Chapter 21
Brownian Motion

In Example 14.45, we constructed a (canonical) process $(X_t)_{t\in[0,\infty)}$ with independent stationary normally distributed increments. For example, such a process can be used to describe the motion of a particle immersed in water or the change of prices in the stock market. We are now interested in properties of this process X that cannot be described in terms of finite-dimensional distributions but reflect the whole *path* $t \mapsto X_t$. For example, we want to compute the distribution of the functional $F(X) := \sup_{t\in[0,1]} X_t$. The first problem that has to be resolved is to show that $F(X)$ is a random variable.

In this chapter, we investigate continuity properties of paths of stochastic processes and show how they ensure measurability of some path functionals. Then we construct a version of X that has continuous paths, the so-called *Wiener process* or *Brownian motion*. Without exaggeration, it can be stated that Brownian motion is *the* central object of probability theory.

For further reading, we recommend, e.g., [86, 118, 144, 151].

21.1 Continuous Versions

A priori the paths of a canonical process are of course not continuous since *every* map $[0,\infty) \to \mathbb{R}$ is possible. Hence, it will be important to find out which paths are $\mathbf{P}$-almost surely negligible.

Definition 21.1 Let X and Y be stochastic processes on $(\Omega, \mathcal{A}, \mathbf{P})$ with time set I and state space E. X and Y are called

(i) *modifications* or *versions* of each other if, for any $t \in I$, we have

$$X_t = Y_t \qquad \mathbf{P}\text{-almost surely},$$

(ii) *indistinguishable* if there exists an $N \in \mathcal{A}$ with $\mathbf{P}[N] = 0$ such that

$$\{X_t \neq Y_t\} \subset N \quad \text{for all } t \in I.$$

A. Klenke, *Probability Theory*, Universitext, DOI 10.1007/978-1-4471-5361-0_21,

Clearly, indistinguishable processes are modifications of each other. Under certain assumptions on the continuity of the paths, however, the two notions coincide.

Definition 21.2 Let (E,d) and (E',d') be metric spaces and $\gamma \in (0,1]$. A map $\varphi : E \to E'$ is called *Hölder-continuous* of order γ (briefly, Hölder-γ-continuous) at the point $r \in E$ if there exist $\varepsilon > 0$ and $C < \infty$ such that, for any $s \in E$ with $d(s,r) < \varepsilon$, we have

$$d'\big(\varphi(r),\varphi(s)\big) \le C d(r,s)^{\gamma}. \tag{21.1}$$

φ is called locally Hölder-continuous of order γ if, for every $t \in E$, there exist $\varepsilon > 0$ and $C = C(t,\varepsilon) > 0$ such that, for all $s, r \in E$ with $d(s,t) < \varepsilon$ and $d(r,t) < \varepsilon$, the inequality (21.1) holds. Finally, φ is called Hölder-continuous of order γ if there exists a C such that (21.1) holds for all $s, r \in E$.

In the case $\gamma = 1$, Hölder continuity is Lipschitz continuity (see Definition 13.8). Furthermore, for $E = \mathbb{R}$ and $\gamma > 1$, every locally Hölder-γ-continuous function is constant. Evidently, a locally Hölder-γ-continuous map is Hölder-γ-continuous at every point. On the other hand, for a function φ that is Hölder-γ-continuous at a given point t, there need not exist an open neighborhood in which φ is continuous. In particular, φ need not be locally Hölder-γ-continuous.

We collect some simple properties of Hölder-continuous functions.

Lemma 21.3 *Let $I \subset \mathbb{R}$ and let $f : I \to \mathbb{R}$ be locally Hölder-continuous of order $\gamma \in (0,1]$. Then the following statements hold.*

(i) *f is locally Hölder-continuous of order γ' for every $\gamma' \in (0,\gamma)$.*
(ii) *If I is compact, then f is Hölder-continuous.*
(iii) *Let I be a bounded interval of length $T > 0$. Assume that there exists an $\varepsilon > 0$ and an $C(\varepsilon) < \infty$ such that, for all $s, t \in I$ with $|t-s| \le \varepsilon$, we have*

$$\big|f(t) - f(s)\big| \le C(\varepsilon)|t-s|^{\gamma}.$$

Then f is Hölder-continuous of order γ with constant $C := C(\varepsilon)\lceil T/\varepsilon\rceil^{1-\gamma}$.

Proof (i) This is obvious since $|t-s|^{\gamma} \le |t-s|^{\gamma'}$ for all $s, t \in I$ with $|t-s| \le 1$.

(ii) For $t \in I$ and $\varepsilon > 0$, let $U_\varepsilon(t) := \{s \in I : |s-t| < \varepsilon\}$. For every $t \in I$, choose $\varepsilon(t) > 0$ and $C(t) < \infty$ such that

$$\big|f(r) - f(s)\big| \le C(t)\cdot|r-s|^{\gamma} \quad \text{for all } r, s \in U_t := U_{\varepsilon(t)}(t).$$

There exists a finite subcovering $\mathfrak{U}' = \{U_{t_1}, \ldots, U_{t_n}\}$ of the covering $\mathfrak{U} := \{U_t, t \in I\}$ of I. Let $\varrho > 0$ be a Lebesgue number of the covering $\mathfrak{U}'$; that is, $\varrho > 0$ is such that, for every $t \in I$, there exists a $U \in \mathfrak{U}$ such that $U_\varrho(t) \subset U$. Define

$$\overline{C} := \max\big\{C(t_1), \ldots, C(t_n), 2\|f\|_\infty \varrho^{-\gamma}\big\}.$$

For $s, t \in I$ with $|t-s| < \varrho$, there is an $i \in \{1, \ldots, n\}$ with $s, t \in U_{t_i}$. By assumption, we have $|f(t) - f(s)| \le C(t_i)|t-s|^\gamma \le \overline{C}|t-s|^\gamma$. Now let $s, t \in I$ with $|s-t| \ge \varrho$. Then

$$\big|f(t) - f(s)\big| \le 2\|f\|_\infty \left(\frac{|t-s|}{\varrho}\right)^\gamma \le \overline{C}|t-s|^\gamma.$$

Hence f is Hölder-continuous of order γ with constant $\overline{C}$.

(iii) Let $n = \lceil \frac{T}{\varepsilon} \rceil$. For $s, t \in I$, by assumption, $\frac{|t-s|}{n} \le \varepsilon$ and thus

$$\begin{aligned}\big|f(t) - f(s)\big| &\le \sum_{k=1}^{n} \left| f\left(s + (t-s)\frac{k}{n}\right) - f\left(s + (t-s)\frac{k-1}{n}\right)\right| \\ &\le C(\varepsilon) n^{1-\gamma} |t-s|^\gamma = C|t-s|^\gamma.\end{aligned}$$

□

Definition 21.4 (Path properties) Let $I \subset \mathbb{R}$ and let $X = (X_t, t \in I)$ be a stochastic process on some probability space $(\Omega, \mathcal{A}, \mathbf{P})$ with values in a metric space (E, d). Let $\gamma \in (0, 1]$. For every $\omega \in \Omega$, we say that the map $I \to E$, $t \mapsto X_t(\omega)$ is a *path* of X.

We say that X has almost surely continuous paths, or briefly that X is a.s. continuous, if for almost all $\omega \in \Omega$, the path $t \mapsto X_t(\omega)$ is continuous. Similarly, we define locally Hölder-γ-continuous paths and so on.

Lemma 21.5 *Let X and Y be modifications of each other. Assume that one of the following conditions holds.*

(i) *I is countable.*
(ii) *$I \subset \mathbb{R}$ is a (possibly unbounded) interval and X and Y are almost surely right continuous.*

Then X and Y are indistinguishable.

Proof Define $N_t := \{X_t \ne Y_t\}$ for $t \in I$ and $\bar{N} = \bigcup_{t \in I} N_t$. By assumption, $\mathbf{P}[N_t] = 0$ for every $t \in I$. We have to show that there exists an $N \in \mathcal{A}$ with $\bar{N} \subset N$ and $\mathbf{P}[N] = 0$.

(i) If I is countable, then $N := \bar{N}$ is measurable and $\mathbf{P}[N] \le \sum_{t \in I} \mathbf{P}[N_t] = 0$.

(ii) Now let $I \subset \mathbb{R}$ be an interval and let X and Y be almost surely right continuous. Define

$$\bar{R} := \{X \text{ and } Y \text{are right continuous}\}$$

and choose an $R \in \mathcal{A}$ with $R \subset \bar{R}$ and $\mathbf{P}[R] = 1$. Define

$$\widetilde{I} := \begin{cases} \mathbb{Q} \cap I, & \text{if } I \text{ is open to the right,} \\ (\mathbb{Q} \cap I) \cup \max I, & \text{if } I \text{ is closed to the right,} \end{cases}$$

and $\widetilde{N} := \bigcup_{r \in \widetilde{I}} N_r$. By (i), we have $\mathbf{P}[\widetilde{N}] = 0$. Furthermore, for every $t \in I$,

$$N_t \cap R \subset \bigcup_{r \geq t, r \in \widetilde{I}} (N_r \cap R) \subset \widetilde{N}.$$

Hence

$$\bar{N} \subset R^c \cup \bigcup_{t \in I} N_t \subset R^c \cup \widetilde{N} =: N,$$

and thus $\mathbf{P}[N] \leq \mathbf{P}[R^c] + \mathbf{P}[\widetilde{N}] = 0$. □

We come to the main theorem of this section.

Theorem 21.6 (Kolmogorov–Chentsov) *Let $X = (X_t, t \in [0, \infty))$ be a real-valued process. Assume for every $T > 0$, there are numbers $\alpha, \beta, C > 0$ such that*

$$\mathbf{E}\big[|X_t - X_s|^\alpha\big] \leq C|t - s|^{1+\beta} \quad \text{for all } s, t \in [0, T]. \tag{21.2}$$

Then the following statements hold.

(i) *There is a modification $\widetilde{X} = (\widetilde{X}_t, t \in [0, \infty))$ of X whose paths are locally Hölder-continuous of every order $\gamma \in \big(0, \frac{\beta}{\alpha}\big)$.*
(ii) *Let $\gamma \in \big(0, \frac{\beta}{\alpha}\big)$. For every $\varepsilon > 0$ and $T < \infty$, there exists a number $K < \infty$ that depends only on $\varepsilon, T, \alpha, \beta, C, \gamma$ such that*

$$\mathbf{P}\big[|\tilde{X}_t - \tilde{X}_s| \leq K|t - s|^\gamma, s, t \in [0, T]\big] \geq 1 - \varepsilon. \tag{21.3}$$

Proof (i) It is enough to show that, for any $T > 0$, the process X on $[0, T]$ has a modification X^T that is locally Hölder-continuous of any order $\gamma \in (0, \beta/\alpha)$. For $S, T > 0$, by Lemma 21.5, two such modifications X^S and X^T are indistinguishable on $[0, S \wedge T]$; hence

$$\Omega_{S,T} := \big\{\text{there is a } t \in [0, S \wedge T] \text{ with } X_t^T \neq X_t^S\big\}$$

is a null set and thus also $\Omega_\infty := \bigcup_{S,T \in \mathbb{N}} \Omega_{S,T}$ is a null set. Therefore, defining $\tilde{X}_t(\omega) := X_t^t(\omega)$, $t \geq 0$, for $\omega \in \Omega \setminus \Omega_\infty$, we get that $\tilde{X}$ is a locally Hölder-continuous modification of X on $[0, \infty)$.

Without loss of generality, assume $T = 1$. We show that X has a continuous modification on $[0, 1]$. By Chebyshev's inequality, for every $\varepsilon > 0$,

$$\mathbf{P}\big[|X_t - X_s| \geq \varepsilon\big] \leq C\varepsilon^{-\alpha}|t - s|^{1+\beta}. \tag{21.4}$$

Hence

$$X_s \xrightarrow{s \to t} X_t \quad \text{in probability.} \tag{21.5}$$

The idea is first to construct $\widetilde{X}$ on the dyadic rational numbers and then to extend it continuously to $[0, 1]$. To this end, we will need (21.5). In particular, for $\gamma > 0$,

$n \in \mathbb{N}$ and $k \in \{1, \ldots, 2^n\}$, we have

$$\mathbf{P}\big[|X_{k2^{-n}} - X_{(k-1)2^{-n}}| \geq 2^{-\gamma n}\big] \leq C 2^{-n(1+\beta-\alpha\gamma)}.$$

Define

$$A_n = A_n(\gamma) := \Big\{\max\big\{|X_{k2^{-n}} - X_{(k-1)2^{-n}}|, k \in \big\{1, \ldots, 2^n\big\}\big\} \geq 2^{-\gamma n}\Big\}$$

and

$$B_n := \bigcup_{m=n}^{\infty} A_m \quad \text{and} \quad N := \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} B_n.$$

It follows that, for every $n \in \mathbb{N}$,

$$\mathbf{P}[A_n] \leq \sum_{k=1}^{2^n} \mathbf{P}\big[|X_{k2^{-n}} - X_{(k-1)2^{-n}}| \geq 2^{-\gamma n}\big] \leq C 2^{-n(\beta-\alpha\gamma)}.$$

Now fix $\gamma \in (0, \beta/\alpha)$ to obtain

$$\mathbf{P}[B_n] \leq \sum_{m=n}^{\infty} \mathbf{P}[A_m] \leq C \frac{2^{-(\beta-\alpha\gamma)n}}{1 - 2^{\alpha\gamma-\beta}} \overset{n\to\infty}{\longrightarrow} 0, \tag{21.6}$$

hence $\mathbf{P}[N] = 0$. Now fix $\omega \in \Omega \setminus N$ and choose $n_0 = n_0(\omega)$ such that $\omega \notin \bigcup_{n=n_0}^{\infty} A_n$. Hence

$$\big|X_{k2^{-n}}(\omega) - X_{(k-1)2^{-n}}(\omega)\big| < 2^{-\gamma n} \quad \text{for } k \in \big\{1, \ldots, 2^n\big\},\ n \geq n_0. \tag{21.7}$$

Define the sets of finite dyadic rationals $D_m = \{k2^{-m}, k = 0, \ldots, 2^m\}$, and let $D = \bigcup_{m\in\mathbb{N}} D_m$. Any $t \in D_m$ has a unique dyadic expansion

$$t = \sum_{i=0}^{m} b_i(t)2^{-i} \quad \text{for some } b_i(t) \in \{0, 1\}, i = 0, \ldots, m.$$

Let $m \geq n \geq n_0$ and $s, t \in D_m$, $s \leq t$ with $|s - t| \leq 2^{-n}$. Let $u := \max(D_n \cap [0, s])$. Then

$$u \leq s < u + 2^{-n} \quad \text{and} \quad u \leq t < u + 2^{1-n}$$

and hence $b_i(t - u) = b_i(s - u) = 0$ for $i < n$. Define

$$t_l = u + \sum_{i=n}^{l} b_i(t - u)2^{-i} \quad \text{for } l = n - 1, \ldots, m.$$

Then, we have $t_{n-1} = u$ and $t_m = t$. Furthermore, $t_l \in D_l$ for $l = n, \ldots, m$ and

$$t_l - t_{l-1} \leq 2^{-l} \quad \text{for } l = n, \ldots, m.$$

Hence, by (21.7),

$$\big|X_t(\omega) - X_u(\omega)\big| \le \sum_{l=n}^{m} \big|X_{t_l}(\omega) - X_{t_{l-1}}(\omega)\big| \le \sum_{l=n}^{m} 2^{-\gamma l} \le \frac{2^{-\gamma n}}{1-2^{-\gamma}}.$$

Analogously, we obtain $|X_s(\omega) - X_u(\omega)| \le 2^{-\gamma n}(1-2^{-\gamma})^{-1}$, and thus

$$\big|X_t(\omega) - X_s(\omega)\big| \le 2\frac{2^{-\gamma n}}{1-2^{-\gamma}}. \tag{21.8}$$

Define $C_0 = 2^{1+\gamma}(1-2^{-\gamma})^{-1} < \infty$. Let $s, t \in D$ with $|s-t| \le 2^{-n_0}$. By choosing the minimal $n \ge n_0$ such that $|t-s| \ge 2^{-n}$, we obtain by (21.8),

$$\big|X_t(\omega) - X_s(\omega)\big| \le C_0 |t-s|^{\gamma}. \tag{21.9}$$

As in the proof of Lemma 21.3(iii), we infer (with $K := C_0 2^{(1-\gamma)n_0}$)

$$\big|X_t(\omega) - X_s(\omega)\big| \le K|t-s|^{\gamma} \quad \text{for all } s, t \in D. \tag{21.10}$$

In other words, for dyadic rationals D, $X(\omega)$ is (globally) Hölder-γ-continuous. In particular, X is uniformly continuous on D; hence it can be extended to [0, 1]. For $t \in D$, define $\widetilde{X}_t := X_t$. For $t \in [0,1] \setminus D$ and $\{s_n, n \in \mathbb{N}\} \subset D$ with $s_n \longrightarrow t$, the sequence $(X_{s_n}(\omega))_{n\in\mathbb{N}}$ is a Cauchy sequence. Hence the limit

$$\widetilde{X}_t(\omega) := \lim_{D \ni s \to t} X_s(\omega) \tag{21.11}$$

exists. Furthermore, the statement analogous to (21.10) holds not only for $s, t \in D$:

$$\big|\widetilde{X}_t(\omega) - \widetilde{X}_s(\omega)\big| \le K|t-s|^{\gamma} \quad \text{for all } s, t \in [0,1]. \tag{21.12}$$

Hence $\widetilde{X}$ is locally Hölder-continuous of order γ. By (21.5) and (21.11), we have $\mathbf{P}[X_t \ne \widetilde{X}_t] = 0$ for every $t \in [0,1]$. Hence $\widetilde{X}$ is a modification of X.

(ii) Let $\varepsilon > 0$ and choose $n \in \mathbb{N}$ large enough that (see (21.6))

$$\mathbf{P}[B_n] \le C\frac{2^{-(\beta-\alpha\gamma)n}}{1-2^{\alpha\gamma-\beta}} < \varepsilon.$$

For $\omega \notin B_n$, we conclude that (21.10) holds. However, this is exactly (21.3) with $T = 1$. For general T, the claim follows by linear scaling. □

Remark 21.7 The statement of Theorem 21.6 remains true if X assumes values in some Polish space (E, ϱ) since in the proof we did not make use of the assumption that the range was in $\mathbb{R}$. However, if we change the time set, then the assumptions have to be strengthened: If $(X_t)_{t\in\mathbb{R}^d}$ is a process with values in E, and if, for certain $\alpha, \beta > 0$, all $T > 0$ and some $C < \infty$, we have

$$\mathbf{E}\big[\varrho(X_t, X_s)^\alpha\big] \le C\|t-s\|_2^{d+\beta} \quad \text{for all } s,t \in [-T,T]^d, \tag{21.13}$$

then for every $\gamma \in (0, \beta/\alpha)$, there is a locally Hölder-γ-continuous version of X. $\Diamond$

Exercise 21.1.1 Show the claim of Remark 21.7.

Exercise 21.1.2 Let $X = (X_t)_{t\ge 0}$ be a real-valued process with continuous paths. Show that, for all $0 \le a < b$, the map $\omega \mapsto \int_a^b X_t(\omega)\, dt$ is measurable.

Exercise 21.1.3 (Optional sampling/stopping) Let $\mathbb{F}$ be a filtration and let $(X_t)_{t\ge0}$ be an $\mathbb{F}$-supermartingale with right continuous paths. Let σ and τ be bounded stopping times with $\sigma \le \tau$. Define $\sigma^n := 2^{-n}\lceil 2^n\sigma\rceil$ and $\tau^n := 2^{-n}\lceil 2^n\tau\rceil$.

(i) Show that $\mathbf{E}[X_{\tau^m} \mid \mathcal{F}_{\sigma^n}] \overset{n\to\infty}{\longrightarrow} \mathbf{E}[X_{\tau^m} \mid \mathcal{F}_\sigma]$ almost surely and in L^1 as well as $X_{\sigma_n} \overset{n\to\infty}{\longrightarrow} X_\sigma$ almost surely and in L^1.
(ii) Infer the optional sampling theorem for right continuous supermartingales by using the analogous statement for discrete time (Theorem 10.11); that is, $X_\sigma \ge \mathbf{E}[X_\tau \mid \mathcal{F}_\sigma]$.
(iii) Show that if Y is adapted, integrable and right continuous, then Y is a martingale if and only if $\mathbf{E}[Y_\tau] = \mathbf{E}[Y_0]$ for every bounded stopping time τ.
(iv) Assume that X is uniformly integrable and that $\sigma \le \tau$ are *finite* (not necessarily bounded) stopping times. Show that $X_\sigma \ge \mathbf{E}[X_\tau \mid \mathcal{F}_\sigma]$.
(v) Now let τ be an arbitrary stopping time. Deduce the optional stopping theorem for right continuous supermartingales: $(X_{\tau\wedge t})_{t\ge0}$ is a right continuous supermartingale.

