. Other references include Zhang and Zhou (editors) [**199**] and Tsoi, Nualart and Yin [**176**].

(11) Another interesting subject is "Ergodic theory" and, correspondingly, invariant measures. We mention some references: Krengel [**99**], Karlin and Taylor [**90**], Meyn and Tweedie [**125**], Eisner and Nagel [**62**], Van Casteren [**184**], Seidler [**155**], Goldys [**77**], [**151**].

(12) Central limit theorems and related results are also relevant. Again we mention some references: Bhattaraya and Waymire [**20**], Nourdin and Peccati [**131**], Barbour and Chen [**13**], Berckmoes, Lowen and Van Casteren [**16, 17, 18, 19**], Tao [**175**], Stein [**161, 162**], Chen, Goldstein and Shao [**41**], Barbour and Hall [**14**].

(13) Investigate Markov processes with a Polish space as state space: see, *e.g.*, Sharpe [**156**], Swart and Winter [**172**], Van Casteren [**184**], Bovier [**33**].

(14) Discuss and make a careful study of the Skorohod space as described in Remark 3.40. Try to include applications to convergence properties of stochastic processes.

(15) Discuss stochastic analysis in the infinite-dimensional context. A nice and relevant survey paper is [**190**] written by van Neerven, Veraar and Weis. A simplified version in Dutch is authored by van Neerven: see [**189**].

(16) Make a (further) study of models in financial mathematics, like the stochastic volatility model (Heston model), stochastic interest rate model (Vasicek model) and others. Some details can be found in Klebaner [**92**], in Cont and Tankov [**46**], Gulisashvili [**79**] and others. The books [**52**] and [**72**] also contains much interesting information. A book with numerical applications is [**200**].

(17) Applications of stochastic processes in cell biology or population dynamics: references include [**36**, cell biology] and [**5**, population dynamics]. A Ph.D. thesis with interesting material is Ferreira [**70**]. It contains a discussion on the stochastic Lotka-Volterra equation which serves as a predator-prey model.

(18) An interesting topic is a combination of functional analysis, operator theory and stochastic processes with applications in among others data processing. A main theorem is the Karhunen-Loève expansion theorem. A relevant book in this context is [**141**] by Papadopoulos and Giovanis. Furthermore there are at least two Ph.D. theses with a theoretical background and applications of these expansions: one by Wang [**191**] and another by Giambartolomei [**76**].

CHAPTER 6

Advanced stochastic processes: a summary of the lectures

0.1. Abstract. The aim of this work is to present some of the main topics which will be explained during the lectures to come. In particular, we will explain the following notions: Brownian motion as a Gaussian process with values in $\mathbb{R}^d$, Brownian motion as a Markov process, and last but not least Brownian motion as a martingale. In order to achieve this we will need the concept of conditional expectation on sub-σ-fields. Since we also want to discuss the strong Markov property, and, more generally, stopped processes, we will introduce stopping times. In addition, we will pay attention to stochastic differential equations (SDEs), and to a lesser extent backward stochastic differential equations (BSDEs). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space endowed with a *filtration* $(\mathcal{F}_t)_{t\geqslant 0}$. This means that the latter collection consists of σ-fields contained in $\mathcal{F}$ and that $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$ for $0 \leqslant t_1 < t_2$. During this course entitled "Advanced stochastic processes" our attention will be focused on adapted processes $\{X(t) : t \geqslant 0\}$ which are called semi-martingales. This means that, for every $t > 0$, $X(t) : \Omega \to \mathbb{R}$ is $\mathcal{F}_t$-$\mathcal{B}_{\mathbb{R}}$-measurable, and that $X(t)$ can be written in the form $X(t) = A(t) + M(t)$. Here, the process $t \mapsto A(t)$ is a real-valued process of bounded variation (i.e. $A(t)$ can be written as a difference of two non-decreasing processes), and $t \mapsto M(t)$ is a martingale, that is $\mathbb{E}\left[M(t_2) \,\middle|\, \mathcal{F}_{t_1}\right] = M(t_1)$, $\mathbb{P}$-almost surely, for all $0 \leqslant t_1 < t_2$. Extensions to other state spaces like $\mathbb{R}^d$ can be made.

Introduction

During the lectures which lie at the basis of this course we will explain, in mathematical details, which concepts and notions are involved and relevant to understand the ideas around stochastic processes. As a leading example, and an important process we will consider Brownian motion and related processes which include Ornstein-Uhlenbeck process, Brownian bridge, Itô process. Several of these processes are Gaussian processes and, in the appropriate context, semi-martingales and/or Markov processes.

However, before we go into the more technical details, we will describe, very briefly, possible contents of papers which can be presented as part of the exam:

1. Martingale representation theorem. An interesting reference is [**147**].
2. Stationary or invariant measure and its relation with ergodic theory. This topic is contained in an interesting book [**126**].

(3) Strong law of large numbers (SLLN), weak law of large numbers (WLLN) and connections with renewal theory. For an account on this see e.g. [**163**].
(4) Central limit theorem using Stein's method, or, more generally, a version of the functional central limit theorem (FCLT). In the context of FCLT the limiting process may coincide with Brownian motion. The reader may consult several papers among which [**111**], [**130**], [**132**], [**18**] and [**19**].
(5) Black-Sholes-Merton model for option pricing: see e.g. [**127**].
(6) Poisson processes, Lévy processes, Markov chains and their connection with invariant measures: see e.g. [**163**].
(7) Stochastic differential equations: concrete examples: see e.g. [**139**].
(8) Ornstein-Uhlenbeck processes and stochastic interest rates: Vasicek model. For example, see [**134**].
(9) Discrete models and their limits for option pricing. For example, see [**157**]
(10) Stochastic volatility models like the Heston model or implied volatility: see e.g. [**79**].
(11) Itô calculus for models in financial mathematics: see e.g. [**157**].
(12) Singular stochastic differential equations: see e.g. [**44**].
(13) Backward stochastic differential equations (BSDEs) and hedging models: a combination a (forward) stochastic differential equations and BSDEs: see e.g. [**142**] and [**65**]. Another interesting source of information is [**48**]. A recent book is [**143**].
(14) Stochastic Partial differential equations: see e.g. [**80**].
(15) Applications of stochastic processes in cell biology or population dynamics: references include [**36**, cell biology] and [**5**, population dynamics].
(16) A combination of analysis, geometry and Markov process theory is another possibility. There are relations with the spectral gap, logarithmic inequalities, Poincaré inequalities iso-perimetric inequalities and others. The interested reader can find details in, e.g., [**9, 110, 26, 11, 10, 69**].

Anyway the idea is that after this course the students should be able to understand and work with the concepts which are employed in the above topics. In the sequel we will be more precise about the kind of techniques and processes we have in mind.

1. Brownian motion as a Gaussian process

In this section we describe Brownian motion as a Gaussian process. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $X = (X_1, \dots, X_N) : \Omega \to \mathbb{R}^N$, with $N \in \mathbb{N}$, be a random or stochastic or random vector. This vector is called Gaussian or (multivariate) normal provided its distribution only depends on its expectation and its covariance. More precisely, if its characteristic function is given by

$$\mathbb{E}\left[e^{-i\sum_{j=1}^N X_j\xi_j}\right] = e^{-i\sum_{j=1}^N \mu_j\xi_j - \frac{1}{2}\sum_{j_1,\,j_2=1}^N \sigma_{j_1,j_2}\xi_{j_1}\xi_{j_2}}, \tag{6.1}$$

where $(\mu_1, \dots, \mu_N)$ is a (deterministic) vector in $\mathbb{R}^N$, and where the matrix

$$\left(\sigma_{j_1,j_2}\right)_{1 \leqslant j_1,\, j_2 \leqslant N}$$

is a symmetric $N \times N$-matrix. It then follows that, for $1 \leqslant j \leqslant N$, X_j belongs to $L^2(\Omega, \mathcal{F}, \mathbb{P}) \subset L^1(\Omega, \mathcal{F}, \mathbb{P})$, that $\mu_j = \mathbb{E}\left[X_j\right]$, and that $\sigma_{j_1,j_2} = \operatorname{cov}\left(X_{j_1}, X_{j_2}\right) = \mathbb{E}\left[X_{j_1}X_{j_2}\right] - \mathbb{E}\left[X_{j_1}\right]\mathbb{E}\left[X_{j_2}\right]$, $1 \leqslant j_1,\, j_2 \leqslant N$. Since the characteristic function of a random vector determines its distribution, the equality in (6.1) entails that the distribution of a Gaussian vector only depends on its expectation and its covariance matrix.

As above, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let, for $t \geqslant 0$, $W(t) : \Omega \to \mathbb{R}^d$ be a random vector with the following properties:

(1) The function $t \mapsto W(t)$ is $\mathbb{P}$-almost surely continuous. This means that there exists an event $\Omega' \subset \Omega$ of full $\mathbb{P}$-measure, i.e. $\mathbb{P}[\Omega'] = 1$, such that the function $t \mapsto W(t)(\omega) = X(t, \omega)$ is continuous for all $\omega \in \Omega'$.

(2) For every finite choice $(t_1, \dots, t_n) \in [0, \infty)^n$ the vector

$$(W(t_1), \dots, W(t_n))$$

is a Gaussian (or a multi-variate normally distributed) vector in $\left(\mathbb{R}^d\right)^n = \underbrace{\mathbb{R}^d \times \cdots \times \mathbb{R}^d}_{n \text{ times}}$ such that $\mathbb{E}[W(t_k)] = x^k = \left(x_1^k, \dots, x_d^k\right) \in \mathbb{R}^d$, and

$$\mathbb{E}\left[\left(W_{j_1}(t_{k_1}) - x_{j_1}^{k_1}\right)\left(W_{j_2}(t_{k_2}) - x_{j_2}^{k_2}\right)\right] = \min(t_{k_1}, t_{k_2})\, \delta_{j_1, j_2},$$

for $1 \leqslant j_1,\ j_2 \leqslant d$, and $1 \leqslant k_1,\ k_2 \leqslant n$.

It was Einstein ([**61**]) who posed the problem whether or not the properties in (1) and (2) are compatible. Indeed, it was proved by Wiener in 1923 that they indeed are: see [**192**], and for more details see Mazliak [**122**]. We observe that

$$\begin{aligned}
&\mathbb{E}\left[\left\langle W(t_{k_1}) - x^{k_1}, \xi^{k_1}\right\rangle \left\langle W(t_{k_2}) - x^{k_2}, \xi^{k_2}\right\rangle\right] \\
&= \mathbb{E}\left[\sum_{j_1, j_2 = 1}^{d} \left(W_{j_1}(t_{k_1}) - x_{j_1}^{k_1}\right) \xi_{j_1}^{k_1} \times \left(W_{j_2}(t_{k_2}) - x_{j_2}^{k_2}\right) \xi_{j_2}^{k_2}\right] \\
&= \sum_{j_1, j_2 = 1}^{d} \mathbb{E}\left[\left(W_{j_1}(t_{k_1}) - x_{j_1}^{k_1}\right) \times \left(W_{j_2}(t_{k_2}) - x_{j_2}^{k_2}\right)\right] \xi_{j_1}^{k_1} \xi_{j_2}^{k_2} \\
&= \sum_{j=1}^{d} \min(t_{k_1}, t_{k_2})\, \xi_j^{k_1} \xi_j^{k_2} = \min(t_{k_1}, t_{k_2}) \left\langle \xi^{k_1}, \xi^{k_2}\right\rangle,
\end{aligned} \tag{6.2}$$

and so the characteristic function of the Gaussian vector

$$(W(t_1), \dots, W(t_n)) \in \left(\mathbb{R}^d\right)^n$$

is given by

$$\mathbb{E}\left[e^{-i\sum_{k=1}^{n}\left\langle W(t_k), \xi^k\right\rangle}\right] = e^{-i\sum_{k=1}^{n}\left\langle x^k, \xi^k\right\rangle - \frac{1}{2}\sum_{k_1, k_2 = 1}^{n} \min\left(t_{k_1}, t_{k_2}\right)\left\langle \xi^{k_1}, \xi^{k_2}\right\rangle}, \tag{6.3}$$

for all $0 \leqslant t_1, \dots, t_n < \infty$ and all $(\xi^1, \dots, \xi^n) \in \left(\mathbb{R}^d\right)^n$. In most cases it is assumed that $x^1 = \cdots = x^n \in \mathbb{R}^d$. In the latter case the process $t \mapsto W(t)$ which starts at x. In particular when $x = 0$ our Brownian motion, or Wiener process, $t \mapsto W(t)$ starts at 0. More details on Gaussian vectors can be found in the main text; see, e.g., Section 4 in Chapter 2, Propositions 3.5 and 3.6 in Chapter 3.

2. Brownian motion as a Markov process

In this section we will give a too short survey of Brownian motion viewed as a Markov process. We will assume that the Brownian motion $t \mapsto W(t)$ can start at any position (initial state) $x \in \mathbb{R}^d$. We will consider the following quadruple

$$\left\{ (\Omega, \mathcal{F}, \mathbb{P}_x)_{x \in \mathbb{R}^d}, (W(t), t \geqslant 0), (\vartheta_t, t \geqslant 0), \left(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d}\right) \right\}. \tag{6.4}$$

In (6.4) for every $x \in \mathbb{R}^d$ the triple $(\Omega, \mathcal{F}, \mathbb{P}_x)$ denotes a probability space. The σ-field $\mathcal{F}$ contains the union of an increasing family of σ-fields $(\mathcal{F}_t)_{t \geqslant 0}$; i.e. $0 \leqslant t_1 \leqslant t_2$ implies $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2} \subset \mathcal{F}$. Moreover it is assumed that $\mathcal{F}$ is the smallest σ-field containing all the σ-fields $\mathcal{F}_t$, $t \geqslant 0$. In addition, for every $t \geqslant 0$ the random vector $W(t) : \Omega \to \mathbb{R}^d$ is supposed to be $\mathcal{F}_t$-$\mathcal{B}_{\mathbb{R}^d}$-measurable. In other words the process $t \mapsto W(t)$ is adapted to the filtration $(\mathcal{F}_t)_{t \geqslant 0}$. Often it is assumed that $\mathcal{F}_t$ coincides with, or is some completion of the smallest σ-field which makes all variables $W(s) : \Omega \to \mathbb{R}^d$, $0 \leqslant s \leqslant t$, $\mathcal{F}_t$-$\mathcal{B}_{\mathbb{R}^d}$-measurable. In other words, we often have $\mathcal{F}_t = \sigma\{W(s) : 0 \leqslant s \leqslant t\}$ or some completion of $\sigma\{W(s) : 0 \leqslant s \leqslant t\}$. For example, it is often assumed that $\mathcal{F}_t$ satisfies the standard assumptions (SDA). This means that all $\mathbb{P}_x$-null sets are contained in $\mathcal{F}_0$. Then it follows that $\mathcal{F}_t$ contains these null sets as well, because $\mathcal{F}_t$ contains $\mathcal{F}_0$. The σ-field $\mathcal{F}_t$ can be identified with the information which one has at time t. If t is considered as the present time, then $\mathcal{F}_t$ is the information from the past. The *time translation* or *time shift* operators $\vartheta_s : \Omega \to \Omega$, $s \geqslant 0$, have the property that $W(t) \circ \vartheta_s = W(t+s)$, $\mathbb{P}_x$-almost surely, for all s, $t \geqslant 0$. In other words, if $f_j : \mathbb{R}^d \to \mathbb{C}$ are bounded Borel measurable functions on $\mathbb{R}^d$, and if $0 \leqslant t_1, \dots, t_n$ and $t > 0$ represent times, then

$$\prod_{j=1}^n f_j\left(W\left(t_j\right)\right) \circ \vartheta_t = \prod_{j=1}^n f_j\left(W\left(t_j + t\right)\right), \tag{6.5}$$

and hence, with $Y = \prod_{j=1}^n f_j\left(W\left(t_j\right)\right)$ the random variable $Y \circ \vartheta_t$ is measurable with respect to the σ-field $\mathcal{F}^t := \sigma\left(W(s) : s \geqslant t\right)$. The σ-field $\mathcal{F}^t$ is called the information from the future. Next let $Y : \Omega \to \mathbb{C}$ be a bounded random variable. This means that the bounded variable Y is $\mathcal{F}$-$\mathcal{B}_{\mathbb{C}}$-measurable. Then Brownian motion has the Markov property, or is a Markov process, provided the following identity holds $\mathbb{P}_x$-almost surely for all bounded random variables Y with values in $\mathbb{R}$ or $\mathbb{C}$:

$$\mathbb{E}_x\left[Y \circ \vartheta_t \,\middle|\, \mathcal{F}_t\right] = \mathbb{E}_{W(t)}\left[Y\right], \quad \text{for all } t \geqslant 0 \text{ and for all } x \in \mathbb{R}^d. \tag{6.6}$$

It requires the monotone class theorem that the Markov property of the Brownian motion as defined in (6.6) is equivalent to the following assertion. The Markov property of Brownian motion is equivalent to the following claim. For every real or complex valued bounded continuous function f defined on $\mathbb{R}^d$, and for every $s \geqslant 0$ the equality

$$\mathbb{E}_x\left[Y \circ \vartheta_t \,\middle|\, \mathcal{F}_t\right] = \mathbb{E}_{W(t)}\left[Y\right], \quad \text{holds for all } t \geqslant 0 \text{ and for all } x \in \mathbb{R}^d. \tag{6.7}$$

Here Y is of the form $Y = f\left(W(s)\right)$. Of course, for a variable Y of the form $Y = f\left(W(s)\right)$, $Y \circ \vartheta_t = f\left(W(s+t)\right)$, and hence (6.7) is equivalent to: for every continuous function $f : \mathbb{R}^d \to \mathbb{C}$ and for every $s \geqslant 0$ the equality

$$\mathbb{E}_x\left[f\left(W(s+t)\right) \,\middle|\, \mathcal{F}_t\right] = \mathbb{E}_{W(t)}\left[f\left(W(s)\right)\right] \tag{6.8}$$

holds $\mathbb{P}_x$-almost surely for all $t \geqslant 0$ and for all $x \in \mathbb{R}^d$. The passage from (6.8) to (6.6) goes via stochastic (or random) variables of the form (6.5). An application of the monotone class theorem then enables us to conclude (6.6). For more details on the monotone class theorem see Theorems 2.42 and 2.43, and the Propositions 2.44 and 2.45 in Chapter 2.

Put, for $t > 0$, x, $y \in \mathbb{R}^d$,

$$p_d(t,x,y) = \frac{1}{(2\pi t)^{d/2}} \exp\left(-\frac{|x-y|^2}{2t}\right) = \frac{1}{(2\pi t)^{d/2}} e^{-\frac{|x-y|^2}{2t}}.$$

Then, by definition, the marginal distributions, or one-dimensional distributions, of Brownian motion are given by

$$\mathbb{E}_x\left[f\left(W(t)\right)\right] = \int_{\mathbb{R}^d} p_d(t,x,y) f(y)\, dy, \quad t > 0, \tag{6.9}$$

where f is a bounded continuous function attaining values in $\mathbb{R}$. From the equality, for s, $t > 0$, x, y, $z \in \mathbb{R}^d$,

$$p_d\left(s,x,z\right) p_d\left(t,z,y\right) = p_d(s+t,x,y) p_d\left(\frac{st}{s+t}, \frac{sy+tx}{s+t}, z\right), \tag{6.10}$$

we infer that collection probability densities $\{p_d(t,x,y) : t > 0\}$ satisfies the following version of the *Chapman-Kolmogorov equation.* For all s, $t > 0$, and all x, $y \in \mathbb{R}^d$, the equalities

$$\begin{aligned}\int_{\mathbb{R}^d} p_d(s,x,z) p_d(t,z,y)\, dz &= p_d(s+t,x,y) \int_{\mathbb{R}^d} p_d\left(\frac{st}{s+t}, \frac{sy+tx}{s+t}, z\right) dz \\ &= p_d(s+t,x,y) \end{aligned}\tag{6.11}$$

are true. Put $P(t)f(x) = \mathbb{E}_x\left[f\left(W(t)\right)\right]$, $t \geqslant 0$, $f \in C_b\left(\mathbb{R}^d\right)$. Then by the Markov property we have

$$\begin{aligned} P(t)P(s)f(x) &= \mathbb{E}_x\left[P(s)f\left(W(t)\right)\right] = \mathbb{E}_x\left[\mathbb{E}_{W(t)}\left[f\left(W(s)\right)\right]\right] \\ &= \mathbb{E}_x\left[\mathbb{E}_x\left[f\left(W(s)\right)\circ\vartheta_t \;\middle|\; \mathcal{F}_t\right]\right] \\ &= \mathbb{E}_x\left[\mathbb{E}_x\left[f\left(W(s+t)\right) \;\middle|\; \mathcal{F}_t\right]\right] \\ &= \mathbb{E}_x\left[f\left(W(s+t)\right)\right] = P(t+s)f(x). \end{aligned}\tag{6.12}$$

