Universität Trier

Stochastic Processes

Dr. Christoph Belak

Winter Term 2017/18

Christoph Belak

Department IV – Mathematics
University of Trier
Universitätsring 19
D–54269 Trier
Germany

belak@uni-trier.de

Contents

Chapter

1 Stochastic Processes

In the course on probability theory, the main object of study was that of a **random variable**. Random variables are often interpreted as mathematical models of **static** random phenomena, that is, random phenomena which occur only once. Think for example of the lottery next weekend, the outcome of which could be modeled as a hypergeometric random variable.

In this course, we direct our focus towards the mathematics behind **dynamic** random phenomena, i.e. random phenomena which have an additional time component. Examples could be the sunshine hours each day next year, stock prices, or even the number of students showing up in each lecture of this course. A natural way of modeling these phenomena is to use sequences of random variables, or, more generally, families of random variables indexed by a suitable time index set. These families will be referred to as **stochastic processes** and are the main object of interest in this course.

To formalize the notion of a stochastic process mathematically, we take as given a **probability space** $(\Omega, \mathfrak{A}, \mathbb{P})$ (the source of uncertainty) as well as a **measurable space** $(S, \mathfrak{S})$ (the space of possible outcomes at each time instant). Moreover, we fix a **time index set** $\mathcal{T}$, which is assumed to be a subset of the extended real line $\overline{\mathbb{R}} \triangleq \mathbb{R} \cup \{-\infty, \infty\}$.

Definition 1.1 (Stochastic Process)**.** *A family $X = \{X(t)\}_{t \in \mathcal{T}}$ of S-valued random variables is called a **stochastic process**.* ◇

Observe that the random variables $\{X(t)\}_{t \in \mathcal{T}}$ are defined on the **same probability space** and take values in the **same measurable space**. Typical choices for the time index set are $\mathcal{T} = \{1, \ldots, N\}$ for $N \in \mathbb{N}$, $\mathcal{T} = \mathbb{N}_0$, $\mathcal{T} = [0, \infty)$, or $\mathcal{T} = [0, T]$ for $0 < T \leq \infty$.

The parameter t in the definition of a stochastic process is interpreted as a **time parameter** and hence the time index set $\mathcal{T}$ can be thought of as the set all possible time points. The random variable $X(t)$ is the model for the dynamic random phenomenon at time t. For example, if we choose $\mathcal{T} = \{1, 2, \ldots, 365\}$ and $S = \{1, 2, \ldots, 24\}$, then $X(1)$ is a model for the number of sunshine hours on the first day next year, $X(2)$ is our model for the number of sunshine hours on the second day next year, and so on.

In Definition 1.1, we take the point of view that for each $t \in \mathcal{T}$ fixed, the time-t value of the stochastic process is described by the random variable

$$X(t) : \Omega \to S, \qquad \omega \mapsto X(t, \omega).$$

On the other hand, we may also turn this around, keep $\omega \in \Omega$ fixed, and think of the stochastic process as a mapping

$$X(\cdot, \omega) : \mathcal{T} \to S, \qquad t \mapsto X(t, \omega).$$

The mapping $X(\cdot, \omega)$ is called a **path of the stochastic process** and describes the dynamic evolution if the state of the world is $\omega \in \Omega$.

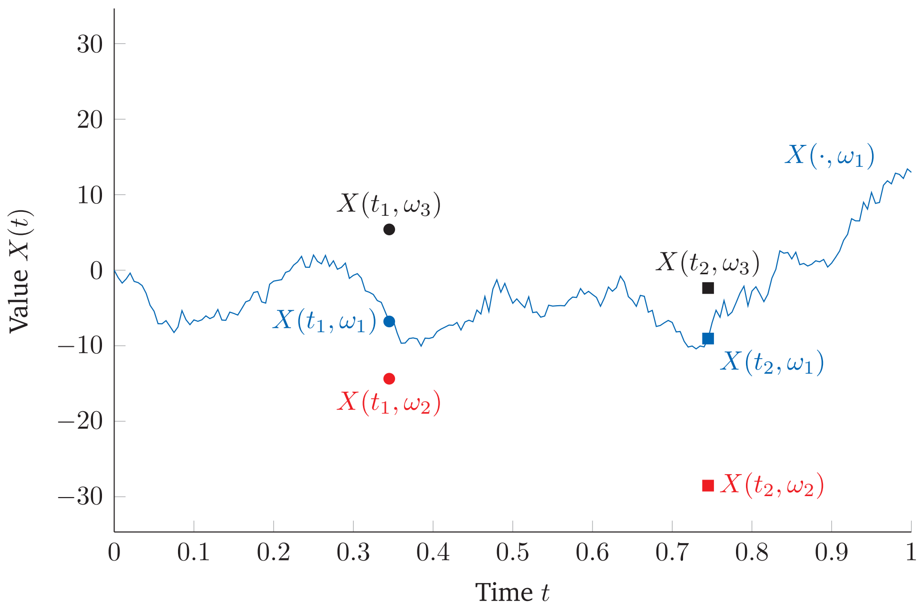

Figure 1.1 Interpreting X as a family of random variables vs. interpreting X as a path-valued random variable.

Taking this interpretation one step further, we may also think of a stochastic process as a random variable taking values in the set of all paths, i.e.

$$X : \Omega \to S^{\mathcal{T}}, \qquad \omega \mapsto X(\cdot, \omega) = \{X(t, \omega)\}_{t \in \mathcal{T}},$$

where $S^{\mathcal{T}} \triangleq \{f : \mathcal{T} \to S\}$ is the set of all functions from $\mathcal{T}$ to S, and $S^{\mathcal{T}}$ is equipped with the product σ-field $\mathfrak{S}^{\mathcal{T}} \triangleq \bigotimes_{t \in \mathcal{T}} \mathfrak{S}$. Indeed, the path-valued function $X : \Omega \to S^{\mathcal{T}}$ is $\mathfrak{A}$-$\mathfrak{S}^{\mathcal{T}}$-measurable if and only if the S-valued function $X(t) : \Omega \to S$ is $\mathfrak{A}$-$\mathfrak{S}$-measurable for all $t \in \mathcal{T}$.

Exercise 1. *Let $X = \{X(t)\}_{t \in \mathcal{T}}$ be a stochastic process. Show that X is $\mathfrak{A}$-$\mathfrak{S}^{\mathcal{T}}$-measurable if and only if $X(t)$ is $\mathfrak{A}$-$\mathfrak{S}$-measurable for all $t \in \mathcal{T}$.* ◇

Since we can regard stochastic processes as path-valued random variables, each stochastic process induces a probability measure on $(S^{\mathcal{T}}, \mathfrak{S}^{\mathcal{T}})$.

Definition 1.2 (Distribution of a Process). *Let $X = \{X(t)\}_{t \in \mathcal{T}}$ be a stochastic process. Then the distribution $\mathbb{P}_X \triangleq \mathbb{P}[X \in \cdot]$ on $(S^{\mathcal{T}}, \mathfrak{S}^{\mathcal{T}})$ of the path-valued random variable*

$$X : \Omega \to S^{\mathcal{T}}, \qquad \omega \mapsto X(\cdot, \omega) = \{X(t, \omega)\}_{t \in \mathcal{T}},$$

*is called the **distribution** of X.* ◇

The distribution of a stochastic process $X = \{X(t)\}_{t \in \mathcal{T}}$ can equivalently be described by the family of all joint distributions of the process evaluated at finitely many time points.

Definition 1.3 (Finite-Dimensional Distributions). *Let $X = \{X(t)\}_{t \in \mathcal{T}}$ be a stochastic process and let $\mathcal{F} = [t_1, \ldots, t_n]$ be a finite subset[1] of $\mathcal{T}$. Then we denote by $\mathbb{P}_{\mathcal{F}}$ the joint distribution of $X(t_1), \ldots, X(t_n)$, i.e.*

$$\mathbb{P}_{\mathcal{F}} \triangleq \mathbb{P}\big[\big(X(t_1), \ldots, X(t_n)\big) \in \cdot\big] \qquad \text{on } \mathfrak{S}^{\mathcal{F}} = \bigotimes_{k=1\ldots,n} \mathfrak{S}.$$

*We call the family $\{\mathbb{P}_{\mathcal{F}} : \mathcal{F} = [t_1, \ldots, t_n] \subset \mathcal{T}, n \in \mathbb{N}\}$ the **finite-dimensional distributions** of the stochastic process X.* ◇

The finite dimensional distributions of a stochastic process **characterize** its distribution in the following sense: two stochastic processes $X = \{X(t)\}_{t \in \mathcal{T}}$

[1] $[t_1, \ldots, t_n]$ denotes the set $\{t_1, \ldots, t_n\}$ with the convention that $t_1 < \ldots < t_n$.

and $Y = \{Y(t)\}_{t\in\mathcal{T}}$ have the same finite-dimensional distributions if and only if X and Y have the same distribution. This is a simple consequence of the fact that the product σ-field $\mathfrak{S}^{\mathcal{T}}$ is generated by the π-system of all finite-dimensional cylinder subsets of $S^{\mathcal{T}}$, i.e. all sets of the form

$$\bigtimes_{t\in\mathcal{T}} B_t, \qquad \text{where } B_t \in \mathfrak{S} \text{ for all } t \in \mathcal{T}$$
$$\text{and } B_t \neq S \text{ for at most finitely many } t \in \mathcal{T}.$$

Since two probability measures coincide if and only if they coincide on a π-system generating the underlying σ-field, the statement follows. Hence, if we are only interested in distributional properties of a stochastic process, it suffices to look at the finite-dimensional distributions.