Exercise 21.1.4 Let $X = (X_t)_{t\ge0}$ be a stochastic process on $(\Omega, \mathcal{F}, \mathbf{P})$ with values in the Polish space E and with right continuous paths. Show the following.

(i) The map $(\omega, t) \mapsto X_t(\omega)$ is measurable with respect to $\mathcal{F} \otimes \mathcal{B}([0,\infty)) - \mathcal{B}(E)$.
(ii) If in addition X is adapted to the filtration $\mathbb{F}$, then for any $t \ge 0$, the map $\Omega \times [0,t] \to E$, $(\omega, s) \mapsto X_s(\omega)$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t]) - \mathcal{B}(E)$ measurable.
(iii) If τ is an $\mathbb{F}$-stopping time and X is adapted, then X_τ is an $\mathcal{F}_\tau$-measurable random variable.

21.2 Construction and Path Properties

Definition 21.8 A real-valued stochastic process $B = (B_t, t \in [0,\infty))$ is called a *Brownian motion* if

(i) $B_0 = 0$,
(ii) B has independent, stationary increments (compare Definition 9.7),
(iii) $B_t \sim \mathcal{N}_{0,t}$ for all $t > 0$, and
(iv) $t \mapsto B_t$ is $\mathbf{P}$-almost surely continuous.

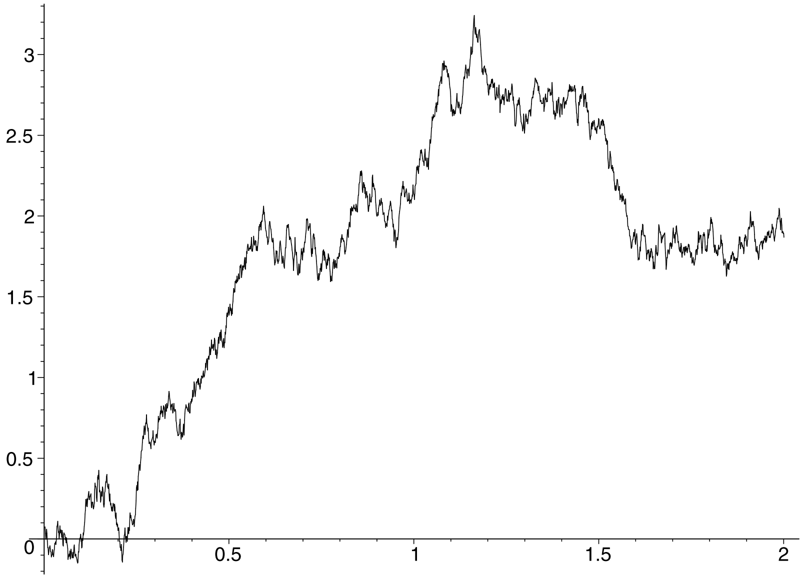

Fig. 21.1 Computer simulation of a Brownian motion

See Fig. 21.1 for a computer simulation of a Brownian motion.

Theorem 21.9 *There exists a probability space* $(\Omega, \mathcal{A}, \mathbf{P})$ *and a Brownian motion* B *on* $(\Omega, \mathcal{A}, \mathbf{P})$. *The paths of* B *are a.s. locally Hölder-*γ*-continuous for every* $\gamma < \frac{1}{2}$.

Proof As in Example 14.45 or Corollary 16.10 there exists a stochastic process X that fulfills (i), (ii) and (iii). Evidently, $X_t - X_s \stackrel{\mathcal{D}}{=} \sqrt{t-s}X_1 \sim \mathcal{N}_{0,t-s}$ for all $t > s \geq 0$. Thus, for every $n \in \mathbb{N}$, writing $C_n := \mathbf{E}[X_1^{2n}] = \frac{(2n)!}{2^n n!} < \infty$, we have

$$\mathbf{E}\big[(X_t - X_s)^{2n}\big] = \mathbf{E}\big[(\sqrt{t-s}X_1)^{2n}\big] = C_n|t-s|^n.$$

Now let $n \geq 2$ and $\gamma \in (0, \frac{n-1}{2n})$. Theorem 21.6 yields the existence of a version B of X that has Hölder-γ-continuous paths. Since all continuous versions of a process are equivalent, B is locally Hölder-γ-continuous for every $\gamma \in (0, \frac{n-1}{2n})$ and every $n \geq 2$ and hence for every $\gamma \in (0, \frac{1}{2})$. $\square$

Recall that a stochastic process $(X_t)_{t \in I}$ is called a *Gaussian process* if, for every $n \in \mathbb{N}$ and for all $t_1, \ldots, t_n \in I$, we have that

$$(X_{t_1}, \ldots, X_{t_n}) \quad \text{is } n\text{-dimensional normally distributed.}$$

X is called *centered* if $\mathbf{E}[X_t] = 0$ for every $t \in I$. The map

$$\Gamma(s,t) := \mathbf{Cov}[X_s, X_t] \quad \text{for } s, t \in I$$

is called the *covariance function* of X.

Remark 21.10 The covariance function determines the finite-dimensional distributions of a centered Gaussian process since a multidimensional normal distribution is determined by the vector of expectations and by the covariance matrix. ◊

Theorem 21.11 *Let $X = (X_t)_{t\in[0,\infty)}$ be a stochastic process. Then the following are equivalent*:

(i) *X is a Brownian motion.*
(ii) *X is a continuous centered Gaussian process with $\mathbf{Cov}[X_s, X_t] = s \wedge t$ for all $s, t \geq 0$.*

Proof By Remark 21.10, X is characterized by (ii). Hence, it is enough to show that, for Brownian motion X, we have $\mathbf{Cov}[X_s, X_t] = \min(s, t)$. This is indeed true since for $t > s$, the random variables X_s and $X_t - X_s$ are independent; hence

$$\mathbf{Cov}[X_s, X_t] = \mathbf{Cov}[X_s, X_t - X_s] + \mathbf{Cov}[X_s, X_s] = \mathbf{Var}[X_s] = s. \qquad \square$$

Corollary 21.12 (Scaling property of Brownian motion) *If B is a Brownian motion and if $K \neq 0$, then $(K^{-1} B_{K^2 t})_{t\geq 0}$ is also a Brownian motion.*

Example 21.13 Another example of a continuous Gaussian process is the so-called *Brownian bridge* $X = (X_t)_{t\in[0,1]}$ that is defined by the covariance function $\Gamma(s, t) = s \wedge t - st$. We construct the Brownian bridge as follows.

Let $B = (B_t, t \in [0, 1])$ be a Brownian motion and let

$$X_t := B_t - tB_1.$$

Clearly, X is a centered Gaussian process with continuous paths. We compute the covariance function Γ of X,

$$\begin{aligned}\Gamma(s, t) &= \mathbf{Cov}[X_s, X_t] = \mathbf{Cov}[B_s - sB_1, B_t - tB_1] \\ &= \mathbf{Cov}[B_s, B_t] - s\mathbf{Cov}[B_1, B_t] - t\mathbf{Cov}[B_s, B_1] + st\mathbf{Cov}[B_1, B_1] \\ &= \min(s, t) - st - st + st = \min(s, t) - st. \end{aligned} \qquad \Diamond$$

Theorem 21.14 *Let $(B_t)_{t\geq 0}$ be a Brownian motion and*

$$X_t = \begin{cases} tB_{1/t}, & \text{if } t > 0, \\ 0, & \text{if } t = 0. \end{cases}$$

Then X is a Brownian motion.

Proof Clearly, X is a Gaussian process. For $s, t > 0$, we have

$$\mathbf{Cov}[X_s, X_t] = ts \cdot \mathbf{Cov}[B_{1/s}, B_{1/t}] = ts \min\bigl(s^{-1}, t^{-1}\bigr) = \min(s, t).$$

Clearly, $t \mapsto X_t$ is continuous at every point $t > 0$. To show continuity at $t = 0$, consider

$$\begin{aligned}\limsup_{t\downarrow 0} X_t &= \limsup_{t\to\infty} \frac{1}{t} B_t \\ &\le \limsup_{n\to\infty} \frac{1}{n} B_n + \limsup_{n\to\infty} \frac{1}{n} \sup\big\{B_t - B_n, t \in [n, n+1]\big\}.\end{aligned}$$

By the strong law of large numbers, we have $\lim_{n\to\infty} \frac{1}{n} B_n = 0$ a.s. Using a generalization of the reflection principle (Theorem 17.15; see also Theorem 21.19), for $x > 0$, we have (using the abbreviation $B_{[a,b]} := \{B_t : t \in [a, b]\}$)

$$\begin{aligned}\mathbf{P}[\sup B_{[n,n+1]} - B_n > x] = \mathbf{P}[\sup B_{[0,1]} > x] &= 2\mathbf{P}[B_1 > x] \\ &= \frac{2}{\sqrt{2\pi}} \int_x^\infty e^{-u^2/2}\, du \le \frac{1}{x} e^{-x^2/2}.\end{aligned}$$

In particular, $\sum_{n=1}^\infty \mathbf{P}[\sup B_{[n,n+1]} - B_n > n^\varepsilon] < \infty$ for every $\varepsilon > 0$. By the Borel–Cantelli lemma (Theorem 2.7), we infer

$$\limsup_{n\to\infty} \frac{1}{n} \sup\big\{B_t - B_n, t \in [n, n+1]\big\} = 0 \quad \text{almost surely.}$$

Hence X is also continuous at 0. □

Theorem 21.15 (Blumenthal's 0–1 law, see [18]) *Let B be a Brownian motion and let $\mathbb{F} = (\mathcal{F}_t)_{t\ge 0} = \sigma(B)$ be the filtration generated by B. Further, let $\mathcal{F}_0^+ = \bigcap_{t>0} \mathcal{F}_t$. Then $\mathcal{F}_0^+$ is a* **P***-trivial σ-algebra.*

Proof Define $Y^n = (B_{2^{-n}+t} - B_{2^{-n}})_{t\in[0,2^{-n}]}$, $n \in \mathbb{N}$. Then $(Y^n)_{n\in\mathbb{N}}$ is an independent family of random variables (with values in $C([0, 2^{-n}])$). By Kolmogorov's 0–1 law (Theorem 2.37), the tail σ-algebra $\mathcal{T} = \bigcap_{n\in\mathbb{N}} \sigma(Y^m, m \ge n)$ is **P**-trivial. On the other hand, $\sigma(Y^m, m \ge n) = \mathcal{F}_{2^{-n+1}}$; hence

$$\mathcal{F}_0^+ = \bigcap_{t>0} \mathcal{F}_t = \bigcap_{n\in\mathbb{N}} \mathcal{F}_{2^{-n+1}} = \mathcal{T}$$

is **P**-trivial. □

Example 21.16 Let B be a Brownian motion. For every $K > 0$, we have

$$\mathbf{P}\Big[\inf\{t > 0 : B_t \ge K\sqrt{t}\} = 0\Big] = 1. \tag{21.14}$$

To check this, define $A_s := \{\inf\{t > 0 : B_t \ge K\sqrt{t}\} \le s\}$ and

$$A := \Big\{\inf\{t > 0 : B_t \ge K\sqrt{t}\} = 0\Big\} = \bigcap_{s>0} A_s \in \mathcal{F}_0^+.$$

Then $\mathbf{P}[A] \in \{0, 1\}$. By the scaling property of Brownian motion,

$$\mathbf{P}[A] = \inf_{s>0} \mathbf{P}[A_s] \geq \mathbf{P}[B_1 \geq K] > 0$$

and thus $\mathbf{P}[A] = 1$. $\Diamond$

The preceding example shows that, for every $t \geq 0$, almost surely B is not Hölder-$\frac{1}{2}$-continuous at t. Note that the order of quantifiers is subtle. We have *not* shown that almost surely B was not Hölder-$\frac{1}{2}$-continuous at any $t \geq 0$ (however, see Remark 22.4). However, it is not too hard to show the following theorem, which for the case $\gamma = 1$ is due to Paley, Wiener and Zygmund [126]. The proof presented here goes back to an idea of Dvoretzky, Erdös and Kakutani (see [40]).

Theorem 21.17 (Paley–Wiener–Zygmund (1933)) *For every $\gamma > \frac{1}{2}$, almost surely the paths of Brownian motion $(B_t)_{t\geq 0}$ are not Hölder-continuous of order γ at any point. In particular, the paths are almost surely nowhere differentiable.*

Proof Let $\gamma > \frac{1}{2}$. It suffices to consider $B = (B_t)_{t\in[0,1]}$. Denote by $H_{\gamma,t}$ the set of maps $[0, 1] \to \mathbb{R}$ that are Hölder-γ-continuous at t and define $H_\gamma := \bigcup_{t\in[0,1]} H_{\gamma,t}$. The aim is to show that almost surely $B \notin H_\gamma$.

If $t \in [0, 1)$ and $w \in H_{\gamma,t}$, then for every $\delta > 0$ there exists a $c = c(\delta, w)$ with the property $|w_s - w_t| \leq c|s - t|^\gamma$ for every $s \in [0, 1]$ with $|s - t| < \delta$. Choose a $k \in \mathbb{N}$ with $k > \frac{2}{2\gamma-1}$. Then, for $n \in \mathbb{N}$ with $n \geq n_0 := \lceil (k + 1)/\delta \rceil$, $i = \lfloor tn \rfloor + 1$ and $l \in \{0, \dots, k - 1\}$, we get

$$\begin{aligned}|w_{(i+l+1)/n} - w_{(i+l)/n}| &\leq |w_{(i+l+1)/n} - w_t| + |w_{(i+l)/n} - w_t| \\ &\leq 2c(k + 1)^\gamma n^{-\gamma}.\end{aligned}$$

Hence, for $N \geq 2c(k + 1)^\gamma$, we have $w \in A_{N,n,i}$, where

$$A_{N,n,i} := \bigcap_{l=0}^{k-1} \big\{ w : |w_{(i+l+1)/n} - w_{(i+l)/n}| \leq N n^{-\gamma} \big\}.$$

Define $A_{N,n} = \bigcup_{i=1}^{n} A_{N,n,i}$, $A_N = \liminf_{n\to\infty} A_{N,n}$ and $A = \bigcup_{N=1}^{\infty} A_N$. Clearly, $H_\gamma \subset A$. Owing to the independence of increments and since the density of the standard normal distribution is bounded by 1, we get

$$\begin{aligned}\mathbf{P}[B \in A_{N,n,i}] &= \mathbf{P}\big[|B_{1/n}| \leq N n^{-\gamma}\big]^k = \mathbf{P}\big[|B_1| \leq N n^{-\gamma+1/2}\big]^k \\ &\leq N^k n^{k(-\gamma+1/2)}.\end{aligned}$$

By the choice of k and since the increments of B are stationary, we have

$$\begin{aligned}\mathbf{P}[B \in A_N] &= \lim_{n\to\infty} \mathbf{P}\Bigg[\bigcap_{m\geq n} A_{N,m}\Bigg] \leq \limsup_{n\to\infty} \mathbf{P}[A_{N,n}] \leq \limsup_{n\to\infty} \sum_{i=1}^{n} \mathbf{P}[A_{N,n,i}] \\ &\leq \limsup_{n\to\infty} n\mathbf{P}[B \in A_{N,n,1}] \leq N^k \limsup_{n\to\infty} n^{1+k(-\gamma+1/2)} = 0.\end{aligned}$$

Thus $\mathbf{P}[B \in A] = 0$. Therefore, we almost surely have $B \notin H_\gamma$. □

Exercise 21.2.1 Let B be a Brownian motion and let λ be the Lebesgue measure on $[0, \infty)$.

(i) Compute the expectation and variance of $\int_0^1 B_s\, ds$. (For the measurability of the integral see Exercise 21.1.2.)
(ii) Show that almost surely $\lambda(\{t : B_t = 0\}) = 0$.
(iii) Compute the expectation and variance of

$$\int_0^1 \left(B_t - \int_0^1 B_s\, ds\right)^2 dt.$$

Exercise 21.2.2 Let B be a Brownian motion. Show that $(B_t^2 - t)_{t\geq 0}$ is a martingale.

Exercise 21.2.3 Let B be a Brownian motion and $\sigma > 0$. Show that the process $(\exp(\sigma B_t - \frac{\sigma^2}{2} t))_{t\geq 0}$ is a martingale.

Exercise 21.2.4 Let B be a Brownian motion, $a < 0 < b$. Define the stopping time $\tau_{a,b} = \inf\{t \geq 0 : B_t \in \{a, b\}\}$.

Show that almost surely $\tau_{a,b} < \infty$ and that $\mathbf{P}[B_{\tau_{a,b}} = b] = -\frac{a}{b-a}$. Furthermore, show (using Exercise 21.2.2) that $\mathbf{E}[\tau_{a,b}] = -ab$.

Exercise 21.2.5 Let B be a Brownian motion, $b > 0$ and $\tau_b = \inf\{t \geq 0 : B_t = b\}$. Show the following.

(i) $\mathbf{E}[e^{-\lambda\tau_b}] = e^{-b\sqrt{2\lambda}}$ for $\lambda \geq 0$.
Hint: Use Exercise 21.2.3 and the optional sampling theorem.
(ii) τ_b has a $\frac{1}{2}$-stable distribution with Lévy measure

$$\nu(dx) = \big(b/(\sqrt{2\pi})\big)x^{-3/2}\mathbb{1}_{\{x>0\}}\, dx.$$

(iii) The distribution of τ_b has density $f_b(x) = \frac{b}{\sqrt{2\pi}} e^{-b^2/(2x)} x^{-3/2}$.

Exercise 21.2.6 Let B be a Brownian motion, $a \in \mathbb{R}$, $b > 0$ and $\tau = \inf\{t \geq 0 : B_t = at + b\}$. For $\lambda \geq 0$, show that

$$\mathbf{E}\big[e^{-\lambda\tau}\big] = \exp\big(-ba - b\sqrt{a^2 + 2\lambda}\big).$$

Conclude that $\mathbf{P}[\tau < \infty] = 1 \wedge e^{-2ba}$.

21.3 Strong Markov Property

Denote by $\mathbf{P}_x$ the probability measure such that $B = (B_t)_{t\geq 0}$ is a Brownian motion started at $x \in \mathbb{R}$. To put it differently, under $\mathbf{P}_x$, the process $(B_t - x)_{t\geq 0}$ is a standard Brownian motion. While the (simple) Markov property of $(B, (\mathbf{P}_x)_{x\in\mathbb{R}})$ is evident, it takes some work to check the strong Markov property.

Theorem 21.18 (Strong Markov property) *Brownian motion B with distributions $(\mathbf{P}_x)_{x\in\mathbb{R}}$ has the strong Markov property.*

Proof Let $\mathbb{F} = \sigma(B)$ be the filtration generated by B and let $\tau < \infty$ be an $\mathbb{F}$-stopping time. We have to show that, for every bounded measurable $F : \mathbb{R}^{[0,\infty)} \to \mathbb{R}$, we have:

$$\mathbf{E}_x\big[F\big((B_{t+\tau})_{t\geq 0}\big) \,\big|\, \mathcal{F}_\tau\big] = \mathbf{E}_{B_\tau}\big[F(B)\big]. \tag{21.15}$$

It is enough to consider continuous bounded functions F that depend on only finitely many coordinates $t_1, \ldots, t_N$ since these functions determine the distribution of $(B_{t+\tau})_{t\geq 0}$. Hence, let $f : \mathbb{R}^N \to \mathbb{R}$ be continuous and bounded and $F(B) = f(B_{t_1}, \ldots, B_{t_N})$. Manifestly, the map $x \mapsto \mathbf{E}_x[F(B)] = \mathbf{E}_0[f(B_{t_1} + x, \ldots, B_{t_N} + x)]$ is continuous and bounded. Now let $\tau^n := 2^{-n}\lfloor 2^n\tau + 1\rfloor$ for $n \in \mathbb{N}$. Then τ^n is a stopping time and $\tau^n \downarrow \tau$; hence $B_{\tau^n} \overset{n\to\infty}{\longrightarrow} B_\tau$ almost surely. Now every Markov process with countable time set (here all positive rational linear combinations of $1, t_1, \ldots, t_N$) is a strong Markov process (by Theorem 17.14); hence we have

$$\begin{aligned}\mathbf{E}_x\big[F\big((B_{\tau^n+t})_{t\geq 0}\big) \,\big|\, \mathcal{F}_{\tau^n}\big] &= \mathbf{E}_x\big[f(B_{\tau^n+t_1}, \ldots, B_{\tau^n+t_N}) \,\big|\, \mathcal{F}_{\tau^n}\big] \\ &= \mathbf{E}_{B_{\tau^n}}\big[f(B_{t_1}, \ldots, B_{t_N})\big] \\ &\overset{n\to\infty}{\longrightarrow} \mathbf{E}_{B_\tau}\big[f(B_{t_1}, \ldots, B_{t_N})\big] = \mathbf{E}_{B_\tau}\big[F(B)\big].\end{aligned} \tag{21.16}$$

As B is right continuous, we have $F((B_{\tau^n+t})_{t\geq 0}) \overset{n\to\infty}{\longrightarrow} F((B_{\tau+t})_{t\geq 0})$ almost surely and in L^1 and thus

$$\begin{aligned}&\mathbf{E}\Big[\Big|\mathbf{E}_x\big[F\big((B_{\tau^n+t})_{t\geq 0}\big)\big|\mathcal{F}_{\tau^n}\big] - \mathbf{E}_x\big[F\big((B_{\tau+t})_{t\geq 0}\big)\big|\mathcal{F}_{\tau^n}\big]\Big|\Big] \\ &\quad\leq \mathbf{E}_x\Big[\Big|F\big((B_{\tau^n+t})_{t\geq 0}\big) - F\big((B_{\tau+t})_{t\geq 0}\big)\Big|\Big] \overset{n\to\infty}{\longrightarrow} 0.\end{aligned} \tag{21.17}$$

Furthermore,

$$\mathcal{F}_{\tau^n} \downarrow \mathcal{F}_{\tau+} := \bigcap_{\sigma>\tau \text{ is a stopping time}} \mathcal{F}_\sigma \supset \mathcal{F}_\tau.$$

By (21.16) and (21.17), using the convergence theorem for backwards martingales (Theorem 12.14), we get that in the sense of L^1-limits

$$\begin{aligned}\mathbf{E}_{B_\tau}\big[F(B)\big] &= \lim_{n\to\infty}\mathbf{E}_x\big[F\big((B_{\tau^n+t})_{t\geq 0}\big)\,\big|\,\mathcal{F}_{\tau^n}\big]\\ &= \lim_{n\to\infty}\mathbf{E}_x\big[F\big((B_{\tau+t})_{t\geq 0}\big)\,\big|\,\mathcal{F}_{\tau^n}\big] = \mathbf{E}_x\big[F\big((B_{\tau+t})_{t\geq 0}\big)\,\big|\,\mathcal{F}_{\tau+}\big].\end{aligned}$$

The left-hand side is $\mathcal{F}_\tau$-measurable. The tower property of conditional expectation thus yields (21.15). □

Using the strong Markov property, we show the reflection principle for Brownian motion.