In other words the family $\{P(t) : t \geqslant 0\}$ has the semigroup property. In addition, we have

$$\begin{aligned} \lim_{t\downarrow 0} P(t)f(x) &= \lim_{t\downarrow 0} \mathbb{E}_x\left[f\left(W(t)\right)\right] \\ &= \lim_{t\downarrow 0} \int_{\mathbb{R}^d} p_d(t,x,y) f(y)\, dy \\ &= \lim_{t\downarrow 0} \int_{\mathbb{R}^d} p_d(1,0,y) f\left(x + \sqrt{t}y\right)\, dy \\ &= \int_{\mathbb{R}^d} p_d(1,0,y) f\left(x\right)\, dy = f(x), \end{aligned}\tag{6.13}$$

for all bounded continuous functions f defined on $\mathbb{R}^d$. Another relevant property is the following one. If $f \in C_b\left(\mathbb{R}^d\right)$ is such that $0 \leqslant f \leqslant \mathbf{1}$, then $0 \leqslant P(t)f \leqslant \mathbf{1}$, $t \geqslant 0$. Let $C_0\left(\mathbb{R}^d\right)$ denote the space of all continuous functions $f : \mathbb{R}^d \to \mathbb{C}$

with the property that $\lim_{|x|\to\infty} f(x) = 0$. Then it is fairly easy to see that $P(t)f$ belongs to $C_0\left(\mathbb{R}^d\right)$ whenever f does so. All this means that, by definition, the family $\{P(t) : t \geqslant 0\}$ is a *Feller* or, even better, a *Dynkin-Feller semigroup*. Such semigroups have the property that

$$\lim_{s\to t,\, s\geqslant 0} \|P(s)f - P(t)f\|_\infty = \lim_{s\to t,\, s\geqslant 0} \sup_{x\in\mathbb{R}^d} |P(s)f(x) - P(t)f(x)| = 0 \tag{6.14}$$

for all $t \geqslant 0$ and for all $f \in C_0\left(\mathbb{R}^d\right)$. A calculation shows the following equalities for $t > 0$ and $\xi \in \mathbb{R}^d$:

$$\begin{aligned}
\mathbb{E}_x\left[e^{-i\langle W(t),\xi\rangle}\right] &= \int_{\mathbb{R}^d} e^{-i\langle y,\xi\rangle} p_d\left(t,x,y\right)\, dy \\
&= e^{-i\langle x,\xi\rangle} \int_{\mathbb{R}^d} e^{-i\langle y,\xi\rangle} p_d(t,0,y)\, dy \\
&= e^{-i\langle x,\xi\rangle} \int_{\mathbb{R}^d} e^{-i\sqrt{t}\langle y,\xi\rangle} p_d(1,0,y)\, dy \\
&= e^{-i\langle x,\xi\rangle} e^{-\frac{1}{2}t|\xi|^2}. \qquad (6.15)
\end{aligned}$$

Since $x = \mathbb{E}_x\left[W(t)\right]$ and $t\left|\xi\right|^2 = \mathbb{E}_x\left[\left(\langle W(t)-x,\xi\rangle\right)^2\right]$, it is clear from (6.15) that $W(t)$, distributed with density $p_d(t,x,y)$, is a Gaussian vector. In fact, let $0 < t_1 < \cdots < t_n$ and let f_j, $1 \leqslant j \leqslant n$, be bounded Borel measurable functions defined on $\mathbb{R}^d$. Then, with $t_0 = 0$ and $x_0 = x \in \mathbb{R}^d$, we have

$$\begin{aligned}
&\mathbb{E}_x\left[\prod_{j=1}^{n} f_j\left(W\left(t_j\right)\right)\right] \\
&= \int_{\mathbb{R}^d}\cdots\int_{\mathbb{R}^d}\prod_{j=1}^{n} f_j\left(x_j\right)\prod_{j=1}^{n} p_d\left(t_j - t_{j-1}, x_{j-1}, x_j\right)\, dx_1\cdots dx_n. \qquad (6.16)
\end{aligned}$$

Fix $f : \mathbb{R}^d \to \mathbb{C}$, let $0 < t_1 < \cdots < t_n < t$, and $g_j : \mathbb{R}^d \to \mathbb{C}$, $1 \leqslant j \leqslant n$, be bounded Borel measurable functions. From (6.16) together with the Chapman-Kolmogorov equation (see (6.11)) we infer the following equality

$$\begin{aligned}
&\mathbb{E}_x\left[\mathbb{E}_{W(t)}\left[f\left(W(s)\right)\right]\prod_{j=1}^{n} g_j\left(W\left(t_j\right)\right)\right] \\
&= \mathbb{E}_x\left[f\left(W(s+t)\right)\prod_{j=1}^{n} g_j\left(W\left(t_j\right)\right)\right]. \qquad (6.17)
\end{aligned}$$

Applying the monotone class theorem, and, possibly, employing some other approximation arguments, the equality in (6.17) implies

$$\mathbb{E}_x\left[f\left(W(s+t)\right) \,\middle|\, \mathcal{F}_t\right] = \mathbb{E}_{W(t)}\left[f\left(W(s)\right)\right], \quad \mathbb{P}_x\text{-almost surely.} \qquad (6.18)$$

Let $f_k : \mathbb{R}^d \to \mathbb{C}$, $1 \leqslant k \leqslant m$, be bounded Borel measurable functions. Let $0 < s_1 < \cdots < s_m$. Repeating the previous arguments leads to

$$\begin{aligned}
&\mathbb{E}_x\left[\mathbb{E}_{W(t)}\left[\prod_{k=1}^{m} f_k\left(W\left(s_k\right)\right)\right]\prod_{j=1}^{n} g_j\left(W\left(t_j\right)\right)\right] \\
&= \mathbb{E}_x\left[\prod_{k=1}^{m} f_k\left(W\left(s_k+t\right)\right)\prod_{j=1}^{n} g_j\left(W\left(t_j\right)\right)\right]. \qquad (6.19)
\end{aligned}$$

Assume that the σ-field $\mathcal{F}$ is generated by $(W(t) : t \geqslant 0)$. In other words $\mathcal{F}$ is the smallest σ-field which makes all variables $W(t) : \Omega \to \mathbb{R}^d$ measurable. Let $Y : \Omega \to \mathbb{C}$ be bounded random variable. This means that Y is $\mathcal{F}$-$\mathcal{B}_{\mathbb{C}}$-measurable. Another application of the monotone class theorem and using (6.19) then implies, for $t \geqslant 0$,

$$\mathbb{E}_x\left[Y \circ \vartheta_t \,\middle|\, \mathcal{F}_t\right] = \mathbb{E}_{W(t)}\left[Y\right], \quad \mathbb{P}_x\text{-almost surely.} \qquad (6.20)$$

The reader should compare the results above with Subsection 4.2 in Chapter 1 and with Theorems 3.29 of Chapter 3. In this chapter can find much more information on Brownian motion.

3. Brownian motion as a martingale

In this section we will give a brief review of the martingale properties of Brownian motion: see Section 6 of Chapter 2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space (sample path space), and let $\{W(t) : t \geqslant 0\}$ be Brownian motion on Ω starting at $0 \in \mathbb{R}^d$. We consider Brownian motion as a Gaussian process in $\mathbb{R}^d$. Then $W(t) - W(s)$ is $\mathbb{P}$-independent of $\mathcal{F}_s$. In other words Brownian motion has independent increments, and consequently it is a martingale in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. In addition, we have $\mathbb{E}\left[W_{j_1}(s)W_{j_2}(t)\right] = \min(s,t)\,\delta_{j_1,j_2}$, $0 \leqslant s,t$, $1 \leqslant j_1, j_2 \leqslant d$. It also follows that the process $t \mapsto W_{j_1}(t)W_{j_2}(t) - t\delta_{j_1,j_2}$ is a $\mathbb{P}$-martingale. The Lévy's characterization theorem says that Brownian motion is the only continuous time martingale with the latter property. For $d = 1$ the process $t \mapsto W(t)^2 - t$ is a martingale. In other words the positive sub-martingale $t \mapsto |W(t)|^2$ can be written in the form $|W(t)|^2 = A(t) + M(t)$ where $A(t) = dt$ is an increasing process, and where $M(t) = |W(t)|^2 - dt$ is a martingale. This is a very special case of the Doob-Meyer decomposition theorem which says among other things that continuous sub-martingales $S(t)$ of class DL (a certain uniform integrability property) can be written in the form $S(t) = A(t) + M(t)$ where the process $A(t)$ is a non-decreasing adapted process, and $M(t)$ is a martingale. In fact throughout the lectures of this course we will frequently encounter semi-martingales $S(t)$. Such processes $S(t)$ are of the form $S(t) = A(t) + M(t)$ where $A(t)$ is a process of bounded variation, and where $M(t)$ is a martingale. An real-valued adapted process $A(t)$ is of bounded variation (BV) whenever it can be written in the form $A(t) = A_+(t) - A_-(t)$ where the processes $t \mapsto A_\pm(t)$ are increasing, or more precisely, non-decreasing adapted processes. The bounded variation part is usually considered as the trend or the drift of the process, and the martingale part is considered as the uncertainty, the risk or volatility part. Let $f : \mathbb{R}^d \to \mathbb{C}$ be a C^2-function with bounded first and second derivatives. Then by Itô's formula we have

$$\begin{aligned} f(W(t)) &= f(W(0)) + \int_0^t \nabla f(W(s))\, dW(s) + \frac{1}{2}\int_0^t \Delta f(W(s))\, ds \\ &= M(t) + A(t), \end{aligned} \tag{6.21}$$

where

$$\begin{aligned} A(t) &= \frac{1}{2}\int_0^t \Delta f(W(s))\, ds \quad \text{and} \\ M(t) &= f(W(0)) + \int_0^t \nabla f(W(s))\, dW(s). \end{aligned} \tag{6.22}$$

Since the Itô integral $\int_0^t \nabla f(W(s))\, dW(s)$ is a martingale, the process $t \mapsto f(W(t))$ is decomposed into a martingale part $M(t)$ and a bounded variation part $A(t)$ which is also called the *Itô correction term.* Other important examples of semi-martingales are solutions to stochastic differential equations (SDEs) $t \mapsto X(t)$. Such processes satisfy an equation of the form

$$X(t) = X(0) + \int_0^t b(s, X(s))\, ds + \int_0^t \sigma(s, X(s))\, dW(s) = A(t) + M(t), \tag{6.23}$$

where $A(t) = \int_0^t b(s, X(s))\, ds$ and $M(t) = X(0) + \int_0^t \sigma(s, X(s))\, dW(s)$. Under appropriate conditions on the coefficients $b(s,x)$ and $\sigma(s,x)$ the process $t \mapsto A(t)$ is an adapted process of bounded variation, and the process $t \mapsto M(t)$ is an adapted process which is a martingale. Observe that solutions to stochastic differential equations are supposed to be adapted (and continuous). Often the equation in (6.23) is written in differential form:

$$dX(t) = b\left(t, X(t)\right)\, dt + \sigma\left(t, X(t)\right)\, dW(t) = dA(t) + dM(t). \tag{6.24}$$

Moreover, the fact that processes of the form $t \mapsto W_{j_1}(t)W_{j_2}(t) - t\delta_{j_1,j_2}$, $1 \leqslant j_1, j_2 \leqslant d$ are martingales forms the basis for stochastic or Itô calculus with respect to Brownian motion. Processes closely related to Brownian motion include

(1) Ornstein-Uhlenbeck processes have the form

$$X(t) = e^{-tA}X(0) + \int_0^t e^{-(t-s)A}\sigma(s)\, dW(s).$$

Here A is a square $d \times d$-matrix. The corresponding SDE reads as follows:

$$dX(t) = -AX(t)\, dt + \sigma(t)\, dW(t).$$

Such processes can be used in the context of models with stochastic interest rates (Vasicek model).

(2) Processes of the form

$$S(t) = e^{\int_0^t \left(\mu - \frac{1}{2}|\sigma(s)|^2\right)\, ds + \int_0^t \sigma(s)\, dW(s)},$$

where $s \mapsto \sigma(s)$ is an adapted $\mathbb{R}^d$-valued process and μ is a (positive) constant, play a crucial role in the modelling of stock prices. The Black-Sholes-Merton model is an important example. The process $t \mapsto S(t)$ satisfies the following SDE:

$$dS(t) = S(t)\left(\mu\, dt + \sigma(t)\, dW(t)\right).$$

If $\mu = 0$ and σ is a constant vector, then the process $t \mapsto S(t)$ is called a *geometric Brownian motion*. Such processes play a relevant role in the construction of the risk neutral measure. These measures have a (stochastic) density of the form $S(T)$ with $\mu = 0$ provided T is the time of maturity. Mathematically speaking the risk neutral measure is the Girsanov transformation of a Brownian motion with drift: it neutralizes the drift.

(3) Another Gaussian process which plays a role in the construction of (deterministic) densities to describe the distribution of certain processes is the Brownian bridge. Such a process takes the form:

$$\begin{aligned} s \mapsto X_t^1(s) &:= \left(1 - \frac{s}{t}\right)x + \frac{s}{t}y + \left(1 - \frac{s}{t}\right)W\left(\frac{st}{t-s}\right), \quad \text{or} \\ s \mapsto X_t^2(s) &:= \left(1 - \frac{s}{t}\right)x + \frac{s}{t}y + W(s) - \frac{s}{t}W\left(t\right), \end{aligned} \tag{6.25}$$

where $t > 0$ is fixed and x and y belong to $\mathbb{R}^d$, and where $W(t)$ is a Brownian motion in $\mathbb{R}^d$ which starts at 0. Observe that these processes are not adapted to Brownian motion, because at time s one needs information from a later time to describe the processes in (6.25). In fact for the first process one needs $W\left(\dfrac{st}{t-s}\right)$, and for the second $W(t)$ is needed. The processes in (6.25) have realizations which connect x (at time 0) with y (at time t. It is not so difficult to see that the expectations and covariances of the processes $X_t^1(s)$ and $X_t^2(s)$ are the same. Since both processes are Gaussian, their distributions are the same. For instance for $d = 1$ and $0 \leqslant s_1,\ s_2 < t$ we have

$$\begin{aligned}
&\mathbb{E}\left[\left(X_t^1\left(s_1\right)-\mathbb{E}\left[X_t^1\left(s_1\right)\right]\right)\left(X_t^1\left(s_2\right)-\mathbb{E}\left[X_t^1\left(s_2\right)\right]\right)\right]\\
&=\mathbb{E}\left[\left(1-\frac{s_1}{t}\right)W\left(\frac{s_1t}{t-s_1}\right)\left(1-\frac{s_2}{t}\right)W\left(\frac{s_2t}{t-s_2}\right)\right]\\
&=\left(1-\frac{s_1}{t}\right)\left(1-\frac{s_2}{t}\right)\mathbb{E}\left[W\left(\frac{s_1t}{t-s_1}\right)W\left(\frac{s_2t}{t-s_2}\right)\right]\\
&=\left(1-\frac{s_1}{t}\right)\left(1-\frac{s_2}{t}\right)\min\left(\frac{s_1t}{t-s_1},\frac{s_2t}{t-s_2}\right)\\
&=\min\left(s_1,s_2\right)-\frac{s_1s_2}{t}. \qquad (6.26)
\end{aligned}$$

By a similar token we get

$$\begin{aligned}
&\mathbb{E}\left[\left(X_t^2\left(s_1\right)-\mathbb{E}\left[X_t^2\left(s_1\right)\right]\right)\left(X_t^2\left(s_2\right)-\mathbb{E}\left[X_t^2\left(s_2\right)\right]\right)\right]\\
&=\mathbb{E}\left[\left(W\left(s_1\right)-\frac{s_1}{t}W\left(t\right)\right)\left(W\left(s_2\right)-\frac{s_2}{t}W\left(t\right)\right)\right]\\
&=\min\left(s_1,s_2\right)-\frac{s_1s_2}{t}. \qquad (6.27)
\end{aligned}$$

Since $\mathbb{E}\left[W(t)\right) = 0$, and since the processes $s \mapsto X_t^1(s)$ and $s \mapsto X_t^2(s)$ are Gaussian, the equalities in (6.26) and (6.27) entail that their distributions are the same.

It is perhaps useful to include a formal definition of a function of bounded variation on a finite interval $[0,T]$. Let $f : [0,T] \to \mathbb{C}$ be a right-continuous function. Then f is said to be of *bounded variation*, BV for short, if

$$\sup\left\{\sum_{j=1}^{n}\left|f\left(t_j\right)-f\left(t_{j-1}\right)\right| : \, 0 \leqslant t_0 < t_1 < \cdots < t_n \leqslant T\right\} < \infty.$$

Here the supremum is taken over all subdivisions $0 \leqslant t_0 < t_1 < \cdots < t_n \leqslant T$ of the interval $[0,T]$. If f is real-valued and of bounded variation, then f can be written as a difference of two increasing (in fact non-decreasing) functions: $f = f_+ - f_-$. If f is continuous from the right and has limits from the left, in other words, if f is càÒdlàg, then the functions f_+ and f_- can be chosen to possess the same properties. It is also mentioned that for a function $f : [0,T] \to \mathbb{C}$ of bounded variation which is continuous from the right and has left limits there exists a unique complex-valued measure μ_f on the Borel subsets of $[0,T]$ such that $\mu_f\left(0,t\right) = f(t) - f(0)$. In other words it makes sense to write $\int_0^T g(t)\,df(t) = \int_0^T g(t)\,d\mu_f(t) = \int_0^T g(t)\,\left(d\mu_{f_+}(t) - d\mu_{f_-}(t)\right)$, $g \in L^\infty\left([0,T],\mathbb{C}\right)$.

4. Some relevant martingales

Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space with filtration $(\mathcal{F}_t)_{t\geqslant 0}$. Let the stochastic variable $Y : \Omega \to \mathbb{R}$ belong to $L^1\left(\Omega,\mathcal{F},\mathbb{P}\right)$, and put $Y(t) = \mathbb{E}\left[Y \mid \mathcal{F}_t\right]$. Then the process $t \mapsto Y(t)$ is a so-called *closed martingale*; i.e. for some $Y \in L^1\left(\Omega,\mathcal{F},\mathbb{P}\right)$ the equality $Y(t) = \mathbb{E}\left[Y \mid \mathcal{F}_t\right]$ holds $\mathbb{P}$-almost surely for all $t \geqslant 0$. If a martingale $\{M(t) : t \geqslant 0\}$ is uniformly integrable, then its pointwise limit $M = \lim\limits_{t\to\infty} M(t)$ exists in $L^1\left(\Omega,\mathcal{F},\mathbb{P}\right)$ and $M(t) = \mathbb{E}\left[M \mid \mathcal{F}_t\right]$ $\mathbb{P}$-almost surely for all $t \geqslant 0$, and so $\{M(t) : t \geqslant 0\}$ is closed.

Again we consider the Markov process of Brownian motion as explained in Section 2. Let $Y : \Omega \to \mathbb{C}$ be a stochastic variable. Then, on the interval $[0,t]$ the process $s \mapsto \mathbb{E}_{W(s)}\left[Y \circ \vartheta_{t-s}\right]$, $0 \leqslant s \leqslant t$, is a $\mathbb{P}_x$-martingale for all $t > 0$ and for all $x \in \mathbb{R}^d$. This is so because by the Markov property we have, for $0 \leqslant s < t$,

$$\mathbb{E}_x\left[Y \circ \vartheta_t \mid \mathcal{F}_s\right] = \mathbb{E}_x\left[\left(Y \circ \vartheta_{t-s}\right) \circ \vartheta_s \mid \mathcal{F}_s\right] = \mathbb{E}_{W(s)}\left[Y \circ \vartheta_{t-s}\right]. \tag{6.28}$$

Fix x, $y \in \mathbb{R}^d$ and $t > 0$. Then the process $s \mapsto p_d\left(t-s, W(s), y\right)$ is martingale on the half-open interval $[0,t)$. It is not a closed martingale in the interval. Let us prove this. To this end fix $0 \leqslant s_1 < s_2 < t$ and employ the Markov property to infer:

$$\begin{aligned}&\mathbb{E}_x\left[p_d\left(t-s_2, W\left(s_2\right), y\right) \mid \mathcal{F}_{s_1}\right]\\ &\quad = \mathbb{E}_x\left[p_d\left(t-s_2, W\left(s_2-s_1\right), y\right) \circ \vartheta_{s_1} \mid \mathcal{F}_{s_1}\right]\end{aligned}$$

(Markov property)

$$= \mathbb{E}_{W(s_1)}\left[p_d\left(t-s_2, W\left(s_2-s_1\right), y\right)\right]$$

(write $w = W(s_1)$)

$$= \mathbb{E}_w \left[p_d \left(t - s_2, W\left(s_2 - s_1\right), y \right) \right]$$

$$= \int_{\mathbb{R}^d} p_d \left(s_2 - s_1, w, z \right) p_d \left(t - s_2, z, y \right) \, dz$$

(Chapman-Kolmogorov)

$$= p_d \left(t - s_1, w, y \right) = p_d \left(t - s_1, W\left(s_1\right), y \right). \tag{6.29}$$

The equality of the first and final expression in (6.29) shows the martingale property of the process $s \mapsto p_d\left(t - s, W(s), y\right)$, $0 \leqslant s < t$. We also observe that $\lim_{s \uparrow t} p_d\left(t - s, W(s), y\right) = 0$ $\mathbb{P}_x$-almost surely on the event $\{W(t) \neq y\}$, and so this limit vanishes $\mathbb{P}_x$-almost surely. Using the equality $\mathbb{E}_x\left[p_d\left(t - s, W(s), y\right)\right] = p_d(t, x, y)$ shows by virtue of Scheffé's theorem that the martingale

$$\left\{ p_d\left(t - s, W(s), y\right) : 0 \leqslant s < t \right\}$$

is not uniformly integrable, and hence it cannot be closed.