Exercise 2. *Let $X = \{X(t)\}_{t\in\mathcal{T}}$ and $Y = \{Y(t)\}_{t\in\mathcal{T}}$ be two stochastic processes. Show that X and Y have the same distribution if and only if X and Y have the same finite-dimensional distributions.* ◇

1.1 Examples of Stochastic Processes

Before we proceed any further, let us take a look at several examples of stochastic processes. We distinguish two different types: **discrete time** and **continuous time** stochastic processes.

Definition 1.4 (Discrete/Continuous Time Processes). *Let $X = \{X(t)\}_{t\in\mathcal{T}}$ be a stochastic process. If the time index set $\mathcal{T}$ is either finite or countably infinite, we say that X is a **discrete time stochastic process**. Otherwise we call X a **continuous time stochastic process**.* ◇

Our first example of a stochastic process is a rather trivial one and, as it turns out, we have encountered it many times before.

Example 1.5 (White Noise). *Let $\{Z_t\}_{t\in\mathcal{T}}$ be a family of independent and identically distributed random variables with values in $S \triangleq \mathbb{R}$. Then this family may be regarded as a stochastic process $X = \{X(t)\}_{t\in\mathcal{T}}$ by setting*

$$X(t) \triangleq Z_t \qquad \text{for all } t \in \mathcal{T}.$$

*The process X is known as the **white noise** process.* ◇

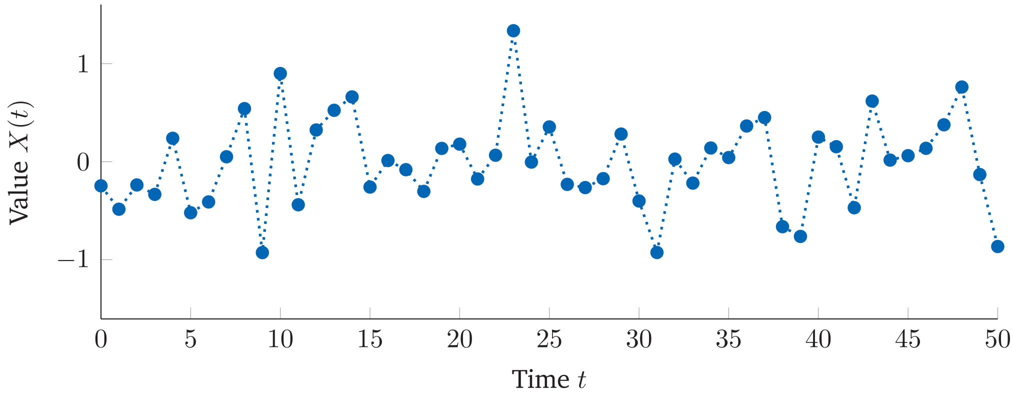

Figure 1.2 Path of a white noise process X constructed from normally distributed random variables.

Observe that the time index set $\mathcal{T}$ of the white noise process can be chosen arbitrarily (in particular, it can both be a discrete time and continuous time process), but is typically chosen to be $\mathbb{N}$, in which case the white noise process is simply a **sequence** of independent and identically distributed random variables. The white noise process has played a very important role in the course on probability theory (as it shows up, e.g., in the law of large numbers and the central limit theorem). From the perspective of stochastic processes, however, things only start getting interesting when there is dependence between the outcomes at two (or more) different time points. Thus, the white noise process will not play any major role in this course.

Example 1.6 ((Classical) Random Walk)**.** *Let $\{Z_n\}_{n\in\mathbb{N}}$ be a sequence of independent and identically distributed random variables with values in $S \triangleq \mathbb{R}$. Choose the time index set as $\mathcal{T} = \mathbb{N}_0$ and define $X = \{X(t)\}_{t\in\mathbb{N}_0}$ by*

$$X(t) \triangleq \sum_{n=1}^{t} Z_n \qquad \text{for all } t \in \mathbb{N}_0.$$

*In this definition, we use the convention that $\sum_{n=1}^{0} \triangleq 0$, i.e. $X(0) \triangleq 0$. The process X is called a **random walk**. If, in addition, it holds that*

$$\mathbb{P}[Z_n = 1] = \mathbb{P}[Z_n = -1] = \tfrac{1}{2} \qquad \text{for all } n \in \mathbb{N},$$

*then X is called a **classical random walk**.* ◇

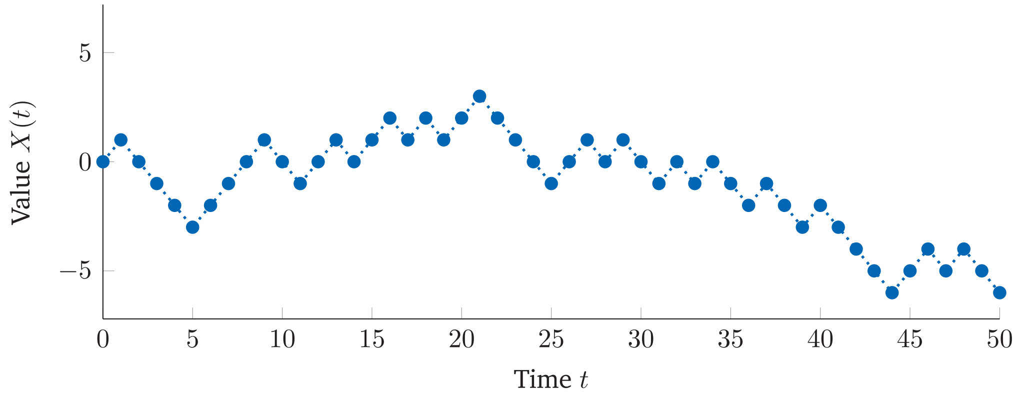

Figure 1.3 Path of a classical random walk X.

Quite obviously, the random walk exhibits dependence between different time points, i.e. unless the random variables $\{Z_n\}_{n\in\mathbb{N}}$ are almost surely constant, the random variables $X(s)$ and $X(t)$ for $s, t \in \mathbb{N}_0$ with $s \neq t$ are **not independent**. The dependence structure of the random walk is, however, relatively simple: It has **independent** and **stationary increments**.

Definition 1.7 (Independent/Stationary Increments). *Let $X = \{X(t)\}_{t\in\mathcal{T}}$ be a stochastic process. We say that X has **independent increments** if*

$$X(t) - X(s) \text{ is independent of } \sigma\big(X(r) : r \in \mathcal{T}, r \leq s\big) \text{ for all } s, t \in \mathcal{T}, s < t.$$

*Moreover, if the time index set is closed under addition, i.e. $s + t \in \mathcal{T}$ whenever $s, t \in \mathcal{T}$, we say that X has **stationary increments** if*

$$X(t_2 + s) - X(t_1 + s) \sim X(t_2) - X(t_1) \quad \text{for all } s, t_1, t_2 \in \mathcal{T}, t_2 > t_1.$$ ◇

Exercise 3. *Let $X = \{X(t)\}_{t\in\mathcal{T}}$ be a stochastic process. Show that X has independent increments if and only if $X(t_n)-X(t_{n-1}), \ldots, X(t_1)-X(t_0), X(t_0)$ are independent for all $[t_0, t_1, \ldots, t_n] \subset \mathcal{T}$, $n \in \mathbb{N}$.* ◇

It is not difficult to verify that the random walk has independent and stationary increments. As a matter of fact, it is the **only** stochastic process on $\mathcal{T} = \mathbb{N}_0$ with $X(0) = 0$ satisfying this property.

Exercise 4. *Let $X = \{X(t)\}_{t\in\mathbb{N}_0}$ be a stochastic process. Show that X is a random walk if and only if $X(0) = 0$ and X has independent and stationary increments.* ◇

Example 1.8 (Markov Chain). *As before, let $\{Z_n\}_{n\in\mathbb{N}}$ be a sequence of independent and identically distributed $\mathbb{R}$-valued random variables. Suppose that*[2]

$$g : \mathbb{R} \times S \to S \qquad \text{is } \mathfrak{B}(\mathbb{R}) \otimes \mathfrak{S}\text{-}\mathfrak{S}\text{-measurable.}$$

Given an S-valued random variable $X(0)$ which is independent of $\{Z_n\}_{n\in\mathbb{N}}$, we construct a stochastic process $X = \{X(t)\}_{t\in\mathbb{N}_0}$ recursively by setting

$$X(t) \triangleq g\big(Z_t, X(t-1)\big) \qquad \text{for all } t \in \mathbb{N}.$$

The process X is called a ***Markov chain****.* ◇

Observe that we have already encountered an example of a Markov chain: The random walk. Indeed, the random walk X can be constructed as in Example 1.8 by setting $X(0) \triangleq 0$ and

$$g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \qquad g(z, x) \triangleq z + x.$$

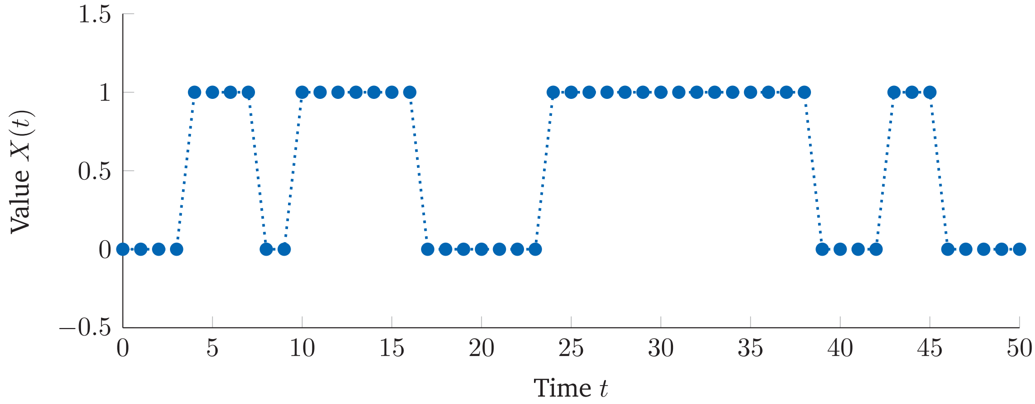

Figure 1.4 Path of a Markov chain X.