Theorem 21.19 (Reflection principle for Brownian motion) *For every* $a > 0$ *and* $T > 0$,

$$\mathbf{P}\big[\sup\big\{B_t : t \in [0, T]\big\} > a\big] = 2\mathbf{P}[B_T > a] \leq \frac{2\sqrt{T}}{\sqrt{2\pi}}\frac{1}{a}e^{-a^2/2T}.$$

Proof By the scaling property of Brownian motion (Corollary 21.12), without loss of generality, we may assume $T = 1$. Let $\tau := \inf\{t \geq 0 : B_t \geq a\} \wedge 1$. By symmetry, we have $\mathbf{P}_a[B_{1-\tau} > a] = \frac{1}{2}$ if $\tau < 1$; hence

$$\begin{aligned}\mathbf{P}[B_1 > a] &= \mathbf{P}[B_1 > a \mid \tau < 1]\mathbf{P}[\tau < 1]\\ &= \mathbf{P}_a[B_{1-\tau} > a]\mathbf{P}[\tau < 1] = \frac{1}{2}\mathbf{P}[\tau < 1].\end{aligned}$$

For the inequality compute

$$\begin{aligned}\mathbf{P}[B_1 > a] &= \frac{1}{\sqrt{2\pi}}\int_a^\infty e^{-x^2/2}\,dx\\ &\leq \frac{1}{\sqrt{2\pi}}\frac{1}{a}\int_a^\infty x e^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}}\frac{1}{a}e^{-a^2/2}.\end{aligned}$$

□

As an application of the reflection principle we derive Paul Lévy's arcsine law [107, p. 216] for the last time a Brownian motion visits zero.

Theorem 21.20 (Lévy's arcsine law) *Let* $T > 0$ *and* $\zeta_T := \sup\{t \leq T : B_t = 0\}$. *Then, for* $t \in [0, T]$,

$$\mathbf{P}[\zeta_T \leq t] = \frac{2}{\pi}\arcsin\big(\sqrt{t/T}\big).$$

Proof Without loss of generality, assume $T = 1$ and $\zeta = \zeta_1$. Let $\widetilde{B}$ be a further, independent Brownian motion. By the reflection principle,

$$\begin{aligned}
\mathbf{P}[\zeta \le t] &= \mathbf{P}\big[B_s \neq 0 \text{ for all } s \in [t,1]\big] \\
&= \int_{-\infty}^{\infty} \mathbf{P}\big[B_s \neq 0 \text{ for all } s \in [t,1] \,\big|\, B_t = a\big]\mathbf{P}[B_t \in da] \\
&= \int_{-\infty}^{\infty} \mathbf{P}_{|a|}\big[\widetilde{B}_s > 0 \text{ for all } s \in [0,1-t]\big]\mathbf{P}[B_t \in da] \\
&= \int_{-\infty}^{\infty} \mathbf{P}_0\big[|\widetilde{B}_{1-t}| \le |a|\big]\mathbf{P}[B_t \in da] \\
&= \mathbf{P}\big[|\widetilde{B}_{1-t}| \le |B_t|\big].
\end{aligned}$$

If X and Y are independent and $\mathcal{N}_{0,1}$-distributed, then

$$(B_t, \widetilde{B}_{1-t}) \stackrel{\mathcal{D}}{=} (\sqrt{t}X, \sqrt{1-t}Y).$$

Hence

$$\begin{aligned}
\mathbf{P}[\zeta \le t] &= \mathbf{P}\big[\sqrt{1-t}|Y| \le \sqrt{t}|X|\big] \\
&= \mathbf{P}\big[Y^2 \le t\big(X^2 + Y^2\big)\big] \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\, e^{-(x^2+y^2)/2}\mathbb{1}_{\{y^2 \le t(x^2+y^2)\}}.
\end{aligned}$$

Passing to polar coordinates, we obtain

$$\mathbf{P}[\zeta \le t] = \frac{1}{2\pi}\int_0^{\infty} r\,dr\,e^{-r^2/2}\int_0^{2\pi} d\varphi\,\mathbb{1}_{\{\sin(\varphi)^2 \le t\}} = \frac{2}{\pi}\arcsin(\sqrt{t}). \qquad \square$$

Exercise 21.3.1 (Hard problem!) Let $\mathbf{P}_x$ be the distribution of Brownian motion started at $x \in \mathbb{R}$. Let $a > 0$ and $\tau = \inf\{t \ge 0 : B_t \in \{0, a\}\}$. Use the reflection principle to show that, for every $x \in (0, a)$,

$$\mathbf{P}_x[\tau > T] = \sum_{n=-\infty}^{\infty} (-1)^n \mathbf{P}_x\big[B_T \in \big[na, (n+1)a\big]\big]. \tag{21.18}$$

If f is the density of a probability distribution on $\mathbb{R}$ with characteristic function φ and $\sup_{x\in\mathbb{R}} x^2 f(x) < \infty$, then the Poisson summation formula holds,

$$\sum_{n=-\infty}^{\infty} f(s+n) = \sum_{k=-\infty}^{\infty} \varphi(k)e^{2\pi i s} \quad \text{for all } s \in \mathbb{R}. \tag{21.19}$$

Use (21.18) and (21.19) (compare also (21.37)) to conclude that

$$\mathbf{P}_x[\tau > T] = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} \exp\left(-\frac{(2k+1)^2\pi^2 T}{2a^2}\right) \sin\left(\frac{(2k+1)\pi x}{a}\right). \quad (21.20)$$

21.4 Supplement: Feller Processes

In many situations, a continuous version of a process would be too much to expect, for instance, the Poisson process is generically discontinuous. However, often there is a version with right continuous paths that have left-sided limits. At this point, we only briefly make plausible the existence theorem for such regular versions of processes in the case of so-called Feller semigroups.

Definition 21.21 Let E be a Polish space. A map $f : [0,\infty) \to E$ is called *RCLL* (*right continuous with left limits*) or *càdlàg* (*continue à droit, limites à gauche*) if $f(t) = f(t+) := \lim_{s\downarrow t} f(s)$ for every $t \ge 0$ and if, for every $t > 0$, the left-sided limit $f(t-) := \lim_{s\uparrow t} f(s)$ exists and is finite.

Definition 21.22 A filtration $\mathbb{F} = (\mathcal{F}_t)_{t\ge 0}$ is called *right continuous* if $\mathbb{F} = \mathbb{F}^+$, where $\mathcal{F}_t^+ = \bigcap_{s>t} \mathcal{F}_s$. We say that a filtration $\mathbb{F}$ satisfies the *usual conditions* (from the French *conditions habituelles*) if $\mathbb{F}$ is right continuous and if $\mathcal{F}_0$ is $\mathbf{P}$-complete.

Remark 21.23 If $\mathbb{F}$ is an arbitrary filtration and $\mathcal{F}_t^{+,*}$ is the completion of $\mathcal{F}_t^+$, then $\mathbb{F}^{+,*}$ satisfies the usual conditions. ◊

Theorem 21.24 (Doob's regularization) *Let $\mathbb{F}$ be a filtration that satisfies the usual conditions and let $X = (X_t)_{t\ge 0}$ be an $\mathbb{F}$-supermartingale such that $t \mapsto \mathbf{E}[X_t]$ is right continuous. Then there exists a modification $\widetilde{X}$ of X with RCLL paths.*

Proof For $a, b \in \mathbb{Q}^+$, $a < b$ and $I \subset [0,\infty)$, let $U_I^{a,b}$ be the number of upcrossings of $(X_t)_{t\in I}$ over $[a,b]$. By the upcrossing inequality (Lemma 11.3), for every $N > 0$ and every finite set $I \subset [0,N]$, we have $\mathbf{E}[U_I^{a,b}] \le (\mathbf{E}[|X_N|] + |a|)/(b-a)$. Define $U_N^{a,b} = U_{\mathbb{Q}^+\cap[0,N]}^{a,b}$. Then $\mathbf{E}[U_N^{a,b}] \le (\mathbf{E}[|X_N|] + |a|)/(b-a)$. By Exercise 11.1.1, for $\lambda > 0$, we have

$$\begin{aligned}
&\lambda \mathbf{P}\big[\sup\big\{|X_t| : t \in \mathbb{Q}^+ \cap [0,N]\big\} > \lambda\big] \\
&\quad = \lambda \sup\big\{\mathbf{P}\big[\sup\big\{|X_t| : t \in I\big\} > \lambda\big] : I \subset \mathbb{Q}^+ \cap [0,N] \text{ finite}\big\} \\
&\quad \le 12\mathbf{E}\big[|X_0|\big] + 9\mathbf{E}\big[|X_N|\big].
\end{aligned}$$

Consider the event

$$A := \bigcap_{N\in\mathbb{N}} \Bigg(\bigcap_{\substack{a,b\in\mathbb{Q}^+ \\ 0\le a<b\le N}} \big\{U_N^{a,b} < \infty\big\} \cap \big\{\sup\big\{|X_t| : t \in \mathbb{Q}^+ \cap [0,N]\big\} < \infty\big\}\Bigg).$$

We have $\mathbf{P}[A]=1$; hence $A\in\mathcal{F}_t$ for every $t\ge 0$ since $\mathbb{F}$ satisfies the usual conditions. For $\omega\in A$, for every $t\ge 0$, the limit

$$\widetilde{X}_t(\omega):=\lim_{\mathbb{Q}^+\ni s\downarrow t,s>t} X_s(\omega)$$

exists and is RCLL. For $\omega\in A^c$, we define $\widetilde{X}_t(\omega)=0$. As $\mathbb{F}$ satisfies the usual conditions, $\widetilde{X}$ is $\mathbb{F}$-adapted. As X is a supermartingale, for every N, the family $(X_s)_{s\le N}$ is uniformly integrable. Hence, by assumption,

$$\mathbf{E}[\widetilde{X}_t]=\lim_{\mathbb{Q}^+\ni s\downarrow t,s>t}\mathbf{E}[X_s]=\mathbf{E}[X_t].$$

However, since X is a supermartingale, for every $s>t$, we have

$$X_t\ge\mathbf{E}[X_s\,|\,\mathcal{F}_t]\overset{\mathbb{Q}^+\ni s\downarrow t,s>t}{\longrightarrow}\mathbf{E}[\widetilde{X}_t\,|\,\mathcal{F}_t]=\widetilde{X}_t\quad\text{in } L^1.$$

Therefore, $X_t=\widetilde{X}_t$ almost surely and hence $\widetilde{X}$ is a modification of X. □

Corollary 21.25 *Let $(\nu_t)_{t\ge 0}$ be a continuous convolution semigroup and assume that $\int|x|\nu_1(dx)<\infty$. Then there exists a Markov process X with RCLL paths and with independent stationary increments $\mathbf{P}_{X_t-X_s}=\nu_{t-s}$ for all $t>s$.*

Let E be a locally compact Polish space and let $C_0(E)$ be the set of (bounded) continuous functions that vanish at infinity. If κ is a stochastic kernel from E to E and if f is measurable and bounded, then we define $\kappa f(x)=\int\kappa(x,dy)f(y)$.

Definition 21.26 A Markov semigroup $(\kappa_t)_{t\ge 0}$ on E is called a *Feller semigroup* if

$$f(x)=\lim_{t\to 0}\kappa_t f(x)\quad\text{for all } x\in E,\ f\in C_0(E)$$

and $\kappa_t f\in C_0(E)$ for every $f\in C_0(E)$.

Let X be a Markov process with transition kernels $(\kappa_t)_{t\ge 0}$ and with respect to a filtration $\mathbb{F}$ that satisfies the usual conditions.

Let $g\in C_0(E)$, $g\ge 0$. Let $h=\int_0^\infty e^{-t}\kappa_t g\,dt$. Then

$$e^{-s}\kappa_s h=e^{-s}\int_0^\infty e^{-t}\kappa_s\kappa_t g\,dt=\int_s^\infty e^{-t}\kappa_t g\,dt\le h.$$

Hence $X^g:=(e^{-t}h(X_t))_{t\ge 0}$ is an $\mathbb{F}$-supermartingale.

The Feller property and Theorem 21.24 ensure the existence of an RCLL version $\widetilde{X}^g$ of X^g. It takes a little more work to show that there exists a countable set $G\subset C_0(E)$ and a process $\widetilde{X}$ that is uniquely determined by $\widetilde{X}^g$, $g\in G$, and is an RCLL version of X. See, e.g., [146, Chapter III.7ff].

Let us take a moment's thought and look back at how we derived the strong Markov property of Brownian motion in Section 21.3. Indeed, there we needed only right continuity of the paths and a certain continuity of the distribution as a function of the starting point, which is exactly the Feller property. With a little more work, one can show the following theorem (see, e.g., [146, Chapter III.8ff] or [144, Chapter III, Theorem 2.7]).

Theorem 21.27 *Let $(\kappa_t)_{t\geq 0}$ be a Feller semigroup on the locally compact Polish space E. Then there exists a strong Markov process $(X_t)_{t\geq 0}$ with RCLL paths and transition kernels $(\kappa_t)_{t\geq 0}$.*

Such a process X is called a Feller process.

Exercise 21.4.1 (Doob's inequality) Let $X = (X_t)_{t\geq 0}$ be a martingale or a nonnegative submartingale with RCLL paths. For $T \geq 0$, let $|X|_T^* = \sup_{t\in[0,T]} |X_t|$. Show Doob's inequalities:

(i) For any $p \geq 1$ and $\lambda > 0$, we have $\lambda^p \mathbf{P}[|X|_T^* \geq \lambda] \leq \mathbf{E}[|X_T|^p]$.
(ii) For any $p > 1$, we have $\mathbf{E}[|X_T|^p] \leq \mathbf{E}[(|X|_T^*)^p] \leq (\frac{p}{p-1})^p \mathbf{E}[|X_T|^p]$.

Construct a counterexample that shows that right continuity of the paths of X is essential.

Exercise 21.4.2 (Martingale convergence theorems) Let X be a stochastic process with RCLL paths. Use Doob's inequality (Exercise 21.4.1) to show that the martingale convergence theorems (a.s. convergence (Theorem 11.4), a.s. and L^1-convergence for uniformly integrable martingales (Theorem 11.7) and the L^p-martingale convergence theorem (Theorem 11.10)) hold for X.

Exercise 21.4.3 Let $p \geq 1$ and let $X^1, X^2, X^3, \ldots$ be L^p-integrable martingales. Assume that, for every $t \geq 0$, there exists an $\widetilde{X}_t \in \mathcal{L}^p(\mathbf{P})$ such that $X_t^n \stackrel{n\to\infty}{\longrightarrow} \widetilde{X}_t$ in L^p.

(i) Show that $(\widetilde{X}_t)_{t\geq 0}$ is a martingale.
(ii) Use Doob's inequality to show the following. If $p > 1$ and if $X^1, X^2, \ldots$ are a.s. continuous, then there is a continuous martingale X with the following properties: X is a modification of $\widetilde{X}$ and $X_t^n \stackrel{n\to\infty}{\longrightarrow} X_t$ in L^p for every $t \geq 0$.

Exercise 21.4.4 Let X be a stochastic process with values in a Polish space E and with RCLL paths. Let $\mathbb{F} = \sigma(X)$ be the filtration generated by X and define $\mathbb{F}^+ := (\mathcal{F}_t^+)_{t\geq 0}$ by $\mathcal{F}_t^+ = \bigcap_{s>t} \mathcal{F}_s$. Let $U \subset E$ be open and let $C \subset E$ be closed. For every set $A \subset E$, define $\tau_A := \inf\{t > 0 : X_t \in A\}$. Show the following.

(i) τ_C is an $\mathbb{F}$-stopping time (and an $\mathbb{F}^+$-stopping time).
(ii) τ_U is an $\mathbb{F}^+$-stopping time but in general (even for continuous X) is not an $\mathbb{F}$-stopping time.

Exercise 21.4.5 Show the statement of Remark 21.23. Conclude that if $\mathbb{F}$ is a filtration and if B is a Brownian motion that is an $\mathbb{F}$-martingale, then B is also an $\mathbb{F}^{+,*}$-martingale.

21.5 Construction via L^2-Approximation

We give an alternative construction of Brownian motion by functional analytic means as an L^2-approximation. For simplicity, as the time interval we take $[0, 1]$ instead of $[0, \infty)$.

Let $H = L^2([0, 1])$ be the Hilbert space of square integrable (with respect to Lebesgue measure λ) functions $[0, 1] \to \mathbb{R}$ with inner product

$$\langle f, g\rangle = \int_{[0,1]} f(x)g(x)\lambda(dx)$$

and with norm $\|f\| = \sqrt{\langle f, f\rangle}$ (compare Section 7.3). Two functions $f, g \in H$ are considered equal if $f = g$ λ-a.e. Let $(b_n)_{n\in\mathbb{N}}$ be an orthonormal basis (ONB) of H; that is, $\langle b_m, b_n\rangle = \mathbb{1}_{\{m=n\}}$ and

$$\lim_{n\to\infty} \left\| f - \sum_{m=1}^{n} \langle f, b_m\rangle b_m \right\| = 0 \quad \text{for all } f \in H.$$

In particular, for every $f \in H$, *Parseval's equation*

$$\|f\|^2 = \sum_{m=1}^{\infty} \langle f, b_m\rangle^2 \tag{21.21}$$

holds and for $f, g \in H$

$$\langle f, g\rangle = \sum_{m=1}^{\infty} \langle f, b_m\rangle\langle g, b_m\rangle. \tag{21.22}$$

Now consider an i.i.d. sequence $(\xi_n)_{n\in\mathbb{N}}$ of $\mathcal{N}_{0,1}$-random variables on some probability space $(\Omega, \mathcal{A}, \mathbf{P})$. For $n \in \mathbb{N}$ and $t \in [0, 1]$, define

$$X_t^n = \int \mathbb{1}_{[0,t]}(s)\left(\sum_{m=1}^{n} \xi_m b_m(s)\right)\lambda(ds) = \sum_{m=1}^{n} \xi_m \langle \mathbb{1}_{[0,t]}, b_m\rangle.$$

Clearly, for $n \geq m$,

$$\begin{aligned}\mathbf{E}\big[\big(X_t^m - X_t^n\big)^2\big] &= \mathbf{E}\left[\left(\sum_{k=m+1}^{n} \xi_k \langle \mathbb{1}_{[0,t]}, b_k\rangle\right)\left(\sum_{l=m+1}^{n} \xi_l \langle \mathbb{1}_{[0,t]}, b_l\rangle\right)\right] \\ &= \sum_{k=m+1}^{n} \langle \mathbb{1}_{[0,t]}, b_k\rangle^2 \leq \sum_{k=m+1}^{\infty} \langle \mathbb{1}_{[0,t]}, b_k\rangle^2.\end{aligned}$$

Since $\sum_{k=1}^{\infty} \langle \mathbb{1}_{[0,t]}, b_k\rangle^2 = \|\mathbb{1}_{[0,t]}\|^2 = t < \infty$, we have $X_t^n \in L^2(\mathbf{P})$ and

$$\lim_{m\to\infty} \sup_{n\geq m} \mathbf{E}\big[\big(X_t^m - X_t^n\big)^2\big] = 0.$$

Hence $(X_t^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\mathbf{P})$ and thus (since $L^2(\mathbf{P})$ is complete, see Theorem 7.3) has an L^2-limit X_t. Thus, for $N \in \mathbb{N}$ and $0 \leq t_1, \ldots, t_N \leq 1$,

$$\lim_{n\to\infty} \mathbf{E}\left[\sum_{i=1}^{N} \big(X_{t_i}^n - X_{t_i}\big)^2\right] = 0.$$

In particular, $(X_{t_1}^n, \ldots, X_{t_N}^n) \stackrel{n\to\infty}{\longrightarrow} (X_{t_1}, \ldots, X_{t_N})$ in $\mathbf{P}$-probability.