Another martingale which is related to Brownian motion as a Markov process and to stochastic calculus can be described as follows. Let $f : \mathbb{R}^d \to \mathbb{C}$ be a bounded C^2-function with bounded first and second order derivatives. Then the process

$$t \mapsto M_f(t) := f\left(W(t)\right) - f\left(W(0)\right) - \frac{1}{2}\int_0^t \Delta f\left(W(s)\right)\, ds$$

is a martingale. Let us prove this. First of all, for $0 \leqslant t_1 < t_2$ we have

$$M_f\left(t_2\right) - M_f\left(t_1\right) = M_f\left(t_2 - t_1\right) \circ \vartheta_{t_1}. \tag{6.30}$$

The equality in (6.30) says that the mapping $t \mapsto M_f(t)$ is *an additive process* which is usually written as

$$M_f\left(t_1 + t_2\right) = M_f\left(t_1\right) + M_f\left(t_2\right) \circ \vartheta_{t_1}, \quad \text{for } 0 \leqslant t_1,\, t_2.$$

Secondly, we observe the following equality:

$$\frac{\partial}{\partial t} p_d\left(t, x, y\right) = \frac{1}{2}\Delta_x p_d\left(t, x, y\right) = \frac{1}{2}\Delta_y p_d\left(t, x, y\right), \tag{6.31}$$

where Δ_x denotes the Laplace operator:

$$\Delta_x f(x) = \Delta f(x) = \sum_{j=1}^d \frac{\partial^2 f(x)}{\left(\partial x_j\right)^2}.$$

Here the function f is differentiable up to order 2. For $w \in \mathbb{R}^d$ and $t > 0$ we have

$$\begin{aligned}
&\mathbb{E}_w\left[M_f(t)\right] \\
&= \int_{\mathbb{R}^d} p_d\left(t, w, y\right) f(y)\, dy - f(w) - \frac{1}{2}\int_0^t \int_{\mathbb{R}^d} p_d\left(s, w, y\right) \Delta f(y)\, dy\, ds
\end{aligned}$$

(integration by parts)

$$= \int_{\mathbb{R}^d} p_d\left(t, w, y\right) f(y)\, dy - f(w) - \frac{1}{2}\int_0^t \int_{\mathbb{R}^d} \Delta_y p_d\left(s, w, y\right) f(y)\, dy\, ds$$

(apply (6.31))

$$= \int_{\mathbb{R}^d} p_d\left(t, w, y\right) f(y)\, dy - f(w) - \int_0^t \int_{\mathbb{R}^d} \frac{\partial}{\partial s} p_d\left(s, w, y\right) f(y)\, dy\, ds$$

(apply Fubini's theorem)

$$\begin{aligned}
&= \int_{\mathbb{R}^d} p_d\left(t, w, y\right) f(y)\, dy - f(w) - \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d} \int_\varepsilon^t \frac{\partial}{\partial s} p_d\left(s, w, y\right)\, ds\, f(y)\, dy \\
&= \int_{\mathbb{R}^d} p_d\left(t, w, y\right) f(y)\, dy - f(w) \\
&\qquad - \lim_{\varepsilon \downarrow 0} \left\{ \int_{\mathbb{R}^d} p_d\left(t, w, y\right) f(y)\, dy - \int_{\mathbb{R}^d} p_d\left(\varepsilon, w, y\right) f(y)\, dy \right\} \\
&= \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d} p_d\left(1, 0, y\right) f\left(w + \sqrt{\varepsilon} y\right)\, dy - f(w)
\end{aligned}$$

$$= f(w) - f(w) = 0. \tag{6.32}$$

By the Markov property of Brownian motion we infer, for $0 \leqslant t_1 < t_2$,

$$\begin{aligned}&\mathbb{E}_x\left[M_f\left(t_2\right) \mid \mathcal{F}_{t_1}\right] - M_f\left(t_1\right)\\ &\quad= \mathbb{E}_x\left[M_f\left(t_2\right) - M_f\left(t_1\right) \mid \mathcal{F}_{t_1}\right]\end{aligned}$$

(employ the equality in (6.30))

$$= \mathbb{E}_x\left[M_f\left(t_2 - t_1\right) \circ \vartheta_{t_1} \mid \mathcal{F}_{t_1}\right]$$

(Markov property of Brownian motion)

$$= \mathbb{E}_{W(t_1)}\left[M_f\left(t_2 - t_1\right)\right]$$

(apply (6.32) with $w = W\left(t_1\right)$ and $t = t_2 - t_1$)

$$= \mathbb{E}_{W(t_1)}\left[M_f(0)\right] = 0. \tag{6.33}$$

The final three equalities hold $\mathbb{P}_x$-almost surely. The equalities in (6.33) show the martingale property of the process $t \mapsto M_f(t)$. A very similar argument yields the following result. Let

$$\left\{\left(\Omega, \mathcal{F}, \mathbb{P}_x\right)_{x\in E}, \left(X(t), t \geqslant 0\right), \left(\vartheta_t, t \geqslant 0\right), \left(E, \mathcal{E}\right)\right\} \tag{6.34}$$

be a Markov process, and let $t \mapsto A(t)$ be an additive process in $L^1\left(\Omega, \mathcal{F}_t, \mathbb{P}_x\right)$ for all $x \in E$ and all $t \geqslant 0$ such that $\mathbb{E}_w\left[A(t)\right] = \mathbb{E}_w\left[A(0)\right]$ for all $w \in E$ and for all $t \geqslant 0$. Then the process $t \mapsto A(t)$ is a $\mathbb{P}_x$-martingale for all $x \in E$. In fact this result is closely related to the so-called *martingale problem.* Again we consider the Markov process in (6.34). Observe that $A\left(t_1\right) + A(0) \circ \vartheta_{t_1} = A\left(t_1\right)$ $\mathbb{P}_w$-almost surely for all $t_1 \geqslant 0$. Whence, for all $w \in E$, $A(0) \circ \vartheta_{t_1} = 0$ $\mathbb{P}_w$-almost surely for all $t_1 \geqslant 0$. A formal proof of the martingale property reads as follows. Let $x \in E$ and $0 \leqslant t_1 < t_2$. Then we have

$$\mathbb{E}_x\left[A\left(t_2\right) \mid \mathcal{F}_{t_1}\right] = \mathbb{E}_x\left[A\left(t_1\right) + A\left(t_2 - t_1\right) \circ \vartheta_{t_1} \mid \mathcal{F}_{t_1}\right]$$

(the variable $A\left(t_1\right)$ is $\mathcal{F}_{t_1}$-measurable)

$$= A\left(t_1\right) + \mathbb{E}_x\left[A\left(t_2 - t_1\right) \circ \vartheta_{t_1} \mid \mathcal{F}_{t_1}\right]$$

(Markov property)

$$\begin{aligned}&= A\left(t_1\right) + \mathbb{E}_{X(t_1)}\left[A\left(t_2 - t_1\right)\right]\\ &= A\left(t_1\right) + \mathbb{E}_{X(t_1)}\left[A\left(0\right)\right]\end{aligned}$$

(again we use the Markov property)

$$\begin{aligned}&= A\left(t_1\right) + \mathbb{E}_x\left[A\left(0\right) \circ \vartheta_{t_1} \mid \mathcal{F}_{t_1}\right]\\ &= \mathbb{E}_x\left[A\left(t_1\right) + A\left(0\right) \circ \vartheta_{t_1} \mid \mathcal{F}_{t_1}\right]\\ &= \mathbb{E}_x\left[A\left(t_1\right) \mid \mathcal{F}_{t_1}\right] = A\left(t_1\right).\end{aligned} \tag{6.35}$$

The equalities in (6.35) show the $\mathbb{P}_x$-martingale property of the process $t \mapsto A(t)$.

An adapted real- or complex-valued process $t \mapsto M(t)$ is called *multiplicative* if $M(t_1+t_2) = M(t_2) \circ \vartheta_{t_1} M(t_1)$ $\mathbb{P}_x$-almost surely for all $0 \leqslant t_1,\, t_2 < \infty$ and for all $x \in \mathbb{E}$. Then $M(t_1) = M(0) \circ \vartheta_{t_1} M(t_1)$ $\mathbb{P}_x$-almost surely for all $0 \leqslant t_1 < \infty$ and for all $x \in \mathbb{E}$. If such a process belongs to $L^1(\Omega, \mathcal{F}, \mathbb{P}_x)$ for all $x \in \mathbb{R}^d$ (or attains only values in $[0,\infty)$) and is such that $\mathbb{E}_w[M(t)] = \mathbb{E}_w[M(0)]$ for all $w \in E$ and all $t \geqslant 0$, then the process $t \mapsto M(t)$ is a $\mathbb{P}_x$-martingale for all $x \in E$. The proof can be patterned after the above proof of the martingale property of the additive process $t \mapsto A(t)$; instead of a plus sign one writes a multiplication sign in (6.35). In case we apply the latter to Brownian motion we see, e.g., that for every $c \in \mathbb{R}^d$ the process $t \mapsto M(t) := e^{-\langle c, W(t)-W(0)\rangle - \frac{1}{2}|c|^2 t}$ is a $\mathbb{P}_x$-martingale for all $x \in \mathbb{R}^d$. Here we employ the following equalities

$$\begin{aligned}\mathbb{E}_x[M(t)] &= \mathbb{E}_x\left[e^{-\langle c, W(t)-W(0)\rangle - \frac{1}{2}|c|^2 t}\right] = \int_{\mathbb{R}^d} p_d(t,x,y)\, e^{-\langle c, y-x\rangle - \frac{1}{2}|c|^2 t}\, dy \\ &= 1 = \mathbb{E}_x[M(0)].\end{aligned}$$

All this is closely related to the concept of exponential martingales, the notion of risk-neutral measure, and the Girsanov transformation of measures. For more details the reader is referred to Subsection 8.1 of Chapter 3 and to Section 3 in Chapter 4 of the main text.

5. Conditional expectation

Properties of stochastic processes are very often described in terms of conditional expectations on sub-σ-fields: martingale property, sub- and super-martingale property, Markov property. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and let $\mathcal{B}$ be a sub-σ-field of $\mathcal{A}$. Let $Y : \Omega \to \mathbb{R}$ be a bounded stochastic variable. This means that the variable Y is $\mathcal{A}$-$\mathcal{B}_{\mathbb{R}}$-measurable. Put $Z = \mathbb{E}\left[Y \mid \mathcal{B}\right]$. By definition this means that Z is $\mathcal{B}$-measurable (qualitative property) and that $\mathbb{E}\left[Z\mathbf{1}_B\right] = \mathbb{E}\left[Y\mathbf{1}_B\right]$ for all $B \in \mathcal{B}$. By comparing the measures $\mathbb{P}_1 = \mathbb{P}\restriction_{\mathcal{B}} : B \mapsto \mathbb{P}[B]$ and $\mathbb{P}_2 : B \mapsto \mathbb{E}\left[Y\mathbf{1}_B\right]$ with $B \in \mathcal{B}$ we see that P_2 is absolutely continuous with respect to $\mathbb{P}_1$. This means that $\mathbb{P}_1[B] = 0$ implies $\mathbb{P}_2[B] = 0$ with $B \in \mathcal{B}$. It is a consequence of the theorem of Radon-Nikodym that a conditional expectation exists. An important example comes in the context of the Markov property. Suppose we want to prove the equality in (6.7). Then we consider the subspace H of $L^{\infty}(\Omega, \mathcal{F}_t, \mathbb{P}_x)$ defined by

$$\mathsf{H} = \left\{G \in L^{\infty}(\Omega, \mathcal{F}_t, \mathbb{P}_x) : \mathbb{E}_x\left[\mathbb{E}_{W(t)}[Y]\, G\right] = \mathbb{E}_x\left[Y \circ \vartheta_t G\right]\right\}. \tag{6.36}$$

Then H is a vector space over $\mathbb{R}$. Moreover, if a sequence $(G_n)_{n\in\mathbb{N}} \subset \mathsf{H}$ is such that $0 \leqslant G_1 \leqslant G_2 \leqslant \cdots \leqslant G_n \leqslant G = \sup_{n\in\mathbb{N}} G_n$ where G is bounded, then G belongs to H. If H contains all variables G of the form $G = \prod_{j=1}^{n} g_j(W(t_j))$, with $0 \leqslant t_1 < t_2 < \cdots < t_n < t$, then by the monotone class theorem the subspace H coincides with $L^{\infty}(\Omega, \mathcal{F}_t, \mathbb{P}_x)$, whence $\mathbb{E}_x\left[Y \circ \vartheta_t \mid \mathcal{F}_t\right] = \mathbb{E}_{W(t)}[Y]$, $\mathbb{P}_x$-almost surely. Another application of the monotone class theorem shows that it suffices to take Y of the form $Y = \prod_{k=1}^{m} f_k(W(s_k))$ where $0 < s_1 < \cdots < s_m$. So we have to prove the equality as described in (6.19). Put $h_j = g_j$, $1 \leqslant j \leqslant n$, $h_j = f_{j-n}$, $n+1 \leqslant j \leqslant n+m$; $\tau_j = t_j$, $1 \leqslant j \leqslant n$, $\tau_j = t + s_{j-n}$, $n+1 \leqslant j \leqslant n+m$. Then the left-hand side of (6.19) is equal to

$$\begin{aligned}
&\mathbb{E}_x\left[\mathbb{E}_{W(t)}[Y]\, G\right] \\
&= \underbrace{\int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d}}_{n\text{ times}} \int_{\mathbb{R}^d} \prod_{j=1}^{n} g_j(x_j) \prod_{j=1}^{n} p_d(t_j - t_{j-1}, x_{j-1}, x_j)\, p_d(t - t_n, x_n, y_0) \\
&\qquad dx_1 \cdots dx_n \\
&\quad \underbrace{\int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d}}_{m\text{ times}} \prod_{k=1}^{m} f_k(y_k) \prod_{k=1}^{m} p_d(s_k - s_{k-1}, y_{k-1}, y_k)\; dy_0\, dy_1 \cdots dy_m
\end{aligned}$$

(Chapman-Kolmogorov: integral with respect to dy_0 can be explicity calculated: it is equal to $p_d(t + s_1 - t_n, x_n, y_1)$)

$$\begin{aligned}
&= \underbrace{\int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d}}_{n+m\text{ times}} \prod_{j=1}^{n+m} h_j(x_j) \prod_{j=1}^{n+m} p_d(\tau_j - \tau_{j-1}, x_{j-1}, x_j)\; dx_1 \cdots dx_{n+m} \\
&= \mathbb{E}_x\left[\prod_{j=1}^{n+m} h_j(W(\tau_j))\right]
\end{aligned}$$

$$= \mathbb{E}_x\left[\prod_{k=1}^{m} f_k\left(W\left(s_k+t\right)\right)\prod_{j=1}^{n} g_j\left(W\left(t_j\right)\right)\right] = \mathbb{E}_x\left[Y\circ\vartheta_t G\right] \tag{6.37}$$

where $G = \prod_{j=1}^{n} g_j\left(W\left(t_j\right)\right)$ and $Y = \prod_{k=1}^{m} f_k\left(W\left(s_k\right)\right)$. The variables y_j were renamed: $y_j = x_{n+j}$, $1 \leqslant j \leqslant m$. From the equality in (6.37) the Markov property of Brownian motion follows after some approximation arguments: see Theorem 3.29 Chapter 3 of the main text. More details on conditional expectations can be found in Theorem 1.4 in Chapter 1 of the main text.

5.1. Exponential martingales. We consider Brownian motion or Wiener process as a Markov process:

$$\left\{\left(\Omega, \mathcal{F}, \mathbb{P}_x\right), \left(W(t), t \geqslant 0\right), \left(\vartheta_t, t \geqslant 0\right), \left(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d}\right)\right\}. \tag{6.38}$$

Here $\Omega = C\left([0,\infty), \mathbb{R}^d\right)$ equipped with its σ-field $\mathcal{F}$ generated by the variables $t \mapsto W(t)$, where $W(t)(\omega) = \omega(t)$, $\omega \in \Omega$. In fact we will mainly use the martingale properties of Brownian motion and the corresponding Itô calculus. Let $x \mapsto c(x)$, $x \in \mathbb{R}^d$, be an $\mathbb{R}^d$-valued function with the property that there exists a (continuous) process $t \mapsto X(t)$ which satisfies the following stochastic integral equation:

$$X(t) = W(t) + \int_0^t c\left(X(s)\right)\, ds. \tag{6.39}$$

Define the process $t \mapsto Z(t)$ by

$$Z(t) = \int_0^t c\left(X(s)\right)\, dW(s) + \frac{1}{2}\int_0^t \left|c\left(X(s)\right)\right|^2\, ds, \tag{6.40}$$

and suppose that $\mathbb{E}_x\left[e^{-Z(t)}\right] = 1$ for $x \in \mathbb{R}^d$ and $t \geqslant 0$. Fix $T > 0$, and let $Y : \Omega \to \mathbb{C}$ be any bounded $\mathcal{F}_T = \sigma\left(W(s), 0 \leqslant s \leqslant T\right)$ measurable stochastic variable, and put

$$\mathbb{E}_x^Z\left[Y\right] = \mathbb{E}_x\left[e^{-Z(T)}Y\right]. \tag{6.41}$$

The corresponding probability measure is denoted by $\mathbb{P}_x^Z$.

6.1. Theorem. *With the above notation the following assertions hold true:*

1. *The process $t \mapsto e^{-Z(t)}$, $t \geqslant 0$, is a $\mathbb{P}_x$-martingale;*
2. *The process $t \mapsto e^{-Z(t)}\left(W(t) + \int_0^t c\left(X(s)\right)\, ds\right)$, $t \geqslant 0$, is a local $\mathbb{P}_x$-martingale;*
3. *The process $t \mapsto W(t) + \int_0^t c\left(X(s)\right)\, ds$, $t \geqslant 0$, is a local $\mathbb{P}_x^Z$-martingale;*
4. *The process $t \mapsto W(t) + \int_0^t c\left(X(s)\right)\, ds$, $t \geqslant 0$, is a Browian motion relative to the probability measure $\mathbb{P}_x^Z$;*
5. *The distribution of the process $t \mapsto X(t)$ is given by the measure*

$$\left\{\left\{\left(X(s)\right)_{0\leqslant s\leqslant T}\right\} \in A\right\} \mapsto \mathbb{P}_x^{\widetilde{Z}}\left[A\right] = \mathbb{E}_x\left[e^{-\widetilde{Z}(T)}, \left\{\left(W(s)\right)_{0\leqslant s\leqslant T}\right\} \in A\right],$$

$A \in \mathcal{F}_T$, *where*

$$\widetilde{Z}(t) = -\int_0^t c(W(s))\cdot dW(s) + \frac{1}{2}\int_0^t \left|c(W(s))\right|^2 ds.$$

6.2. Remark. Let the bounded stochastic variable Y be measurable with respect to $\mathcal{F}_t = \sigma(W(s), 0 \leqslant s \leqslant t)$, $0 \leqslant t \leqslant T$. Then (1) implies: $\mathbb{E}_x\left[e^{-Z(T)}Y\right] = \mathbb{E}_x\left[e^{-Z(t)}Y\right]$.

Proof. (1) Itô's formula, applied to the function $y \mapsto e^{-y}$ implies:

$$\begin{aligned} e^{-Z(t)} - e^{-Z(0)} &= -\int_0^t e^{-Z(s)}\,dZ(s) + \frac{1}{2}\int_0^t e^{-Z(s)}\,d\langle Z, Z\rangle(s) \\ &= -\int_0^t e^{-Z(s)}\,dW(s) - \frac{1}{2}\int_0^t |c(X(s))|^2\,ds \\ &= -\int_0^t e^{-Z(s)}\,dW(s). \end{aligned} \tag{6.42}$$

Since $\mathbb{E}_x\left[e^{-Z(t)}\right] = 1$ for all $0 \leqslant t \leqslant T$, the equality in (6.42) shows that the assertion in (1) is true. For more explanation see Proposition 6.4 below.

(2) Another application of Itô's formula, now applied to the function $(x, y) \mapsto e^{-x}y$ yields the following equalities:

$$\begin{aligned} &e^{-Z(t)}X(t) - e^{-Z(0)}X(0) \\ &= -\int_0^t e^{-Z(s)}X(s)\,dZ(s) + \int_0^t e^{-Z(s)}\,dX(s) \\ &\quad + \frac{1}{2}\int_0^t e^{-Z(s)}X(s)\,d\langle Z, Z\rangle(s) - \int_0^t e^{-Z(s)}\,d\langle Z, X\rangle(s) \end{aligned}$$

(employ the equalities
$\langle Z, Z\rangle(t) = \int_0^t |c(X(s))|^2\,ds$ and $\langle Z, X\rangle(t) = \int_0^t c(X(s))\,ds$)

$$= -\int_0^t e^{-Z(s)}X(s)c(X(s))\,dW(s) + \int_0^t e^{-Z(s)}\,dW(s). \tag{6.43}$$

From (6.43) it follows that the process $t \mapsto e^{-Z(t)}X(t)$ is a local martingale.

(3) By a stopping argument we may assume that the process $t \mapsto e^{-Z(t)}X(t)$ is a genuine $\mathbb{P}_x$-martingale. Let $0 \leqslant t_1 < t_2 \leqslant T$. In order to show the equality

$$\mathbb{E}_x^Z\left[X(t_2) \mid \mathcal{F}_{t_1}\right] = X(t_1) \tag{6.44}$$

it suffices to prove the equality

$$\mathbb{E}_x^Z\left[X(t_2)\mathbf{1}_A\right] = \mathbb{E}_x^Z\left[X(t_1)\mathbf{1}_A\right] \tag{6.45}$$

for all events $A \in \mathcal{F}_{t_1}$. Fix $A \in \mathcal{F}_{t_1}$. The equality in (6.45) can be achieved as follows:

$$\mathbb{E}_x^Z\left[X(t_2)\mathbf{1}_A\right] = \mathbb{E}_x\left[e^{-Z(T)}X(t_2)\mathbf{1}_A\right]$$

(the process $t \mapsto e^{-Z(t)}$ is a $\mathbb{P}_x$-martingale)

$$= \mathbb{E}_x\left[e^{-Z(t_2)}X(t_2)\mathbf{1}_A\right]$$

(the process $t \mapsto e^{-Z(t)}X(t)$ is a $\mathbb{P}_x$-martingale, the event A is $\mathcal{F}_{t_1}$-measurable)

$$= \mathbb{E}_x\left[e^{-Z(t_1)}X(t_1)\,\mathbf{1}_A\right]$$

(again use the fact that the process $t \mapsto e^{-Z(t)}$ is a $\mathbb{P}_x$-martingale)

$$= \mathbb{E}_x\left[e^{-Z(T)}X(t_1)\,\mathbf{1}_A\right]. \tag{6.46}$$

The equalities in (6.46) entail (6.45). Hence, (6.44) follows and therefore item (3) is proved.