Figure 1.4 shows another example of a Markov chain in which $X(0)$ is Bernoulli distributed with $\mathbb{P}[X(0) = 0] = \mathbb{P}[X(0) = 1] = 1/2$, each Z_n is uniformly distributed on $[0, 1]$, and, for $p \in [0, 1]$, the mapping g is chosen as

$$g : \mathbb{R} \times \{0, 1\} \to \{0, 1\}, \qquad g(z, x) \triangleq \begin{cases} x & \text{if } z < p, \\ 1 - x & \text{if } z \geq p. \end{cases}$$

[2] Here and subsequently, $\mathfrak{B}(E)$ always denotes the Borel σ-field on a topological space E, i.e. the smallest σ-field containing the open sets in E.

Exercise 5. *Show that the Markov chain X depicted in Figure 1.4 is stationary, i.e. show that for every $s \in \mathbb{N}$, the process $X(\cdot + s) = \{X(t+s)\}_{t \in \mathbb{N}_0}$ has the same distribution as X.* ◇

Let us conclude this section with an example of a continuous time process: The **renewal process**.

Example 1.9 (Renewal Process). *Let $\{Z_n\}_{n \in \mathbb{N}}$ be a sequence of independent and identically distributed random variables taking values in $[0, \infty)$. We define a stochastic process $X = \{X(t)\}_{t \in [0,\infty)}$ by setting*

$$X(t) \triangleq \sup\{n \in \mathbb{N} : \textstyle\sum_{k=1}^{n} Z_k \leq t\} \qquad \textit{for all } t \in [0, \infty).$$

Then X is called a ***renewal process****.* ◇

The name renewal process stems from the following interpretation: Thinking of Z_n as the lifetime of the n-th replacement of some component (say, a light bulb), $X(t)$ is the number of replacements by time t. Every time a component breaks down and has to be renewed, X jumps by one.

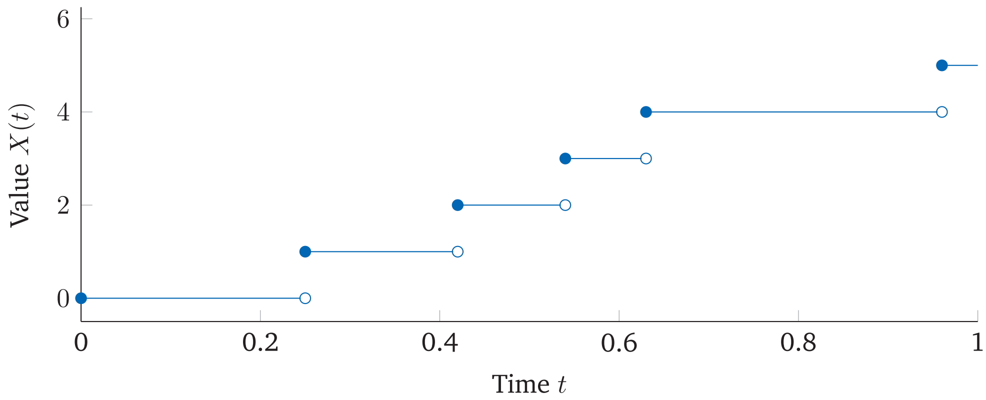

Figure 1.5 Path of a renewal process X.

It should be noted that the renewal process is well defined since the supremum in the definition of $X(t)$ is only taken over a countable set and hence $X(t)$ is indeed a random variable. Moreover, X takes values in $S = \mathbb{N}_0 \cup \{\infty\}$ and each path of X is **nondecreasing** and **right continuous**. A process with these properties is also referred to as a **counting process**.

Definition 1.10 (Path Properties of a Process). *Let $X = \{X(t)\}_{t\in\mathcal{T}}$ be a stochastic process and let ♣ be a property of functions $f : \mathcal{T} \to S$. We say that* ***X is ♣*** *if each path of X is ♣, i.e.*

$$X(\cdot,\omega) : \mathcal{T} \to S, \qquad t \mapsto X(t,\omega) \qquad \text{is ♣ for all } \omega \in \Omega. \qquad \diamond$$

This rather generic definition needs some explanation: An example of ♣ could for example be *right continuous*. Thus a process X is said to be right continuous if **each path** is right continuous. Thus, e.g., the renewal process is **nondecreasing** and **right continuous**.

Observe furthermore that the paths of the renewal process have **existing left limits**. We recall that if S is a metric space (with metric d), then a function $f : \mathcal{T} \to S$ is said to have existing left limits if

$$f(t-) \triangleq \lim_{s\uparrow t} f(s) \text{ exists for all } t \in \mathcal{T}.$$

If f has existing left limits and is furthermore right continuous, i.e.

$$f(t) = \lim_{s\downarrow t} f(s) \qquad \text{for all } t \in \mathcal{T},$$

we say that the function f is **càdlàg** (an acronym from the French *continue à droite, limite à gauche*). Finally, we can associate with f the **total variation** $V_f : \mathcal{T} \to [0,\infty]$ defined for each $t \in \mathcal{T}$ as

$$V_f(t) \triangleq \sup\Big\{\textstyle\sum_{k=1}^n d\big(f(t_{k-1}), f(t_k)\big) : [t_0,\dots,t_n] \subset \mathcal{T}, n \in \mathbb{N}\Big\}$$

We say that f is of **finite variation** if $V_f(t) < \infty$ for all $t \in \mathcal{T}$.

1.2 Filtrations, Stopping Times, and Hitting Times

We now turn to the mathematical concepts relating to the **flow of information** associated with a stochastic process and to the issue of **stopping** of stochastic processes.

Consider the following **example**: Suppose there is a group of criminals that regularly sets out to rob banks. We model their success by some $\{0,1\}$-valued stochastic process $X = \{X(t)\}_{t\in\mathbb{N}}$, where $X(t) = 1$ means that the criminals were successful and did not get caught during the t^{th} bank robbery whereas $X(t) = 0$ means that the group did get caught. What kind of information can we associate with this process and how do these information evolve with time? Clearly, this depends. A person who follows the news regularly will know at any time $t \in \mathbb{N}$ which ones of the previous robberies where successful, i.e. this person knows the outcomes of $X(1), \ldots, X(t)$ at any time $t \in \mathbb{N}$. On the other hand, a person who does not follow the news and is not aware of the existence of the group of criminals cannot see the process X at all. A witness of the 5^{th} bank robbery, who was previously unaware of the criminals, clearly knows the outcome of $X(5)$ at time $t = 5$, but does not know $X(1), \ldots, X(4)$. If after the 9^{th} robbery the police sets out a trap which guarantees that the criminals will get caught on their next attempt, then the police officers know $X(1), \ldots, X(10)$ at time $t = 9$, whereas the criminals only know $X(1), \ldots, X(9)$ at this point in time. Finally, at each time $t \in \mathbb{N}$, the criminals will have additional information on the next bank robbery at time $t + 1$ such as location of the bank, calender date of the robbery, number of security guards in the bank, and so on, whereas this information becomes available to the police only at time $t + 1$.

How can we **formalize this mathematically**? What does it mean to know the outcome of a random variable? It means that for every event involving the random variable, we can decide whether the event has occurred or not. In other words: Knowing the random variable means knowing the σ-field generated by this random variable. In the above example, the regular news observer knows $\sigma(X(1), \ldots, X(5))$ at time $t = 5$, whereas the eye witness only knows $\sigma(X(5))$ at time $t = 5$, and the person not following the news merely knows $\sigma(\emptyset)$. As time evolves, the regular news observer will gather new information, i.e. his or her σ-field increases with time. The flow of information available to an observer can therefore be modeled as a nondecreasing family of σ-fields, indexed by the time index set $\mathcal{T}$. Such a family of σ-fields is called a **filtration**.