Manifestly, $(X_{t_1}^n, \ldots, X_{t_N}^n)$ is normally distributed and centered. For $s, t \in [0, 1]$, we have

$$\begin{aligned}\mathbf{Cov}\big[X_s^n, X_t^n\big] &= \mathbf{E}\left[\left(\sum_{k=1}^{n} \xi_k \langle \mathbb{1}_{[0,s]}, b_k\rangle\right)\left(\sum_{l=1}^{n} \xi_l \langle \mathbb{1}_{[0,t]}, b_l\rangle\right)\right] \\ &= \sum_{k,l=1}^{n} \mathbf{E}[\xi_k \xi_l] \langle \mathbb{1}_{[0,s]}, b_k\rangle \langle \mathbb{1}_{[0,t]}, b_l\rangle \\ &= \sum_{k=1}^{n} \langle \mathbb{1}_{[0,s]}, b_k\rangle \langle \mathbb{1}_{[0,t]}, b_k\rangle \\ &\stackrel{n\to\infty}{\longrightarrow} \langle \mathbb{1}_{[0,s]}, \mathbb{1}_{[0,t]}\rangle = \min(s, t).\end{aligned}$$

Hence $(X_t)_{t\in[0,1]}$ is a centered Gaussian process with

$$\mathbf{Cov}[X_s, X_t] = \min(s, t). \tag{21.23}$$

Lévy Construction of Brownian Motion

Up to continuity of paths, X is thus a Brownian motion. A continuous version of X can be obtained via the Kolmogorov–Chentsov theorem (Theorem 21.6). However,

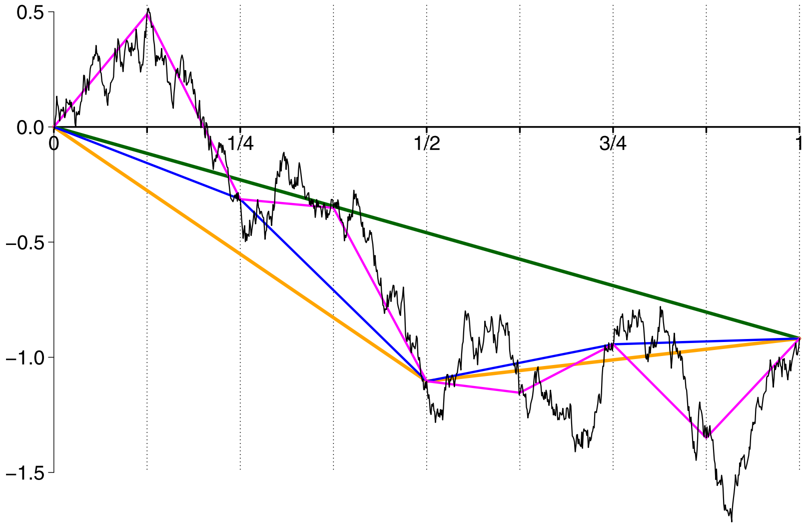

Fig. 21.2 The processes X^n, $n = 0, 1, 2, 3, 10$ of the Lévy construction of Brownian motion

by a clever choice of the ONB $(b_n)_{n\in\mathbb{N}}$, we can construct X directly as a continuous process. The *Haar functions* $b_{n,k}$ are one such choice: Let $b_{0,1} \equiv 1$ and for $n \in \mathbb{N}$ and $k = 1, \ldots, 2^{n-1}$, let

$$b_{n,k}(t) = \begin{cases} 2^{(n-1)/2}, & \text{if } \frac{2k-2}{2^n} \le t < \frac{2k-1}{2^n}, \\ -2^{(n-1)/2}, & \text{if } \frac{2k-1}{2^n} \le t < \frac{2k}{2^n}, \\ 0, & \text{else.} \end{cases}$$

Then $(b_{n,k})$ is an orthonormal system: $\langle b_{m,k}, b_{n,l}\rangle = \mathbb{1}_{\{(m,k)=(n,l)\}}$. It is easy to check that $(b_{n,k})$ is a basis (exercise!). Define the *Schauder functions* by

$$B_{n,k}(t) = \int_{[0,t]} b_{n,k}(s)\lambda(ds) = \langle \mathbb{1}_{[0,t]}, b_{n,k}\rangle.$$

Let $\xi_{0,1}$, $(\xi_{n,k})_{n\in\mathbb{N},k=1,\ldots,2^{n-1}}$ be independent and $\mathcal{N}_{0,1}$-distributed. Let

$$X^n := \xi_{0,1} B_{0,1} + \sum_{m=1}^{n} \sum_{k=1}^{2^{m-1}} \xi_{m,k} B_{m,k},$$

and define $\tilde{X}_t$ as the $L^2(\mathbf{P})$-limit $\tilde{X}_t = L^2 - \lim_{n\to\infty} X_t^n$. See Fig. 21.2 for a computer simulation of X^n, $n = 0, 1, 2, 3, 10$.

Theorem 21.28 (Brownian motion, L^2-approximation) *There is a continuous version X of $\tilde{X}$. X is a Brownian motion and we have*

$$\lim_{n\to\infty} \left\| X^n - X \right\|_\infty = 0 \quad \mathbf{P}\text{-almost surely.} \tag{21.24}$$

Proof By (21.24), we have $X_t = \tilde{X}_t$ a.s. for all $t \in [0, 1]$. As uniform limits of continuous functions are continuous, (21.24) implies that X is continuous. Hence, by (21.23) (and Theorem 21.11), X is a Brownian motion. Therefore, it is enough to prove the existence of an X such that (21.24) holds.

Since $(C([0,1]), \|\cdot\|_\infty)$ is complete, it suffices to show that **P**-almost surely (X^n) is a Cauchy sequence in $(C([0,1]), \|\cdot\|_\infty)$. Note that $\|B_{n,k}\|_\infty \le 2^{-(n+1)/2}$ if $n \in \mathbb{N}$ and $B_{n,k}B_{n,l} = 0$ if $k \neq l$. Hence

$$\left\| X^n - X^{n-1} \right\|_\infty \le 2^{-(n+1)/2} \max\left\{ |\xi_{n,k}|,\ k = 1, \ldots, 2^{n-1} \right\}.$$

Therefore,

$$\begin{aligned}
\mathbf{P}\left[\left\| X^n - X^{n-1} \right\|_\infty > 2^{-n/4} \right] &\le \sum_{k=1}^{2^{n-1}} \mathbf{P}\left[|\xi_{n,k}| > 2^{(n+2)/4} \right] \\
&= 2^{n-1} \frac{2}{\sqrt{2\pi}} \int_{2^{(n+2)/4}}^{\infty} e^{-x^2/2}\, dx \\
&\le 2^n \exp\left(-2^{n/2} \right).
\end{aligned}$$

Evidently, $\sum_{n=1}^{\infty} \mathbf{P}[\|X^n - X^{n-1}\|_\infty > 2^{-n/4}] < \infty$; hence, by the Borel–Cantelli lemma,

$$\mathbf{P}\left[\|X^n - X^{n-1}\|_\infty > 2^{-n/4} \quad \text{only finitely often} \right] = 1.$$

We conclude that $\lim_{n\to\infty} \sup_{m \ge n} \|X^m - X^n\|_\infty = 0$ **P**-almost surely. □

Brownian Motion and White Noise

The construction of Brownian motion via Haar functions has the advantage that continuity of the paths is straightforward. For some applications, however, a decomposition in trigonometric functions is preferable. Here as the orthonormal basis of $L^2([0,1])$ we use $b_0 = 1$ and

$$b_n(x) = \sqrt{2} \cos(n\pi x) \quad \text{for } n \in \mathbb{N}.$$

For $t \in [0, 1]$ and $n \in \mathbb{N}_0$, define

$$B_n(t) = \int_0^t b_n(s) \lambda(ds);$$

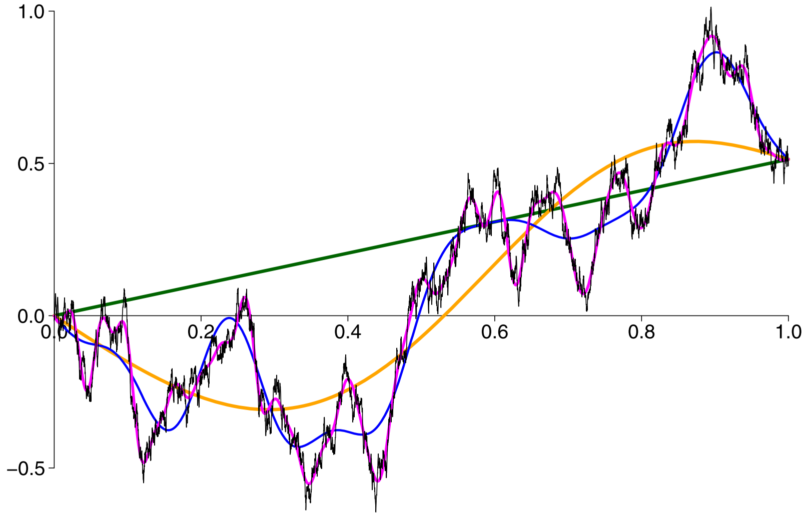

Fig. 21.3 The processes X^n, $n = 0, 1, 4, 64, 8192$ from the Fourier Construction of Brownian motion

that is, $B_0(t) = t$ and

$$B_n(t) = \frac{\sqrt{2}}{n\pi}\sin(n\pi t) \quad \text{for } n \in \mathbb{N}.$$

Let ξ_n, $n \in \mathbb{N}_0$, be independent standard normally distributed random variables. Define $A_0 = \xi_0$ and

$$A_n := \frac{\sqrt{2}}{\pi n}\xi_n \quad \text{for } n \in \mathbb{N}.$$

Finally, let

$$X^n := \sum_{k=0}^{n} \xi_k B_k;$$

that is,

$$X^n(t) = \xi_0 t + \sum_{k=1}^{n} A_k \sin(k\pi t).$$

See Fig. 21.3 for a computer simulation of X^n, $n = 0, 1, 4, 64, 8192$.

As shown above, the sequence (X^n) converges in $L^2([0, 1])$ towards a process X, which (up to continuity of paths) has all properties of Brownian motion:

$$X_t = \xi_0 t + \sum_{n=1}^{\infty} \frac{\sqrt{2}}{n\pi}\xi_n \sin(n\pi t).$$

This representation of the Brownian motions goes back to Paley and Wiener who also show that along a suitable subsequence the series converges uniformly almost surely and hence the limit X is indeed continuous, see [125, Theorem XLIII, p. 148]. The representation is also sometimes called Karhunen–Loève expansion. More precisely, up to the first summand, it is the Karhunen–Loève expansion of the Brownian bridge $(X_t - tX_1)_{t\in[0,1]}$ (see, e.g., [1, Chapter 3.3]).

Taking the formal derivative

$$\dot{X}_t := \frac{d}{dt}X_t = \xi_0 + \sqrt{2}\sum_{n=1}^{\infty} \xi_n \cos(n\pi t)$$

we get independent identically distributed Fourier coefficients for all frequencies. Hence, the formal object $\dot{X}$ is often referred to as *white noise* as opposed to colored noise where the coefficients for the different frequencies have different distributions.

The Fourier basis is not too well suited to showing continuity of paths. For example, the sufficient criterion of absolute summability of coefficients (A_n) fails (see Exercise 21.5.5).

Example 21.29 (Stochastic integral à la Paley–Wiener) Assume that $(\xi_n)_{n\in\mathbb{N}}$ is an i.i.d. sequence of $\mathcal{N}_{0,1}$-distributed random variables. Let $(b_n)_{n\in\mathbb{N}}$ be an orthonormal basis of $L^2([0,1])$ such that $W_t := \lim_{n\to\infty}\sum_{k=1}^{n} \xi_k \langle \mathbb{1}_{[0,t]}, b_k\rangle$, $t \in [0,1]$, is a Brownian motion. For $f \in L^2([0,1])$, define

$$I(f) := \sum_{n=1}^{\infty} \xi_n \langle f, b_n\rangle.$$

By Parseval's equation and the Bienaymé formula, we have

$$\|f\|_2^2 = \sum_{n=1}^{\infty} \langle f, b_n\rangle^2 = \mathbf{Var}\big[I(f)\big] = \mathbf{E}\big[I(f)^2\big].$$

Hence

$$I : L^2\big([0,1]\big) \to L^2(\mathbf{P}), \quad f \mapsto I(f) \quad \text{is an isometry.} \tag{21.25}$$

We call

$$\int_0^t f(s)\, dW_s := I(f\mathbb{1}_{[0,t]}), \quad t \in [0,1],\ f \in L^2\big([0,1]\big),$$

the *stochastic integral* of f with respect to W. In the special case of the Fourier basis $b_0(x) = 1$ and $b_n(x) = \sqrt{2}\cos(n\pi x)$, $n \in \mathbb{N}$, this construction goes back to Paley and Wiener [125, Theorem XLV, p. 154].

The process $X_t := \int_0^t f(s)\,dW_s$, $t \in [0,1]$, is centered Gaussian with covariance function

$$\mathbf{Cov}[X_s, X_t] = \int_0^{s\wedge t} f^2(u)\,du.$$

In fact, it is obvious that X is centered and Gaussian (since it is a limit of the Gaussian processes of partial sums) and has the given covariance function. Furthermore, the existence of a continuous version can be obtained as for Brownian motion by employing the fourth moments of the increments, which for normal random variables can be computed from the variances (compare Theorem 21.9). In the following we will assume for the stochastic integral that such a continuous version is chosen.

In the special case, $f = \sum_{i=1}^n \alpha_i \mathbb{1}_{(t_{i-1},t_i]}$ for some $n \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_n$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, we obtain

$$\int_0^1 f(s)\,dW_s = \sum_{i=1}^n \alpha_i (W_{t_i} - W_{t_{i-1}}).$$

◊

Exercise 21.5.1 Use the representation of Brownian motion $(W_t)_{t\in[0,1]}$ as a random linear combination of the Schauder functions $(B_{n,k})$ to show that the Brownian bridge $Y = (Y_t)_{t\in[0,1]} = (W_t - tW_1)_{t\in[0,1]}$ is a continuous, Gaussian process with covariance function $\mathbf{Cov}[Y_t, Y_s] = (s \wedge t) - st$. Further, show that

$$\mathbf{P}_Y = \lim_{\varepsilon\downarrow 0} \mathbf{P}\big[W \in \cdot \mid W_1 \in (-\varepsilon, \varepsilon)\big].$$

Exercise 21.5.2 (Compare Example 8.32) Fix $T \in (0,1)$. Use an orthonormal basis $b_{0,1}$, $(c_{n,k})$, $(d_{n,k})$ of suitably modified Haar functions (such that the $c_{n,k}$ have support $[0,T]$ and the $d_{n,k}$ have support $[T,1]$) to show that a regular conditional distribution of W_T given W_1 is defined by

$$\mathbf{P}[W_T \in \cdot \mid W_1 = x] = \mathcal{N}_{Tx,T}.$$

Exercise 21.5.3 Define $Y := (Y_t)_{t\in[0,1]}$ by $Y_1 = 0$ and

$$Y_t = (1-t)\int_0^t (1-s)^{-1}\,dW_s \quad \text{for } t \in [0,1).$$

Show that Y is a Brownian bridge.

Hint: Show that Y is a continuous Gaussian process with the correct covariance function. In particular, it has to be shown that $\lim_{t\uparrow 1} Y_t = 0$ almost surely.

Exercise 21.5.4 Let $d \in \mathbb{N}$. Use a suitable orthonormal basis on $[0,1]^d$ to show:

(i) There is a Gaussian process $(W_t)_{t\in[0,1]^d}$ with covariance function

$$\mathbf{Cov}[W_t, W_s] = \prod_{i=1}^d (t_i \wedge s_i).$$

(ii) There is a modification of W such that $t \mapsto W_t$ is almost surely continuous (see Remark 21.7).

A process W with properties (i) and (ii) is called a *Brownian sheet*.

Exercise 21.5.5 Consider the coefficients $(A_n)_{n\in\mathbb{N}_0}$ of the Fourier basis of the construction of Brownian motion. Show the following statements:

(i) $\sum_{n=0}^{\infty} A_n^2 < \infty$ almost surely.
(ii) $\sum_{n=0}^{\infty} |A_n| = \infty$ almost surely.
(iii) $\sum_{k=0}^{n} A_k$, $n \in \mathbb{N}$ converges almost surely.

Hint: Kolmogorov's three-series theorem (Theorem 15.50).

Exercise 21.5.6 Let $t \in (0, 1)$ and $f_0(x) := t$ as well as

$$f_n(x) := \frac{2\sin(n\pi t)}{n\pi} \cos(n\pi x) \quad \text{for } n \in \mathbb{N},\ x \in [0, 1].$$

Show that $\sum_{n=0}^{\infty} f_n(x) = \mathbb{1}_{[0,t]}(x)$ for $x \in (0, 1) \setminus \{t\}$.

21.6 The Space $C([0, \infty))$

Are functionals that depend on the whole path of a Brownian motion measurable? For example, is $\sup\{X_t, t \in [0, 1]\}$ measurable? For general stochastic processes, this is false since the supremum depends on more than countably many coordinates. However, for processes with continuous paths, this is true, as we will show in this section in a somewhat more general framework.

We may consider Brownian motion as the canonical process on the space $\Omega := C([0, \infty))$ of continuous paths.

We start by collecting some properties of the space $\Omega = C([0, \infty)) \subset \mathbb{R}^{[0,\infty)}$. Define the *evaluation map*

$$X_t : \Omega \to \mathbb{R}, \quad \omega \mapsto \omega(t), \tag{21.26}$$

that is, the restriction of the canonical projection $\mathbb{R}^{[0,\infty)} \to \mathbb{R}$ to Ω.

For $f, g \in C([0, \infty))$ and $n \in \mathbb{N}$, let $d_n(f, g) := \|(f - g)\big|_{[0,n]}\|_\infty \wedge 1$ and

$$d(f, g) = \sum_{n=1}^{\infty} 2^{-n} d_n(f, g). \tag{21.27}$$

Theorem 21.30 *d is a complete metric on $\Omega := C([0, \infty))$ that induces the topology of uniform convergence on compact sets. The space (Ω, d) is separable and hence Polish.*

Proof Clearly, every d_n is a complete metric on $(C([0,n]), \|\cdot\|_\infty)$. Thus, for every Cauchy sequence (f_N) in (Ω, d) and every $n \in \mathbb{N}$, there exists a $g_n \in \Omega$ with $d_n(f_N, g_n) \overset{N\to\infty}{\longrightarrow} 0$. Evidently, $g_n(x) = g_m(x)$ for every $x \le m \wedge n$; hence there exists a $g \in \Omega$ with $g(x) = g_n(x)$ for every $x \le n$ for every $n \in \mathbb{N}$. Hence, clearly, $d(f_N, g) \overset{N\to\infty}{\longrightarrow} 0$, and thus d is complete.

The set of polynomials with rational coefficients is countable and by the Weierstraß theorem, it is dense in any $(C([0,n]), \|\cdot\|_\infty)$; hence it is dense in (Ω, d). $\square$

Theorem 21.31 *With respect to the Borel σ-algebra $\mathcal{B}(\Omega, d)$, the canonical projections X_t, $t \in [0,\infty)$ are measurable. On the other hand, the X_t generate $\mathcal{B}(\Omega, d)$. Hence*

$$\big(\mathcal{B}(\mathbb{R})\big)^{\otimes[0,\infty)}\Big|_\Omega = \sigma\big(X_t, t \in [0,\infty)\big) = \mathcal{B}(\Omega, d).$$

Proof The first equation holds by definition. For the second one, we must show the mutual inclusions.