(4) In order to prove the assertion in (4) we will employ Lévy's characterization of Brownian motion: see Theorem 6.5 below with the measure $\mathbb{P}_x^Z$ instead of $\mathbb{P}$. From assertion (3) it follows that the process $t \mapsto X(t)$ is a local $\mathbb{P}_x^Z$-martingale. It is also clear that $\langle X_{j_1}, X_{j_2}\rangle(t) = \delta_{j_1,j_2}t$, $1 \leqslant j_1,\, j_2 \leqslant d$. From Theorem 6.5 it then follows that the process $t \mapsto X(t)$ is a Brownian motion relative to $\mathbb{P}_x^Z$ on $(\Omega, \mathcal{F})$, where $\mathcal{F}$ is the σ-field generated by the process $t \mapsto W(t)$, $t \geqslant 0$.

(5) Let $Y = Y\left((X(s))_{0\leqslant s\leqslant T}\right)$ be a bounded stochastic variable which depends on the path $s \mapsto X(s)$, $0 \leqslant s \leqslant T$. Then by the observation in (4) we infer

$$\begin{aligned}\mathbb{E}_x[Y] &= \mathbb{E}_x\left[Y\left((X(s))_{0\leqslant s\leqslant T}\right)\right]\\ &= \mathbb{E}_x\left[e^{-Z(T)}e^{Z(T)}Y\left((X(s))_{0\leqslant s\leqslant T}\right)\right]\\ &= \mathbb{E}_x\left[e^{-Z(T)}e^{\int_0^T c(X(s))\,dX(s)-\frac{1}{2}\int_0^T |c(X(s))|^2\,ds}Y\left((X(s))_{0\leqslant s\leqslant T}\right)\right]\end{aligned}$$

(the process $t \mapsto X(t)$ is a Brownian motion relative to the measure $\mathbb{P}_x^Z$)

$$\begin{aligned}&= \mathbb{E}_x\left[e^{\int_0^T c(W(s))\,dW(s)-\frac{1}{2}\int_0^T |c(W(s))|^2\,ds}Y\left((W(s))_{0\leqslant s\leqslant T}\right)\right]\\ &= \mathbb{E}_x\left[e^{-\widetilde{Z}(T)}Y\left((W(s))_{0\leqslant s\leqslant T}\right)\right]\\ &= \mathbb{E}_x^{\widetilde{Z}}\left[Y\left((W(s))_{0\leqslant s\leqslant T}\right)\right].\end{aligned} \tag{6.47}$$

The equality in (6.47) shows assertion (5).

Altogether this completes the proof of Theorem 6.1. ☐

6.3. Remark. The equality in (6.47) is a version of Girsanov's theorem. If in (6.47) we take $Y = f(X(t))$, then we see

$$\mathbb{E}_x[f(X(t))] = \mathbb{E}_x\left[e^{-\widetilde{Z}(T)}f(W(t))\right]. \tag{6.48}$$

The equality in (6.48) is in agreement with the example below.

The following proposition was used in the proof of assertion (1) of Theorem 6.1 with $M(t) = \int_0^t c(X(s))\,dW(s)$.

6.4. Proposition. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space with filtration* $(\mathcal{F}_t)_{t\geqslant 0}$. *Let the process* $t \mapsto M(t)$ *be continuous local martingale. Put* $Z_M(t) = M(t) + \frac{1}{2}\langle M, M\rangle(t)$. *Suppose that for* $t \geqslant 0$ *the equality* $\mathbb{E}\left[e^{-Z_M(t)}\right] = 1$ *holds. Then the processt* $\mapsto e^{-Z_M(t)}$ *is a martingale.*

In what follows we write $t_1 \wedge t_2 = \min(t_1, t_2)$ whenever t_1 and t_2 are real numbers.

PROOF OF PROPOSITION 6.4. Fix $T > 0$, and choose an increasing sequence of stopping times $(T_n)_{n\in\mathbb{N}}$ such that $T_n \uparrow \infty$ and such that, for every $n \in \mathbb{N}$, the processes $t \mapsto M_n(t) := M(t \wedge T_n)$ and

$$t \mapsto \langle M_n, M_n \rangle (t) = \langle M, M \rangle (t \wedge T_n)$$

are uniformly bounded on a fixed interval $[0, T]$. Put

$$Z_n(t) = M(t \wedge T_n) + \frac{1}{2} \langle M, M \rangle (t \wedge T_n).$$

An application of Itô's formula yields:

$$e^{-Z_n(t)} - e^{-Z_n(0)} = -\int_0^t e^{-Z_n(s)}\, dM_n(s), \tag{6.49}$$

and hence, for each $n \in \mathbb{N}$, the process $t \mapsto e^{-Z_n(t)}$ is a martingale on the interval $[0, T]$. It follows that

$$\mathbb{E}\left[e^{-Z_n(t)}\right] = \mathbb{E}_x\left[e^{-Z_n(0)}\right] = 1 = \mathbb{E}\left[e^{-Z(t)}\right], \quad t \in [0, T]. \tag{6.50}$$

For $t \in [0, T]$ the sequence of non-negative stochastic variables

$$\left(e^{-Z_n(t)}\right)_{n\in\mathbb{N}} \subset L^1(\Omega, \mathcal{F}, \mathbb{P})$$

converges pointwise to $e^{-Z(t)}$. By Scheffé's theorem and by the equalities in (6.50) it follows that this convergence also takes place in the space $L^1(\Omega, \mathcal{F}, \mathbb{P})$. Consequently, the process $t \mapsto e^{-Z(t)}$ is a genuine martingale.

This completes the proof of Proposition 6.4. ☐

The following theorem is the same as Theorem 4.5 in the regular text.

6.5. THEOREM. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space with filtration (or reference system)* $(\mathcal{F}_t)_{t\geqslant 0}$. *Suppose* $\mathcal{F}$ *is the* σ*-algebra generated by* $\cup_{t\geqslant 0}\mathcal{F}_t$ *augmented with the* $\mathbb{P}$*-zero sets, and suppose* $\mathcal{F}_t$ *is continuous from the right:* $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$ *for all* $t \geqslant 0$. *Let*

$$\{M(t) = (M_1(t), \ldots, M_d(t)) : t \geqslant 0\}$$

be an $\mathbb{R}^d$*-valued local* $\mathbb{P}$*-almost surely continuous martingale with the property that the quadratic covariation processes* $t \mapsto \langle M_i, M_j \rangle (t)$ *satisfy*

$$\langle M_i, M_j \rangle (t) = \delta_{i,j} t, \quad 1 \leqslant i, j \leqslant d. \tag{6.51}$$

Then $\{M(t) : t \geqslant 0\}$ *is d-dimensional Brownian motion with initial distribution given by* $\mu(B) = \mathbb{P}[M(0) \in B]$, $B \in \mathcal{B}_{\mathbb{R}^d}$, *the Borel field of* $\mathbb{R}^d$.

The following result is Example 6 in Chapter 3. The idea of proof goes back to the method of Cameron-Martin [**37**]. Related material can be found in [**39**] and in [**38**]. More recent material can be consulted in [**86**] and [**87**].

Example. Let U be an open subset of $\mathbb{R}^d$, let f belong to $C_0(U)$ and let $c : U \to \mathbb{R}^d$ be an appropriate vector field on U and let $u : [0,\infty) \times U \to \mathbb{R}$ be a solution to the following problem:

$$\begin{cases} \dfrac{\partial u}{\partial t} & = \frac{1}{2}\Delta u + c.\nabla u \text{ in } [0,\infty)\times U; \\ u & \text{is continuous on } [0,\infty)\times U \text{ and } u(0,x) = f(x). \end{cases}$$

Moreover we want that $\lim_{x\to b, x\in U} u(t,x) = 0$ if b belongs to ∂U. Then

$$u(t,x) = \mathbb{E}_x\left[e^{-\widetilde{Z}(t)} f(W(t)) : \tau > t\right],$$

where $\widetilde{Z}(t) = \int_0^t c(W(r))\cdot dW(r) - \frac{1}{2}\int_0^t |c(W(r))|^2\, dr$. For a proof we fix $t > 0$ and we consider the process $\{M(s) : 0 \leqslant s \leqslant t\}$ defined by $M(s) = u\left(t-s, W(s)\right) e^{-\widetilde{Z}(s)} 1_{\{\tau > s\}}$. An application of Itô's formula to the function $f(s,x,y) = u(t-s,x)\exp(-y)$ will yield the following result

$$\begin{aligned} M(s) - M(0) &= -\int_0^s \frac{\partial u}{\partial t}(t-r, W(r)) e^{-\widetilde{Z}(r)} \mathbf{1}_{\{\tau>r\}} dr \\ &\quad + \int_0^s \nabla u(t-r, W(r)) e^{-\widetilde{Z}(r)} \mathbf{1}_{\{\tau>r\}} \cdot dW(r) \\ &\quad - \int_0^s u(t-r, W(r)) e^{-\widetilde{Z}(r)}\, d\widetilde{Z}(r) \\ &\quad + \frac{1}{2}\int_0^s \Delta u(t-r, W(r)) e^{-\widetilde{Z}(r)} \mathbf{1}_{\{\tau>r\}} \mathbf{1}_{\{\tau>r\}}\, dr \\ &\quad - \sum_{j=1}^d \int_0^s D_j u(t-r, W(r)) e^{-\widetilde{Z}(r)} \mathbf{1}_{\{\tau>r\}} d\left\langle W_j, \widetilde{Z}\right\rangle(r) \\ &\quad + \frac{1}{2}\int_0^s u(t-r, W(r)) e^{\widetilde{Z}(r)} \mathbf{1}_{\{\tau>r\}}\, d\left\langle \widetilde{Z}, \widetilde{Z}\right\rangle(r) \\ &= \int_0^s \left\{ -\frac{\partial u}{\partial t}(t-r, W(r)) + \frac{1}{2}\Delta u(t-r, W(r)) \right. \\ &\qquad \left. + c(W(r)).\nabla u(t-r, W(r)) \right\} e^{-\widetilde{Z}(r)} \mathbf{1}_{\{\tau>r\}} dr \\ &\quad + \int_0^s \left\{ \nabla u(t-r, W(r)) + u(t-r, W(r)) c(W(r)) \right\} \\ &\qquad e^{-\widetilde{Z}(r)} \mathbf{1}_{\{\tau>r\}} \cdot dW(r) \\ &\quad - \frac{1}{2}\int_0^s u(t-r, W(r)) e^{-\widetilde{Z}(r)} \left|c(W(r))\right|^2 \mathbf{1}_{\{\tau>r\}}\, dr \\ &\quad + \frac{1}{2}\int_0^s u(t-r, W(r)) e^{-\widetilde{Z}(r)} \left|c(W(r))\right|^2 \mathbf{1}_{\{\tau>r\}} dr \\ &= \int_0^s \left\{ \nabla u(t-r, W(r)) + u(t-r, W(r)) c(W(r)) \right\} \\ &\qquad e^{-\widetilde{Z}(r)} \mathbf{1}_{\{\tau>r\}} \cdot dW(r). \end{aligned}$$

It will follow that $u(t,x) = \mathbb{E}_x\left[e^{-\tilde{Z}(t)} f(W(t)) : \tau > t\right]$. Also notice that the previous equalities we could have written $\int_0^{\min(\tau,s)} \dots dr$ instead of $\int_0^s \dots \mathbf{1}_{\{\tau > r\}}\, dr$.

In the following theorem we present a result on the Markov process generated by the operator

$$Lf(x) = c(x) \cdot \nabla f(x) + \frac{1}{2} \Delta f(x).$$

We suppose that its domain

$$D(L) := \left\{ f \in C^2\left(\mathbb{R}^d\right),\ f,\ Lf \in C_0(E),\ \nabla f \text{ uniformly bounded} \right\}$$

is dense in $C_0\left(\mathbb{R}^d\right)$.

6.6. Theorem. *Suppose that for every $x \in \mathbb{R}^d$ there exists a Brownian motion with state space $\mathbb{R}^d$*

$$\left\{(W(t))_{t\geqslant 0}, \mathbb{P}\right\}, \quad W(0)=0,$$

together with an adapted $\mathbb{R}^d$-valued process $(X^x(t))_{t\geqslant 0}$ such that

$$X^x(t) = x + W(t) + \int_0^t c\left(X^x(s)\right) dW(s) \tag{6.52}$$

Define, for $x \in \mathbb{R}^d$, the process $t \mapsto Z(x,t)$ by

$$Z(x,t) = \int_0^t c\left(X^x(s)\right) dW(s) + \frac{1}{2}\int_0^t \left|c\left(X^x(s)\right)\right|^2 ds. \tag{6.53}$$

and assume that $\mathbb{E}\left[e^{-Z(x,t)}\right] = 1$ for all $t \geqslant 0$. In addition, put

$$\widetilde{Z}(x,t) = -\int_0^t c\left(x+W(s)\right) dW(s) + \frac{1}{2}\int_0^t \left|c\left(x+W(s)\right)\right|^2 ds. \tag{6.54}$$

Then the operator L is closable and its closure generates a strong Markov process

$$\left\{(\Omega, \mathcal{F}, \mathbb{P}_x), (X(t), t\geqslant 0), (\vartheta_t, t \geqslant 0), \left(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d}\right)\right\} \tag{6.55}$$

with the property that, for all bounded Borel measurable functions $f : \mathbb{R}^d \to \mathbb{C}$, the equality

$$\mathbb{E}_x\left[f\left(X(t)\right)\right] = \mathbb{E}\left[e^{-\widetilde{Z}(x,t)} f\left(x+W(t)\right)\right] \tag{6.56}$$

holds. In particular it follows that the stochastic differential equation in (6.53) has unique weak solutions.

6.7. Remark. It can be proved that a process like in (6.52) exists whenever the coefficient $x \mapsto c(x)$ is bounded and continuous. In this case the process $t \mapsto e^{-Z(t)}$ is a genuine martingale.

Proof. Fix $T > 0$ and $x \in \mathbb{R}^d$. From assertion (5) in Theorem 6.1 it follows that the process $t \mapsto Z(x,t)$, $t \in [0,T]$, has a distribution determined by

$$\mathbb{E}\left[Y\left(\left(X^x(s)\right)_{0\leqslant s\leqslant T}\right)\right] = \mathbb{E}\left[e^{-\widetilde{Z}(x,T)} Y\left(\left(x+W(s)\right)_{0\leqslant s\leqslant T}\right)\right]. \tag{6.57}$$

Here $Y : C\left([0,T], \mathbb{R}^d\right) \to \mathbb{C}$ is a bounded measurable function. Let $\Omega = C\left([0,\infty), \mathbb{R}^d\right)$, and define $X(t) : \Omega \to \mathbb{R}^d$ by $X(t)(\omega) = \omega(t)$. Define the probability measures $\mathbb{P}_x$, $x \in \mathbb{R}^d$, by the formula in (6.57), i.e.

$$\mathbb{E}_x\left[Y\left(\left(X(s)\right)_{0\leqslant s\leqslant T}\right)\right] = \mathbb{E}\left[e^{-\widetilde{Z}(x,T)} Y\left(\left(x+W(s)\right)_{0\leqslant s\leqslant T}\right)\right]. \tag{6.58}$$

In addition we write $\vartheta_t\omega(s) = \omega(s+t)$. We will show that the process

$$\left\{(\Omega, \mathcal{F}, \mathbb{P}_x), (X(t), t\geqslant 0), (\vartheta_t, t \geqslant 0), \left(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d}\right)\right\} \tag{6.59}$$

constitutes a Markov process. Since, by definition, the paths are continuous it is then a strong Markov process. Let f be an arbitrary bounded Borel measurable function on $\mathbb{R}^d$. In order to prove that the process in (6.59) is a Markov process we need to prove the equality

$$\begin{aligned}&\mathbb{E}_x\left[F\left(X(t_1), \dots, X(t_n)\right) f\left(X(t+s)\right)\right] \\ &\quad = \mathbb{E}_x\left[F\left(X(t_1), \dots, X(t_n)\right) \mathbb{E}_{X(t)}\left[f\left(X(s)\right)\right]\right],\end{aligned} \tag{6.60}$$

where $F : \left(\mathbb{R}^d\right)^n \to \mathbb{C}$ is any bounded measurable function on $\left(\mathbb{R}^d\right)^n$ with $0 \leqslant t_1 < \cdots < t_n \leqslant t$ for arbitrary $n \in \mathbb{R}^d$ and arbitrary t and s positive real numbers. With such data we have

$$\begin{aligned}
&\mathbb{E}_x\left[F\left(X\left(t_1\right),\dots,X\left(t_n\right)\right)\mathbb{E}_{X(t)}\left[f\left(X(s)\right)\right]\right]\\
&= \mathbb{E}\left[e^{-\widetilde{Z}(x,t)}F\left(x+W\left(t_1\right),\dots,x+W\left(t_n\right)\right)\mathbb{E}_{x+W(t)}\left[f\left(X(s)\right)\right]\right]\\
&= \mathbb{E}\left[e^{-\widetilde{Z}(x,t)}F\left(x+W\left(t_1\right),\dots,x+W\left(t_n\right)\right)\right.\\
&\qquad\left.\mathbb{E}\left[e^{-\widetilde{Z}(x+W(t),s)}f\left(x+W(t)+W(s)\right)\bigm| W(t)\right]\right]. \qquad (6.61)
\end{aligned}$$

Using the Markov property of Brownian motion we may rewrite

$$\mathbb{E}\left[e^{-\widetilde{Z}(x+W(t),s)}f\left(x+W(t)+W(s)\right)\bigm| W(t)\right]$$

as follows:

$$\begin{aligned}
&\mathbb{E}\left[e^{-\widetilde{Z}(x+W(t),s)}f\left(x+W(t)+W(s)\right)\bigm| W(t)\right]\\
&= \mathbb{E}\left[e^{\int_0^s c(x+W(t)+W(\rho))\,dW(\rho)-\frac{1}{2}\int_0^s|c(x+W(t)+W(\rho))|^2\,d\rho}\right.\\
&\qquad\left. f\left(x+W(t)+W(s)\right)\bigm| W(t)\right]
\end{aligned}$$

(for given $W(t)$ the distribution of $\rho \mapsto W(\rho)$ is the same as that of $\rho \mapsto W\left(t+\rho\right)-W(t)$)

$$\begin{aligned}
&= \mathbb{E}\left[e^{\int_0^s c(x+W(t+\rho))\,dW(t+\rho)-\frac{1}{2}\int_0^s|c(x+W(t+\rho))|^2\,d\rho}f\left(x+W(t+s)\right)\bigm| W(t)\right]\\
&= \mathbb{E}\left[e^{\int_t^{t+s} c(x+W(\rho))\,dW(\rho)-\frac{1}{2}\int_t^{t+s}|c(x+W(\rho))|^2\,d\rho}f\left(x+W(t+s)\right)\bigm| W(t)\right]
\end{aligned}$$

(Brownian motion is a Markov process)

$$= \mathbb{E}\left[e^{\int_t^{t+s} c(x+W(\rho))\,dW(\rho)-\frac{1}{2}\int_t^{t+s}|c(x+W(\rho))|^2\,d\rho}f\left(x+W(t+s)\right)\bigm| \mathcal{F}_t^W\right] \qquad (6.62)$$

where $\mathcal{F}_t^W = \sigma\left(W(\rho), 0 \leqslant \rho \leqslant t\right)$. Inserting the equality in (6.62) into (6.61) yields

$$\begin{aligned}
&\mathbb{E}_x\left[F\left(X\left(t_1\right),\dots,X\left(t_n\right)\right)\mathbb{E}_{X(t)}\left[f\left(X(s)\right)\right]\right]\\
&= \mathbb{E}\left[e^{\int_0^t c(X(\rho))\,dW(\rho)-\frac{1}{2}\int_0^t|c(X(\rho))|^2\,d\rho}F\left(x+W\left(t_1\right),\dots,x+W\left(t_n\right)\right)\right.\\
&\qquad\left.\mathbb{E}\left[e^{\int_t^{t+s} c(x+W(\rho))\,dW(\rho)-\frac{1}{2}\int_t^{t+s}|c(x+W(\rho))|^2\,d\rho}f\left(x+W(t+s)\right)\bigm| \mathcal{F}_t^W\right]\right]\\
&= \mathbb{E}\left[\mathbb{E}\left[e^{\int_0^{t+s} c(X(\rho))\,dW(\rho)-\frac{1}{2}\int_0^{t+s}|c(X(\rho))|^2\,d\rho}\right.\right.\\
&\qquad\left.\left.F\left(x+W\left(t_1\right),\dots,x+W\left(t_n\right)\right)f\left(x+W(t+s)\right)\bigm| \mathcal{F}_t^W\right]\right]
\end{aligned}$$

(elementary property of conditional expectation)

$$\begin{aligned}
&= \mathbb{E}\left[e^{-\widetilde{Z}(x,t+s)}F\left(x+W\left(t_1\right),\dots,x+W\left(t_n\right)\right)f\left(x+W(t+s)\right)\right]\\
&= \mathbb{E}_x\left[F\left(x+W\left(t_1\right),\dots,x+W\left(t_n\right)\right)f\left(X\left(s+t\right)\right)\right]. \qquad (6.63)
\end{aligned}$$

The equality in (6.60) is a consequence of (6.63), which in turn implies the Markov property of the process in (6.55). Let $f \in D(L)$ be such that the vector valued function $x \mapsto \nabla f(x)$ is bounded. Then by Itô calculus we have

$$\begin{aligned} & f\left(X^x(t)\right) - f\left(X^x(0)\right) \\ & = \int_0^t \nabla f\left(X^x(s)\right)\, dX^x(s) + \frac{1}{2}\int_0^t \Delta f\left(X(s)\right)\, ds \\ & = \int_0^t Lf\left(X^x(s)\right)\, ds + \int_0^t \nabla f\left(X^x(s)\right)\, dW(s). \end{aligned} \tag{6.64}$$

Taking expectations in (6.64), dividing by $t > 0$, and taking the limit for $t \downarrow 0$ shows that the (closure of the) operator L generates the Markov process in (6.55).