Definition 1.11 (Filtration; Filtered Probability Space). *Let* $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$ *be a family of sub-*σ*-fields of* $\mathfrak{A}$ *satisfying*

$$\mathfrak{F}(s) \subset \mathfrak{F}(t) \qquad \textit{for all } s,t \in \mathcal{T} \textit{ with } s \le t.$$

*Then $\mathfrak{F}$ is called a **filtration** of $(\Omega, \mathfrak{A})$. The quadruple $(\Omega, \mathfrak{A}, \mathfrak{F}, \mathbb{P})$ is called a **filtered probability space**.* ◇

The filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$ represents the **flow of information** and has (for now) nothing to do with any stochastic processes. $\mathfrak{F}(t)$ is simply the information available at time $t \in \mathcal{T}$. In the bank robbery example, the filtration of the news observer is given by $\mathfrak{F}_1 = \{\mathfrak{F}_1(t)\}_{t\in\mathbb{N}}$ with

$$\mathfrak{F}_1(t) = \sigma\big(X(1), \ldots, X(t)\big) \qquad \text{for all } t \in \mathbb{N}.$$

On the other hand, the person who does not follow the news has the filtration $\mathfrak{F}_2 = \{\mathfrak{F}_2(t)\}_{t\in\mathbb{N}}$ given by

$$\mathfrak{F}_2(t) = \sigma(\emptyset) = \big\{\emptyset, \Omega\big\} \qquad \text{for all } t \in \mathbb{N}.$$

Moreover, the witness of the 5^{th} robbery (who does not follow the news) has the filtration $\mathfrak{F}_3 = \{\mathfrak{F}_3(t)\}_{t\in\mathbb{N}}$ with

$$\mathfrak{F}_3(t) = \begin{cases} \sigma(\emptyset) & \text{if } t < 5 \\ \sigma(X(5)) & \text{if } t \geq 5 \end{cases} \qquad \text{for all } t \in \mathbb{N}.$$

The criminals themselves have a filtration $\mathfrak{F}_4 = \{\mathfrak{F}_4(t)\}_{t\in\mathcal{T}}$ which potentially contains more information than the filtration of the news observer, i.e.

$$\mathfrak{F}_4(t) \supset \mathfrak{F}_3(t) \qquad \text{for all } t \in \mathbb{N}.$$

Finally, being all-knowing, God has the filtration $\mathfrak{F}_5 = \{\mathfrak{F}_5(t)\}_{t\in\mathcal{T}}$ given by

$$\mathfrak{F}_5(t) = \mathfrak{A} \qquad \text{for all } t \in \mathbb{N}.$$

How does this **link to stochastic processes**? Recall that knowing a random variable means knowing the σ-field it generates. Hence knowing a stochastic process $X = \{X(t)\}_{t\in\mathcal{T}}$ means knowing $X(t)$ at any time $t \in \mathcal{T}$. Since our flow of information is modeled by a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$, this is to say that $\sigma(X(t))$ is a subset of $\mathfrak{F}(t)$, i.e. $X(t)$ is $\mathfrak{F}(t)$-measurable for all $t \in \mathcal{T}$.

Definition 1.12 (Adaptedness)**.** *Let $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$ be a filtration. Then we say that a stochastic process $X = \{X(t)\}_{t\in\mathcal{T}}$ is **adapted** to the filtration $\mathfrak{F}$ if*

$$X(t) \text{ is } \mathfrak{F}(t)\text{-measurable for all } t \in \mathcal{T}.$$ ◇

Continuing with our example, the process X modeling the success of the bank robbers is adapted to the filtrations $\mathfrak{F}_1$, $\mathfrak{F}_4$, and $\mathfrak{F}_5$, but not to $\mathfrak{F}_2$ and $\mathfrak{F}_3$.[3] The filtration $\mathfrak{F}_1$ in this example plays a special role, since it contains just enough information to make the process X adapted, i.e. it contains just the bare minimum of information to know X.

Definition 1.13 (Natural Filtration). *Let $X = \{X(t)\}_{t\in\mathcal{T}}$ be a stochastic process and consider the filtration $\mathfrak{F}^X = \{\mathfrak{F}^X(t)\}_{t\in\mathcal{T}}$ given by*

$$\mathfrak{F}^X(t) \triangleq \sigma\big(X(s) : s \in \mathcal{T}, s \leq t\big) \qquad \text{for all } t \in \mathcal{T}.$$

Then $\mathfrak{F}^X$ is called the ***natural filtration*** *of X.* ◇

Clearly, every stochastic process X is adapted to its natural filtration $\mathfrak{F}^X$. It is furthermore easy to see that $\mathfrak{F}^X$ is the smallest filtration X is adapted to.

Exercise 6. *Let $X = \{X(t)\}_{t\in\mathcal{T}}$ be a stochastic process and let $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$ be a filtration. Show that X is adapted to $\mathfrak{F}$ if and only if*

$$\mathfrak{F}^X(t) \subset \mathfrak{F}(t) \qquad \text{for all } t \in \mathcal{T}.$$ ◇

Exercise 7. *Let $Z = \{Z(t)\}_{t\in\mathbb{N}}$ be a white noise process and let $X = \{X(t)\}_{t\in\mathbb{N}_0}$ be the corresponding random walk given by*

$$X(t) \triangleq \textstyle\sum_{n=1}^{t} Z(n), \qquad t \in \mathbb{N}_0.$$

Show that X is $\mathfrak{F}^Z$-adapted, where $\mathfrak{F}^Z(0) \triangleq \{\emptyset, \Omega\}$ and $\{\mathfrak{F}^Z(t)\}_{t\in\mathbb{N}}$ denotes the natural filtration of Z. ◇

Suppose that the group of criminals decides at some point τ to stop their career as bank robbers. Clearly, τ can both be **deterministic** (e.g. *'Stop after the third robbery.'*, i.e. $\tau = 3$) or **random** (e.g. *'Stop after three successful robberies in a row.'*, i.e. $\tau = \inf\{t \geq 3 : X(t) = X(t-1) = X(t-2) = 1\}$). It should furthermore be possible for the criminals to decide never to stop, i.e. $\tau = \infty$. On the other hand, we would like to rule out times which look into the future such as stopping exactly one robbery before the first unsuccessful one: $\tau = \inf\{t \in \mathbb{N} : X(t+1) = 0\}$. These ideas are formalized in the notion of a **stopping time**.

[3] Unless, of course, X is deterministic.

Definition 1.14 (Stopping Time). *Let $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$ be a filtration and let $\tau : \Omega \to \mathcal{T} \cup \{\infty\}$ be a mapping such that*

$$\{\tau \leq t\} \in \mathfrak{F}(t) \qquad \text{for all } t \in \mathcal{T}.$$

Then τ is called a ***stopping time*** *with respect to $\mathfrak{F}$.* ◇

The notion of a stopping time does exactly the job that we hoped for. It can be both deterministic and random and it takes values in the time index set but also allows the value infinity. Moreover, since the filtration $\mathfrak{F}$ models the information available at any point in time, the condition

$$\{\tau \leq t\} \in \mathfrak{F}(t) \qquad \text{for all } t \in \mathcal{T}$$

simply means that at any time t, we can decide if τ has already occurred or not. E.g., in the bank robbery example, the stopping time $\tau = \inf\{t \in \mathbb{N} : X(t+1) = 0\}$ is in general not a stopping time with respect to the natural filtration $\mathfrak{F}^X$ since τ requires the knowledge of $X(t+1)$ at time t.

Exercise 8 (Galmarino's Test). *Suppose that $\Omega = \mathbb{R}^{\mathcal{T}} = \{\omega : \mathcal{T} \to \mathbb{R}\}$ is the set of all functions from $\mathcal{T}$ to $\mathbb{R}$ (such that, in particular, the stopped path $\omega(\cdot \wedge t)$ is a member of Ω for all $\omega \in \Omega$ and $t \in \mathcal{T}$). Denote by $X = \{X(t)\}_{t\in\mathcal{T}}$ the so-called* ***canonical process*** *given by*

$$X(t,\omega) \triangleq \omega(t), \qquad t \in \mathcal{T}, \omega \in \Omega.$$

Denote by $\mathfrak{F}^X = \{\mathfrak{F}^X(t)\}_{t\in\mathcal{T}}$ the natural filtration of X and define $\mathfrak{F}^X(\infty) \triangleq \sigma(\mathfrak{F}^X(t) : t \in \mathcal{T})$. Show that the following statements hold:

(a) Let $F \subset \Omega$ and $t \in \mathcal{T}$. Then $F \in \mathfrak{F}^X(t)$ if and only if $F \in \mathfrak{F}^X(\infty)$ and the following implication holds:

$$\text{if } \omega \in F \text{ and } X(s,\omega) = X(s,\bar{\omega}) \text{ for all } s \in \mathcal{T} \text{ with } s \leq t, \text{ then } \bar{\omega} \in F.$$

(b) A mapping $\tau : \Omega \to \mathcal{T} \cup \{\infty\}$ is a stopping time with respect to $\mathfrak{F}^X$ if and only if τ is an $\mathfrak{F}^X(\infty)$-measurable random variable and the following implication holds for all $t \in \mathcal{T}$:

$$\text{if } \tau(\omega) \leq t \text{ and } X(s,\omega) = X(s,\bar{\omega}) \text{ for all } s \in \mathcal{T} \text{ with } s \leq t, \text{ then } \tau(\bar{\omega}) \leq t.$$

◇

We make two more important observations. First, observe that the property of being a stopping time **depends on the filtration**. While we have already convinced ourselves that $\tau = \inf\{t \in \mathbb{N} : X(t+1) = 0\}$ is in general not a stopping time with respect to $\mathfrak{F}^X$, it is a stopping time with respect to the God filtration $\mathfrak{F}_5$ (which was given by $\mathfrak{F}_5(t) = \mathfrak{A}$, $t \in \mathcal{T}$). Second, observe that we do not assume that τ is a random variable. It is however straightforward to verify that the condition $\{\tau \leq t\} \in \mathfrak{F}(t)$ for all $t \in \mathcal{T}$ implies that **every stopping time is a random variable**. In the case when τ takes at most countably many values (which holds, in particular, if $\mathcal{T}$ is at most countable), there is a weaker condition to check if a random time is a stopping time.