"$\subset$" Clearly, every $X_t : \Omega \longrightarrow \mathbb{R}$ is continuous and hence $(\mathcal{B}(\Omega, d) - \mathcal{B}(\mathbb{R}))$-measurable. Thus $\sigma(X_t, t \in [0,\infty)) \subset \mathcal{B}(\Omega, d)$.

"$\supset$" We have to show that open subsets of (Ω, d) are in $\mathcal{A} := (\mathcal{B}(\mathbb{R}))^{\otimes[0,\infty)}$. Since (Ω, d) is separable (Theorem 21.30), every open set is a countable union of ε-balls. Hence it suffices to show that for fixed $\omega_0 \in \Omega$, the map $\omega \mapsto d(\omega_0, \omega)$ is $\mathcal{A}$-measurable. To this end it is enough to show that for any $n \in \mathbb{N}$, the map $\omega \mapsto Y_n(\omega) := d_n(\omega_0, \omega)$ (see (21.27)) is $\mathcal{A}$-measurable. However, the map

$$\omega \mapsto Z_t(\omega) := \big|X_t(\omega) - X_t(\omega_0)\big| \wedge 1$$

is $\mathcal{A}$-measurable. Since each ω is continuous, Y_n is a countable supremum

$$Y_n = \sup_{t \in [0,n]\cap\mathbb{Q}} Z_t$$

and is hence $\mathcal{A}$-measurable. $\square$

In the following, let $\mathcal{A} := \sigma(X_t, t \in [0,\infty))$.

Corollary 21.32 *The map $F_1 : \Omega \to [0,\infty)$, $\omega \mapsto \sup\{\omega(t) : t \in [0,1]\}$ is $\mathcal{A}$-measurable.*

Proof F_1 is continuous with respect to d and hence $\mathcal{B}(\Omega, d)$-measurable. $\square$

If B is a Brownian motion (on some probability space $(\widetilde{\Omega}, \widetilde{\mathcal{A}}, \widetilde{\mathbf{P}})$), then there exists an $\overline{\Omega} \in \widetilde{\mathcal{A}}$ with $\widetilde{\mathbf{P}}[\overline{\Omega}] = 1$ and $B(\omega) \in C([0,\infty))$ for every $\omega \in \overline{\Omega}$. Let $\overline{\mathcal{A}} = \widetilde{\mathcal{A}}\big|_{\overline{\Omega}}$ and $\overline{\mathbf{P}} = \widetilde{\mathbf{P}}\big|_{\overline{\mathcal{A}}}$. Then $B : \overline{\Omega} \longrightarrow C([0,\infty))$ is measurable with respect to $(\overline{\mathcal{A}}, \mathcal{A})$. With respect to the image measure $\mathbf{P} = \overline{\mathbf{P}} \circ B^{-1}$ on $\Omega = C([0,\infty))$, the canonical process $X = (X_t, t \in [0,\infty))$ on $C([0,\infty))$ is a Brownian motion.

Definition 21.33 Let $\mathbf{P}$ be the probability measure on $\Omega = C([0,\infty))$ with respect to which the canonical process X is a Brownian motion. Then $\mathbf{P}$ is called the *Wiener measure*. The triple $(\Omega, \mathcal{A}, \mathbf{P})$ is called the *Wiener space*, and X is called the *canonical Brownian motion* or the *Wiener process*.

Remark 21.34 Sometimes we want a Brownian motion to start not at $X_0 = 0$ but at an arbitrary point x. Denote by $\mathbf{P}_x$ that measure on $C([0,\infty))$ for which $\widetilde{X} = (X_t - x, t \in [0,\infty))$ is a Brownian motion (with $\widetilde{X}_0 = 0$). ◊

Exercise 21.6.1 Show that the map

$$F_\infty : \Omega \to [0,\infty], \quad \omega \mapsto \sup\big\{\omega(t) : t \in [0,\infty)\big\},$$

is $\mathcal{A}$-measurable.

21.7 Convergence of Probability Measures on $C([0,\infty))$

Let X and $(X^n)_{n\in\mathbb{N}}$ be random variables with values in $C([0,\infty))$ (i.e., continuous stochastic processes) with distributions $\mathbf{P}_X$ and $(\mathbf{P}_{X^n})_{n\in\mathbb{N}}$.

Definition 21.35 We say that the finite-dimensional distributions of (X^n) converge to those of X if, for every $k \in \mathbb{N}$ and $t_1, \ldots, t_k \in [0,\infty)$, we have

$$\big(X^n_{t_1}, \ldots, X^n_{t_k}\big) \overset{n\to\infty}{\Longrightarrow} (X_{t_1}, \ldots, X_{t_k}).$$

In this case, we write

$$X^n \underset{\text{fdd}}{\overset{n\to\infty}{\Longrightarrow}} X \quad \text{or} \quad \mathbf{P}_{X^n} \underset{\text{fdd}}{\overset{n\to\infty}{\longrightarrow}} \mathbf{P}_X.$$

Lemma 21.36 *$P_n \underset{\text{fdd}}{\overset{n\to\infty}{\longrightarrow}} P$ and $P_n \underset{\text{fdd}}{\overset{n\to\infty}{\longrightarrow}} Q$ imply $P = Q$.*

Proof By Theorem 14.12(iii), the finite-dimensional distributions determine P uniquely. □

Theorem 21.37 *Weak convergence in $\mathcal{M}_1(\Omega, d)$ implies fdd-convergence*:

$$P_n \overset{n\to\infty}{\longrightarrow} P \quad \Longrightarrow \quad P_n \underset{\text{fdd}}{\overset{n\to\infty}{\longrightarrow}} P.$$

Proof Let $k \in \mathbb{N}$ and $t_1, \ldots, t_k \in [0,\infty)$. The map

$$\varphi : C\big([0,\infty)\big) \to \mathbb{R}^k, \quad \omega \mapsto \big(\omega(t_1), \ldots, \omega(t_k)\big)$$

is continuous. By the continuous mapping theorem (Theorem 13.25 on p. 257), we have $P_n \circ \varphi^{-1} \overset{n\to\infty}{\longrightarrow} P \circ \varphi^{-1}$; hence $P_n \underset{\text{fdd}}{\overset{n\to\infty}{\longrightarrow}} P$. □

The converse statement in the preceding theorem does not hold. However, we still have the following.

Theorem 21.38 *Let $(P_n)_{n\in\mathbb{N}}$ and P be probability measures on $C([0,\infty))$. Then the following are equivalent*:

(i) $P_n \underset{\text{fdd}}{\overset{n\to\infty}{\longrightarrow}} P$ *and $(P_n)_{n\in\mathbb{N}}$ is tight.*
(ii) $P_n \overset{n\to\infty}{\longrightarrow} P$ *weakly.*

Proof "(ii) $\Longrightarrow$ (i)" This is a direct consequence of Prohorov's theorem (Theorem 13.29 with $E = C([0,\infty))$).

"(i) $\Longrightarrow$ (ii)" By Prohorov's theorem, $(P_n)_{n\in\mathbb{N}}$ is relatively sequentially compact. Let Q be a limit point for $(P_{n_k})_{k\in\mathbb{N}}$ along some subsequence (n_k). Then $P_{n_k} \overset{\text{fdd}}{\longrightarrow} Q$, $k\to\infty$. By Lemma 21.36, we have $P = Q$. □

Next we derive a useful criterion for tightness of sets $\{P_n\} \subset \mathcal{M}_1(C([0,\infty)))$. We start by recalling the Arzelà–Ascoli characterization of relatively compact sets in $C([0,\infty))$ (see, e.g., [37, Theorem 2.4.7] or [173, Theorem III.3]).

For $N, \delta > 0$ and $\omega \in C([0,\infty))$, let

$$V^N(\omega,\delta) := \sup\bigl\{\bigl|\omega(t)-\omega(s)\bigr| : |t-s| \le \delta,\ s,t \le N\bigr\}.$$

Theorem 21.39 (Arzelà–Ascoli) *A set $A \subset C([0,\infty))$ is relatively compact if and only if the following two conditions hold.*

(i) $\{\omega(0), \omega \in A\} \subset \mathbb{R}$ *is bounded.*
(ii) *For every N, we have* $\lim_{\delta\downarrow 0} \sup_{\omega\in A} V^N(\omega,\delta) = 0$.

Theorem 21.40 *A family $(P_i, i \in I)$ of probability measures on $C([0,\infty))$ is weakly relatively compact if and only if the following two conditions hold.*

(i) $(P_i \circ X_0^{-1}, i \in I)$ *is tight*; *that is, for every $\varepsilon > 0$, there is a $K > 0$ such that*

$$P_i\bigl(\bigl\{\omega : \bigl|\omega(0)\bigr| > K\bigr\}\bigr) \le \varepsilon \quad \textit{for all } i \in I. \tag{21.28}$$

(ii) *For all $\eta, \varepsilon > 0$ and $N \in \mathbb{N}$, there is a $\delta > 0$ such that*

$$P_i\bigl(\bigl\{\omega : V^N(\omega,\delta) > \eta\bigr\}\bigr) \le \varepsilon \quad \textit{for all } i \in I. \tag{21.29}$$

Proof "$\Longrightarrow$" By Prohorov's theorem (Theorem 13.29), weak relative compactness of $(P_i, i \in I)$ implies tightness of this family. Thus, for every $\varepsilon > 0$, there exists a compact set $A \subset C([0,\infty))$ with $P_i(A) > 1-\varepsilon$ for every $i \in I$. Using the Arzelà–Ascoli characterization of the compactness of A, we infer (i) and (ii).

"$\Longleftarrow$" Now assume that (i) and (ii) hold. Then, for $\varepsilon > 0$ and $k, N \in \mathbb{N}$, choose numbers K_ε and $\delta_{N,k,\varepsilon}$ such that

$$\sup_{i\in I} P_i\big(\{\omega : |\omega(0)| > K_\varepsilon\}\big) \le \frac{\varepsilon}{2}$$

and

$$\sup_{i\in I} P_i\left(\left\{\omega : V^N(\omega, \delta_{N,k,\varepsilon}) > \frac{1}{k}\right\}\right) \le 2^{-N-k-1}\varepsilon.$$

Define

$$C_{N,\varepsilon} = \left\{\omega : |\omega(0)| \le K_\varepsilon,\ V^N(\omega, \delta_{N,k,\varepsilon}) \le \frac{1}{k} \text{ for all } k \in \mathbb{N}\right\}.$$

By the Arzelà–Ascoli theorem, $C_\varepsilon := \bigcap_{N\in\mathbb{N}} C_{N,\varepsilon}$ is relatively compact in $C([0,\infty))$ and we have

$$P_i\big(C_\varepsilon^c\big) \le \frac{\varepsilon}{2} + \sum_{k,N=1}^{\infty} P_i\big(\{\omega : V^N(\omega, \delta_{N,k,\varepsilon}) > 1/k\}\big) \le \varepsilon \quad \text{for all } i \in I.$$

Hence the claim follows. □

Corollary 21.41 *Let $(X_i, i \in I)$ and $(Y_i, i \in I)$ be families of random variables in $C([0,\infty))$. Assume that $(\mathbf{P}_{X_i}, i \in I)$ and $(\mathbf{P}_{Y_i}, i \in I)$ are tight. Then $(\mathbf{P}_{X_i+Y_i}, i \in I)$ is tight.*

Proof Apply the triangle inequality in order to check (i) and (ii) in the preceding theorem. □

The following is an important tool to check weak relative compactness.

Theorem 21.42 (Kolmogorov's criterion for weak relative compactness) *Let $(X^i, i \in I)$ be a sequence of continuous stochastic processes. Assume that the following conditions are satisfied.*

(i) *The family $(\mathbf{P}[X_0^i \in \cdot], i \in I)$ of initial distributions is tight.*
(ii) *For any $N > 0$ there are numbers $C, \alpha, \beta > 0$ such that, for all $s, t \in [0, N]$ and every $i \in I$, we have*

$$\mathbf{E}\big[|X_s^i - X_t^i|^\alpha\big] \le C|s-t|^{\beta+1}.$$

Then the family $(\mathbf{P}_{X^i}, i \in I) = (\mathcal{L}[X^i], i \in I)$ of distributions of X^i is weakly relatively compact in $\mathcal{M}_1(C([0,\infty)))$.

Proof We check the conditions of Theorem 21.40. The first condition of Theorem 21.40 is exactly (i).

Let $N > 0$. By the Kolmogorov–Chentsov theorem (Theorem 21.6(ii)), for $\varepsilon > 0$ and $\gamma \in (0, \beta/\alpha)$, there exists a K such that, for every $i \in I$, we have

$$\mathbf{P}\big[\big|X_t^i - X_s^i\big| \le K|t-s|^\gamma \text{ for all } s,t \in [0,N]\big] \ge 1-\varepsilon.$$

This clearly implies (21.29) with $\delta = (\eta/K)^{1/\gamma}$. □

21.8 Donsker's Theorem

Let $Y_1, Y_2, \ldots$ be i.i.d. random variables with $\mathbf{E}[Y_1] = 0$ and $\mathbf{Var}[Y_1] = \sigma^2 > 0$. For $t > 0$, let $S_t^n = \sum_{i=1}^{\lfloor nt \rfloor} Y_i$ and $\widetilde{S}_t^n = \frac{1}{\sqrt{\sigma^2 n}} S_t^n$. By the central limit theorem, for $t > s \ge 0$, we have $\mathcal{L}[\widetilde{S}_t^n - \widetilde{S}_s^n] \overset{n\to\infty}{\longrightarrow} \mathcal{N}_{0,t-s}$.

Let $B = (B_t, t \ge 0)$ be a Brownian motion. Then

$$\mathcal{L}\big[\widetilde{S}_t^n - \widetilde{S}_s^n\big] \overset{n\to\infty}{\longrightarrow} \mathcal{L}[B_t - B_s] \quad \text{for any } t > s \ge 0.$$

For $N \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_N$, the random variables $\widetilde{S}_{t_i}^n - \widetilde{S}_{t_{i-1}}^n$, $i = 1, \ldots, N$, are independent, and hence, we have

$$\mathcal{L}\big[\big(\widetilde{S}_{t_1}^n - \widetilde{S}_{t_0}^n, \ldots, \widetilde{S}_{t_N}^n - \widetilde{S}_{t_{N-1}}^n\big)\big] \overset{n\to\infty}{\longrightarrow} \mathcal{L}\big[(B_{t_1} - B_{t_0}, \ldots, B_{t_N} - B_{t_{N-1}})\big].$$

We infer that

$$\mathcal{L}\big[\big(\widetilde{S}_{t_1}^n, \ldots, \widetilde{S}_{t_N}^n\big)\big] \overset{n\to\infty}{\longrightarrow} \mathcal{L}\big[(B_{t_1}, \ldots, B_{t_N})\big]. \tag{21.30}$$

We now define $\bar{S}^n$ as $\widetilde{S}^n$ but linearly interpolated:

$$\bar{S}_t^n = \frac{1}{\sqrt{\sigma^2 n}} \sum_{i=1}^{\lfloor nt \rfloor} Y_i + \frac{tn - \lfloor tn \rfloor}{\sqrt{\sigma^2 n}} Y_{\lfloor nt \rfloor + 1}. \tag{21.31}$$

Then, for $\varepsilon > 0$,

$$\begin{aligned}
\mathbf{P}\big[\big|\widetilde{S}_t^n - \bar{S}_t^n\big| > \varepsilon\big] &\le \varepsilon^{-2}\mathbf{E}\big[\big(\widetilde{S}_t^n - \bar{S}_t^n\big)^2\big] \\
&\le \frac{1}{\varepsilon^2 n}\,\frac{1}{\sigma^2}\mathbf{E}\big[Y_1^2\big] = \frac{1}{\varepsilon^2 n} \overset{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

By Slutzky's theorem (Theorem 13.18), we thus have convergence of the finite-dimensional distributions to the Wiener measure $\mathbf{P}_W$:

$$\mathbf{P}_{\bar{S}^n} \;\underset{\text{fdd}}{\overset{n\to\infty}{\Longrightarrow}}\; \mathbf{P}_W. \tag{21.32}$$

The aim of this section is to strengthen this convergence statement to weak convergence of probability measures on $C([0,\infty))$. The main theorem of this section is the *functional central limit theorem*, which goes back to Donsker [35]. Theorems of this type are also called *invariance principles* since the limiting distribution is the same for all distributions Y_i with expectation 0 and the same variance.

Theorem 21.43 (Donsker's invariance principle) *In the sense of weak convergence on $C([0,\infty))$, the distributions of $\bar{S}^n$ converge to the Wiener measure,*

$$\mathcal{L}\big[\bar{S}^n\big] \overset{n\to\infty}{\longrightarrow} \mathbf{P}_W. \tag{21.33}$$

Proof Owing to (21.32) and Theorem 21.38, it is enough to show that $(\mathcal{L}[\bar{S}^n], n \in \mathbb{N})$ is tight. To this end, we want to apply Kolmogorov's moment criterion. However, as in the proof of existence of Brownian motion, second moments are not enough; rather we need fourth moments in order that we can choose $\beta > 0$. Hence the strategy is to truncate the Y_i to obtain fourth moments.

For $K > 0$, define

$$\begin{aligned} Y_i^K &:= Y_i \mathbb{1}_{\{|Y_i|\le K/2\}} - \mathbf{E}[Y_i \mathbb{1}_{\{|Y_i|\le K/2\}}] \quad \text{and} \\ Z_i^K &:= Y_i - Y_i^K \quad \text{for } i \in \mathbb{N}. \end{aligned}$$

Then $\mathbf{E}[Y_i^K] = \mathbf{E}[Z_i^K] = 0$ and $\mathbf{Var}[Z_i^K] \overset{K\to\infty}{\longrightarrow} 0$ as well as $\mathbf{Var}[Y_i^K] \le \sigma^2$, $i \in \mathbb{N}$. Clearly, $|Y_i^K| \le K$ for every i. Define

$$T_n^K := \sum_{i=1}^n Y_i^K \quad \text{and} \quad U_n^K := \sum_{i=1}^n Z_i^K \quad \text{for } n \in \mathbb{N}.$$

Let $\bar{T}_t^{K,n}$ and $\bar{U}_t^{K,n}$ be the linearly interpolated versions of

$$\widetilde{T}_t^{K,n} := \frac{1}{\sqrt{\sigma^2 n}} T_{\lfloor nt\rfloor}^K \quad \text{and} \quad \widetilde{U}_t^{K,n} := \frac{1}{\sqrt{\sigma^2 n}} U_{\lfloor nt\rfloor}^K \quad \text{for } t \ge 0.$$

Evidently, $\bar{S}^n = \bar{T}^{K,n} + \bar{U}^{K,n}$. By Corollary 21.41, it is enough to show that, for a sequence $(K_n)_{n\in\mathbb{N}}$ (chosen later), the families $(\mathcal{L}[\bar{U}^{K_n,n}], n \in \mathbb{N})$ and $(\mathcal{L}[\bar{T}^{K_n,n}], n \in \mathbb{N})$ are tight.

We consider first the remainder term. As U^K is a martingale, Doob's inequality (Theorem 11.2) yields

$$\mathbf{P}\Big[\sup_{l=1,\ldots,n} \big|U_l^K\big| > \varepsilon\sqrt{n}\Big] \le \varepsilon^{-2}\mathbf{Var}\big[Z_1^K\big] \quad \text{for every } \varepsilon > 0.$$

Now, if $K_n \uparrow \infty$, $n \to \infty$, then for every $N > 0$, we have

$$\mathbf{P}\Big[\sup_{t\in[0,N]} \big|\bar{U}_t^{K_n,n}\big| > \varepsilon\Big] \le \frac{N}{\varepsilon^2\sigma^2}\mathbf{Var}\big[Z_1^{K_n}\big] \overset{n\to\infty}{\longrightarrow} 0,$$

hence $\bar{U}^{K_n,n} \overset{n\to\infty}{\Longrightarrow} 0$ in $C([0,\infty))$. In particular, $(\mathcal{L}[\bar{U}^{K_n,n}], n \in \mathbb{N})$ is tight.