This completes the proof of Theorem 6.6. ☐

Assume that the $\mathbb{R}^d$-valued function $c : [0,T] \times \mathbb{R}^d$ and the $\mathbb{R}^d \times \mathbb{R}^d$-valued function $\sigma = (\sigma_{i,j})_{1\leqslant i\leqslant d, 1\leqslant j\leqslant n} : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^n$ are continuous and have at most linear growth. The latter means that there exist finite constants C_1 and C_2 such that

$$|c(t,x)| + |\sigma(t,x)| \leqslant C_1 + C_2 |x|, \quad 0 \leqslant t \leqslant T, \ x \in \mathbb{R}^d. \tag{6.65}$$

Put $a(t,x) = \sigma(t,x)\,\sigma(t,x)^*$ and assume that

$$\sum_{i,j=1}^d a_{i,j}(t,x) y_i y_j > 0, \quad \text{for all } x \in \mathbb{R}^d \text{ and } y = (y_1, \dots, y_d) \neq 0. \tag{6.66}$$

What follows is mainly taken from Klebaner [**92**] Section 10 in Chapter 6. Some of these results are in fact based on work by Revuz and Yor [**149**] and Stroock and Varadhan [**171**]. Let the process $t \mapsto X(t)$ be a diffusion in d dimensions, described by the multi-dimensional SDE

$$dX(t) = c\left(t, X(t)\right)\, dt + \sigma\left(t, X(t)\right)\, dB(t), \tag{6.67}$$

where σ is an $n \times d$ matrix valued function, $t \mapsto B(t)$ is n-dimensional Brownian motion, $t \mapsto X(t)$ and $(t,x) \mapsto c(t,x)$ are d-dimensional vector valued functions. In coordinate form this reads

$$dX_i(t) = c_i\left(t, X(t)\right)\, dt + \sum_{j=1}^d \sigma_{i,j}\left(t, X(t)\right)\, dB_j(t), \quad i = 1, \dots, d, \tag{6.68}$$

and it means that for all $t > 0$ and $i = 1, \dots, d$

$$X_i(t) = X_i(0) + \int_0^t c_i\left(s, X(s)\right)\, ds + \sum_{j=1}^n \int_0^t \sigma_{i,j}\left(s, X(s)\right)\, dB_j(s). \tag{6.69}$$

The coefficients of the SDE are: the vector $c(t,x)$ and the matrix $\sigma(t,x)$. An existence and uniqueness result for strong solutions, under the assumption of locally Lipschitz coefficients, holds in the following form. The norm of the vector $c(t,x)$ is its (Euclidean) length, $|c(t,x)| = \sqrt{\sum_{i=1}^n c_i(t,x)^2}$. The norm of the matrix $\sigma(t,x)$ is defined by $|\sigma(t,x)|^2 = \operatorname{trace}\left(\sigma(t,x)\sigma^*(t,x)\right) = \operatorname{trace}\left(a(t,x)\right)$, with σ^* being the transposed of σ. The trace of the matrix $a(t,x)$ is defined by $\operatorname{trace}\left(a(t,x)\right) = \sum_{i=1}^{d} a_{ii}(t,x)$. The matrix $a(t,x) = \sigma(t,x)\sigma^*(t,x)$ is called the diffusion matrix.

6.8. THEOREM. *If the coefficients are locally Lipschitz in x with a constant independent of t, that is, for every N, there is a constant K depending only on T and N such that for all $|x|$, $|y| \leqslant N$ and all $0 \leqslant t \leqslant T$*

$$|c(t,x) - c(t,y)| + |\sigma(x,t) - \sigma(y,t)| \leqslant K|x-y|, \tag{6.70}$$

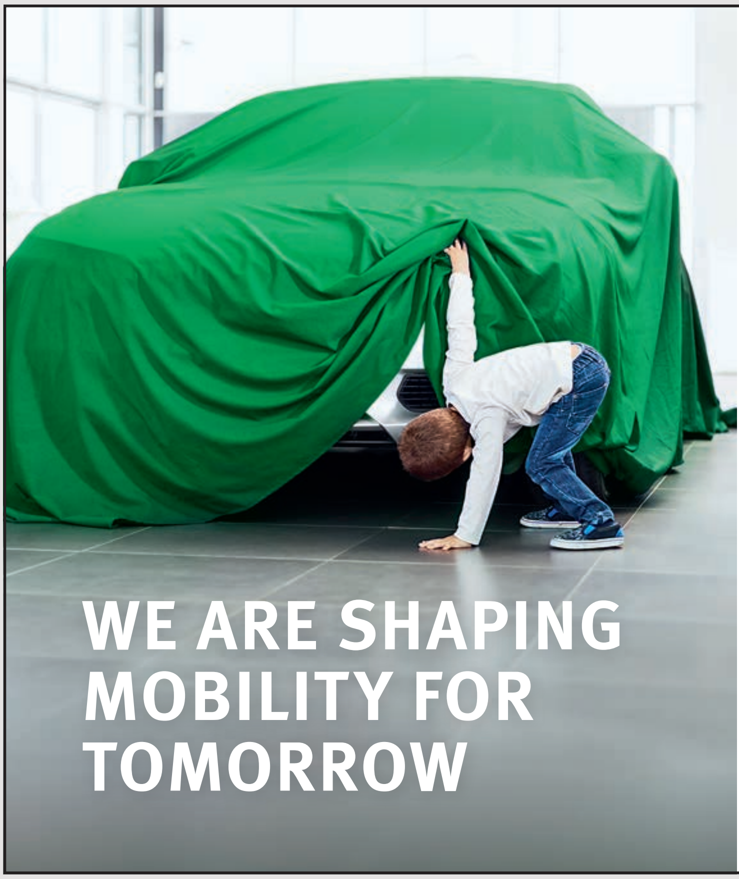
then for any given $X(0)$ the strong solution to SDE (6.67) is unique. If in addition to condition (6.70) the linear growth condition holds

$$|c(t,x)| + |\sigma(t,x)| \leqslant K_T(1+|x|),$$

$X(0)$ is independent of the Brownian motion $B(t)$, and $\mathbb{E}\left[|X(0)|^2\right] < \infty$, then the strong solution exists and is unique on $[0,T]$, moreover

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} |X(t)|^2\right] \leqslant C\left(1 + \mathbb{E}\left[|X(0)|^2\right]\right), \tag{6.71}$$

where the constant C depends only on K_T and T.



Note that unlike in the one-dimensional case, the Lipschitz condition on σ can not be weakened in general to a Hölder condition, i.e. there is no Yamada-Watanabe-type result for multi-dimensional SDEs. The quadratic covariation is easy to work out from (6.68), by taking into account that independent Brownian motions have zero quadratic covariation.

$$\langle X_i, X_j \rangle (t) = \int_0^t dX_i(s)\, dX_j(s) = \int_0^t a_{i,j} \left(s, X(s)\right) ds. \tag{6.72}$$

It can be shown that if $t \mapsto X(t)$ is a solution to (6.67) then

$$\begin{aligned}
&\mathbb{E}\left[X_i(t+\triangle) - x_i \mid X(t) = x\right] = c_i(t,x)\triangle + o(\triangle),\\
&\mathbb{E}\left[\left(X_i(t+\triangle) - x_i\right)\left(X_j(t+\triangle) - x_j\right) \mid X(t) = x\right]\\
&= a_{i,j}(t,x)\triangle + o(\triangle),
\end{aligned}$$

as $\triangle \to 0$. Thus $c(t,x)$ is the coefficient in the infinitesimal mean of the displacement from point x at time t, and $a(t,x)$ is approximately the coefficient in the infinitesimal covariance of the displacement. Weak solutions can be defined as solutions to the martingale problem. Let the operator L_t, acting on twice continuously differentiable functions from $\mathbb{R}^d$ to $\mathbb{R}$, be given by

$$L_t = \sum_{i=1}^d c_i(t,x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^d a_{i,j}(t,x)\frac{\partial^2}{\partial x_i \partial x_j}. \tag{6.73}$$

Note that L_t depends on σ only through a. Then $X(t)$ is a weak solution started at x at time s with respect $\mathbb{P}_{s,x}$ and a filtration based on the process $t \mapsto X(t)$, provided the process

$$t \mapsto f(X(t)) - \int_s^t \left(L_\rho f\right)(X(\rho))\, d\rho \tag{6.74}$$

is a martingale for any twice continuously differentiable function f vanishing outside a compact set in $\mathbb{R}^d$. This process is called a diffusion with generator L_t. In the case of time-independent coefficients, the process is a time-homogeneous diffusion with generator L.

6.9. THEOREM. *Assume that $a(t,x)$ is continuous and satisfies condition (6.66), i.e. $\sum_{i,j}^d a_{i,j}(t,x)\, y_i y_j > 0$, for all $x \in \mathbb{R}^d$ and $y \neq= 0$ and $c(t,x)$ is bounded on bounded sets. Then there exists a unique weak solution up to the time of explosion to the equation in (6.69). If, in addition, a quadratic growth condition is satisfied, that is, for any $T > 0$ there is a constant K_T such that for all $x \in \mathbb{R}^d$*

$$|c(t,x)|^2 + |a(t,x)| \leqslant K_T\left(1 + |x|^2\right), \tag{6.75}$$

then there exists a unique weak solution to the martingale problem (6.74) starting at any point $x \in \mathbb{R}^d$ at any time $s \geqslant 0$, moreover this solution has the strong Markov property.

Here the time of explosion τ_∞ is explained as follows. Let, for $n = 1, 2, \ldots$, $B_n = \left\{x \in \mathbb{R}^d : |x| < n\right\} = nB_1$ be the open ball of radius n in $\mathbb{R}^d$. The stopping time $\tau_n = \tau_{B_n} = \inf\left\{s > 0 : |X(s)| \geqslant n\right\}$ is the first time the process exceeds or equals length n. Since a diffusion process is continuous, it must reach level n

before it reaches level $n+1$. Therefore the times τ_n are non-decreasing, hence they converge to a limit $\tau_\infty = \lim_{n\to\infty} \tau_n$. Explosion occurs on the set $\{\tau_\infty < \infty\}$, because on this set, by continuity of $t \mapsto X(t)$, $X(\tau_\infty) = \lim_{n\to\infty} X(\tau_n)$. Thus

$$|X(\tau_\infty)| = \lim_{n\to\infty} |X(\tau_n)| = \lim_{n\to\infty} n = \infty,$$

and infinity is reached in finite time on this set.

Since the weak solution is defined in terms of the generator, which itself depends on $\sigma(t,x)$ only through $a(t,x)$, the weak solution to (6.67) can be constructed using a single d-valued Brownian motion provided the matrix $a(t,x)$ remains the same. If a single SDE is equivalent to a number of SDEs, heuristically, it means that there is as much randomness in a d-dimensional Brownian motion as there is in a single Brownian motion. Replacement of a system of SDEs by a single one is shown in detail for the Bessel process. Note that the equation $\sigma(t,x)\sigma^*(t,x) = a(t,x)$ has many solutions for $\sigma(t,x)$, the matrix square root is non-unique. However, if $a(t,x)$ is non-negative definite for all x and t, and has for entries twice continuously differentiable functions of $x \in \mathbb{R}^d$ and t, then it has a locally Lipschitz square root $\sigma(t,x)$ of the same dimension as $a(t,x)$ (see for example Friedman [**73**] (1975 Theorem 6.1.2 in Volume 1).

The following theorem gives a nice result for unique weak solutions to stochastic integral equation of the form

$$X(t) = X(0) + \int_0^t c(s, X(s))\, ds + \int_0^t \sigma(s, X(s))\, dW(s) \tag{6.76}$$

with $t \in [0,T]$. A weak solution to this equation is an $\mathbb{R}^d$-valued process $t \mapsto X(t)$, $t \in [0,T]$, on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a Brownian motion $t \mapsto W(t)$, $t \in [0,T]$. A unique weak solution means that its distribution does not depend on the particular Brownian motion for which it satisfies (6.76). More details are exhibited in the regular text. A more general version, where the coefficients $c(t,x)$ and $\sigma(t,x)$ have certain discontinuities, can be found in [**195**] Theorem 3.1. In the following theorem we suppose that the equation in (6.76) possesses unique weak solutions. A proof is based, among other things, on the Burkholder-Davis-Gundy inequality: see Theorem 6.16 below.

6.10. THEOREM. *If $\mathbb{E}\left[|X(0)|^4\right] < \infty$, $\lim_{n\to\infty} \triangle_n = 0$ and c and σ have at most linear growth, then the Euler scheme defined in (6.77) weakly converges to the unique weak solution of the equation in (6.76) as $n \to \infty$.*

Here the (continuous) Euler scheme for the equation (6.76) is defined as follows: $X^n(0) = X(0)$, and

$$\begin{aligned} X^n(t) &= X^n(\tau_k^n) + c(\tau_k^n, X^n(\tau_k^n))(t - \tau_k^n) \\ &\quad + \sigma(\tau_k^n, X^n(\tau_k^n))(W(t) - W(\tau_k^n)), \end{aligned} \tag{6.77}$$

for $\tau_k^n < t \leqslant \tau_{k+1}^n$, $k = 0, 1, \dots, n$, where $0 = \tau_0^n \leqslant \tau_1^n \leqslant \cdots \leqslant \tau_n^n = T$ is a sequence of random or deterministic partitions of $[0,T]$. The times τ_k^n should be stopping times. For example $\tau_k^n = k2^{-n}T$, $0 \leqslant k \leqslant 2^n$. We also wrote

$\triangle_n = \max_{0 \leqslant k \leqslant n-1} \left|\tau^n_{k+1} - \tau^n_k\right|$. If we define $\overline{X}^n(t) = X^n\left(\tau^n_k\right)$, for $\tau^n_k \leqslant t < \tau^n_{k+1}$, then $\left\{\overline{X}^n(t) : 0 \leqslant t \leqslant T\right\}$ is called a discretized Euler scheme. By defining

$$\eta^n(s) = \tau^n_k \quad \text{for} \quad \tau^n_k < s \leqslant \tau^n_{k+1}, \; 0 \leqslant k \leqslant n-1, \tag{6.78}$$

we have

$$\begin{aligned} X^n(t) = X(0) &+ \int_0^t c\left(\eta^n(s), X^n\left(\eta^n(s)\right)\right) ds \\ &+ \int_0^t \sigma\left(\eta^n(s), X^n\left(\eta^n(s)\right)\right) dW(s). \end{aligned} \tag{6.79}$$

If $\tau^n_k = k2^{-n}T$, then $\eta^n(s) = 2^{-n}T\left\lfloor 2^n s/T\right\rfloor$. The convergence in Theorem 6.10 is meant in the following sense. There exists a probability space $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ with a Brownian motion $t \mapsto \widetilde{B}(t)$, $0 \leqslant t \leqslant T$, such that

$$\lim_{n\to\infty} \mathbb{E}\left[F\left(X^n\right)\right] = \widetilde{\mathbb{E}}\left[F\left(\widetilde{X}\right)\right]$$

for all bounded continuous functions $F : C\left([0,T], \mathbb{R}^d\right) \to \mathbb{R}$. Here the process $t \mapsto \widetilde{X}(t)$, $0 \leqslant t \leqslant T$, satisfies:

$$\widetilde{X}(t) = \widetilde{X}(0) + \int_0^t c\left(s, \widetilde{X}(s)\right) ds + \int_0^t \sigma\left(s, \widetilde{X}(s)\right) d\widetilde{B}(s). \tag{6.80}$$

In fact for $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ we take the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the original Brownian motion $t \mapsto B(t)$, $0 \leqslant t \leqslant T$.

For the proof of Theorem 6.10 we need some lemmas.

6.11. Lemma. *If* $\mathbb{E}\left[|X(0)|^4\right] < \infty$, *and the coefficients* $c(t,x)$ *and* $\sigma(t,x)$ *have at most linear growth, that is, there exist two constants* C_1 *and* C_2 *such that, for all* $t \in [0,T]$ *and* $x \in \mathbb{R}^d$, $|c(t,x)| + |\sigma(t,x)| \leqslant C_1 + C_2|x|$, *then*

$$\sup_{n\geqslant 1} \mathbb{E}\left[|(X^{n,*}(T))|^4\right] < \infty \quad \textit{and} \quad \sup_{n\geqslant 1} E\left[|\langle X^n, X^n\rangle (T)|^2\right] < \infty, \tag{6.81}$$

where $|\cdot|$ *stands for the Euclidean norm in the appropriate space.*

Here $|(X^{n,*}(T))| = \sup_{0\leqslant t\leqslant T} |X^n(t)|$.

Proof. Define the sequence of $\eta^n(s)$, $n \in \mathbb{N}$, as in (6.78). Then the processes $X^n(t)$, $n \in \mathbb{N}$, satisfy (6.79). By the inequality $(x+y+z)^4 \leqslant 27\left(x^4+y^4+z^4\right)$ for any real numbers x, y, z and Hölder's inequality together with Burkholder-Davis-Gundy's inequality for $p = 4$, there exist two positive constants a and b such that

$$\mathbb{E}\left[|X^n(t)|^4\right] \leqslant a + b\int_0^t \mathbb{E}\left[|X^n\left(\eta^n(s)\right)|^4\right] ds. \tag{6.82}$$

Put $f_n(t) = \mathbb{E}\left[|X^n\left(\eta^n(t)\right)|^4\right]$. Then by (6.82), and since $\eta^n(t) \leqslant t$, we see $f_n(t) \leqslant a + b\int_0^t f_n(s)\,ds$. By Gronwall's lemma (see main lecture notes), we infer $f_n(t) \leqslant ae^{bt}$ for $0 \leqslant t \leqslant T$. By (6.82) and Burkholder-Davis-Gundy's inequality the assertions in Lemma 6.11 follow. ☐

6.12. Lemma. *If the conditions of Theorem 6.10 are satisfied, then the Euler scheme* $\{X^n : n \geqslant 1\}$ *is tight in* $C\left([0,T],\mathbb{R}^d\right)$.

Proof. We will use the criterion in Theorem 8.3 on page 56 of Billingsley [**22**] to prove the tightness. First, since $\mathbb{E}\left[|X(0)|^4\right] < \infty$, the sequence $\{X^n(0) = X(0) : n \geqslant 1\}$ is tight. Secondly, put

$$\begin{aligned} c_1(T) &= \sup_{n\geqslant 1}\mathbb{E}\left[\sup_{0\leqslant s\leqslant T} |c\left(\eta^n(s), X^n\left(\eta^n(s)\right)\right)|^4\right], \\ c_2(T) &= \sup_{n\geqslant 1}\mathbb{E}\left[\sup_{0\leqslant s\leqslant T} |\sigma\left(\eta^n(s), X^n\left(\eta^n(s)\right)\right)|^4\right]. \end{aligned}$$

Since $c(t,x)$ and $\sigma(t,x)$ have at most linear growth in x, the constants $c_1(T)$ and $c_2(T)$ are finite by Lemma 6.11. For a (fixed) random time $t \in [0,T]$, let

$$\begin{aligned} Z_t^n(s) &= X^n(s+t) - X^n(t) \\ C_t^n(s) &= \int_t^{t+s} c\left(\eta^n(\rho), X^n\left(\eta^n(\rho)\right)\right) d\rho, \\ M_t^n(s) &= \int_t^{t+s} \sigma\left(\eta^n(\rho), X^n\left(\eta^n(\rho)\right)\right) dW(\rho), \quad \text{so that} \\ Z_t^n(s) &= C_t^n(s) + M_t^n(s). \end{aligned}$$

Since, for x and y real numbers, $(x+y)^4 \leqslant 8\left(x^4 = y^4\right)$, we have

$$|Z_t^n(s)|^4 \leqslant 8\,|C_t^n(s)|^4 + 8\,|M_t^n(s)|^4. \tag{6.83}$$

By the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\leqslant t\leqslant T-s}\left|\left\langle Z_t^n, Z_t^n\right\rangle(s)\right|^2\right] \\
&= \mathbb{E}\left[\sup_{0\leqslant t\leqslant T-s}\left(\int_t^{t+s}\left|\sigma\left(\eta^n(\rho), X^n\left(\eta^n(\rho)\right)\right)\right|^2 d\rho\right)^2\right] \\
&\leqslant s\mathbb{E}\left[\sup_{0\leqslant t\leqslant T-s}\int_t^{t+s}\left|\sigma\left(\eta^n(\rho), X^n\left(\eta^n(\rho)\right)\right)\right|^4 d\rho\right] \leqslant c_2(T)s^2. \qquad (6.84)
\end{aligned}$$

Since the process

$$s \mapsto M_t^n(s) = \int_t^{t+s} \sigma\left(\eta^n(\rho), X^n\left(\eta^n(\rho)\right)\right)\, dW(\rho)$$

is a continuous martingale for which

$$\left\langle M_t^n, M_t^n\right\rangle(s) = \left\langle Z_t^n, Z_t^n\right\rangle(s)$$

by Burkholder's inequality and by (6.84) there exists a constant C_4 such that

$$\mathbb{E}\left[\sup_{0\leqslant s\leqslant\delta}\left|M_t^n(s)\right|^4\right] \leqslant C_4\mathbb{E}\left[\left|\left\langle Z_t^n, Z_t^n\right\rangle(\delta)\right|^2\right] \leqslant C_4 c_2(T)\delta^2,$$

for $\delta > 0$. Next, for every $\varepsilon > 0$ and $\eta > 0$ there exists δ, which does not depend on t and n, such that, for some random time $0 \leqslant \tau \leqslant T-\delta$,

$$\begin{aligned}
&\mathbb{P}\left[\sup_{0\leqslant t\leqslant T-\delta}\sup_{t\leqslant s\leqslant t+\delta}\left|X^n(s)-X^n(t)\right| \geqslant \varepsilon\right] \\
&\leqslant \frac{T}{\delta}\sup_{0\leqslant t\leqslant T-\delta}\mathbb{P}\left[\sup_{t\leqslant s\leqslant t+\delta}\left|X^n(s)-X^n(t)\right| \geqslant \varepsilon\right] \\
&\leqslant \frac{T}{\delta\varepsilon^4}\sup_{0\leqslant t\leqslant T-\delta}\mathbb{E}\left[\sup_{0\leqslant s\leqslant\delta}\left|Z_t^n(s)\right|^4\right] \\
&\leqslant \frac{8T}{\delta\varepsilon^4}\left\{\mathbb{E}\left[\sup_{0\leqslant s\leqslant\delta}\left|C_t^n(s)\right|^4\right] + \mathbb{E}\left[\sup_{0\leqslant s\leqslant\delta}\left|M_t^n(s)\right|^4\right]\right\} \\
&\leqslant \frac{8T}{\delta\varepsilon^4}\left\{\delta^4 c_1(T) + \delta^2 C_4 c_2(T)\right\} \leqslant \eta. \qquad (6.85)
\end{aligned}$$

In fact, the choice $\delta = \min\left(1, \dfrac{\varepsilon^4}{8T\left(c_1(T)+C_4c_2(T)\right)}\right)$ will achieve the final inequality in (6.85). From Arzela-Ascoli's theorem it follows that the sequence $\{X^n, n \geqslant 1\}$ is relatively compact in the Banach space $C\left([0,T], \mathbb{R}^d\right)$ which is separable. Prohorov's theorem then implies that the sequence $\{X^n, n \geqslant 1\}$ is tight in $C\left([0,T], \mathbb{R}^d\right)$: see Theorem 6.21. So the proof Lemma 6.12 is complete now. ☐

6.13. Remark. Since, for some random time $0 \leqslant \tau \leqslant T-\delta$, the equality

$$\sup_{0\leqslant t\leqslant T-\delta}\sup_{t\leqslant s\leqslant t+\delta}\left|X^n(s)-X^n(t)\right| = \sup_{\tau\leqslant s\leqslant\tau+\delta}\left|X^n(s)-X^n(\tau)\right|$$

holds $\mathbb{P}$-almost surely, it could be that the inequalities in (6.85) could be sharpened somewhat.