Lemma 1.15 (Characterization of Discrete Stopping Times). *Let $\tau : \Omega \to \mathcal{T} \cup \{\infty\}$ be such that $\tau(\Omega) \triangleq \{\tau(\omega) : \omega \in \Omega\}$ is at most countable and let $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$ be a filtration. Then τ is a stopping time with respect to $\mathfrak{F}$ if and only if*

$$\{\tau = t\} \in \mathfrak{F}(t) \qquad \text{for all } t \in \tau(\Omega) \text{ with } t < \infty. \qquad \diamond$$

Proof. Step 1: Suppose that $\{\tau = t\} \in \mathfrak{F}(t)$ for all $t \in \tau(\Omega)$ with $t < \infty$ and let $s \in \mathcal{T}$. We have to check that $\{\tau \leq s\} \in \mathfrak{F}(s)$, in which case we conclude that τ is a stopping time. If $s = \infty$, then $\{\tau \leq s\} = \Omega \in \mathfrak{F}(s)$, so suppose that $s < \infty$. In that case, we have

$$\{\tau \leq s\} = \bigcup\nolimits_{t\in\tau(\Omega), t\leq s} \{\tau = t\}.$$

Now for every $t \in \tau(\Omega)$ with $t \leq s$, we have $\{\tau = t\} \in \mathfrak{F}(t)$ by assumption and hence, since $t \leq s$ implies $\mathfrak{F}(t) \subset \mathfrak{F}(s)$, we find that $\{\tau = t\} \in \mathfrak{F}(s)$. But now, since $\tau(\Omega)$ is countable, this implies that $\{\tau \leq s\} \in \mathfrak{F}(s)$.

Step 2: Suppose that τ is a stopping time and let $s \in \tau(\Omega)$ with $s < \infty$. We have to show that $\{\tau = s\} \in \mathfrak{F}(s)$. For this, we first observe that, for every $t \in \mathcal{T}$ with $t \leq s$, we have $\{\tau \leq t\} \in \mathfrak{F}(t) \subset \mathfrak{F}(s)$. But then, again since $\tau(\Omega)$ is countable, we conclude since

$$\{\tau = s\} = \{\tau \leq s\} \setminus \bigcup\nolimits_{t\in\tau(\Omega), t<s} \{\tau \leq t\} \in \mathfrak{F}(s). \qquad \square$$

It is of course immediately clear that if $\tau : \Omega \to \mathcal{T} \cup \{\infty\}$ is constant, i.e. $\tau \equiv t$ for some $t \in \mathcal{T}$, then τ is a stopping time. By the previous lemma, we only have to verify that $\{\tau = t\} \in \mathfrak{F}(t)$, which is obvious since $\{\tau = t\} = \Omega$. Hence **all deterministic times are stopping times**. The property of being a stopping time is furthermore stable under several operations, for example taking pointwise minima and maxima.

Lemma 1.16 (Operations on Stopping Times). *Let τ and σ be two stopping times with respect to a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$. Then*

$$\tau \wedge \sigma : \Omega \to \mathcal{T} \cup \{\infty\}, \qquad \tau \wedge \sigma(\omega) \triangleq \min\{\tau(\omega), \sigma(\omega)\},$$
$$\tau \vee \sigma : \Omega \to \mathcal{T} \cup \{\infty\}, \qquad \tau \vee \sigma(\omega) \triangleq \max\{\tau(\omega), \sigma(\omega)\},$$

are stopping times with respect to $\mathfrak{F}$ as well. ◇

Proof. For all $t \in \mathcal{T}$ it holds that

$$\{\tau \wedge \sigma \le t\} = \{\tau \le t\} \cup \{\sigma \le t\} \in \mathfrak{F}(t)$$

as well as

$$\{\tau \vee \sigma \le t\} = \{\tau \le t\} \cap \{\sigma \le t\} \in \mathfrak{F}(t).$$

Thus both $\tau \wedge \sigma$ and $\tau \vee \sigma$ are stopping times. □

Exercise 9. *Let $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in[0,\infty)}$ be a filtration and let $\{\tau_n\}_{n\in\mathbb{N}}$ be a sequence of stopping times. Show that*

$$\tau : \Omega \to [0, \infty], \qquad \tau(\omega) \triangleq \sup_{n\in\mathbb{N}} \tau_n(\omega)$$

is a stopping time. ◇

Exercise 10. *Let $\tau : \Omega \to \mathcal{T} \cup \{\infty\}$ be a function and fix a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$. Show that τ is a stopping time if and only if the indicator process $X = \{X(t)\}_{t\in\mathcal{T}}$ given by*

$$X(t) \triangleq \mathbb{1}_{\{\tau\le t\}}, \qquad t \in \mathcal{T}$$

is adapted to $\mathfrak{F}$. ◇

A natural question arising now is to ask **which information are available** up to the stopping time τ. More precisely, if $F \in \mathfrak{A}$ is some event, we would like to know if F is known at time τ. In other words, at any given time

$t \in \mathcal{T}$, we would like to know if F is known at time t, provided of course that τ has already occurred. Mathematically, this means that F is known at time τ if and only if $F \cap \{\tau \leq t\} \in \mathfrak{F}(t)$ for all $t \in \mathcal{T}$.

Definition 1.17 (σ-Field of the τ-Past)**.** *Let τ be a stopping time with respect to a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t \in \mathcal{T}}$. Then*

$$\mathfrak{F}(\tau) \triangleq \big\{F \in \mathfrak{A} : F \cap \{\tau \leq t\} \in \mathfrak{F}(t) \textit{ for all } t \in \mathcal{T}\big\}$$

*is called the **σ-field of the τ-past**.* ◇

We have to be a bit careful with the previous definition as it is quite suggestive. First, the name '*σ-field of the τ-past*' suggests that $\mathfrak{F}(\tau)$ is a σ-field. This is indeed true and can easily be verified. Moreover, the notation $\mathfrak{F}(\tau)$ suggests some sort of **compatibility with the filtration** $\mathfrak{F}$. Of course, we should **never ever** think of $\mathfrak{F}(\tau)$ as a mapping $\omega \mapsto \mathfrak{F}(\tau(\omega))$. Nevertheless, $\mathfrak{F}(\tau)$ is compatible with the filtration in that $\mathfrak{F}(\tau) = \mathfrak{F}(t)$ if τ is constant equal to t.

Exercise 11. *Let $\tau : \Omega \to \mathcal{T} \cup \{\infty\}$ be a stopping time with respect to a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t \in \mathcal{T}}$. Show that $\mathfrak{F}(\tau)$ is a sub-σ-field of $\mathfrak{A}$. Moreover, if for some $t \in \mathcal{T}$ we have $\tau(\omega) = t$ for all $\omega \in \Omega$, show that $\mathfrak{F}(\tau) = \mathfrak{F}(t)$.* ◇

Another property of the σ-field of the τ-past which justifies the notation $\mathfrak{F}(\tau)$ is that the mapping $\tau \mapsto \mathfrak{F}(\tau)$ respects the order on $\mathfrak{F}$ induced by the order on $\mathcal{T}$, i.e. if σ and τ are two stopping times with $\sigma \leq \tau$, then $\mathfrak{F}(\sigma) \subset \mathfrak{F}(\tau)$.

Lemma 1.18 (Properties of the σ-Field of the τ-Past)**.** *Let σ and τ be stopping times with respect to a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t \in \mathcal{T}}$. Then*

$$\mathfrak{F}(\tau \wedge \sigma) = \mathfrak{F}(\tau) \cap \mathfrak{F}(\sigma).$$

In particular, if $\sigma \leq \tau$, then $\mathfrak{F}(\sigma) \subset \mathfrak{F}(\tau)$. Finally, we have

$$\{\sigma \leq \tau\}, \{\sigma < \tau\}, \{\sigma = \tau\} \in \mathfrak{F}(\tau \wedge \sigma).$$ ◇

Proof. Step 1: We show that $\mathfrak{F}(\tau \wedge \sigma) \subset \mathfrak{F}(\tau) \cap \mathfrak{F}(\sigma)$. Let $F \in \mathfrak{F}(\tau \wedge \sigma)$ and $t \in \mathcal{T}$. Since $\{\sigma \leq t\} = \{\sigma \wedge \tau \leq t\} \cap \{\sigma \leq t\}$, it follows that

$$F \cap \{\sigma \leq t\} = \big(F \cap \{\sigma \wedge \tau \leq t\}\big) \cap \{\sigma \leq t\}.$$

As $F \in \mathfrak{F}(\tau \wedge \sigma)$, we have $F \cap \{\sigma \wedge \tau \leq t\} \in \mathfrak{F}(t)$, and, since σ is a stopping time, we see moreover that $\{\sigma \leq t\} \in \mathfrak{F}(t)$. Thus $F \cap \{\sigma \leq t\} \in \mathfrak{F}(t)$, implying that $F \in \mathfrak{F}(\sigma)$ since $t \in \mathcal{T}$ was chosen arbitrarily. Interchanging the roles of τ and σ furthermore yields $F \in \mathfrak{F}(\tau)$, i.e. $F \in \mathfrak{F}(\tau) \cap \mathfrak{F}(\sigma)$. This in turn implies $\mathfrak{F}(\tau \wedge \sigma) \subset \mathfrak{F}(\tau) \cap \mathfrak{F}(\sigma)$.