Next, for $N > 0$ and $s, t \in [0, N]$, we compute the fourth moments of the differences $\bar{T}^{K_n,n}_{t+s} - \bar{T}^{K_n,n}_s$ for the main term. In the following, let $K_n = n^{1/4}$. Fix $n \in \mathbb{N}$. We distinguish two cases:

Case 1: $t < n^{-1}$. Let $k := \lfloor (t+s)n \rfloor$. If $sn \geq k$, then

$$\bar{T}^{K_n,n}_{t+s} - \bar{T}^{K_n,n}_s = \frac{tn}{\sqrt{n\sigma^2}} Y^{K_n}_{k+1}.$$

If $sn < k$, then

$$\bar{T}^{K_n,n}_{t+s} - \bar{T}^{K_n,n}_s = \frac{1}{\sqrt{n\sigma^2}}\Big(\big((t+s)n - k\big)Y^{K_n}_{k+1} + (k - sn)Y^{K_n}_k\Big).$$

In either case, we have

$$\big|\bar{T}^{K_n,n}_{t+s} - \bar{T}^{K_n,n}_s\big| \leq \frac{t\sqrt{n}}{\sigma}\big(\big|Y^{K_n}_k\big| + \big|Y^{K_n}_{k+1}\big|\big),$$

hence

$$\begin{aligned}\mathbf{E}\Big[\big(\bar{T}^{K_n,n}_{t+s} - \bar{T}^{K_n,n}_s\big)^4\Big] &\leq \frac{n^2t^4}{\sigma^4}(2K_n)^2\mathbf{E}\Big[\big(\big|Y^{K_n}_1\big| + \big|Y^{K_n}_2\big|\big)^2\Big] \\ &\leq \frac{16n^{5/2}t^4}{\sigma^4}\mathbf{Var}\big[Y^{K_n}_1\big] \leq \frac{16}{\sigma^2}t^{3/2}. \end{aligned} \tag{21.34}$$

Case 2: $t \geq n^{-1}$. Using the binomial theorem, we get (note that the mixed terms with odd moments vanish since $\mathbf{E}[Y^{K_n}_1] = 0$)

$$\begin{aligned}\mathbf{E}\Big[\big(T^{K_n}_n\big)^4\Big] &= n\mathbf{E}\Big[\big(Y^{K_n}_1\big)^4\Big] + 3n(n-1)\mathbf{E}\Big[\big(Y^{K_n}_1\big)^2\Big]^2 \\ &\leq nK_n^2\sigma^2 + 3n(n-1)\sigma^4. \end{aligned} \tag{21.35}$$

Note that, for independent real random variables X, Y with $\mathbf{E}[X] = \mathbf{E}[Y] = 0$ and $\mathbf{E}[X^4], \mathbf{E}[Y^4] < \infty$ and for $a \in [-1, 1]$, we have

$$\begin{aligned}\mathbf{E}\big[(aX+Y)^4\big] &= a^4\mathbf{E}\big[X^4\big] + 6a^2\mathbf{E}\big[X^2\big]\mathbf{E}\big[Y^2\big] + \mathbf{E}\big[Y^4\big] \\ &\leq \mathbf{E}\big[X^4\big] + 6\mathbf{E}\big[X^2\big]\mathbf{E}\big[Y^2\big] + \mathbf{E}\big[Y^4\big] = \mathbf{E}\big[(X+Y)^4\big]. \end{aligned}$$

We apply this twice (with $a = \lceil (t+s)n \rceil - (t+s)n$ and $a = sn - \lfloor sn \rfloor$) and obtain (using the estimate $\lceil (t+s)n \rceil - \lfloor sn \rfloor \leq tn + 2 \leq 3tn$) from (21.35) (since $t \leq N$)

$$
\begin{aligned}
\mathbf{E}\big[\big(\bar{T}^{K_n,n}_{t+s}-\bar{T}^{K_n,n}_{s}\big)^4\big] &\le n^{-2}\sigma^{-4}\mathbf{E}\big[\big(T^{K_n}_{\lceil (t+s)n\rceil}-T^{K_n}_{\lfloor sn\rfloor}\big)^4\big] \\
&= n^{-2}\sigma^{-4}\mathbf{E}\big[\big(T^{K_n}_{\lceil (t+s)n\rceil-\lfloor sn\rfloor}\big)^4\big] \\
&\le \frac{3tnK_n^2}{n^2\sigma^2}+18t^2=\frac{3}{\sigma^2}tn^{-1/2}+18t^2 \\
&\le \frac{3}{\sigma^2}t^{3/2}+18t^2\le\left(\frac{3}{\sigma^2}+18\sqrt{N}\right)t^{3/2}. \qquad (21.36)
\end{aligned}
$$

By (21.34) and (21.36), for every $N>0$, there exists a $C=C(N,\sigma^2)$ such that, for every $n\in\mathbb{N}$ and all $s,t\in[0,N]$, we have

$$\mathbf{E}\big[\big(\bar{T}^{K_n,n}_{t+s}-\bar{T}^{K_n,n}_{s}\big)^4\big]\le Ct^{3/2}.$$

Hence, by Kolmogorov's moment criterion (Theorem 21.42 with $\alpha=4$ and $\beta=1/2$), $(\mathcal{L}[\bar{T}^{K_n,n}], n\in\mathbb{N})$ is tight in $\mathcal{M}_1(C([0,\infty)))$. □

Exercise 21.8.1 Let $X_1, X_2, \ldots$ be i.i.d. random variables with continuous distribution function F. Let $G_n:[0,1]\to\mathbb{R}$, $t\mapsto n^{-1/2}\sum_{i=1}^n(\mathbb{1}_{[0,t]}(F(X_i))-t)$ and $M_n:=\|G_n\|_\infty$. Further, let $M=\sup_{t\in[0,1]}|B_t|$, where B is a Brownian bridge.

(i) Show that $\mathbf{E}[G_n(t)]=0$ and $\mathbf{Cov}[G_n(s),G_n(t)]=s\wedge t-st$ for $s,t\in[0,1]$.
(ii) Show that $\mathbf{E}[(G_n(t)-G_n(s))^4]\le C((t-s)^2+|t-s|/n)$ for some $C>0$.
(iii) Conclude that a suitable continuous version of G_n converges weakly to B. For example, choose

$$H_n(t)=n^{-1/2}\sum_{i=1}^n\big(h_n\big(F(X_i)-t\big)-g_n(t)\big),$$

where h_n is a suitable smoothed version of $\mathbb{1}_{(-\infty,0]}$, for example, $h_n(s)=1-(s/\varepsilon_n\vee 0)\wedge 1$ for some sequence $\varepsilon_n\downarrow 0$, and $g_n(t):=\int_0^1 h_n(t-u)\,du$.
(iv) Finally, show that $M_n\overset{n\to\infty}{\Longrightarrow}M$.

Remark: The distribution of M can be expressed by the Kolmogorov–Smirnov formula ([101] and [156]; see, e.g., [133])

$$\mathbf{P}[M>x]=2\sum_{n=1}^\infty(-1)^{n-1}e^{-2n^2x^2}. \qquad (21.37)$$

Compare (21.20). Using the statistic M_n, one can test if random variables of a known distribution are independent. Let $X_1, X_2, \ldots$ and $\tilde{X}_1, \tilde{X}_2, \ldots$ be independent random variables with unknown continuous distribution functions F and $\tilde{F}$ and with empirical distribution functions F_n and $\tilde{F}_n$. Further, let

$$D_n:=\sup_{t\in\mathbb{R}}\big|F_n(t)-\tilde{F}_n(t)\big|.$$

Under the assumption that $F = \tilde{F}$ holds, $\sqrt{n/2}D_n$ converges in distribution to M. This fact is the basis for nonparametric tests on the equality of distributions.

21.9 Pathwise Convergence of Branching Processes*

In this section, we investigate the convergence of rescaled Galton–Watson processes (branching processes). As for sums of independent random variables, we first show convergence for a fixed time point to the distribution of a certain limiting process. The next step is to show convergence of finite-dimensional distributions. Finally, using Kolmogorov's moment criterion for tightness, we show convergence in the path space $C([0,\infty))$.

Consider a Galton–Watson process $(Z_n)_{n\in\mathbb{N}_0}$ with geometric offspring distribution

$$p(k) = 2^{-k-1} \quad \text{for } k \in \mathbb{N}_0.$$

That is, let $X_{n,i}$, $n, i \in \mathbb{N}_0$ be i.i.d. random variables on $\mathbb{N}_0$ with $\mathbf{P}[X_{n,i} = k] = p(k)$, $k \in \mathbb{N}_0$, and based on the initial state Z_0 define inductively

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i}.$$

Thus Z is a Markov chain with transition probabilities $p(i, j) = p^{*i}(j)$, where p^{*i} is the ith convolution power of p. In other words, if $Z, Z^1, \ldots, Z^i$ are independent copies of our Galton–Watson process, with $Z_0 = i$ and $Z_0^1 = \ldots = Z_0^i = 1$, then

$$Z \stackrel{\mathcal{D}}{=} Z^1 + \ldots + Z^i. \tag{21.38}$$

We consider now the probability generating function of $X_{1,1}$, $\psi^{(1)}(s) := \psi(s) := \mathbf{E}[s^{X_{1,1}}]$, $s \in [0, 1]$. Denote by $\psi^{(n)} := \psi^{(n-1)} \circ \psi$ its nth iterate for $n \in \mathbb{N}$. Then, by Lemma 3.10,

$$\mathbf{E}_i\left[s^{Z_n}\right] = \mathbf{E}_1\left[s^{Z_n}\right]^i = \left(\psi^{(n)}(s)\right)^i.$$

For the geometric offspring distribution, $\psi^{(n)}$ can be computed explicitly.

Lemma 21.44 *For the branching process with critical geometric offspring distribution, the nth iterate of the probability generating function is*

$$\psi^{(n)}(s) = \frac{n - (n-1)s}{n + 1 - ns}.$$

Proof Compute

$$\psi(s) = \sum_{k=0}^{\infty} 2^{-k-1} s^k = \frac{1}{-s+2}.$$

In order to compute the iterated function, first consider general linear rational functions of the form $f(x)=\frac{ax+b}{cx+d}$. For such f, define the matrix $M_f=\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)$. For two linear rational functions f and g, we have $M_{f\circ g}=M_f\cdot M_g$. The powers of M are easy to compute:

$$M_\psi=\begin{pmatrix} 0 & 1 \\ -1 & 2\end{pmatrix},\qquad M_\psi^2=\begin{pmatrix} -1 & 2 \\ -2 & 3\end{pmatrix},\qquad M_\psi^3=\begin{pmatrix} -2 & 3 \\ -3 & 4\end{pmatrix},$$

and inductively

$$M_\psi^n=\begin{pmatrix} -(n-1) & n \\ -n & n+1\end{pmatrix}. \qquad \square$$

If we let $s=e^{-\lambda}$, then we get the Laplace transform of Z_n,

$$\mathbf{E}_i\left[e^{-\lambda Z_n}\right]=\psi^{(n)}\left(e^{-\lambda}\right)^i.$$

By Example 6.29, we can compute the moments of Z_n by differentiating the Laplace transform. That is, we obtain the following lemma.

Lemma 21.45 *The moments of Z_n are*

$$\mathbf{E}_i\left[Z_n^k\right]=(-1)^k\frac{d^k}{d\lambda^k}\left(\psi^{(n)}\left(e^{-\lambda}\right)^i\right)_{|\lambda=0}. \tag{21.39}$$

In particular, the first six moments are

$$\begin{aligned}
\mathbf{E}_i[Z_n]&=i,\\
\mathbf{E}_i\left[Z_n^2\right]&=2in+i^2,\\
\mathbf{E}_i\left[Z_n^3\right]&=6in^2+6i^2n+i^3,\\
\mathbf{E}_i\left[Z_n^4\right]&=24in^3+36i^2n^2+\left(12i^3+2i\right)n+i^4,\\
\mathbf{E}_i\left[Z_n^5\right]&=120in^4+240i^2n^3+\left(120i^3+30i\right)n^2+\left(20i^4+10i^2\right)n+i^5,\\
\mathbf{E}_i\left[Z_n^6\right]&=720in^5+1800i^2n^4+\left(1200i^3+360i\right)n^3,\\
&\quad+\left(300i^4+240i^2\right)n^2+\left(30i^5+30i^3+2i\right)n+i^6.
\end{aligned} \tag{21.40}$$

Hence, Z is a martingale, and the first six centered moments are

$$\begin{aligned}
\mathbf{E}_i\big[(Z_n-i)^2\big] &= 2in,\\
\mathbf{E}_i\big[(Z_n-i)^3\big] &= 6in^2,\\
\mathbf{E}_i\big[(Z_n-i)^4\big] &= 24in^3+12i^2n^2+2in,\\
\mathbf{E}_i\big[(Z_n-i)^5\big] &= 120in^4+120i^2n^3+30in^2,\\
\mathbf{E}_i\big[(Z_n-i)^6\big] &= 720in^5+1080i^2n^4+\big(120i^3+360i\big)n^3\\
&\qquad +60i^2n^2+2in.
\end{aligned} \tag{21.41}$$

Proof The exact formulas for the first six moments are obtained by tenaciously computing the right-hand side of (21.39). □

Now consider the following rescaling: Fix $x \ge 0$, start with $Z_0 = \lfloor nx \rfloor$ individuals and consider $\tilde{Z}_t^n := \frac{Z_{\lfloor tn \rfloor}}{n}$ for $t \ge 0$. We abbreviate

$$\mathcal{L}_x\big[\tilde{Z}^n\big] := \mathcal{L}_{\lfloor nx \rfloor}\big[\big(n^{-1}Z_{\lfloor nt \rfloor}\big)_{t\ge 0}\big]. \tag{21.42}$$

Evidently, $\mathbf{E}_x[\tilde{Z}_t^n] = \frac{\lfloor nx \rfloor}{n} \le x$ for every n; hence $(\mathcal{L}_x[\tilde{Z}_t^n], n \in \mathbb{N})$ is tight. By considering Laplace transforms, we obtain that, for every $\lambda \ge 0$, the sequence of distributions converges:

$$\begin{aligned}
\lim_{n\to\infty} \mathbf{E}_x\big[e^{-\lambda \tilde{Z}_t^n}\big] &= \lim_{n\to\infty}\big(\psi^{(\lfloor tn \rfloor)}\big(e^{-\lambda/n}\big)\big)^{nx}\\
&= \lim_{n\to\infty}\left(\frac{nt-(nt-1)e^{-\lambda/n}}{nt+1-nte^{-\lambda/n}}\right)^{nx}\\
&= \lim_{n\to\infty}\left(1-\frac{1-e^{-\lambda/n}}{n(1-e^{-\lambda/n})t+1}\right)^{nx}\\
&= \exp\left(-\lim_{n\to\infty}\frac{xn(1-e^{-\lambda/n})}{n(1-e^{-\lambda/n})t+1}\right)\\
&= \exp\left(-\frac{\lambda}{\lambda+1/t}(x/t)\right) := \psi_t(\lambda)^x.
\end{aligned} \tag{21.43}$$

However, the function ψ_t^x is the Laplace transform of the compound Poisson distribution $\mathrm{CPoi}_{(x/t)\exp_{1/t}}$ (see Definition 16.3).

Consider the stochastic kernel $\kappa_t(x,dy) := \mathrm{CPoi}_{(x/t)\exp_{1/t}}(dy)$. This is the kernel on $[0,\infty)$ whose Laplace transform is given by

$$\int_0^\infty \kappa_t(x,dy)e^{-\lambda y} = \psi_t(\lambda)^x. \tag{21.44}$$

Lemma 21.46 *$(\kappa_t)_{t\geq 0}$ is a Markov semigroup and there exists a Markov process $(Y_t)_{t\geq 0}$ with transition kernels $\mathbf{P}_x[Y_t \in dy] = \kappa_t(x, dy)$.*

Proof It suffices to check that the Chapman–Kolmogorov equation $\kappa_t \cdot \kappa_s = \kappa_{s+t}$ holds. We compute the Laplace transform for these kernels. For $\lambda \geq 0$, applying (21.44) twice yields

$$\begin{aligned}\int\int \kappa_t(x, dy)\kappa_s(y, dz)e^{-\lambda z} &= \int \kappa_t(x, dy) \exp\left(-\frac{\lambda y}{\lambda s + 1}\right)\\ &= \exp\left(-\frac{\frac{\lambda}{\lambda s+1}}{\frac{\lambda}{\lambda s+1}t + 1}x\right)\\ &= \exp\left(-\frac{\lambda x}{\lambda(t+s)+1}\right)\\ &= \int \kappa_{t+s}(x, dz)e^{-\lambda z}.\end{aligned}$$

□

Next we show that Y has a continuous version. To this end, we compute some of its moments and then use the Kolmogorov–Chentsov theorem (Theorem 21.6).

Lemma 21.47 *The first k moments of Y_t can be computed by differentiating the Laplace transform,*

$$\mathbf{E}_x\left[Y_t^k\right] = (-1)^k \frac{d^k}{d\lambda^k}\left(\psi(\lambda)^x\right)_{|\lambda=0},$$

where

$$\psi_t(\lambda) = \exp\left(-\frac{\lambda}{\lambda t + 1}\right).$$

In particular, we have

$$\begin{aligned}\mathbf{E}_x[Y_t] &= x,\\ \mathbf{E}_x\left[Y_t^2\right] &= 2xt + x^2,\\ \mathbf{E}_x\left[Y_t^3\right] &= 6xt^2 + 6x^2t + x^3,\\ \mathbf{E}_x\left[Y_t^4\right] &= 24xt^3 + 36x^2t^2 + 12x^3t + x^4,\\ \mathbf{E}_x\left[Y_t^5\right] &= 120xt^4 + 240x^2t^3 + 120x^3t^2 + 20x^4t + x^5,\\ \mathbf{E}_x\left[Y_t^6\right] &= 720xt^5 + 1800x^2t^4 + 1200x^3t^3 + 300x^4t^2 + 30x^5t + x^6.\end{aligned} \tag{21.45}$$

Hence Y is a martingale, and the first centered moments are

$$\begin{aligned}
\mathbf{E}_x\big[(Y_t-x)^2\big] &= 2xt,\\
\mathbf{E}_x\big[(Y_t-x)^3\big] &= 6xt^2,\\
\mathbf{E}_x\big[(Y_t-x)^4\big] &= 24xt^3+12x^2t^2,\\
\mathbf{E}_x\big[(Y_t-x)^5\big] &= 120xt^4+120x^2t^3,\\
\mathbf{E}_x\big[(Y_t-x)^6\big] &= 720xt^5+1080x^2t^4+120x^3t^3.
\end{aligned} \tag{21.46}$$

Theorem 21.48 *There is a continuous version of the Markov process Y with transition kernels $(\kappa_t)_{t\geq 0}$ given by (21.44). This version is called* Feller's (continuous) branching diffusion.

See Fig. 26.4 for a computer simulation of Feller's branching diffusion.