In the proof below the notion $\mathcal{L}(Y)$ means the law or distribution of the process Y. The proof below is designed for an equation like the in (6.76) with discontinuities of a certain kind. In our case, when the coefficients are continuous, a shorter proof can be found at the end of these notes.

Proof of Theorem 6.10. Since the Banach space $C\left([0,T],\mathbb{R}^d\right)$ is separable, Prohorov's theorem in conjunction with Lemma 6.12 then implies that the sequence $\{X^n,\, n \geqslant 1\}$ is tight in $C\left([0,T],\mathbb{R}^d\right)$. Thus for any subsequence of the sequence $\{X^n,\, n \geqslant 1\}$ there exists a further subsequence $\{X^{n_k},\, k \geqslant 1\}$ and a process $t \mapsto X(t)$, $t \in [0,T]$, in $C\left([0,T],\mathbb{R}^d\right)$ such that the sequence $\{X^{n_k},\, k \geqslant 1\}$ converges to the process $X(t)$ weakly. Let $\mathbb{P}^n$ be the distribution (or law) of the process X^n. Since the equation in (6.52) is supposed to have unique weak solutions, it follows that the sequence $\{\mathbb{P}^n,\, n \geqslant 1\}$ itself converges weakly to a probability measure $\mathbb{P}^\infty$. By the Skorohod's almost sure representation theorem (see Theorem 6.17 which is a simplified version of Theorem 1.10.4 on page 59 of van der Vaart and Wellner [**187**]) there exist a probability space $\left(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}}\right)$ together with a sequence of processes $\{Y^n, n \in \mathbb{N}\}$ and a process Y with distribution $\widetilde{P}$, defined on $\overline{\Omega}$ taking values in $C\left([0,T],\mathbb{R}^d\right)$ with $\mathcal{L}\left(Y^n\right) = \mathcal{L}\left(X^n\right)$ for all $n \geqslant 1$, $\mathbb{P}^\infty = \mathcal{L}(Y)$, and $\lim_{n\to\infty} Y^n = Y$ almost surely in $C\left([0,T],\mathbb{R}^d\right)$. Moreover, for each $n \in \mathbb{N}$, there exists a measurable map $\phi_n : \overline{\Omega} \to \Omega$ such that $Y^n\left(\overline{\omega}\right) = X^n\left(\phi_n\left(\overline{\omega}\right)\right)$, $\overline{\omega} \in \overline{\Omega}$. Furthermore, for Y^n we can choose a Brownian motion W^n on $\overline{\Omega}$ such that

$$\begin{aligned} Y^n(t) = Y(0) &+ \int_0^t c\left(\eta^n(s), Y^n\left(\eta^n(s)\right)\right) ds \\ &+ \int_0^t \sigma\left(\eta^n(s), Y^n\left(\eta^n(s)\right)\right) dW^n(s), \end{aligned} \tag{6.86}$$

which implies that

$$\left\langle Y_i^n, Y_j^n\right\rangle(t) = \sum_{k=1}^r \int_0^t \sigma_{i,k}\left(\eta^n(s), Y^n\left(\eta^n(s)\right)\right) \sigma_{j,k}\left(\eta^n(s), Y^n\left(\eta^n(s)\right)\right) ds. \tag{6.87}$$

The maps ϕ_n are measurable, and $\mathbb{P}^n = \mathbb{P} \circ \phi_n^{-1}$, for $n = 1, 2, \dots$, where $\mathbb{P}^n$ is the probability measure describing the distribution of the process X^n on the original probability space where X^n lives on. Since we build up our Euler scheme $\{X^n : n \geqslant 1\}$ on the same probability space, $\mathbb{P}^n$ actually does not depend on n. If we define $W^n(\overline{\omega}) = B\left(\phi^n\left(\overline{\omega}\right)\right)$, then W^n is a Brownian motion on $\left(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}}\right)$. This is because for any Borel set D,

$$\begin{aligned} \overline{P}\left[W^n \in D\right] &= \overline{P}\left[\overline{\omega} \in \overline{\Omega},\, B\left(\phi^n\left(\overline{\omega}\right)\right) \in D\right] = \overline{P}\left[\phi_n^{-1}\left(B^{-1}\left(D\right)\right)\right] \\ &= \mathbb{P}^n\left[B^{-1}\left(D\right)\right] = \mathbb{P}\left[B \in D\right]. \end{aligned} \tag{6.88}$$

Since Y^n converges to Y almost surely and $\sup\limits_{n\geqslant 1} \overline{\mathbb{E}}\left[\sup\limits_{0\leqslant s\leqslant T} \left|Y^n\right|^4\right] < \infty$, by Theorem 2.2 of Kurtz and Protter [**102**], for all $0 \leqslant t \leqslant T$,

$$L^1\text{-}\lim_{n\to\infty} \left\langle Y_i^n, Y_j^n\right\rangle(t) = \left\langle Y_i, Y_j\right\rangle(t). \tag{6.89}$$

This completes the outline of the first part of the proof of Theorem 6.10. □

6.14. LEMMA. *The sequence* $\left\{\left\langle Y_i^n, Y_j^n\right\rangle(t),\, n \in \mathbb{N}\right\}$ *converges not only* $\overline{\mathbb{P}}$*-almost surely, but also in the space* $L^1\left(\overline{\Omega}, \overline{F}, \overline{\mathbb{P}}\right)$.

PROOF. The equality in (6.89) is a consequence of the fact that, for $t \in [0, T]$ fixed, uniform integrability of the sequence $\left\{\left\langle Y_i^n, Y_j^n\right\rangle(t),\, n \in \mathbb{N}\right\}$ is uniform L^1-integrable. This property is based on the linear growth property of the coefficients $c(t,x)$ and $\sigma(t,x)$ which yields the finiteness of

$$\sup_{n\geqslant 1} \overline{\mathbb{E}}\left[\max_{0\leqslant t\leqslant T} \left|\langle Y^n, Y^n\rangle(t)\right|^2\right] = \sup_{n\geqslant 1} \mathbb{E}\left[\max_{0\leqslant t\leqslant T} \left|\langle X^n, X^n\rangle(t)\right|^2\right]. \tag{6.90}$$

For the latter see (6.81) in Lemma 6.11. The upshot of the finiteness of the expressions in (6.90) is that, for $N > 0$ large,

$$\begin{aligned}
&\overline{\mathbb{E}}\left[\left|\left\langle Y_i^n, Y_j^n\right\rangle(t)\right|, \left|\left\langle Y_i^n, Y_j^n\right\rangle(t)\right| \geqslant N\right]\\
&= \mathbb{E}\left[\left|\left\langle X_i^n, X_j^n\right\rangle(t)\right|, \left|\left\langle X_i^n, X_j^n\right\rangle(t)\right| \geqslant N\right]\\
&\leqslant \frac{1}{N}\mathbb{E}\left[\left|\left\langle X_i^n, X_j^n\right\rangle(t)\right|^2\right] \leqslant \frac{1}{N}\sup_{n\geqslant 1}\mathbb{E}\left[\left|\langle X^n, X^n\rangle(t)\right|^2\right].
\end{aligned} \tag{6.91}$$

The inequality in (6.91) shows the uniform integrability of the sequence

$$\left\{\left\langle Y_i^n, Y_j^n\right\rangle(t),\, n \in \mathbb{N}\right\}.$$

Since, in addition, this sequence converges $\overline{\mathbb{P}}$-almost surely, it converges in $L^1\left(\overline{\Omega}, \overline{F}, \overline{\mathbb{P}}\right)$. This ends the proof of Lemma 6.14. ☐

We also need the following lemma.

6.15. Lemma. *Put* $\triangle^n = \max_{0 \leqslant k \leqslant n} \left|\tau_{k+1}^n - \tau_k^n\right|$. *If the conditions in Theorem 6.10 are satisfied, and* $\lim_{n\to\infty} \triangle^n = 0$, *then for all* i, j *and* $0 \leqslant t \leqslant T$,

$$\begin{aligned} &L^2\text{-}\lim_{n\to\infty} \int_0^t c_i\left(\eta^n(s), Y^n\left(\eta^n(s)\right)\right) ds = \int_0^t c_i\left(s, Y\left(s\right)\right) ds, \quad \textit{and} \\ &L^1\text{-}\lim_{n\to\infty} \int_0^t \sigma_{i,k}\left(\eta^n(s), Y^n\left(\eta^n(s)\right)\right) \sigma_{k,j}\left(\eta^n(s), Y^n\left(\eta^n(s)\right)\right) ds \\ &= \int_0^t \sigma_{i,k}\left(s, Y\left(s\right)\right) \sigma_{k,j}\left(s, Y\left(s\right)\right) ds. \end{aligned} \tag{6.92}$$

Continuation of the proof of Theorem 6.10. We introduce the following sequence of processes $\left\{\overline{M}^n : n \geqslant 1\right\}$ which are in fact martingales relative to the $\left(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}\right)$ endowed with the filtration $\left(\overline{\mathcal{F}}_t^n\right)_{0\leqslant t\leqslant T}$, where $\overline{\mathcal{F}}_t^n$ is the σ-field generated by $\left(W^n(s) : 0 \leqslant s \leqslant t\right) \cup \mathcal{N}$. Here $\mathcal{N}$ denotes the collection of $\overline{\mathbb{P}}$-null sets. In fact we introduce:

$$\begin{aligned} \overline{M}(t) &= Y(t) - \int_0^t c\left(s, Y(s)\right) ds, \\ \overline{M}^n(t) &= Y^n(t) - \int_0^t c\left(\eta^n(s), Y^n\left(\eta^n(s)\right)\right) ds \\ &= Y(0) + \int_0^t \sigma\left(\eta^n(s), Y^n\left(\eta^n(s)\right)\right) dW^n(s). \end{aligned} \tag{6.93}$$

Since $Y(t)$ is the pointwise and weak limit of the sequence $\{Y^n, n \geqslant 1\}$ it can be proved that the process $\overline{M}(t)$ is a martingale relative to the filtration determined by the process $t \mapsto Y(t)$, $0 \leqslant t \leqslant T$. Its quadratic covariation is given by

$$\left\langle \overline{M}_i, \overline{M}_j \right\rangle(t) = \sum_{k=1}^n \int_0^t \sigma_{i,k}\left(s, Y(s)\right) \sigma_{j,k}\left(s, Y(s)\right) ds. \tag{6.94}$$

From this martingale property and from (6.94) it follows that there exists a Brownian motion $t \mapsto B(t)$, $0 \leqslant t \leqslant T$, possibly on an enlarged sample path space such that

$$\overline{M}(t) = \overline{M}(0) + \int_0^t \sigma\left(s, Y(s)\right) dB(s). \tag{6.95}$$

The equality in (6.95) implies:

$$Y(t) = Y(0) + \int_0^t c\left(s, Y(s)\right) ds + \int_0^t \sigma\left(s, Y(s)\right) dB(s). \tag{6.96}$$

The equality in (6.96) means that the process $t \mapsto Y(t)$, $0 \leqslant t \leqslant T$, is the (unique) weak solution to the equation in (6.76). Here we employ the equality in (6.80) with $\widetilde{X}(t) = Y(t)$ and $\widetilde{B}(t) = B(t)$. The proof of the existence of the Brownian motion with the property exhibited in (6.95) can be found in Theorem 7.1' Ikeda and Watanabe [**82**]. This completes the outline of the proof of Theorem 6.10. □

Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P}\right)$ be a filtered probability space. We begin with some notations. For a, $b \in \mathbb{R}$ we write $a \wedge b = \min(a,b)$ For a process $t \mapsto X(t)$, $t \geqslant 0$, we denote $X^*(t) := \sup_{0\leqslant s\leqslant t} |X(s)|$. For a continuous adapted process $t \mapsto X(t)$ and a stopping time τ we let $X(\tau)$ be as usual and $X^\tau(t)$ is the continuous adapted process defined by $X^\tau(t) := X(\tau \wedge t)$. For a continuous local martingale $t \mapsto M(t)$, $t \geqslant 0$, also written as $M = (M(t))_{t>0}$ we denote by $t \mapsto \langle M, M\rangle(t)$ the quadratic variation process. If M and N are both continuous local martingales we let $t \mapsto \langle M, N\rangle(t)$, $t \geqslant 0$, be the covariation process or cross-variation.

The next result is known as the Burkholder-Davis-Gundy inequalities. It was first proved for discrete martingales and $p > 1$ by Burkholder in 1966. In 1968, Millar extended the result to continuous martingales. In 1970, Davis extended the result for discrete martingales to $p = 1$. The extension to $p > 0$ was obtained independently by Burkholder and Gundy in 1970 and Novikov in 1971. We write $\langle M, M\rangle(\infty)$ for the pointwise limit, as $t \to \infty$, of $\langle M, M\rangle(t)$. This pointwise limit always exists in $[0,\infty) \cup \{\infty\}$.

6.16. Theorem (BDG-inequalities). *For each $p > 0$ there exist positive finite constants c_p, $C_p \in (0,\infty)$ such that, for any continuous local martingale $t \mapsto M(t)$, $t \geqslant 0$, with $M(0) = 0$,*

$$c_p\mathbb{E}\left[(\langle M, M\rangle(t))^{p/2}\right] \leqslant \mathbb{E}\left[(M^*(t))^p\right] \leqslant C_p\mathbb{E}\left[(\langle M, M\rangle(t))^{p/2}\right]. \tag{6.97}$$

Of course, in (6.97) we may let $t \to \infty$ to get the result also for $t = \infty$. The BDG-inequalities are also true if $t \mapsto M(t)$ is an $\mathbb{R}^d$-valued martingale with $|M(t)|$ the (Euclidean) norm of $M(t)$. A proof can for example be found in Kallenberg [**88**].

A more general result of the following theorem can be found in Van der Vaart and Wellner [**187**] Theorem 1.10.4.

6.17. Theorem (Almost sure representations). *Let $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)_{n\in\mathbb{N}\cup\{\infty\}}$ be a sequence of probability spaces and let $X_n : \Omega_n \to C\left([0,T], \mathbb{R}^d\right)$, $n \in \mathbb{N} \cup \{\infty\}$, be a sequence of stochastic continuous processes. Suppose that $\lim_{n\to\infty} \mathbb{E}_n\left[F(X_n)\right] = \mathbb{E}_\infty\left[F(X_\infty)\right]$ for all bounded continuous functions $F : C\left([0,T], \mathbb{R}^d\right) \to \mathbb{R}$. Then there exists a probability space $\left(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}\right)$ together with a sequence of stochastic processes $\overline{X}_n : \overline{\Omega} \to C\left([0,T], \mathbb{R}^d\right)$, $n \in \mathbb{N} \cup \{\infty\}$, with the following properties:*

1. $\lim_{n\to\infty} \overline{X}_n = \overline{X}_\infty$, *$\overline{\mathbb{P}}$-almost surely;*
2. $\overline{\mathbb{E}}\left[F\left(\overline{X}_n\right)\right] = \mathbb{E}_n\left[F(X_n)\right]$ *for all $n \in \mathbb{N} \cup \{\infty\}$ and for all bounded continuous functions $F : C\left([0,T], \mathbb{R}^d\right) \to \mathbb{R}$.*

In addition to (1) and (2), for each $n \in \mathbb{N}$, the process $\overline{X}_n$ can be chosen according to the following diagram

$$\begin{tikzcd} \Omega_n \arrow[r, "X_n"] & C\left([0,T],\mathbb{R}^d\right) \\ \overline{\Omega} \arrow[u, "\phi_n"] \arrow[ur, "\overline{X}_n = X_n\circ\phi_n"'] & \end{tikzcd}$$

where the maps $\phi_n : \overline{\Omega} \to \Omega_n$ are measurable, and where $\mathbb{P}_n = \overline{P} \circ \phi_n^{-1}$.

The following result was also used in the proof of Theorem 6.10. We owe it to Ikeda and Watanabe [**82**] Theorem 7.1'.

6.18. THEOREM. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $(\mathcal{F}_t)_{t\geqslant 0}$ and let $t \mapsto M_i(t)$, $1 \leqslant i \leqslant d$, be local martingales with covariation matrix $\Phi(t) = (\Phi_{i,j}(t))_{1\leqslant i,j\leqslant d}$, i.e. $\langle M_i, M_j\rangle(t) = \int_0^t \Phi_{i,j}(s)\,ds$, $1 \leqslant i,j \leqslant d$. In addition, let the predictable matrix process $\Psi(t) = (\Psi_{i,j}(t))_{1\leqslant i\leqslant d, 1\leqslant j\leqslant n}$ be such that $\Phi_{i,j}(t) = \sum_{k=1}^{n} \Psi_{i,k}(t)\Psi_{k,j}(t)$, $1 \leqslant i,\ j \leqslant d$, and such that $\int_0^T |\Psi_{i,j}(s)|^2\,ds < \infty$ for $1 \leqslant i$, $1 \leqslant j \leqslant n$. Then, on some extension $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}\right)$ and $\left(\widetilde{\mathcal{F}}_t\right)_{t\geqslant 0}$ of $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\mathcal{F}_t)_{t\geqslant 0}$, respectively, there exists an n-dimensional $\left(\widetilde{\mathcal{F}}_t\right)_{t\geqslant 0}$-Brownian motion $\widetilde{B}(t)$ such that*

$$M_i(t) = \sum_{k=1}^{n} \int_0^t \Psi_{i,k}(s)\,d\widetilde{B}_k(s).$$

Prohorov's theorem. In the proof of Theorem 6.10 we applied Prohorov's tightness theorem in compination with the Arzela-Ascoli characterization of compact subsets of $C[0,T]$. It was applied to the sequence $X^n(t)$ (Euler scheme) defined in (6.77).

6.19. THEOREM (Prohorov theorem). *Let $(\mathbb{P}_n : n \in \mathbb{N})$ be a sequence of probability measures on a separable complete metrizable topological space S with Borel σ-field $\mathcal{S}$. Then the following assertions are equivalent:*

(i) *For every $\varepsilon > 0$ there exists a compact subset K_ε of S such that $\mathbb{P}_n[K_\varepsilon] \geqslant 1-\varepsilon$ for all $n \in \mathbb{N}$.*
(ii) *Every subsequence of $(\mathbb{P}_n : n \in \mathbb{N})$ has a subsequence which converges weakly to a probability measure on $(S, \mathcal{S})$.*

A sequence $(\mathbb{P}_n)_n$ satisfying (i) (or (ii)) in Theorem 6.19 is called a Prohorov set. In the sequel we write $C[0,T] = C\left([0,T], \mathbb{R}^d\right)$.

6.20. THEOREM (Arzela-Ascoli). *Endow $C[0,T]$ with the topology of uniform convergence. A subset A of $C[0,T]$ has compact closure if and only if it has the following properties:*

(i) $\sup\limits_{\omega\in A} |\omega(0)| < \infty$;
(ii) *The subset A is equi-continuous in the sense that*

$$\lim_{\delta\downarrow 0} \sup_{0\leqslant s,t\leqslant T,\, |s-t|\leqslant\delta} \sup_{\omega\in A} |\omega(s)-\omega(t)| = 0.$$

From (i) and (ii) it follows that $\sup\limits_{\omega\in A} \sup\limits_{s\in[0,T]} |\omega(s)| < \infty$, and hence A is uniformly bounded. The result which is relevant here reads as follows. It is the same as Theorem T.8.4 in Bhattacharaya and Waymire [**20**].