Step 2: We show that $\mathfrak{F}(\tau) \cap \mathfrak{F}(\sigma) \subset \mathfrak{F}(\tau \wedge \sigma)$. Let $F \in \mathfrak{F}(\tau) \cap \mathfrak{F}(\sigma)$ and $t \in \mathcal{T}$. Using that $\{\tau \wedge \sigma \leq t\} = \{\tau \leq t\} \cup \{\sigma \leq t\}$ and $F \in \mathfrak{F}(\tau)$ and $F \in \mathfrak{F}(\sigma)$, we find that

$$F \cap \{\tau \wedge \sigma \leq t\} = \big(F \cap \{\tau \leq t\}\big) \cup \big(F \cap \{\sigma \leq t\}\big) \in \mathfrak{F}(t).$$

However, this implies that $F \in \mathfrak{F}(\tau \wedge \sigma)$, i.e. $\mathfrak{F}(\tau) \cap \mathfrak{F}(\sigma) \subset \mathfrak{F}(\tau \wedge \sigma)$.

Step 3: We are left with showing that $\{\sigma < \tau\} \in \mathfrak{F}(\tau \wedge \sigma)$ since, from this, it follows that

$$\{\sigma \leq \tau\} = \{\tau < \sigma\}^c \in \mathfrak{F}(\tau \wedge \sigma) \text{ and } \{\sigma = \tau\} = \{\sigma \leq \tau\} \setminus \{\sigma < \tau\} \in \mathfrak{F}(\tau \wedge \sigma).$$

To see that $\{\sigma < \tau\} \in \mathfrak{F}(\tau \wedge \sigma)$, let $t \in \mathcal{T}$ be arbitrary and denote by $\mathcal{Q}$ a countable and dense subset of $\mathcal{T}$ with $t \in \mathcal{Q}$. Then it holds that

$$\begin{aligned} \{\sigma < \tau\} \cap \{\tau \leq t\} &= \textstyle\bigcup_{q \in \mathcal{Q}, q < t} \{\sigma \leq q < \tau \leq t\} \\ &= \textstyle\bigcup_{q \in \mathcal{Q}, q < t} \{\sigma \leq q\} \cap \big(\{\tau \leq t\} \setminus \{\tau \leq q\}\big). \end{aligned}$$

Since $\mathcal{Q}$ is countable, we must have $\{\sigma < \tau\} \cap \{\tau \leq t\} \in \mathfrak{F}(t)$, which implies that $\{\sigma < \tau\} \in \mathfrak{F}(\tau)$. Conversely, we have

$$\{\sigma < \tau\} \cap \{\sigma \leq t\} = \textstyle\bigcup_{q \in \mathcal{Q}, q \leq t} \{\sigma \leq q < \tau\} = \bigcup_{q \in \mathcal{Q}, q \leq t} \{\sigma \leq q\} \cap \{\tau \leq q\}^c.$$

By countability of $\mathcal{Q}$, this implies that $\{\sigma < \tau\} \cap \{\sigma \leq t\} \in \mathfrak{F}(t)$ and hence $\{\sigma < \tau\} \in \mathfrak{F}(\sigma)$. But then $\{\sigma < \tau\} \in \mathfrak{F}(\tau) \cap \mathfrak{F}(\sigma) \subset \mathfrak{F}(\tau \wedge \sigma)$ by step 2. □

Recalling that each deterministic time $t \in \mathcal{T}$ is also a stopping time and using that $\mathfrak{F}(\tau \wedge t) \subset \mathfrak{F}(\tau)$ and $\mathfrak{F}(\tau \wedge t) \subset \mathfrak{F}(t)$, Lemma 1.18 implies that the events $\{\tau \leq t\}$, $\{\tau < t\}$, $\{\tau = t\}$ are contained in both $\mathfrak{F}(\tau)$ and $\mathfrak{F}(t)$.

Exercise 12. *Let τ and σ be two stopping times with respect to a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t \in \mathcal{T}}$. Show that*

$$\mathfrak{F}(\tau \vee \sigma) = \sigma\big(\mathfrak{F}(\tau), \mathfrak{F}(\sigma)\big).$$ ◇

Exercise 13. *Let τ and σ be two stopping times with respect to a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$ and let X be a random variable with $\mathrm{E}[|X|] < \infty$. Show that*

$$\mathbb{1}_{\{\tau=\sigma\}}\mathrm{E}\big[X\big|\mathfrak{F}(\tau)\big] = \mathbb{1}_{\{\tau=\sigma\}}\mathrm{E}\big[X\big|\mathfrak{F}(\sigma)\big] \text{ almost surely.}$$ ◇

Thus far, we have not seen any nontrivial example of a stopping time and we have not drawn any connection to stochastic processes yet. This is about to change: The prime example of a stopping time, which we shall encounter over and over again, is the first time that a stochastic process enters into a measurable subset of its state space.

Definition 1.19 (Hitting Time)**.** *Let $X = \{X(t)\}_{t\in\mathcal{T}}$ be a stochastic process and let $B \in \mathfrak{S}$. Then the mapping*

$$\tau_B : \Omega \to \mathcal{T} \cup \{\infty\}, \qquad \tau_B(\omega) \triangleq \inf\big\{t \in \mathcal{T} : X(t,\omega) \in B\big\},$$

is called the (first) ***hitting time*** *of B by X.* ◇

A priori, it is **not clear** if every hitting time is a stopping time with respect to a filtration $\mathfrak{F}$, even if we assume that X is adapted to $\mathfrak{F}$. This turns out to be a very deep question which goes far beyond the scope of this course (and is in general not true). Under some additional assumptions on X, $\mathfrak{F}$, and/or B, we can however give a positive answer to this question, and this will be sufficient for our purposes. The first, most basic, case is when X is a discrete time process.

Lemma 1.20 (Hitting Times of Discrete Time Processes)**.** *Let $X = \{X(t)\}_{t\in\mathcal{T}}$ be a discrete time stochastic process adapted to a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$. Let moreover $B \in \mathfrak{S}$ and assume that*

$$\text{for each } t \in \mathcal{T}, \text{ the set } \{s \in \mathcal{T} : s \leq t\} \text{ is finite.}$$

Then the first hitting time τ_B of B by X is a stopping time. Similarly, if σ is a stopping time, then the first hitting time τ_B^σ of B by X after σ given by

$$\tau_B^\sigma \triangleq \inf\big\{t \in \mathcal{T} \cap (\sigma,\infty) : X(t) \in B\big\}$$

is a stopping time as well. ◇

Proof. The condition that $\{s \in \mathcal{T} : s \leq t\}$ is finite for all $t \in \mathcal{T}$ implies that, for all $\omega \in \Omega$, we either have $\tau_B(\omega) = \infty$ or

$$\tau_B(\omega) = \inf\big\{t \in \mathcal{T} : X(t,\omega) \in B\big\} = \min\big\{t \in \mathcal{T} : X(t,\omega) \in B\big\} \in \mathcal{T},$$

and the same can be said about τ_B^σ. Let now $t \in \mathcal{T}$ with $t < \infty$. Then

$$\{\tau_B = t\} = \{X(t) \in B\} \cap \Big(\bigcap\nolimits_{s\in\mathcal{T}, s<t} \{X(s) \notin B\}\Big).$$

and, similarly,

$$\{\tau_B^\sigma = t\} = \{\sigma < t\} \cap \{X(t) \in B\} \cap \Big(\bigcap\nolimits_{s\in\mathcal{T}, s<t} \{X(s) \notin B\} \cup \{\sigma < s\}^c\Big).$$

Since $\mathcal{T}$ is discrete, X is adapted, and σ is a stopping time, this shows that $\{\tau_B = t\} \in \mathfrak{F}(t)$ and $\{\tau_B^\sigma = t\} \in \mathfrak{F}(t)$ for all $t \in \mathcal{T}$ with $t < \infty$, and we conclude by Lemma 1.15 (Characterization of Discrete Stopping Times). □

In Lemma 1.20, the finiteness of the sets $\{s \in \mathcal{T} : s \leq t\}$ is needed to guarantee that the hitting time τ_B takes values in $\mathcal{T} \cup \{\infty\}$ (in particular $\mathcal{T}$ has a minimal element). Indeed, if for example $\mathcal{T} = \{1/n : n \in \mathbb{N}\}$ and we choose $B = S$, then $\tau_B(\omega) = 0$ for all $\omega \in \Omega$, but $0 \notin \mathcal{T}$.

Exercise 14. *Let $X = \{X(t)\}_{t\in\mathbb{N}}$ be a white noise process and let $B \in \mathfrak{B}(\mathbb{R})$. Determine the distribution of the first hitting time τ_B of B by X and show that $\tau_B < \infty$ almost surely if and only if $\mathbb{P}[Z_1 \in B] > 0$.* ◇

For continuous time processes, the issue of whether hitting times are stopping times is more delicate. Imagine, e.g., that X is a continuous process and B is an open set. Then at time τ_B, the process X will be on the boundary of B, but not in B itself. So, whenever the process X is on the boundary of B, we need to be able to look an infinitesimal amount of time into the future for τ_B to be a stopping time (compare with Figure 1.6), since otherwise we will not be able to decide whether we have to stop or not.