Proof For fixed $N>0$ and $s,t\in[0,N]$, we have

$$\begin{aligned}
\mathbf{E}_x\big[(Y_{t+s}-Y_s)^4\big] &= \mathbf{E}_x\big[\mathbf{E}_{Y_s}\big[(Y_t-Y_0)^4\big]\big]\\
&= \mathbf{E}_x\big[24Y_st^3+12Y_s^2t^2\big]\\
&= 24xt^3+12\big(2sx+x^2\big)t^2\\
&\leq \big(48Nx+12x^2\big)t^2.
\end{aligned}$$

Thus Y satisfies the condition of Theorem 21.6 (Kolmogorov–Chentsov) with $\alpha=4$ and $\beta=1$. □

Remark 21.49

(i) By using higher moments, it can be shown that the paths of Y are Hölder-continuous of any order $\gamma\in(0,\frac{1}{2})$.
(ii) It can be shown that Y is the (unique strong) solution of the stochastic (Itô-) differential equation (see Examples 26.11 and 26.31)

$$dY_t=\sqrt{2Y_t}\,dW_t, \tag{21.47}$$

where W is a Brownian motion. ◊

Theorem 21.50 *We have $\mathcal{L}_x[\tilde{Z}^n]\xrightarrow[\text{fdd}]{n\to\infty}\mathcal{L}_x[Y]$.*

Proof As in (21.43) for $0 \le t_1 \le t_2$, $\lambda_1, \lambda_2 \ge 0$ and $x \ge 0$, we get

$$\begin{aligned}
&\lim_{n\to\infty} \mathbf{E}_x\big[e^{-(\lambda_1 \tilde{Z}^n_{t_1} + \lambda_2 \tilde{Z}^n_{t_2})}\big] \\
&\quad = \lim_{n\to\infty} \mathbf{E}_x\big[\mathbf{E}_x\big[e^{-\lambda_2 \tilde{Z}^n_{t_2}} | \tilde{Z}^n_{t_1}\big] e^{-\lambda_1 \tilde{Z}^n_{t_1}}\big] \\
&\quad = \lim_{n\to\infty} \mathbf{E}_x\left[\exp\left(-\frac{\lambda_2}{\lambda_2(t_2-t_1)+1}\tilde{Z}^n_{t_1}\right) e^{-\lambda_1 \tilde{Z}^n_{t_1}}\right] \\
&\quad = \exp\left(-\frac{\left(\frac{\lambda_2}{\lambda_2(t_2-t_1)+1} + \lambda_1\right)x}{\left(\frac{\lambda_2}{\lambda_2(t_2-t_1)+1} + \lambda_1\right)t_1 + 1}\right) \\
&\quad = \mathbf{E}_x\big[\exp\big(-(\lambda_1 Y_{t_1} + \lambda_2 Y_{t_2})\big)\big].
\end{aligned}$$

Hence, we obtain

$$\mathcal{L}_x\big[\lambda_1 \tilde{Z}^n_{t_1} + \lambda_2 \tilde{Z}^n_{t_2}\big] \overset{n\to\infty}{\longrightarrow} \mathcal{L}_x[\lambda_1 Y_{t_1} + \lambda_2 Y_{t_2}].$$

Using the Cramér–Wold device (Theorem 15.56), this implies

$$\mathcal{L}_x\big[\big(\tilde{Z}^n_{t_1}, \tilde{Z}^n_{t_2}\big)\big] \overset{n\to\infty}{\longrightarrow} \mathcal{L}_x\big[(Y_{t_1}, Y_{t_2})\big].$$

Iterating the argument, for every $k \in \mathbb{N}$ and $0 \le t_1 \le t_2 \le \ldots \le t_k$, we get

$$\mathcal{L}_x\big[\big(\tilde{Z}^n_{t_i}\big)_{i=1,\ldots,k}\big] \overset{n\to\infty}{\longrightarrow} \mathcal{L}_x\big[(Y_{t_i})_{i=1,\ldots,k}\big].$$

However, this was the claim. □

The final step is to show convergence in path space. To this end, we have to modify the rescaled processes so that they become continuous. Assume that $(Z^n_i)_{i\in\mathbb{N}_0}$, $n \in \mathbb{N}$ is a sequence of Galton–Watson processes with $Z^n_0 = \lfloor nx \rfloor$. Define the linearly interpolated processes

$$\bar{Z}^n_t := \big(t - n^{-1}\lfloor tn\rfloor\big)\big(Z^n_{\lfloor tn\rfloor+1} - Z^n_{\lfloor tn\rfloor}\big) + \frac{1}{n} Z^n_{\lfloor tn\rfloor}.$$

Theorem 21.51 (Lindvall (1972), see [109]) *As $n \to \infty$, in the sense of weak convergence in $\mathcal{M}_1(C([0,\infty)))$, the rescaled Galton–Watson processes $\bar{Z}^n$ converge to Feller's diffusion Y:*

$$\mathcal{L}_x\big[\bar{Z}^n\big] \overset{n\to\infty}{\longrightarrow} \mathcal{L}_x[Y].$$

Proof We have shown already the convergence of the finite-dimensional distributions. By Theorem 21.38, it is thus enough to show tightness of $(\mathcal{L}_x[\bar{Z}^n], n \in \mathbb{N})$ in $\mathcal{M}_1(C([0,\infty)))$. To this end, we apply Kolmogorov's moment criterion (Theorem 21.42 with $\alpha = 4$ and $\beta = 1$). Hence, for fixed $N > 0$, we compute the fourth moments $\mathbf{E}_x[(\bar{Z}^n_{t+s} - \bar{Z}^n_s)^4]$ for $s, t \in [0, N]$. We distinguish two cases:

Case 1: $t < \frac{1}{n}$. Let $k = \lfloor (t+s)n \rfloor$. First assume that $\lfloor sn \rfloor = k$. Then (by Lemma 21.45)

$$\begin{aligned}
\mathbf{E}_x\big[\big(\bar{Z}^n_{t+s} - \bar{Z}^n_s\big)^4\big] &= n^{-4}(tn)^4 \mathbf{E}_{\lfloor nx \rfloor}\big[\big(Z^n_{k+1} - Z^n_k\big)^4\big] \\
&= t^4 \mathbf{E}_{\lfloor nx \rfloor}\big[24 Z^n_k + 12\big(Z^n_k\big)^2 + 2 Z^n_k\big] \\
&= t^4\big(26\lfloor nx \rfloor + 24\lfloor nx \rfloor k + \lfloor nx \rfloor^2\big) \\
&\le 26 x t^3 + 24 x s t^2 + x^2 t^2 \\
&\le \big(50 N x + x^2\big) t^2.
\end{aligned}$$

In the case $\lfloor sn \rfloor = k-1$, we get a similar estimate. Therefore, there is a constant $C = C(N,x)$ such that

$$\mathbf{E}_x\big[\big(\bar{Z}^n_{s+t} - \bar{Z}^n_s\big)^4\big] \le C t^2 \quad \text{for all } s,t \in [0,N] \text{ with } t < \frac{1}{n}. \tag{21.48}$$

Case 2: $t \ge \frac{1}{n}$. Define $k := \lceil (t+s)n \rceil - \lfloor sn \rfloor \le tn + 1 \le 2tn$. Then (by Lemma 21.45)

$$\begin{aligned}
&\mathbf{E}_x\big[\big(\bar{Z}^n_{t+s} - \bar{Z}^n_s\big)^4\big] \\
&\quad \le n^{-4} \mathbf{E}_{\lfloor nx \rfloor}\big[\big(Z^n_{\lfloor (t+s)n \rfloor} - Z^n_{\lfloor sn \rfloor}\big)^4\big] \\
&\quad = n^{-4} \mathbf{E}_{\lfloor nx \rfloor}\big[\mathbf{E}_{Z^n_{\lfloor sn \rfloor}}\big[\big(Z^n_k - Z^n_0\big)^4\big]\big] \\
&\quad = n^{-4} \mathbf{E}_{\lfloor nx \rfloor}\big[24 Z^n_{\lfloor sn \rfloor} k^3 + 12\big(Z^n_{\lfloor sn \rfloor}\big)^2 k^2 + 2 Z^n_{\lfloor sn \rfloor} k\big] \\
&\quad \le n^{-4}\big(24 x n (2tn)^3 + \big(24 x n s n + 12 x^2 n^2\big)(2tn)^2 + 4 x t n^2\big) \\
&\quad \le 192 x t^3 + \big(96 x s + 48 x^2\big) t^2 + 4 x n^{-1} t^2 \\
&\quad \le \big(292 N x + 48 x^2\big) t^2.
\end{aligned} \tag{21.49}$$

Combining the estimates (21.48) and (21.49), the assumptions of Kolmogorov's moment criterion for tightness (Theorem 21.42) are fulfilled with $\alpha = 4$ and $\beta = 1$. Hence the sequence $(\mathcal{L}_x[\bar{Z}^n], n \in \mathbb{N})$ is tight. □

21.10 Square Variation and Local Martingales

By the Paley–Wiener-Zygmund theorem (Theorem 21.17), the paths $t \mapsto W_t$ of Brownian motion are almost surely nowhere differentiable and hence have locally infinite variation. In particular, the stochastic integral $\int_0^1 f(s)\,dW_s$ that we introduced in Example 21.29 cannot be understood as a Lebesgue–Stieltjes integral. However, as a preparation for the construction of integrals of this type for larger classes of integrands and integrators (in Chapter 25), here we investigate the path properties of Brownian motion and, somewhat more generally, of continuous local martingales in more detail.

Definition 21.52 Let $G : [0, \infty) \to \mathbb{R}$ be a continuous function. For any $t \geq 0$, define the *variation* up to t by

$$V_t^1(G) := \sup\left\{\sum_{i=0}^{n-1} |G_{t_{i+1}} - G_{t_i}| : 0 = t_0 \leq t_1 \leq \ldots \leq t_n = t,\ n \in \mathbb{N}\right\}.$$

We say that G has locally finite variation if $V_t^1(G) < \infty$ for all $t \geq 0$. We write $\mathcal{C}_{\mathrm{v}}$ for the vector space of continuous functions G with continuous variation $t \mapsto V_t^1(G)$.

Remark 21.53 Clearly, $V^1(F + G) \leq V^1(F) + V^1(G)$ and $V^1(\alpha G) = |\alpha| V^1(G)$ for all continuous $F, G : [0, \infty) \to \mathbb{R}$ and for all $\alpha \in \mathbb{R}$. Hence $\mathcal{C}_{\mathrm{v}}$ is indeed a vector space. ◊

Remark 21.54

(i) If G is of the form $G_t = \int_0^t f(s)\, ds$ for some locally integrable function f, then we have $G \in \mathcal{C}_{\mathrm{v}}$ with $V_t^1(G) = \int_0^t |f(s)|\, ds$.

(ii) If $G = G^+ - G^-$ is the difference of two continuous monotone increasing functions G^+ and G^-, then

$$V_t^1(G) - V_s^1(G) \leq \left(G_t^+ - G_s^+\right) + \left(G_t^- - G_s^-\right) \quad \text{for all } t > s, \tag{21.50}$$

hence we have $G \in \mathcal{C}_{\mathrm{v}}$. In (21.50), equality holds if G^- and G^+ "do not grow on the same sets"; that is, more formally, if G^- and G^+ are the distribution functions of mutually singular measures μ^- and μ^+. The measures μ^- and μ^+ are then the Jordan decomposition of the signed measure $\mu = \mu^+ - \mu^-$ whose distribution function is G. Then the *Lebesgue–Stieltjes integral* is defined by

$$\int_0^t F(s)\, dG_s := \int_{[0,t]} F\, d\mu^+ - \int_{[0,t]} F\, d\mu^-. \tag{21.51}$$

(iii) If $G \in \mathcal{C}_{\mathrm{v}}$, then clearly

$$G_t^+ := \frac{1}{2}\left(V_t^1(G) + G_t\right) \quad \text{and} \quad G_t^- := \frac{1}{2}\left(V_t^1(G) - G_t\right)$$

is a decomposition of G as in (ii). ◊

The fact that the paths of Brownian motion are nowhere differentiable can be used to infer that the paths have infinite variation. However, there is also a simple direct argument.

Theorem 21.55 *Let W be a Brownian motion. Then $V_t^1(W) = \infty$ almost surely for every $t > 0$.*

Proof It is enough to consider $t = 1$ and to show

$$Y_n := \sum_{i=1}^{2^n} |W_{i2^{-n}} - W_{(i-1)2^{-n}}| \overset{n\to\infty}{\longrightarrow} \infty \quad \text{a.s.} \tag{21.52}$$

We have $\mathbf{E}[Y_n] = 2^{n/2}\mathbf{E}[|W_1|] = 2^{n/2}\sqrt{2/\pi}$ and $\mathbf{Var}[Y_n] = 1 - 2/\pi$. By Chebyshev's inequality,

$$\sum_{n=1}^{\infty} \mathbf{P}\left[Y_n \le \frac{1}{2}2^{n/2}\sqrt{2/\pi}\right] \le \sum_{n=1}^{\infty} \frac{2\pi - 4}{2^n} = 2\pi - 4 < \infty.$$

Using the Borel–Cantelli lemma, this implies (21.52). □

Evidently, the variation is too crude a measure to quantify essential path properties of Brownian motion. Hence, instead of the increments (in the definition of the variation), we will sum up the (smaller) *squared* increments. For the definition of this *square* variation, more care is needed than in Definition 21.52 for the variation.

Definition 21.56 A sequence $\mathcal{P} = (\mathcal{P}^n)_{n\in\mathbb{N}}$ of countable subsets of $[0, \infty)$,

$$\mathcal{P}^n := \{t_0, t_1, t_2, \ldots\} \quad \text{with } 0 = t_0 < t_1 < t_2 < \ldots,$$

is called an *admissible partition sequence* if

(i) $\mathcal{P}^1 \subset \mathcal{P}^2 \subset \ldots$,
(ii) $\sup \mathcal{P}^n = \infty$ for every $n \in \mathbb{N}$, and
(iii) the *mesh size*

$$\left|\mathcal{P}^n\right| := \sup_{t\in\mathcal{P}^n} \min_{s\in\mathcal{P}^n, s\neq t} |s - t|$$

tends to 0 as $n \to \infty$.

If $0 \le S < T$, then define

$$\mathcal{P}^n_{S,T} := \mathcal{P}^n \cap [S, T) \quad \text{and} \quad \mathcal{P}^n_T := \mathcal{P}^n \cap [0, T).$$

If $t = t_k \in \mathcal{P}^n_T$, then let $t' := t_{k+1} \wedge T = \min\{s \in \mathcal{P}^n_T \cup \{T\} : s > t\}$.

Example 21.57 $\mathcal{P}^n = \{k2^{-n} : k = 0, 1, 2, \ldots\}$. ◊

Definition 21.58 For continuous $F, G : [0, \infty) \to \mathbb{R}$ and for $p \geq 1$, define the p-variation of G (along $\mathcal{P}$) by

$$V_T^p(G) := V_T^{\mathcal{P},p}(G) := \lim_{n\to\infty} \sum_{t\in\mathcal{P}_T^n} |G_{t'} - G_t|^p \quad \text{for } T \geq 0$$

if the limit exists. In particular, $\langle G\rangle := V^2(G)$ is called the *square variation* of G. If $T \mapsto V_T^2(G)$ is continuous, then we write $G \in \mathcal{C}_{\mathrm{qv}} := \mathcal{C}_{\mathrm{qv}}^{\mathcal{P}}$.

If, for every $T \geq 0$, the limit

$$V_T^{\mathcal{P},2}(F, G) := \lim_{n\to\infty} \sum_{t\in\mathcal{P}_T^n} (F_{t'} - F_t)(G_{t'} - G_t)$$

exists, then we call $\langle F, G\rangle := V^2(F, G) := V^{\mathcal{P},2}(F, G)$ the *quadratic covariation* of F and G (along $\mathcal{P}$).

Remark 21.59 If $p' > p$ and $V_T^p(G) < \infty$, then $V_T^{p'}(G) = 0$. In particular, we have $\langle G\rangle \equiv 0$ if G has locally finite variation. ◊

Remark 21.60 By the triangle inequality, we have

$$\sum_{t\in\mathcal{P}_T^{n+1}} |G_{t'} - G_t| \geq \sum_{t\in\mathcal{P}_T^n} |G_{t'} - G_t| \quad \text{for all } n \in \mathbb{N}, T \geq 0.$$

Hence in the case $p = 1$, the limit always exists and coincides with $V^1(G)$ from Definition 21.52 (and is hence independent of the particular choice of $\mathcal{P}$). A similar inequality does not hold for V^2 and thus the limit need not exist or may depend on the choice of $\mathcal{P}$. In the following, we will, however, show that, for a large class of continuous stochastic processes, V^2 exists almost surely along a suitable subsequence of partitions and is almost surely unique. ◊

Remark 21.61

(i) If $\langle F + G\rangle_T$ and $\langle F - G\rangle_T$ exist, then the covariation $\langle F, G\rangle_T$ exists and the *polarization formula* holds:

$$\langle F, G\rangle_T = \frac{1}{4}\big(\langle F + G\rangle_T - \langle F - G\rangle_T\big).$$

(ii) If $\langle F\rangle_T$, $\langle G\rangle_T$ and $\langle F, G\rangle_T$ exist, then by the Cauchy–Schwarz inequality, we have for the approximating sums

$$V_T^1\big(\langle F, G\rangle\big) \leq \sqrt{\langle F\rangle_T \langle G\rangle_T}.$$ ◊

Remark 21.62 If $f \in C^1(\mathbb{R})$ and $G \in \mathcal{C}_{\mathrm{qv}}$, then (exercise!) in the sense of the Lebesgue–Stieltjes integral

$$\langle f(G)\rangle_T = \int_0^T \left(f'(G_s)\right)^2 d\langle G\rangle_s. \qquad \Diamond$$

Corollary 21.63 *If F has locally finite square variation and if $\langle G\rangle \equiv 0$ (hence, in particular, if G has locally finite variation), then $\langle F, G\rangle \equiv 0$ and $\langle F+G\rangle = \langle F\rangle$.*

Theorem 21.64 *For Brownian motion W and for every admissible sequence of partitions, we have*

$$\langle W\rangle_T = T \quad \textit{for all } T \ge 0 \quad \textit{a.s.}$$

Proof We prove this only for the case where

$$\sum_{n=1}^{\infty} \left|\mathcal{P}^n\right| < \infty. \tag{21.53}$$

For the general case, we only sketch the argument.

Accordingly, assume (21.53). If $\langle W\rangle$ exists, then $T \mapsto \langle W\rangle_T$ is monotone increasing. Hence, it is enough to show that $\langle W\rangle_T$ exists for every $T \in \mathbb{Q}^+ = \mathbb{Q} \cap [0, \infty)$ and that almost surely $\langle W\rangle_T = T$. Since $(\widetilde{W}_t)_{t\ge 0} = (T^{-1/2} W_{tT})_{t\ge 0}$ is a Brownian motion, and since $\langle \widetilde{W}\rangle_1 = T^{-1}\langle W\rangle_T$, it is enough to consider the case $T = 1$.

Define

$$Y_n := \sum_{t\in\mathcal{P}_1^n} (W_{t'} - W_t)^2 \quad \text{for all } n \in \mathbb{N}.$$

Then $\mathbf{E}[Y_n] = \sum_{t\in\mathcal{P}_1^n} (t' - t) = 1$ and

$$\mathbf{Var}[Y_n] = \sum_{t\in\mathcal{P}_1^n} \mathbf{Var}\left[(W_{t'} - W_t)^2\right] = 2 \sum_{t\in\mathcal{P}_1^n} \left(t' - t\right)^2 \le 2|\mathcal{P}^n|.$$

By assumption (21.53), we thus have $\sum_{n=1}^{\infty} \mathbf{Var}[Y_n] \le 2\sum_{n=1}^{\infty} |\mathcal{P}^n| < \infty$; hence $Y_n \overset{n\to\infty}{\longrightarrow} 1$ almost surely.

If we drop the assumption (21.53), then we still have $\mathbf{Var}[Y_n] \overset{n\to\infty}{\longrightarrow} 0$; hence $Y_n \overset{n\to\infty}{\longrightarrow} 1$ in probability. However, it is not too hard to show that $(Y_n)_{n\in\mathbb{N}}$ is a backwards martingale (see, e.g., [140, Theorem I.28]) and thus converges almost surely to 1. □

Corollary 21.65 *If W and $\widetilde{W}$ are independent Brownian motions, then we have $\langle W, \widetilde{W}\rangle_T = 0$.*

Proof The continuous processes $((W+\widetilde{W})/\sqrt{2})$ and $(W-\widetilde{W})/\sqrt{2})$ have independent normally distributed increments. Hence they are Brownian motions. By Remark 21.61(i), we have

$$\begin{aligned}4\langle W,\widetilde{W}\rangle_T &= \langle W+\widetilde{W}\rangle_T - \langle W-\widetilde{W}\rangle_T\\ &= 2\big\langle (W+\widetilde{W})/\sqrt{2}\big\rangle_T - 2\big\langle (W-\widetilde{W})/\sqrt{2}\big\rangle_T\\ &= 2T-2T=0.\end{aligned}$$

□

Clearly, $(W_t\widetilde{W}_t)_{t\geq 0}$ is a continuous martingale. Now, by Exercise 21.4.2, the process $(W_t^2-t)_{t\geq 0}$ is also a continuous martingale. Thus, as shown above, the processes $W^2-\langle W\rangle$ and $W\widetilde{W}-\langle W,\widetilde{W}\rangle$ are martingales. We will see (Theorem 21.70) that the square variation $\langle M(\omega)\rangle$ of a square integrable continuous martingale M always exists (for almost all ω) and that the process $\langle M\rangle$ is uniquely determined by the property that $M^2-\langle M\rangle$ is a martingale.

In order to obtain a similar statement for continuous martingales that are not square integrable, we make the following definition.

Definition 21.66 (Local martingale) Let $\mathbb{F}$ be a filtration on $(\Omega,\mathcal{F},\mathbf{P})$ and let τ be an $\mathbb{F}$-stopping time. An adapted real-valued stochastic process $M=(M_t)_{t\geq 0}$ is called a *local martingale* up to time τ if there exists a sequence $(\tau_n)_{n\in\mathbb{N}}$ of stopping times such that $\tau_n\uparrow\tau$ almost surely and such that, for every $n\in\mathbb{N}$, the stopped process $M^{\tau_n}=(M_{\tau_n\wedge t})_{t\geq 0}$ is a uniformly integrable martingale. Such a sequence $(\tau_n)_{n\in\mathbb{N}}$ is called a *localising sequence* for M. M is called a local martingale if M is a local martingale up to time $\tau\equiv\infty$. Denote by $\mathcal{M}_{\mathrm{loc},c}$ the space of continuous local martingales.

Remark 21.67 Let M be a continuous adapted process and let τ be a stopping time. Then the following are equivalent:

(i) M is a local martingale up to time τ.
(ii) There is a sequence $(\tau_n)_{n\in\mathbb{N}}$ of stopping times with $\tau_n\uparrow\tau$ almost surely and such that every M^{τ_n} is a martingale.
(iii) There is a sequence $(\tau_n)_{n\in\mathbb{N}}$ of stopping times with $\tau_n\uparrow\tau$ almost surely and such that every M^{τ_n} is a bounded martingale.