6.21. THEOREM. *Let $(\mathbb{P}_n)_n$ be a sequence of probability measures on $C[0,T]$. Then $(\mathbb{P}_n)_n$ is tight if and only if the following two conditions hold.*

(i) *For each $\eta > 0$ there is a number B such that*

$$\mathbb{P}_n\left[\omega \in C[0,T] : |\omega(0)| > B\right] < \eta, \quad n = 1, 2, \ldots$$

(ii) *For each $\varepsilon > 0$, $\eta > 0$, there is a $0 < \delta < 1$ such that*

$$\mathbb{P}_n\left[\omega \in C[0,T] : \sup_{0\leqslant s,t\leqslant T,\, |s-t|\leqslant\delta} |\omega(s)-\omega(t)| \geqslant \varepsilon\right] \leqslant \eta, \quad n = 1, 2, \ldots$$

PROOF. If the sequence $(\mathbb{P}_n)_n$ is tight, then given $\eta > 0$ there is a compact subset K of $C([0,T])$ such that $\mathbb{P}_n(K) > 1-\eta$ for all n. By the Arzela-Ascoli theorem (Theorem 6.20), if $B > \sup_{\omega\in K} |\omega(0)|$, then

$$\mathbb{P}_n\left[\omega \in C[0,T] : |\omega(0)| \geqslant B\right] \leqslant \mathbb{P}_n\left[K^c\right] \leqslant 1-(1-\eta) = \eta.$$

Also given $\varepsilon > 0$ select $\delta > 0$ such that $\sup\limits_{\omega\in K}\sup\limits_{0\leqslant s,t\leqslant T,\,|s-t|\leqslant\delta}|\omega(s)-\omega(t)| < \varepsilon$. Then

$$\mathbb{P}_n\left[\omega\in C[0,T] : \sup_{0\leqslant s,t\leqslant T,\,|s-t|\leqslant\delta}|\omega(s)-\omega(t)| \geqslant \varepsilon\right] \leqslant \mathbb{P}_n\left[K^c\right] < \eta$$

for all $n \geqslant 1$. The converse goes as follows. Given $\eta > 0$, first select B using (i) such that $\mathbb{P}_n\left[\omega\in C\left([0,T]\right) : |\omega(0)|\leqslant B\right] \geqslant 1-\frac{1}{2}\eta$, for $n\geqslant 1$. Select $\delta_r > 0$ using (ii) such that

$$\mathbb{P}_n\left[\omega\in C\left([0,T]\right) : \sup_{0\leqslant s,t\leqslant T,\,|s-t|\leqslant\delta_r}|\omega(s)-\omega(t)| < \frac{1}{r}\right] \geqslant 1-2^{-(r+1)}\eta$$

for all $n \geqslant 1$. Next we take K to be the uniform closure of

$$\bigcap_{r=1}^{\infty}\left\{\omega\in C\left([0,T]\right) : |\omega(0)|\leqslant B, \quad \sup_{0\leqslant s,t\leqslant T,\,|s-t|\leqslant\delta_r}|\omega(s)-\omega(t)| < \frac{1}{r}\right\}.$$

Then $\mathbb{P}_n(K) > 1-\eta$ for $n \geqslant 1$, and K is compact by the Arzela-Ascoli theorem. This completes the proof Theorem 6.21. □

Shorter proof of Theorem 6.10. The Banach space $C\left([0,T],\mathbb{R}^d\right)$ being separable, Prohorov's theorem (6.19) in conjunction with Lemma 6.12 then implies that the sequence $\{X^n,\, n\geqslant 1\}$ is tight in $C\left([0,T],\mathbb{R}^d\right)$. Thus for any subsequence of the sequence $\{X^n,\, n\geqslant 1\}$ there exists a further subsequence $\{X^{n_k},\, k\geqslant 1\}$ and a process $t\mapsto X(t)$, $t\in[0,T]$, in $C\left([0,T],\mathbb{R}^d\right)$ such that the sequence $\{X^{n_k},\, k\geqslant 1\}$ converges to the process $t\mapsto X(t)$, $0\leqslant t\leqslant T$, not only weakly, but also $\mathbb{P}$-almost surely uniformly on $[0,T]$. Let $\mathbb{P}^n$ be the distribution (or law) of the process X^n. Since the equation in (6.52) is supposed to have unique weak solutions, it follows that the sequence $\{\mathbb{P}^n,\, n\geqslant 1\}$ itself converges weakly to a probability measure on $C\left([0,T],\mathbb{R}^d\right)$. Let us show this. Take any subsequence $\{X^{n_k},\, k\geqslant 1\}$ which converges, $\mathbb{P}$-almost surely, uniformly on $[0,T]$ to a process X. Let $\mathbb{P}^\infty$ be the distribution of X. Then the sequence $\{X^n : n\in\mathbb{N}\}$ converges weakly to X. If this were not the case, then there would exist $\varepsilon > 0$ and a bounded continuous function $F : C\left([0,T],\mathbb{R}^d\right)\to\mathbb{R}$ such that

$$\left|\mathbb{E}^{n_m}\left[F\right]-\mathbb{E}^\infty\left[F\right]\right| = \left|\mathbb{E}\left[F\left(X^{n_m}\right)\right]-\mathbb{E}\left[F\left(X\right)\right]\right| \geqslant \varepsilon \tag{6.98}$$

for some subsequence $\{X^{n_m},\, m\geqslant 1\}$. By passing to another subsequence we may assume that the subsequence itself converges $\mathbb{P}$-almost surely to a process $\widetilde{X}$. Hence it converges weakly, and by (6.98) we get

$$\lim_{m\to\infty}\left|\mathbb{E}^{n_m}\left[F\right]-\mathbb{E}^\infty\left[F\right]\right| = \left|\mathbb{E}\left[F\left(\widetilde{X}\right)\right]-\mathbb{E}\left[F\left(X\right)\right]\right| \geqslant \varepsilon \tag{6.99}$$

Since the coefficients of the equation in (6.76) are continuous it follows that both processes X and $\widetilde{X}$ satisfy this equation (by L^1-convergence). Weak uniqueness implies that the distribution of X and $\widetilde{X}$ are the same, which contradicts (6.99). This completes the proof of Theorem 6.10. □

Acknowledgment. The author is grateful to those students who took this course for their patience and alertness.

Bibliography

1. Sergio Albeverio and Sonia Mazzucchi, *Theory and applications of infinite dimensional oscillatory integrals*, Stochastic analysis and applications, Abel Symp., vol. 2, Springer, Berlin, 2007, pp. 73–91. MR 2397784 (2008m:58019)
2. ______, *A survey on mathematical Feynman path integrals: construction, asymptotics, applications*, Quantum field theory, Birkhäuser, Basel, 2009, pp. 49–66. MR 2742748 (2011k:58009)
3. Charalambos D. Aliprantis and Kim C. Border, *Infinite dimensional analysis*, third ed., Springer, Berlin, 2006, A hitchhiker's guide. MR 2378491 (2008m:46001)
4. Linda J. S. Allen, *An introduction to stochastic processes with applications to biology*, Pearson Education, Inc., Upper Saddle River, NJ 07458, April 2003, Textbook, Hardcover.
5. ______, *An introduction to stochastic processes with applications to biology*, second ed., CRC Press, Boca Raton, FL, 2011. MR 2560499
6. David Applebaum, *Lévy processes and stochastic calculus*, second ed., Cambridge Studies in Advanced Mathematics, vol. 116, Cambridge University Press, Cambridge, 2009. MR 2512800
7. Robert B. Ash, *Probability and measure theory*, second ed., Harcourt/Academic Press, Burlington, MA, 2000, With contributions by Catherine Doléans-Dade. MR 1810041 (2001j:28001)
8. Louis Bachelier, *Théorie de la spéculation*, Les Grands Classiques Gauthier-Villars. [Gauthier-Villars Great Classics], Éditions Jacques Gabay, Sceaux, 1995, Théorie mathématique du jeu. [Mathematical theory of games], Reprint of the 1900 original (English translation available in: The random character of stock market prices (ed. P.H. Cootner), pp. 17–78, MIT Press, Cambridge 1964).
9. D. Bakry and M. Ledoux, *Lévy-Gromov's isoperimetric inequality for an infinite-dimensional diffusion generator*, Invent. Math. **123** (1996), no. 2, 259–281. MR 1374200
10. Dominique Bakry, Ivan Gentil, and Michel Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348, Springer, Cham, 2014. MR 3155209
11. Dominique Bakry, Michel Ledoux, and Laurent Saloff-Coste, *Markov semigroups at Saint-Flour*, Probability at Saint-Flour, Springer, Heidelberg, 2012. MR 3075390
12. Erik J. Balder, *Infinite-dimensional extension of a theorem of Komlós*, Probab. Theory Related Fields **81** (1989), no. 2, 185–188. MR 982652 (90a:60008)
13. A. D. Barbour and Louis H. Y. Chen (eds.), *An introduction to Stein's method*, Lecture Notes Series. Institute for Mathematical Sciences. National University of Singapore, vol. 4, Singapore University Press, Singapore, 2005, Lectures from the Meeting on Stein's Method and Applications: a Program in Honor of Charles Stein held at the National University of Singapore, Singapore, July 28–August 31, 2003. MR 2235447 (2007j:60001)
14. A. D. Barbour and Peter Hall, *Stein's method and the Berry-Esseen theorem*, Austral. J. Statist. **26** (1984), no. 1, 8–15. MR 746011 (85k:60032)
15. Heinz Bauer, *Probability theory and elements of measure theory*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1981, Second edition of the translation by R. B. Burckel from the third German edition, Probability and Mathematical Statistics. MR 636091 (82k:60001)

16. B. Berckmoes, R. Lowen, and J. Van Casteren, *Approach theory meets probability theory*, Topology Appl. **158** (2011), no. 7, 836–852. MR 2783139 (2012d:60003)
17. ______, *Distances on probability measures and random variables*, J. Math. Anal. Appl. **374** (2011), no. 2, 412–428. MR 2729231 (2011k:60006)
18. B. Berckmoes, R. Lowen, and J. Van Casteren, *An isometric study of the Lindeberg-Feller central limit theorem via Stein's method*, J. Math. Anal. Appl. **405** (2013), no. 2, 484–498. MR 3061027
19. ______, *Stein's method and a quantitative Lindeberg CLT for the Fourier transforms of random vectors*, J. Math. Anal. Appl. **433** (2016), no. 2, 1441–1458. MR 3398773
20. Rabi Bhattacharya and Edward C. Waymire, *A basic course in probability theory*, Universitext, Springer, New York, 2007. MR MR2331066 (2009e:60001)
21. Rabi N. Bhattacharya and Edward C. Waymire, *Stochastic processes with applications*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons Inc., New York, 1990, A Wiley-Interscience Publication. MR MR1054645 (91m:60001)
22. Patrick Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, 1999, A Wiley-Interscience Publication. MR 1700749 (2000e:60008)
23. Jean-Michel Bismut, *Large deviations and the Malliavin calculus*, Progress in Mathematics, vol. 45, Birkhäuser Boston Inc., Boston, MA, 1984. MR MR755001 (86f:58150)
24. Fischer Black and Myron Scholes, *The pricing of options and corporate liabilities*, The Journal of Political Economy **81** (1973), no. 3, 637–654, The University of Chicago Press.
25. R. M. Blumenthal and R. K. Getoor, *Markov processes and potential theory*, Pure and Applied Mathematics, Vol. 29, Academic Press, New York, 1968. MR MR0264757 (41 #9348)
26. Sergey G. Bobkov and Michel Ledoux, *On weighted isoperimetric and Poincaré-type inequalities*, High dimensional probability V: the Luminy volume, Inst. Math. Stat. (IMS) Collect., vol. 5, Inst. Math. Statist., Beachwood, OH, 2009, pp. 1–29. MR 2797936
27. V. I. Bogachev, *Measure theory. Vol. I, II*, Springer-Verlag, Berlin, 2007. MR 2267655 (2008g:28002)
28. Frank F. Bonsall and John Duncan, *Complete normed algebras*, Springer-Verlag, New York, 1973, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 80. MR 0423029 (54 #11013)
29. M. Bossy and N. Champagnat, *Encyclopedia of quantitative finance*, ch. Markov processes, Wiley, 2010, DOI: 10.1002/9780470061602.eqf02016.
30. B. Boufoussi and J. van Casteren, *An approximation result for a nonlinear Neumann boundary value problem via BSDEs*, Stochastic Process. Appl. **114** (2004), no. 2, 331–350. MR MR2101248 (2005j:60130)
31. ______, *An approximation result for a nonlinear Neumann boundary value problem via BSDEs*, Stochastic Process. Appl. **114** (2004), no. 2, 331–350. MR MR2101248 (2005j:60130)
32. Brahim Boufoussi, Jan Van Casteren, and N. Mrhardy, *Generalized backward doubly stochastic differential equations and SPDEs with nonlinear Neumann boundary conditions*, Bernoulli **13** (2007), no. 2, 423–446. MR MR2331258 (2008h:60245)
33. Anton Bovier, *Markov processes*, internet: pdf-file, November 2012, Lecture Notes Summer 2012, Bonn.
34. Stephen Boyd and Lieven Vandenberghe, *Convex optimization*, Cambridge University Press, Cambridge, 2004. MR 2061575 (2005d:90002)
35. Leo Breiman, *Probability*, Classics in Applied Mathematics, vol. 7, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992, Corrected reprint of the 1968 original. MR 1163370 (93d:60001)
36. Paul C. Bressloff, *Stochastic processes in cell biology*, Interdisciplinary Applied Mathematics, vol. 41, Springer, Cham, 2014. MR 3244328

37. R. H. Cameron and W. T. Martin, *Transformations of Wiener integrals under translations*, Ann. of Math. (2) **45** (1944), 386–396. MR 0010346
38. R. H. Cameron and D. A. Storvick, *Change of scale formulas for Wiener integral*, Rend. Circ. Mat. Palermo (2) Suppl. (1987), no. 17, 105–115 (1988), Functional integration with emphasis on the Feynman integral (Sherbrooke, PQ, 1986). MR 950411
39. ______, *Relationships between the Wiener integral and the analytic Feynman integral*, Rend. Circ. Mat. Palermo (2) Suppl. (1987), no. 17, 117–133 (1988), Functional integration with emphasis on the Feynman integral (Sherbrooke, PQ, 1986). MR 950412
40. S. D. Chatterji, *A principle of subsequences in probability theory: the central limit theorem*, Advances in Math. **13** (1974), 31–54; correction, ibid. 14 (1974), 266–269. MR 0341564 (49 #6312)
41. Louis H. Y. Chen, Larry Goldstein, and Qi-Man Shao, *Normal approximation by Stein's method*, Probability and its Applications (New York), Springer, Heidelberg, 2011. MR 2732624 (2012b:60103)
42. Steve Cheng, *Conditional expectation under change of measure, version 9*, Planetmath (2013), 2 pages.
43. N.N. Chentsov, *Weak convergence of stochastic processes whose trajectories have no discontinuities of the second kind and the "heuristic" approach to the Kolmogorov-Smirnov tests*, Theory Probab. Appl. **1** (1956), no. 1, 140–144.
44. Alexander S. Cherny and Hans-Jürgen Engelbert, *Singular stochastic differential equations*, Lecture Notes in Mathematics, vol. 1858, Springer-Verlag, Berlin, 2005. MR 2112227
45. Anna Chojnowska-Michalik and Beniamin Goldys, *Generalized Ornstein-Uhlenbeck semigroups: Littlewood-Paley-Stein inequalities and the P. A. Meyer equivalence of norms*, J. Funct. Anal. **182** (2001), no. 2, 243–279. MR 1828795 (2002e:47050)
46. Rama Cont and Peter Tankov, *Financial modelling with jump processes*, Chapman & Hall/CRC Financial Mathematics Series, Chapman & Hall/CRC, Boca Raton, FL, 2004. MR 2042661
47. Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.) **27** (1992), no. 1, 1–67. MR 1118699 (92j:35050)
48. Stéphane Crépey, *Financial modeling*, Springer Finance, Springer, Heidelberg, 2013, A backward stochastic differential equations perspective, Springer Finance Textbooks. MR 3154654
49. Ana Bela Cruzeiro and Paul Malliavin, *Renormalized stochastic calculus of variations for a renormalized infinite-dimensional Brownian motion*, Stochastics **81** (2009), no. 3-4, 385–399. MR 2549495 (2011g:60096)
50. Ana Bela Cruzeiro and Jean-Claude Zambrini, *Malliavin calculus and Euclidean quantum mechanics. I. Functional calculus*, J. Funct. Anal. **96** (1991), no. 1, 62–95. MR 1093507 (92i:81149)
51. ______, *Malliavin calculus and Euclidean quantum mechanics. II. Variational principle for infinite-dimensional processes*, J. Funct. Anal. **130** (1995), no. 2, 450–476. MR 1335388 (96g:60072)
52. Giulia Di Nunno, Bernt Ø ksendal, and Frank Proske, *Malliavin calculus for Lévy processes with applications to finance*, Universitext, Springer-Verlag, Berlin, 2009. MR 2460554
53. Jacques Dixmier, *von Neumann algebras*, North-Holland Mathematical Library, vol. 27, North-Holland Publishing Co., Amsterdam, 1981, With a preface by E. C. Lance, Translated from the second French edition by F. Jellett. MR 641217 (83a:46004)
54. Bruce Driver, *Math 280 (Probability Theory) Lecture Notes*, Lecture notes, University of California, San Diego, Department of Mathematics, 0112 University of California, San Diego 9500 Gilman Drive, La Jolla, California 92093-0112 USA, June 10 2010, File:prob.tex.

55. Richard M. Dudley, *Real analysis and probability*, The Wadsworth & Brooks/Cole Mathematics Series, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1989. MR MR982264 (91g:60001)
56. ———, *Uniform central limit theorems*, Cambridge Studies in Advanced Mathematics, vol. 63, Cambridge University Press, Cambridge, 1999. MR 1720712 (2000k:60050)
57. Richard Durrett, *Brownian motion and martingales in analysis*, Wadsworth Mathematics Series, Wadsworth International Group, Belmont, CA, 1984. MR MR750829 (87a:60054)
58. ———, *Brownian motion and martingales in analysis*, Wadsworth Mathematics Series, Wadsworth International Group, Belmont, CA, 1984. MR 750829 (87a:60054)
59. ———, *Probability: theory and examples*, second ed., Duxbury Press, Belmont, CA, 1996. MR MR1609153 (98m:60001)
60. Rick Durrett, *Probability: theory and examples*, fourth ed., Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2010. MR 2722836 (2011e:60001)
61. Albert Einstein, *Investigations on the theory of the Brownian movement*, Dover Publications, Inc., New York, 1956, Edited with notes by R. Fürth, Translated by A. D. Cowper. MR 0077443 (17,1035g)
62. Tanja Eisner and Rainer Nagel, *Arithmetic progressions – an operator theoretic view*, Discrete and Continuous Dynamical Systems - Series S (DCDS-S) **6** (2013), no. 3, 657–667.
63. N. El Karoui, E. Pardoux, and M. C. Quenez, *Reflected backward SDEs and American options*, Numerical methods in finance, Publ. Newton Inst., vol. 13, Cambridge Univ. Press, Cambridge, 1997, pp. 215–231. MR MR1470516 (99b:60078)
64. N. El Karoui and M. C. Quenez, *Imperfect markets and backward stochastic differential equations*, Numerical methods in finance, Publ. Newton Inst., vol. 13, Cambridge Univ. Press, Cambridge, 1997, pp. 181–214. MR MR1470515 (98e:90057)
65. Nicole El Karoui and Laurent Mazliak (eds.), *Backward stochastic differential equations*, Pitman Research Notes in Mathematics Series, vol. 364, Longman, Harlow, 1997, Papers from the study group held at the University of Paris VI, Paris, 1995–1996. MR MR1752671 (2000k:60003)
66. K. D. Elworthy, *Stochastic differential equations on manifolds*, London Mathematical Society Lecture Note Series, vol. 70, Cambridge University Press, Cambridge, 1982. MR MR675100 (84d:58080)
67. ———, *Stochastic differential equations on manifolds*, Probability towards 2000 (New York, 1995), Lecture Notes in Statist., vol. 128, Springer, New York, 1998, pp. 165–178. MR 1632635 (99g:58131)
68. Stewart N. Ethier and Thomas G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1986, Characterization and convergence. MR MR838085 (88a:60130)
69. F. Feo, M. R. Posteraro, and C. Roberto, *Quantitative isoperimetric inequalities for log-convex probability measures on the line*, J. Math. Anal. Appl. **420** (2014), no. 2, 879–907. MR 3240055
70. Marina Filipa Amado Ferreira, *Stochastic differential equation models in population dynamics*, Ph.d. thesis, Universidade de Coimbra, September 2014.
71. Gerald B. Folland, *A course in abstract harmonic analysis*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1995. MR 1397028 (98c:43001)
72. Jean-Pierre Fouque, George Papanicolaou, and K. Ronnie Sircar, *Derivatives in financial markets with stochastic volatility*, Cambridge University Press, Cambridge, 2000. MR 1768877
73. Avner Friedman, *Stochastic differential equations and applications*, Dover Publications, Inc., Mineola, NY, 2006, Two volumes bound as one, Reprint of the 1975 and 1976 original published in two volumes. MR 2295424