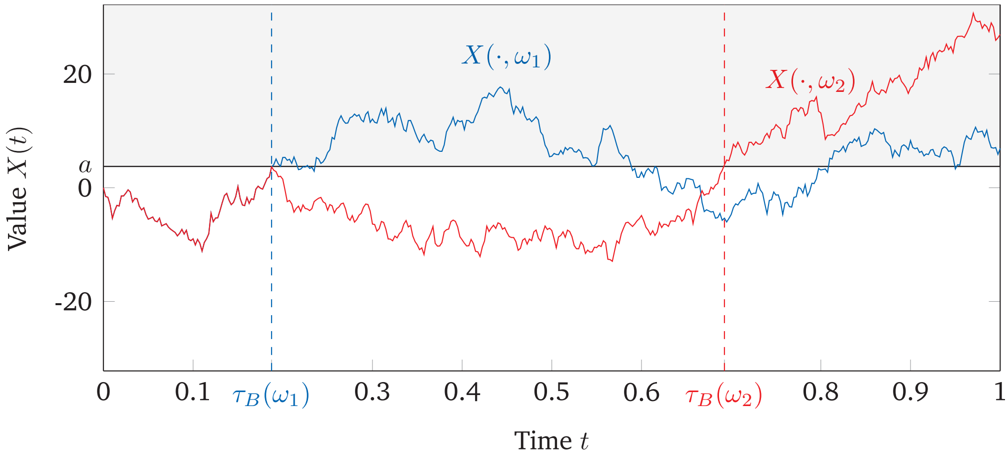

Figure 1.6 Hitting time of the open set $B = (a, \infty)$ by a continuous process X. Observe that while both paths take the value a at time $\tau_B(\omega_1)$, only the path $X(\cdot, \omega_1)$ enters B at this time, whereas $X(\cdot, \omega_2)$ moves away from B. Hence, in order to decide if we should stop at any given time, we need to be able to look an infinitesimal amount of time into the future.

Definition 1.21 (Right-Continuous Filtration). *We say that a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$ is* ***right continuous*** *if*

$$\mathfrak{F}(t+) \triangleq \bigcap_{s\in\mathcal{T}, s>t} \mathfrak{F}(s) = \mathfrak{F}(t) \qquad \text{for all } t \in \mathcal{T}. \qquad \diamond$$

Right continuity of a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$ is only interesting if the time index set is sufficiently rich. Indeed, if each $t \in \mathcal{T}$ has a distinct next element $t+1$ (which holds, e.g., for $\mathcal{T} = \mathbb{N}$ but not for $\mathcal{T} = \mathbb{Q}$), then $\mathfrak{F}(t+) = \mathfrak{F}(t+1)$ for all $t \in \mathcal{T}$ and hence the notion of right continuity of $\mathfrak{F}$ boils down to the requirement $\mathfrak{F}(t) = \mathfrak{F}(t+1)$ for all $t \in \mathcal{T}$.

Proposition 1.22 (Hitting Times of Continuous Time Processes). *Let $X = \{X(t)\}_{t\in[0,\infty)}$ be a stochastic process adapted to a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in[0,\infty)}$. Suppose that S is a metric space with metric d and assume that $\mathfrak{S} = \mathfrak{B}(S)$ is the Borel σ-field on S. Then the hitting time τ_B of a set $B \in \mathfrak{S}$ by X is a stopping time with respect to $\mathfrak{F}$ in each of the following cases:*

(i) X is continuous and B is closed.

(ii) X is right continuous, B is open, and $\mathfrak{F}$ is right continuous. $\diamond$

Proof. Step 1: Assume that X is continuous and B is closed. For every $t \in [0,\infty)$, we claim that

$$\{\tau_B \leq t\} = \{X(t) \in B\} \cup \Big(\bigcup_{s\in[0,t)} \{X(s) \in B\}\Big) \tag{1.1}$$

$$= \{X(t) \in B\} \cup \Big(\bigcap_{k=1}^{\infty} \bigcup_{q\in[0,t)\cap\mathbb{Q}} \big\{d\big(X(q), B\big) \leq 1/k\big\}\Big). \tag{1.2}$$

From this and the adaptedness of X, it follows that $\{\tau_B \leq t\} \in \mathfrak{F}(t)$, and hence τ_B is a stopping time. To see that this representation of $\{\tau_B \leq t\}$ indeed holds, let us first fix $\omega \in \Omega$ with $\tau_B(\omega) < \infty$. Then the continuity of X and the closedness of B imply that $X(\tau_B(\omega), \omega) \in B$. But then $\omega \in \{\tau_B \leq t\}$ implies that there exists $s \in [0,t]$ such that $X(s,\omega) \in B$. On the other hand, if $\omega \in \{X(s) \in B\}$ for some $s \in [0,t]$, then $\tau_B(\omega) \leq s \leq t$. Thus Equation (1.1) is argued for. To see that Equation (1.2) holds, let us first assume that $\omega \in \{X(s) \in B\}$ for some $s \in [0,t)$. Then the continuity of $X(\cdot,\omega)$ implies that there exists a sequence $\{q_k\}_{k\in\mathbb{N}}$ in $[0,t) \cap \mathbb{Q}$ such that $d(X(q_k,\omega), X(s,\omega)) \leq 1/k$. But then, since $X(s,\omega) \in B$, we must necessarily have $d(X(q_k,\omega), B) \leq 1/k$. This shows that the set on the right hand side of Equation (1.1) is contained in the set in Equation (1.2). For the reverse containment, assume that $\omega \in \Omega$ is such that, for each $k \in \mathbb{N}$, there exists $q_k \in [0,t) \cap \mathbb{Q}$ such that $d(X(q_k,\omega), B) \leq 1/k$. Since the sequence $\{q_k\}_{k\in\mathbb{N}} \subset [0,t]$ and $[0,t]$ is compact, we may without loss of generality assume that $s \triangleq \lim_{k\to\infty} q_k$ exists in $[0,t]$ (drop to a subsequence if $\{q_k\}_{k\in\mathbb{N}}$ does not converge). But then $X(s,\omega) \in B$ since by continuity of X we have

$$0 \leq d\big(X(s,\omega), B\big) = \lim_{k\to\infty} d\big(X(q_k,\omega), B\big) \leq \lim_{k\to\infty} 1/k = 0.$$

Step 2: Assume that X and $\mathfrak{F}$ are right continuous and B is open. Observe that, in this situation, we do not necessarily have $X(\tau_B) \in B$ on $\{\tau_B < \infty\}$, so we must be careful with events of the form $\{\tau_B = t\}$. Nevertheless, if $t \in [0,\infty)$, we have

$$\{\tau_B < t\} = \bigcup_{q\in[0,t)\cap\mathbb{Q}} \{X(q) \in B\} \in \mathfrak{F}(t). \tag{1.3}$$

Indeed, if $\omega \in \{X(q) \in B\}$ for some $q \in [0,t) \cap \mathbb{Q}$, then clearly $\tau_B(\omega) \leq q < t$ and hence $\omega \in \{\tau_B < t\}$. On the other hand, if $\omega \in \{\tau_B < t\}$, then by definition of $\tau_B(\omega)$ there exists $s \in [0,t)$ such that $X(s,\omega) \in B$. Now let $\{q_k\}_{k\in\mathbb{N}}$ be a sequence in $[0,t) \cap \mathbb{Q}$ with $q_k > s$ and $q_k \downarrow s$ as $k \to \infty$. Then $X(q_k,\omega) \to X(s,\omega)$ by the right continuity of X. But then, since B is open and $X(s,\omega) \in B$, we must have $X(q_k,\omega) \in B$ for all $k \in \mathbb{N}$ sufficiently large, i.e. $\omega \in \bigcup_{q\in[0,t)\cap\mathbb{Q}}\{X(q) \in B\}$, which establishes Equation (1.3). But then

$$\{\tau_B \leq t\} = \bigcap_{q\in(t,\infty)\cap\mathbb{Q}}\{\tau_B < q\} = \bigcap_{q\in(t,s]\cap\mathbb{Q}}\{\tau_B < q\} \in \mathfrak{F}(s) \quad \text{for all } s > t$$

since the events $\{\tau_B < q\}$ are increasing in q. But then $\{\tau_B \leq t\} \in \mathfrak{F}(t+) = \mathfrak{F}(t)$ by right continuity of $\mathfrak{F}$ and hence τ_B is a stopping time. □

The previous proposition extends immediately to processes $X = \{X(t)\}_{t\in\mathcal{T}}$ with time index set of the form $\mathcal{T} = [0,T]$ for some $T > 0$. Indeed, given X, we can simply define a new process $Y = \{Y(t)\}_{t\in[0,\infty)}$ by setting $Y(t) \triangleq X(t \wedge T)$ for all $t \in [0,\infty)$ and extend the filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in[0,T]}$ to a filtration on $[0,\infty)$ by setting $\mathfrak{F}(t) = \mathfrak{F}(T)$ for all $t \geq T$. Since X and Y coincide on $[0,T]$ and Y is constant outside of $[0,T]$, any hitting time of the process X coincides with any hitting time of the process Y. Thus Proposition 1.22 remain valid also for the process X.