Indeed, (iii)$\Longrightarrow$(i)$\Longrightarrow$(ii) is trivial. Hence assume that (ii) holds, and define τ_n' by

$$\tau_n' := \inf\big\{t\geq 0 : |M_t|\geq n\big\}\quad\text{for all } n\in\mathbb{N}.$$

Since M is continuous, we have $\tau_n'\uparrow\infty$. Hence $(\sigma_n)_{n\in\mathbb{N}} := (\tau_n\wedge\tau_n')_{n\in\mathbb{N}}$ is a localising sequence for M such that every M^{σ_n} is a bounded martingale. ◊

Remark 21.68 A bounded local martingale M is a martingale. Indeed, assume that $|M_t|\leq C<\infty$ almost surely for all $t\geq 0$ and that $(\tau_n)_{n\in\mathbb{N}}$ is a localising sequence for M. Let $t>s\geq 0$ and $A\in\mathcal{F}_s$. Then $A\cap\{\tau_n\leq s\}\in\mathcal{F}_{\tau_n\wedge s}$ and hence

$$\mathbf{E}[M_{\tau_n\wedge t}\mathbb{1}_{A\cap\{\tau_n\leq s\}}] = \mathbf{E}[M_{\tau_n\wedge s}\mathbb{1}_{A\cap\{\tau_n\leq s\}}].$$

Since $\tau_n \uparrow \infty$, the dominated convergence theorem (Corollary 6.26) yields

$$\mathbf{E}[M_t \mathbb{1}_A] = \mathbf{E}[M_s \mathbb{1}_A].$$

Hence $\mathbf{E}[M_t \mid \mathcal{F}_s] = M_s$ and thus M is a martingale. ◊

Example 21.69

(i) Every martingale is a local martingale.
(ii) In Remark 21.68, we saw that bounded local martingales are martingales. On the other hand, even a uniformly integrable local martingale need not be a martingale: Let $W = (W^1, W^2, W^3)$ be a three-dimensional Brownian motion (that is, W^1, W^2 and W^3 are independent Brownian motions) that starts at $W_0 = x \in \mathbb{R}^3 \setminus \{0\}$. Let

$$u(y) = \|y\|^{-1} \quad \text{for } y \in \mathbb{R}^3 \setminus \{0\}.$$

It is easy to check that u is harmonic; that is, $\Delta u(y) = 0$ for all $y \neq 0$. We will see later (Corollary 25.34) that this implies that $M := (u(W_t))_{t\geq 0}$ is a local martingale. Define a localising sequence for M by

$$\tau_n := \inf\{t > 0 : M_t \geq n\} = \inf\big\{t > 0 : \|W_t\| \leq 1/n\big\}, \quad n \in \mathbb{N}.$$

An explicit computation with the three-dimensional normal distribution shows $\mathbf{E}[M_t] \leq t^{-1/2} \overset{t\to\infty}{\longrightarrow} 0$; hence M is integrable but is not a martingale. Since $M_t \overset{t\to\infty}{\longrightarrow} 0$ in L^1, M is uniformly integrable. ◊

Theorem 21.70 *Let M be a continuous local martingale.*

(i) *There exists a unique continuous, monotone increasing, adapted process $\langle M\rangle = (\langle M\rangle_t)_{t\geq 0}$ with $\langle M\rangle_0 = 0$ such that*

$$\big(M_t^2 - \langle M\rangle_t\big)_{t\geq 0} \quad \textit{is a continuous local martingale.}$$

(ii) *If M is a continuous square integrable martingale, then $M^2 - \langle M\rangle$ is a martingale.*
(iii) *For every admissible sequence of partitions $\mathcal{P} = (\mathcal{P}^n)_{n\in\mathbb{N}}$, we have*

$$U_T^n := \sum_{t\in\mathcal{P}_T^n} (M_{t'} - M_t)^2 \overset{n\to\infty}{\longrightarrow} \langle M\rangle_T \quad \textit{in probability for all } T \geq 0.$$

The process $\langle M\rangle$ is called the square variation process *of M.*

Remark 21.71 By possibly passing in (iii) to a subsequence $\mathcal{P}'$ (that might depend on T), we may assume that $U_T^n \overset{n\to\infty}{\longrightarrow} \langle M\rangle_T$ almost surely. Using the diagonal sequence argument, we obtain (as in the proof of Helly's theorem) a sequence of partitions such that $U_T^n \overset{n\to\infty}{\longrightarrow} \langle M\rangle_T$ almost surely for all $T \in \mathbb{Q}^+$. Since both $T \mapsto U_T^n$

and $T \mapsto \langle M \rangle_T$ are monotone and continuous, we get $U_T^n \overset{n\to\infty}{\longrightarrow} \langle M \rangle_T$ for all $T \geq 0$ almost surely. Hence, for this sequence of partitions, the pathwise square variation almost surely equals the square variation process:

$$\big\langle M(\omega)\big\rangle = V^2\big(M(\omega)\big) = \langle M \rangle(\omega). \qquad \Diamond$$

Proof of Theorem 21.70 *Step 1*. First let $|M_t| \leq C$ almost surely for all $t \geq 0$ for some $C < \infty$. Then, in particular, M is a martingale (by Remark 21.68). Write $U_T^n = M_T^2 - M_0^2 - N_T^n$, where

$$N_T^n = 2 \sum_{t \in \mathcal{P}_T^n} M_t (M_{t'} - M_t), \quad T \geq 0,$$

is a continuous martingale. Assume that we can show that, for every $T \geq 0$, $(U_T^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{L}^2(\mathbf{P})$. Then also $(N_T^n)_{n\in\mathbb{N}}$ is a Cauchy sequence, and we can define $\widetilde{N}_T$ as the L^2-limit of $(N_T^n)_{n\in\mathbb{N}}$. By Exercise 21.4.3, $\widetilde{N}$ has a continuous modification N, and we have $N_T^n \overset{n\to\infty}{\longrightarrow} N_T$ in L^2 for all $T \geq 0$. Thus there exists a continuous process $\langle M \rangle$ with

$$U_T^n \overset{n\to\infty}{\longrightarrow} \langle M \rangle_T \quad \text{in } L^2 \quad \text{for all } T \geq 0, \tag{21.54}$$

and $N = M^2 - M_0^2 - \langle M \rangle$ is a continuous martingale.

It remains to show that, for all $T \geq 0$,

$$\big(U_T^n\big)_{n\in\mathbb{N}} \quad \text{is a Cauchy sequence in } L^2. \tag{21.55}$$

For $m \in \mathbb{N}$, let

$$Z_m := \max\big\{(M_t - M_s)^2 : s \in \mathcal{P}_T^m, t \in \mathcal{P}_{s,s'}^n, n \geq m\big\}.$$

Since M is almost surely uniformly continuous on $[0, T]$, we have $Z_m \overset{m\to\infty}{\longrightarrow} 0$ almost surely. As $Z_m \leq 4C^2$, we infer

$$\mathbf{E}\big[Z_m^2\big] \overset{m\to\infty}{\longrightarrow} 0. \tag{21.56}$$

For $n \in \mathbb{N}$ and numbers $a_0, \ldots, a_n$, we have

$$(a_n - a_0)^2 - \sum_{k=0}^{n-1} (a_{k+1} - a_k)^2 = 2 \sum_{k=0}^{n-1} (a_k - a_0)(a_{k+1} - a_k).$$

In the following computation, we apply this to each summand in the outer sum to obtain for $m \in \mathbb{N}$ and $n \geq m$

$$\begin{aligned} U_T^m - U_T^n &= \sum_{s \in \mathcal{P}_T^m} \Bigg((M_{s'} - M_s)^2 - \sum_{t \in \mathcal{P}^n_{s,s'}} (M_{t'} - M_t)^2 \Bigg) \\ &= 2 \sum_{s \in \mathcal{P}_T^m} \sum_{t \in \mathcal{P}^n_{s,s'}} (M_t - M_s)(M_{t'} - M_t). \end{aligned} \tag{21.57}$$

Since M is a martingale, for $s_1, s_2 \in \mathcal{P}_T^m$ and $t_1 \in \mathcal{P}^n_{s_1,s_1'}$, $t_2 \in \mathcal{P}^n_{s_2,s_2'}$ with $t_1 < t_2$, we have

$$\begin{aligned} &\mathbf{E}\big[(M_{t_1} - M_{s_1})(M_{t_1'} - M_{t_1})(M_{t_2} - M_{s_2})(M_{t_2'} - M_{t_2})\big] \\ &\quad = \mathbf{E}\big[(M_{t_1} - M_{s_1})(M_{t_1'} - M_{t_1})(M_{t_2} - M_{s_2})\mathbf{E}[M_{t_2'} - M_{t_2} \mid \mathcal{F}_{t_2}]\big] = 0. \end{aligned}$$

If we use (21.57) to compute the expectation of $(U_T^m - U_T^n)^2$, then the mixed terms vanish and we get (using the Cauchy–Schwarz inequality in the third line)

$$\begin{aligned} \mathbf{E}\big[\big(U_T^n - U_T^m\big)^2\big] &= 4\mathbf{E}\Bigg[\sum_{s \in \mathcal{P}_T^m} \sum_{t \in \mathcal{P}^n_{s,s'}} (M_t - M_s)^2 (M_{t'} - M_t)^2\Bigg] \\ &\le 4\mathbf{E}\Bigg[Z_m \sum_{t \in \mathcal{P}_T^n} (M_{t'} - M_t)^2\Bigg] \\ &\le 4\mathbf{E}\big[Z_m^2\big]^{1/2} \mathbf{E}\Bigg[\Bigg(\sum_{t \in \mathcal{P}_T^n} (M_{t'} - M_t)^2\Bigg)^2\Bigg]^{1/2}. \end{aligned} \tag{21.58}$$

For the second factor,

$$\begin{aligned} &\mathbf{E}\Bigg[\Bigg(\sum_{t \in \mathcal{P}_T^n} (M_{t'} - M_t)^2\Bigg)^2\Bigg] \\ &\quad = \mathbf{E}\Bigg[\sum_{t \in \mathcal{P}_T^n} (M_{t'} - M_t)^4\Bigg] \\ &\qquad + 2\mathbf{E}\Bigg[\sum_{s \in \mathcal{P}_T^n} (M_{s'} - M_s)^2 \sum_{t \in \mathcal{P}^n_{s',T}} (M_{t'} - M_t)^2\Bigg]. \end{aligned} \tag{21.59}$$

The first summand in (21.59) is bounded by

$$4C^2 \mathbf{E}\Bigg[\sum_{t \in \mathcal{P}_T^n} (M_{t'} - M_t)^2\Bigg] = 4C^2 \mathbf{E}\big[(M_T - M_0)^2\big] \le 16C^4.$$

The second summand in (21.59) equals

$$\begin{aligned}
&2\mathbf{E}\Bigg[\sum_{s\in\mathcal{P}_T^n}(M_{s'}-M_s)^2\mathbf{E}\Bigg[\sum_{t\in\mathcal{P}_{s',T}^n}(M_{t'}-M_t)^2\,\Big|\,\mathcal{F}_{s'}\Bigg]\Bigg]\\
&\quad=2\mathbf{E}\Bigg[\sum_{s\in\mathcal{P}_T^n}(M_{s'}-M_s)^2\mathbf{E}\big[(M_T-M_{s'})^2\,\big|\,\mathcal{F}_{s'}\big]\Bigg]\\
&\quad\le 8C^2\mathbf{E}\big[(M_T-M_0)^2\big]\le 32C^4.
\end{aligned}$$

Together with (21.58) and (21.56), we obtain

$$\sup_{n\ge m}\mathbf{E}\big[\big(U_T^n-U_T^m\big)^2\big]\le 16\sqrt{3}C^2\mathbf{E}\big[Z_m^2\big]^{1/2}\overset{m\to\infty}{\longrightarrow}0.$$

This shows (21.55).

Step 2. Now let $M\in\mathcal{M}_{\mathrm{loc},c}$ and let $(\tau_N)_{N\in\mathbb{N}}$ be a localising sequence such that every M^{τ_N} is a bounded martingale (see Remark 21.67). By Step 1, for $T\ge 0$ and $N\in\mathbb{N}$, we have

$$U_T^{N,n}:=\sum_{t\in\mathcal{P}_T^n}\big(M_{t'}^{\tau_N}-M_t^{\tau_N}\big)^2\overset{n\to\infty}{\longrightarrow}\big\langle M^{\tau_N}\big\rangle_T\ \text{ in } L^2.$$

Since $U_T^{N,n}=U_T^{N+1,n}$ if $T\le\tau_N$, there is a continuous process U with $U_T^{N,n}\overset{n\to\infty}{\longrightarrow}U_T$ in probability if $T\le\tau_N$. Thus $\langle M^{\tau_N}\rangle_T=\langle M\rangle_T:=U_T$ if $T\le\tau_N$. Since $\tau_N\uparrow\infty$ almost surely, for all $T\ge 0$,

$$U_T^n\overset{n\to\infty}{\longrightarrow}\langle M\rangle_T\quad\text{ in probability.}$$

As $((M_T^{\tau_N})^2-\langle M^{\tau_N}\rangle_T)_{T\ge 0}$ is a continuous martingale and since $\langle M^{\tau_N}\rangle=\langle M\rangle^{\tau_N}$, we have $M^2-\langle M\rangle\in\mathcal{M}_{\mathrm{loc},c}$.

Step 3. It remains to show (ii). Let M be a continuous square integrable martingale and let $(\tau_n)_{n\in\mathbb{N}}$ be a localising sequence for the local martingale $M^2-\langle M\rangle$. Let $T>0$ and let $\tau\le T$ be a stopping time. Then $M_{\tau_n\wedge\tau}^2\le\mathbf{E}[M_T^2\,|\,\mathcal{F}_{\tau_n\wedge\tau}]$ since M^2 is a nonnegative submartingale. Hence $(M_{\tau_n\wedge\tau}^2)_{n\in\mathbb{N}}$ is uniformly integrable and thus (using the monotone convergence theorem in the last step)

$$\begin{aligned}
\mathbf{E}\big[M_\tau^2\big]&=\lim_{n\to\infty}\mathbf{E}\big[M_{\tau_n\wedge\tau}^2\big]=\lim_{n\to\infty}\mathbf{E}\big[\langle M\rangle_{\tau_n\wedge\tau}\big]+\mathbf{E}\big[M_0^2\big]\\
&=\mathbf{E}\big[\langle M\rangle_\tau\big]+\mathbf{E}\big[M_0^2\big].
\end{aligned}$$

Thus, by the optional sampling theorem, $M^2-\langle M\rangle$ is a martingale.

Step 4 (Uniqueness). Let A and A' be continuous, monotone increasing, adapted processes with $A_0=A_0'$ such that M^2-A and M^2-A' are local martingales. Then also $N=A-A'$ is a local martingale, and for almost all ω, the path $N(\omega)$ has locally finite variation. Thus $\langle N\rangle\equiv 0$ and hence $N^2-\langle N\rangle=N^2$ is a continuous local martingale with $N_0 = 0$. Let $(\tau_n)_{n\in\mathbb{N}}$ be a localising sequence for N^2. Then $\mathbf{E}[N^2_{\tau_n\wedge t}] = 0$ for any $n \in \mathbb{N}$ and $t \geq 0$; hence $N^2_{\tau_n\wedge t} = 0$ almost surely and thus $N_t^2 = \lim_{n\to\infty} N^2_{\tau_n\wedge t} = 0$ almost surely. We conclude $A = A'$. □

Corollary 21.72 *Let M be a continuous local martingale with $\langle M\rangle \equiv 0$. Then $M_t = M_0$ for all $t \geq 0$ almost surely. In particular, this holds if the paths of M have locally finite variation.*

Corollary 21.73 *Let $M, N \in \mathcal{M}_{\mathrm{loc},c}$. Then there exists a unique continuous adapted process $\langle M, N\rangle$ with almost surely locally finite variation and $\langle M, N\rangle_0 = 0$ such that*

$$MN - \langle M, N\rangle \text{ is a continuous local martingale.}$$

$\langle M, N\rangle$ is called the quadratic covariation process *of M and N. For every admissible sequence of partitions $\mathcal{P}$ and for every $T \geq 0$, we have*

$$\langle M, N\rangle_T = \lim_{n\to\infty} \sum_{t\in\mathcal{P}^n_T} (M_{t'} - M_t)(N_{t'} - N_t) \quad \textit{in probability.} \tag{21.60}$$

Proof *Existence*. Manifestly, $M + N, M - N \in \mathcal{M}_{\mathrm{loc},c}$. Define

$$\langle M, N\rangle := \frac{1}{4}\big(\langle M + N\rangle - \langle M - N\rangle\big).$$

As the difference of two monotone increasing functions, $\langle M, N\rangle$ has locally finite variation. Using Theorem 21.70(iii), we get (21.60). Furthermore,

$$MN - \langle M, N\rangle = \frac{1}{4}\big((M + N)^2 - \langle M + N\rangle\big) - \frac{1}{4}\big((M - N)^2 - \langle M - N\rangle\big)$$

is a local martingale.

Uniqueness. Let A and A' with $A_0 = A'_0 = 0$ be continuous, adapted and with locally finite variation such that $MN - A$ and $MN - A'$ are in $\mathcal{M}_{\mathrm{loc},c}$. Then $A - A' \in \mathcal{M}_{\mathrm{loc},c}$ have locally finite variation; hence $A - A' = 0$. □

Corollary 21.74 *If $M \in \mathcal{M}_{\mathrm{loc},c}$ and A are continuous and adapted with $\langle A\rangle \equiv 0$, then $\langle M + A\rangle = \langle M\rangle$.*

If M is a continuous local martingale up to the stopping time τ, then $M^\tau \in \mathcal{M}_{\mathrm{loc},c}$, and we write $\langle M\rangle_t := \langle M^\tau\rangle_t$ for $t < \tau$.

Theorem 21.75 *Let τ be a stopping time, M be a continuous local martingale up to τ and $\tau_0 < \tau$ a stopping time with $\mathbf{E}[\langle M\rangle_{\tau_0}] < \infty$. Then $\mathbf{E}[M_{\tau_0}] = \mathbf{E}[M_0]$, and M^{τ_0} is an L^2-bounded martingale if $\mathbf{E}[M_0^2] < \infty$.*

Proof Let $\tau_n \uparrow \tau$ be a localising sequence of stopping times for M such that every M^{τ_n} is even a bounded martingale (see Remark 21.67). Then $M^{\tau_0 \wedge \tau_n}$ is also a bounded martingale, and for every $t \geq 0$, we have

$$\mathbf{E}\big[M^2_{\tau_0 \wedge \tau_n \wedge t}\big] = \mathbf{E}\big[M_0^2\big] + \mathbf{E}\big[\langle M \rangle_{\tau_0 \wedge \tau_n \wedge t}\big] \leq \mathbf{E}\big[M_0^2\big] + \mathbf{E}\big[\langle M \rangle_{\tau_0}\big] < \infty. \quad (21.61)$$

Hence $((M_{\tau_0 \wedge \tau_n \wedge t}), n \in \mathbb{N}, t \geq 0)$ is bounded in L^2 and is thus uniformly integrable. Therefore, by the optional sampling theorem for uniformly integrable martingales,

$$\mathbf{E}[M_{\tau_0}] = \lim_{n \to \infty} \mathbf{E}[M_{\tau_0 \wedge \tau_n}] = \mathbf{E}[M_0],$$

and, for $t > s$,

$$\begin{aligned}
\mathbf{E}\big[M_t^{\tau_0} \,\big|\, \mathcal{F}_s\big] &= \mathbf{E}\Big[\lim_{n \to \infty} M_t^{\tau_0 \wedge \tau_n} \,\big|\, \mathcal{F}_s\Big] \\
&= \lim_{n \to \infty} \mathbf{E}\big[M_t^{\tau_0 \wedge \tau_n} \,\big|\, \mathcal{F}_s\big] \\
&= \lim_{n \to \infty} M_s^{\tau_0 \wedge \tau_n} = M_s^{\tau_0}.
\end{aligned}$$

Hence M^{τ_0} is a martingale. □

Corollary 21.76 *If $M \in \mathcal{M}_{\mathrm{loc},c}$ with $\mathbf{E}[M_0^2] < \infty$ and $\mathbf{E}[\langle M \rangle_t] < \infty$ for every $t \geq 0$, then M is a square integrable martingale.*

Exercise 21.10.1 Show that the random variables $(Y_n)_{n \in \mathbb{N}}$ from the proof of Theorem 21.64 form a backwards martingale.

Exercise 21.10.2 Let $f : [0, \infty) \to \mathbb{R}$ be continuous and let $X \in \mathcal{C}_{\mathrm{qv}}^{\mathcal{P}}$ for the admissible sequence of partitions $\mathcal{P}$. Show that

$$\int_0^T f(s)\, d\langle X \rangle_s = \lim_{n \to \infty} \sum_{t \in \mathcal{P}_T^n} f(t)(X_{t'} - X_t)^2 \quad \text{for all } T \geq 0.$$

Exercise 21.10.3 Show by a counterexample that if M is a continuous local martingale with $M_0 = 0$ and if τ is a stopping time with $\mathbf{E}[\langle M \rangle_\tau] = \infty$, then this does not necessarily imply $\mathbf{E}[M_\tau^2] = \infty$.