74. Peter K. Friz, *An introduction to Malliavin calculus*, internet, August 2002, Courant Institute of Mathematical Sciences, New York University.
75. Paul Garrett, *Measurable choice functions*, Lecture notes, University of Minnesota, December 2004.
76. Giordano Giambartolomei, *The Karhunen-Loève theorem*, Ph.d. thesis, Università di Bologna, 2015.
77. B. Goldys and B. Maslowski, *Lower estimates of transition densities and bounds on exponential ergodicity for stochastic PDE's*, Ann. Probab. **34** (2006), no. 4, 1451–1496. MR MR2257652 (2007j:60091)
78. Alexander Grigor'yan, *Heat kernel and analysis on manifolds*, AMS/IP Studies in Advanced Mathematics, vol. 47, American Mathematical Society, Providence, RI, 2009. MR 2569498 (2011e:58041)
79. Archil Gulisashvili, *Analytically tractable stochastic stock price models*, Springer Finance, Springer, Heidelberg, 2012. MR 2976192
80. Helge Holden, Bernt Øksendal, Jan Ubøe, and Tusheng Zhang, *Stochastic partial differential equations*, second ed., Universitext, Springer, New York, 2010, A modeling, white noise functional approach. MR 2571742
81. N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, 2 ed., North-Holland Mathematical Library, vol. 24, North-Holland, Amsterdam, 1998.
82. Nobuyuki Ikeda and Shinzo Watanabe, *Stochastic differential equations and diffusion processes*, second ed., North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989. MR 1011252
83. N. Jacob, *Pseudo differential operators and Markov processes. Vol. I*, Imperial College Press, London, 2001, Fourier analysis and semigroups. MR 1873235 (2003a:47104)
84. ______, *Pseudo differential operators & Markov processes. Vol. II*, Imperial College Press, London, 2002, Generators and their potential theory. MR 1917230 (2003k:47077)
85. ______, *Pseudo differential operators and Markov processes. Vol. III*, Imperial College Press, London, 2005, Markov processes and applications. MR 2158336 (2006i:60001)
86. Gerald W. Johnson and Michel L. Lapidus, *The Feynman integral and Feynman's operational calculus*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 2000, Oxford Science Publications. MR 1771173
87. Gerald W. Johnson, Michel L. Lapidus, and Lance Nielsen, *Feynman's operational calculus and beyond*, Oxford Mathematical Monographs, Oxford University Press, Oxford, 2015, Noncommutativity and time-ordering. MR 3381096
88. Olav Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR MR1876169 (2002m:60002)
89. Ioannis Karatzas and Steven E. Shreve, *Methods of mathematical finance*, Applications of Mathematics (New York), vol. 39, Springer-Verlag, New York, 1998. MR 1640352 (2000e:91076)
90. Samuel Karlin and Howard M. Taylor, *A first course in stochastic processes*, second ed., Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. MR MR0356197 (50 #8668)
91. Alexander S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995. MR 1321597 (96e:03057)
92. Fima C. Klebaner, *Introduction to stochastic calculus with applications*, third ed., Imperial College Press, London, 2012. MR 2933773
93. Hagen Kleinert, *Path integrals in quantum mechanics, statistics, polymer physics, and financial markets*, fifth ed., World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2009. MR 2518082 (2010i:81250)
94. A. N. Kolmogorov, *On the Skorohod convergence*, Teor. Veroyatnost. i Primenen. **1** (1956), 239–247. MR 0085638 (19,69i)

95. Vassili N. Kolokoltsov, *Measure-valued limits of interacting particle systems with k-nary interactions. II. Finite-dimensional limits*, Stoch. Stoch. Rep. **76** (2004), no. 1, 45–58. MR MR2038028 (2004k:60268)
96. ______, *On Markov processes with decomposable pseudo-differential generators*, Stoch. Stoch. Rep. **76** (2004), no. 1, 1–44. MR MR2038027 (2005b:60193)
97. J. Komlós, *A generalization of a problem of Steinhaus*, Acta Math. Acad. Sci. Hungar. **18** (1967), 217–229. MR 0210177 (35 #1071)
98. P. E. Kopp, *Martingales and stochastic integrals*, Cambridge University Press, Cambridge, 1984. MR 774050 (86i:60004)
99. Ulrich Krengel, *Ergodic theorems*, de Gruyter Studies in Mathematics, vol. 6, Walter de Gruyter & Co., Berlin, 1985, With a supplement by Antoine Brunel. MR MR797411 (87i:28001)
100. N.V. Krylov, *A simple proof of a result of A. Novikov*, arXiv:math.PR/0207013 v1 (2002), no. 1, 3.
101. Hui Hsiung Kuo, *Gaussian measures in Banach spaces*, Lecture Notes in Mathematics, Vol. 463, Springer-Verlag, Berlin, 1975. MR 0461643 (57 #1628)
102. Thomas G. Kurtz and Philip E. Protter, *Weak convergence of stochastic integrals and differential equations. II. Infinite-dimensional case*, Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), Lecture Notes in Math., vol. 1627, Springer, Berlin, 1996, pp. 197–285. MR 1431303
103. S. Kusuoka and D. Stroock, *Applications of the Malliavin calculus. II*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **32** (1985), no. 1, 1–76. MR MR783181 (86k:60100b)
104. ______, *Applications of the Malliavin calculus. III*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **34** (1987), no. 2, 391–442. MR MR914028 (89c:60093)
105. Shigeo Kusuoka and Daniel Stroock, *Applications of the Malliavin calculus. I*, Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, vol. 32, North-Holland, Amsterdam, 1984, pp. 271–306. MR MR780762 (86k:60100a)
106. Peter D. Lax, *Functional analysis*, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2002. MR 1892228 (2003a:47001)
107. Rémi Léandre, *Malliavin calculus of Bismut type without probability*, Proc. Indian Acad. Sci. Math. Sci. **116** (2006), no. 4, 507–518. MR 2349206 (2009a:60057)
108. ______, *Malliavin calculus of Bismut type in semi-group theory*, Far East J. Math. Sci. (FJMS) **30** (2008), no. 1, 1–26. MR 2483411 (2010a:60195)
109. ______, *Malliavin calculus of Bismut type for fractional powers of Laplacians in semi-group theory*, Int. J. Differ. Equ. (2011), Art. ID 575383, 26. MR 2832507 (2012g:60181)
110. Michel Ledoux, *A short proof of the Gaussian isoperimetric inequality*, High dimensional probability (Oberwolfach, 1996), Progr. Probab., vol. 43, Birkhäuser, Basel, 1998, pp. 229–232. MR 1652328
111. Michel Ledoux, Ivan Nourdin, and Giovanni Peccati, *Stein's method, logarithmic Sobolev and transport inequalities*, Geom. Funct. Anal. **25** (2015), no. 1, 256–306. MR 3320893
112. Jorge A. León, Josep L. Solé, Frederic Utzet, and Josep Vives, *On Lévy processes, Malliavin calculus and market models with jumps*, Finance Stoch. **6** (2002), no. 2, 197–225. MR 1897959 (2003b:60062)
113. Thomas M. Liggett, *Interacting particle systems*, Classics in Mathematics, Springer-Verlag, Berlin, 2005, Reprint of the 1985 original.
114. Kian Guan Lim, *Probability and finance theory*, World Scientific Publishing Co. Pte. Ltd., Singapore, May 2011.
115. József Lőrinczi, Fumio Hiroshima, and Volker Betz, *Feynman-Kac-type theorems and Gibbs measures on path space*, de Gruyter Studies in Mathematics, vol. 34, Walter de Gruyter & Co., Berlin, 2011, With applications to rigorous quantum field theory. MR 2848339 (2012h:58014)
116. George Lowther, *Almost sure: Continuous processes with independent increments*, blog, `https://almostsure.wordpress.com/2010/06/16`, June 2010.

117. Dorothy Maharam, *From finite to countable additivity*, Portugal. Math. **44** (1987), no. 3, 265–282. MR 911841 (88j:28001)
118. A. G. Malliaris, *Stochastic methods in economics and finance*, Advanced Textbooks in Economics, vol. 17, North-Holland Publishing Co., Amsterdam, 1982, With a foreword and with contributions by W. A. Brock. MR 642839 (83k:90003)
119. Paul Malliavin, *Stochastic calculus of variation and hypoelliptic operators*, Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976) (New York), Wiley, 1978, pp. 195–263. MR MR536013 (81f:60083)
120. Paul Malliavin and Anton Thalmaier, *Stochastic calculus of variations in mathematical finance*, Springer Finance, Springer-Verlag, Berlin, 2006. MR MR2189710 (2007b:91002)
121. Rogemar S. Mamon, *Three ways to solve for bond prices in the Vasicek model*, J. Appl. Math. Decis. Sci. **8** (2004), no. 1, 1–14. MR 2042166 (2004j:91142)
122. Laurent Mazliak, *The Ghosts of the École Normale*, Statist. Sci. **30** (2015), no. 3, 391–412. MR 3383887
123. Sonia Mazzucchi, *Mathematical Feynman path integrals and their applications*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2009. MR 2537928 (2010g:46123)
124. Michel Métivier and Jean Pellaumail, *Stochastic integration*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1980, Probability and Mathematical Statistics. MR 578177 (82b:60060)
125. S. P. Meyn and R. L. Tweedie, *Markov chains and stochastic stability*, Communications and Control Engineering Series, Springer-Verlag London Ltd., London, 1993, (new version September 2005: http://probability.ca/MT/; Second edition, Cambridge University Press, 2009.). MR MR1287609 (95j:60103)
126. Sean Meyn and Richard L. Tweedie, *Markov chains and stochastic stability*, second ed., Cambridge University Press, Cambridge, 2009, With a prologue by Peter W. Glynn. MR 2509253
127. Thomas Mikosch, *Elementary stochastic calculus—with finance in view*, Advanced Series on Statistical Science & Applied Probability, vol. 6, World Scientific Publishing Co. Inc., River Edge, NJ, 1998. MR MR1728093 (2001c:60001)
128. Nikolai Nadirashvili, *Nonuniqueness in the martingale problem and the Dirichlet problem for uniformly elliptic operators*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **24** (1997), no. 3, 537–549. MR MR1612401 (99b:35042)
129. James Norris, *Simplified Malliavin calculus*, Séminaire de Probabilités, XX, 1984/85, Lecture Notes in Math., vol. 1204, Springer, Berlin, 1986, pp. 101–130. MR MR942019 (89f:60058)
130. Ivan Nourdin and Giovanni Peccati, *Stein's method meets Malliavin calculus: a short survey with new estimates*, Recent development in stochastic dynamics and stochastic analysis, Interdiscip. Math. Sci., vol. 8, World Sci. Publ., Hackensack, NJ, 2010, pp. 207–236. MR 2807823
131. ———, *Normal approximations with Malliavin calculus*, Cambridge Tracts in Mathematics, vol. 192, Cambridge University Press, Cambridge, 2012, From Stein's method to universality. MR 2962301
132. Ivan Nourdin, Giovanni Peccati, and Gesine Reinert, *Stein's method and stochastic analysis of Rademacher functionals*, Electron. J. Probab. **15** (2010), no. 55, 1703–1742. MR 2735379
133. David Nualart, *Analysis on Wiener space and anticipating stochastic calculus*, Lectures on probability theory and statistics (Saint-Flour, 1995), Lecture Notes in Math., vol. 1690, Springer, Berlin, 1998, pp. 123–227. MR MR1668111 (99k:60144)
134. David Nualart, *Stochastic processes*, 2000.
135. David Nualart, *The Malliavin calculus and related topics*, second ed., Probability and its Applications (New York), Springer-Verlag, Berlin, 2006. MR MR2200233 (2006j:60004)

136. ______, *Malliavin calculus and its applications*, CBMS Regional Conference Series in Mathematics, vol. 110, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2009. MR MR2498953 (2010b:60164)
137. David Nualart and Wim Schoutens, *Backward stochastic differential equations and Feynman-Kac formula for Lévy processes, with applications in finance*, Bernoulli **7** (2001), no. 5, 761–776. MR 1867081 (2002h:60125)
138. B. K. Øksendal, *Stochastic differential equations: An introduction with applications*, Universitext, Springer, September 22 2010.
139. Bernt Øksendal and Kristin Reikvam, *Viscosity solutions of optimal stopping problems*, Stochastics Stochastics Rep. **62** (1998), no. 3-4, 285–301. MR 99a:60044
140. Horst Osswald, *Malliavin calculus for Lévy processes and infinite-dimensional Brownian motion*, Cambridge Tracts in Mathematics, vol. 191, Cambridge University Press, Cambridge, 2012, An introduction. MR 2918805
141. Vissarion Papadopoulos and Dimitris G. Giovanis, *Stochastic finite element methods*, Mathematical Engineering, Springer, Cham, 2018, An introduction. MR 3726875
142. É. Pardoux, *Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order*, Stochastic analysis and related topics, VI (Geilo, 1996), Progr. Probab., vol. 42, Birkhäuser Boston, Boston, MA, 1998, pp. 79–127. MR 99m:35279
143. Etienne Pardoux and Aurel Răşcanu, *Stochastic differential equations, backward SDEs, partial differential equations*, Stochastic Modelling and Applied Probability, vol. 69, Springer, Cham, 2014. MR 3308895
144. Vygantas Paulauskas, *Encyclopaedia of mathematics. D–Feynman measure*, vol. 3, ch. Skorohod space, pp. 370–371, Kluwer Academic Publishers, Dordrecht, 1989.
145. David Pollard, *Convergence of stochastic processes*, Springer Series in Statistics, Springer-Verlag, New York, 1984. MR 762984 (86i:60074)
146. Yu. V. Prohorov, *Convergence of random processes and limit theorems in probability theory*, Teor. Veroyatnost. i Primenen. **1** (1956), 177–238. MR 0084896 (18,943b)
147. Philip E. Protter, *Stochastic integration and differential equations*, Stochastic Modelling and Applied Probability, vol. 21, Springer-Verlag, Berlin, 2005, Second edition. Version 2.1, Corrected third printing. MR MR2273672
148. Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999. MR MR1725357 (2000h:60050)
149. ______, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999. MR MR1725357 (2000h:60050)
150. Marianito R. Rodrigo and Rogemar S. Mamon, *An alternative approach to the calibration of the Vasicek and CIR interest rate models via generating functions*, Quant. Finance **14** (2014), no. 11, 1961–1970. MR 3267769
151. Dan Romik, *Math 235B (Probability Theory) Lecture Notes*, Lecture notes, University of California, Davis, Department of Mathematics, UC Davis, USA, March 15 2012.
152. Sheldon M. Ross, *Introduction to probability models*, 10 ed., Academic Press, imprint of Elsevier, Burlington, MA 01803, USA, December 2010.
153. Walter Rudin, *Functional analysis*, second ed., International Series in Pure and Applied Mathematics, McGraw-Hill Inc., New York, 1991. MR MR1157815 (92k:46001)
154. Marta Sanz-Solé, *Malliavin calculus*, Fundamental Sciences, EPFL Press, Lausanne, 2005, With applications to stochastic partial differential equations. MR MR2167213 (2006h:60005)
155. Jan Seidler, *Ergodic behaviour of stochastic parabolic equations*, Czechoslovak Math. J. **47(122)** (1997), no. 2, 277–316. MR MR1452421 (98e:60099)
156. Michael Sharpe, *General theory of Markov processes*, Pure and Applied Mathematics, vol. 133, Academic Press Inc., Boston, MA, 1988. MR MR958914 (89m:60169)

157. Steven E. Shreve, *Stochastic calculus for finance. I*, Springer Finance, Springer-Verlag, New York, 2004, The binomial asset pricing model. MR MR2049045 (2004m:91003)
158. Barry Simon, *Functional integration and quantum physics*, Pure and Applied Mathematics, vol. 86, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1979. MR MR544188 (84m:81066)
159. A. V. Skorohod, *Limit theorems for stochastic processes*, Teor. Veroyatnost. i Primenen. **1** (1956), 289–319. MR 0084897 (18,943c)
160. R. T. Smythe, *Strong laws of large numbers for r-dimensional arrays of random variables*, Ann. Probability **1** (1973), no. 1, 164–170. MR 0346881 (49 #11602)
161. Charles Stein, *A bound for the error in the normal approximation to the distribution of a sum of dependent random variables*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory (Berkeley, Calif.), Univ. California Press, 1972, pp. 583–602. MR 0402873 (53 #6687)
162. ———, *Approximate computation of expectations*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, 7, Institute of Mathematical Statistics, Hayward, CA, 1986. MR 882007 (88j:60055)
163. David Stirzaker, *Stochastic processes and models*, Oxford University Press, Oxford, 2005. MR 2169515 (2006k:60004)
164. Daniel W. Stroock, *The Malliavin calculus, a functional analytic approach*, J. Funct. Anal. **44** (1981), no. 2, 212–257. MR MR642917 (83h:60076)
165. ———, *The Malliavin calculus and its application to second order parabolic differential equations. I*, Math. Systems Theory **14** (1981), no. 1, 25–65. MR 603973 (84d:60092a)
166. ———, *The Malliavin calculus and its application to second order parabolic differential equations. II*, Math. Systems Theory **14** (1981), no. 2, 141–171. MR 616961 (84d:60092b)
167. ———, *A concise introduction to the theory of integration*, third ed., Birkhäuser Boston Inc., Boston, MA, 1999. MR MR1658777 (99i:28003)
168. ———, *Probability theory, an analytic view*, Cambridge University Press, Cambridge, 2000.
169. ———, *An introduction to Markov processes*, Graduate Texts in Mathematics, vol. 230, Springer-Verlag, Berlin, 2005. MR MR2126069 (2005k:60003)
170. ———, *Elements of stochastic calculus and analysis*, CRM Short Courses, Springer, Cham, 2018. MR 3823207
171. Daniel W. Stroock and S. R. Srinivasa Varadhan, *Multidimensional diffusion processes*, Classics in Mathematics, Springer-Verlag, Berlin, 2006, Reprint of the 1997 edition. MR 2190038 (2006f:60005)
172. Jan Swart and Anita Winter, *Markov processes: theory and examples*, internet: ps-file, March 2005.
173. Atsushi Takeuchi, *Bismut-Elworthy-Li-type formulae for stochastic differential equations with jumps*, J. Theoret. Probab. **23** (2010), no. 2, 576–604. MR 2644877 (2012a:60174)
174. Terence Tao, *An epsilon of room, I: real analysis*, Graduate Studies in Mathematics, vol. 117, American Mathematical Society, Providence, RI, 2010, Pages from year three of a mathematical blog. MR 2760403 (2012b:42002)
175. ———, *An introduction to measure theory*, Graduate Studies in Mathematics, vol. 126, American Mathematical Society, Providence, RI, 2011. MR 2827917 (2012h:28003)
176. Allanus Tsoi, David Nualart, and George Yin (eds.), *Stochastic analysis, stochastic systems, and applications to finance*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2011. MR 2882737 (2012h:60006)
177. A. Süleyman Üstünel and Moshe Zakai, *Transformation of measure on Wiener space*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2000. MR MR1736980 (2001g:60137)
178. J. A. Van Casteren, *Some problems in stochastic analysis and semigroup theory*, Semigroups of operators: theory and applications (Newport Beach, CA, 1998), Birkhäuser, Basel, 2000, pp. 43–60. MR 2001k:47064

179. Jan A. Van Casteren, *Generators of strongly continuous semigroups*, Research Notes in Mathematics, vol. 115, Pitman, 1985, Pitman Advanced Publishing Program.
180. Jan A. van Casteren, *On martingales and Feller semigroups*, Results Math. **21** (1992), no. 3-4, 274–288. MR MR1157331 (93h:60115)
181. Jan A. Van Casteren, *On the Korovkin property and Feller semigroups*, Stochastic analysis and mathematical physics (Santiago, 1998), Trends Math., Birkhäuser Boston, Boston, MA, 2000, pp. 123–154. MR MR1764791 (2001i:47067)
182. ______, *Feynman-Kac semigroups, martingales and wave operators*, J. Korean Math. Soc. **38** (2001), no. 2, 227–274. MR MR1817618 (2002b:47086)
183. ______, *Markov processes and Feller semigroups*, Conf. Semin. Mat. Univ. Bari (2002), no. 286, 1–75 (2003). MR MR1988245 (2004e:47064)
184. ______, *Markov processes, Feller semigroups and evolution equations*, Series on Concrete and Applicable Mathematics, vol. 12, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2011. MR 2779929
185. ______, *On backward stochastic differential equations in infinite dimensions*, Discrete and Continuous Dynamical Systems - Series S (DCDS-S) **6** (2013), no. 3, 803–824.
186. ______, *Partial differential equations and operators fundamental solutions and semigroups*, 3rd ed., Bookboon, 2018, e-book.
187. Aad W. van der Vaart and Jon A. Wellner, *Weak convergence and empirical processes*, Springer Series in Statistics, Springer-Verlag, New York, 1996, With applications to statistics. MR 1385671
188. Jan van Neerven, *The Doob-Meyer decomposition theorem*, Electronically: `http://en.wikipedia.org/wiki/Doob-Meyer_decomposition_theorem`, 2004, Seminar Lectures Technical University Delft: Lecture 3.
189. ______, *Stochastische integratie in Banachruimten*, Nieuw Archief voor Wiskunde, vijfde serie **15** (2014), no. 3, 180–187.
190. Jan van Neerven, Mark Veraar, and Lutz Weis, *Stochastic integration in Banach spaces–a survey*, Proceedings of a stochastic analysis semester in Lausanne (2014), 36 pages.
191. Limin Wang, *Karhunen-Loève expansions and their applications*, Ph.D. thesis, The London School of Economics and Political Science, March 16 2008.
192. Norbert Wiener, *Differential-space*, J. Math. Phys. **2** (1923), 131–174.
193. ______, *Collected works. Vol. I*, MIT Press, Cambridge, Mass., 1976, Mathematical philosophy and foundations; potential theory; Brownian movement, Wiener integrals, ergodic and chaos theories, turbulence and statistical mechanics, With commentaries, Edited by P. Masani, Mathematicians of Our Time, 10.
194. Stephen Willard, *General topology*, Dover Publications Inc., Mineola, NY, 2004, Reprint of the 1970 original [Addison-Wesley, Reading, MA; MR0264581]. MR 2048350
195. Liqing Yan, *The Euler scheme with irregular coefficients*, Ann. Probab. **30** (2002), no. 3, 1172–1194. MR 1920104
196. J. Yeh, *Martingales and stochastic analysis*, Series on Multivariate Analysis, vol. 1, World Scientific Publishing Co. Inc., River Edge, NJ, 1995. MR MR1412800 (97j:60002)
197. Kōsaku Yosida, *Functional analysis*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the sixth (1980) edition. MR MR1336382 (96a:46001)
198. Jianfeng Zhang, *Backward stochastic differential equations*, Probability Theory and Stochastic Modelling, vol. 86, Springer, New York, 2017, From linear to fully nonlinear theory. MR 3699487
199. Tusheng Zhang and Xunyu Zhou (eds.), *Stochastic analysis and applications to finance*, Interdisciplinary Mathematical Sciences, vol. 13, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2012, Essays in honour of Jia-an Yan. MR 2976662
200. Zhongqiang Zhang and George Em Karniadakis, *Numerical methods for stochastic partial differential equations with white noise*, Applied Mathematical Sciences, vol. 196, Springer, Cham, 2017. MR 3700670

Index