Exercise 15. *Let $\tau : \Omega \to [0,\infty]$ and fix a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in[0,\infty)}$. Setting $\mathfrak{F}(\infty) \triangleq \sigma(\mathfrak{F}(t), t \in [0,\infty))$, we say that τ is an* ***optional time*** *with respect to $\mathfrak{F}$ if*

$$\{\tau < t\} \in \mathfrak{F}(t) \qquad \textit{for all } t \in [0,\infty].$$

Show that every stopping time is an optional time. Moreover, show that τ is an optional time if and only if τ is a stopping time with respect to the filtration $\mathfrak{F}(\cdot+) = \{\mathfrak{F}(t+)\}_{t\in[0,\infty)}$. ◇

There are, of course, more cases in which hitting times turn out to be stopping times, but the two cases in Proposition 1.22 are sufficient for our purposes. The following exercise gives another sufficient set of conditions.

Exercise 16. *Let $X = \{X(t)\}_{t\in[0,\infty)}$ be a real valued, right continuous, nondecreasing process which is adapted to a filtration $\mathfrak{F}$. Given $a \in \mathbb{R}$, show that the hitting time $\tau_{[a,\infty)}$ of the set $[a,\infty)$ by the process X is a stopping time.* ◇

Now what are stopping times good for? Well, as the name suggests, they can be used to stop stochastic processes.

Definition 1.23 (Stopped Process)**.** *Let $X = \{X(t)\}_{t\in\mathcal{T}}$ be a stochastic process and let τ be a $\mathcal{T}$-valued stopping time with respect to a filtration $\mathfrak{F}$. We define*

$$X(\tau) : \Omega \to S, \qquad \omega \mapsto X(\tau,\omega) \triangleq X(\tau(\omega),\omega). \tag{1.4}$$

With this, if τ is an arbitrary $\mathcal{T} \cup \{\infty\}$-valued stopping time, we refer to $X^\tau = X(\cdot \wedge \tau) = \{X(t \wedge \tau)\}_{t \in \mathcal{T}}$ as the ***process X stopped at time τ****, or* ***stopped process*** *for short.* ◇

We observe that Equation (1.4) makes sense for $\tau(\omega) = \infty$ only if $\infty \in \mathcal{T}$, which is why we restrict this part of the definition to stopping times taking values in $\mathcal{T}$ instead of $\mathcal{T} \cup \{\infty\}$. Next, for any $t \in \mathcal{T}$ and any stopping time τ (with values in $\mathcal{T} \cup \{\infty\}$), Lemma 1.16 (Operations on Stopping Times) implies that $t \wedge \tau$ is again a stopping time and $t \wedge \tau$ can only take the value ∞ if $\infty \in \mathcal{T}$. The stopped process X^τ simply freezes the original process X at time τ, see Figure 1.7.

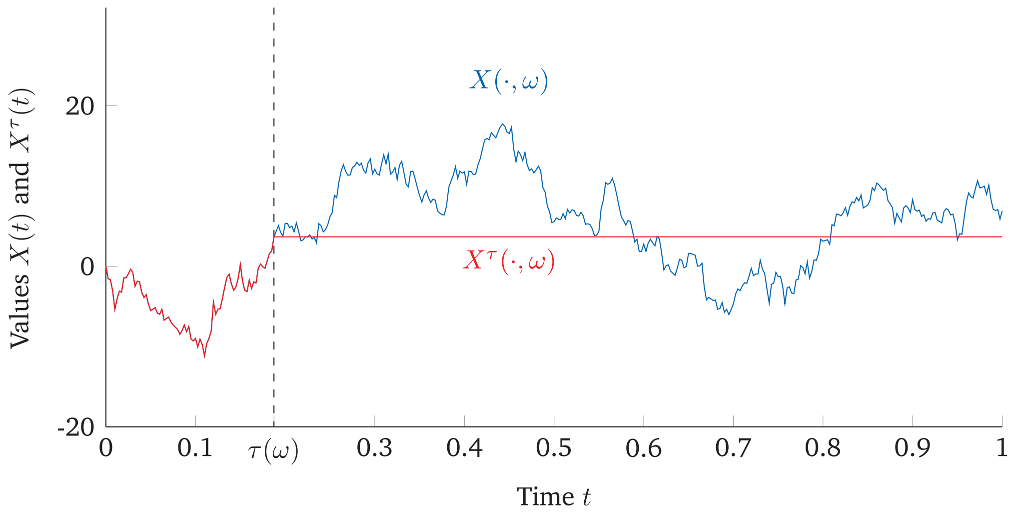

Figure 1.7 A process X and the corresponding stopped process X^τ.

Now take another close look at Equation (1.4). Since $\omega \mapsto X(\tau(\omega), \omega)$ has a double dependence on ω, **it is not clear** without any additional conditions if $X(\tau) : \Omega \to S$ is a random variable and hence it is not clear if X^τ is a stochastic process in the sense of Definition 1.1. If X is adapted to $\mathfrak{F}$, then one would naturally expect $X^\tau(t) = X(t \wedge \tau)$ to be $\mathfrak{F}(t \wedge \tau)$-measurable for all $t \in \mathcal{T}$, i.e. one would expect the stopped process $X^\tau = \{X^\tau(t)\}_{t \in \mathcal{T}}$ to be adapted to the **stopped filtration** $\mathfrak{F}^\tau = \mathfrak{F}(\cdot \wedge \tau) = \{\mathfrak{F}(t \wedge \tau)\}_{t \in \mathcal{T}}$ (and hence also adapted to $\mathfrak{F}$ as $\mathfrak{F}(t \wedge \tau) \subset \mathfrak{F}(t)$, $t \in \mathcal{T}$). As the next exercise shows, however, this is not true in general!

Exercise 17. *Let $\Omega = [0,1]$, $\mathfrak{A} = \mathfrak{B}([0,1])$, and let $\mathbb{P}$ be the Lebesgue measure on $(\Omega, \mathfrak{A})$. Moreover, fix a nonmeasurable subset A of Ω and define $X = \{X(t)\}_{t\in[0,\infty)}$ by*

$$X(t,\omega) \triangleq \begin{cases} t+\omega & \text{if } t \in A, \\ -t-\omega & \text{if } t \notin A, \end{cases} \qquad \text{for all } t \in [0,\infty), \omega \in \Omega.$$

Show that $\sigma(X(t)) = \mathfrak{A}$ and hence X is a stochastic process with natural filtration given by $\mathfrak{F}^X(t) = \mathfrak{A}$ for all $t \in [0,\infty)$. Now define $\tau : \Omega \to [0,\infty]$ by

$$\tau(\omega) \triangleq \inf\big\{t \geq 0 : 2t \geq |X(t)|\big\}, \qquad \omega \in \Omega.$$

Show that τ is an $\mathfrak{F}^X$-stopping time, that the σ-field of the τ-past is given by $\mathfrak{F}^X(\tau) = \mathfrak{A}$, but $\{X(\tau) > 0\} \notin \mathfrak{F}^X(\tau)$, i.e. $X(\tau)$ is not $\mathfrak{F}^X(\tau)$ measurable. ◇

So under which conditions is the stopped process adapted? The answer is once again easy in discrete time, or, more generally, if the stopping time takes only countably many values.

Lemma 1.24 (Measurability of Stopped Processes (Discrete Case))**.** *Let $X = \{X(t)\}_{t\in\mathcal{T}}$ be a stochastic process adapted to a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$ and let τ be a stopping time with respect to $\mathfrak{F}$. If $\tau(\Omega)$ is countable, the stopped process X^τ is adapted to the stopped filtration $\mathfrak{F}^\tau = \{\mathfrak{F}^\tau(t)\}_{t\in\mathcal{T}}$ given by $\mathfrak{F}^\tau(t) \triangleq \mathfrak{F}(t\wedge\tau)$ for all $t \in \mathcal{T}$. In particular, X^τ is adapted to $\mathfrak{F}$ and, if τ is $\mathcal{T}$-valued, $X(\tau)$ is $\mathfrak{F}(\tau)$-measurable.* ◇

Proof. Assume that τ is $\mathcal{T}$-valued. We argue that in this case $X(\tau)$ is $\mathfrak{F}(\tau)$-measurable. From this, it follows immediately that for arbitrary stopping times τ, the random variable $X^\tau(t) = X(t\wedge\tau)$ is $\mathfrak{F}(t\wedge\tau) = \mathfrak{F}^\tau(t)$-measurable, and hence X^τ is $\mathfrak{F}^\tau$-adapted. Since $\mathfrak{F}^\tau(t) \subset \mathfrak{F}(t)$ for all $t \in \mathcal{T}$, this furthermore implies that X^τ is $\mathfrak{F}$-adapted.

Let $B \in \mathfrak{S}$ and observe that, for any $s, t \in \mathcal{T}$ with $s \leq t$, we have

$$\{X(\tau) \in B\} \cap \{\tau = s\} = \{X(s) \in B\} \cap \{\tau = s\} \in \mathfrak{F}(s) \subset \mathfrak{F}(t)$$

by adaptedness of X and Lemma 1.18 (Properties of the σ-Field of the τ-Past). But then, using that $\tau(\Omega)$ is countable, we find

$$\{X(\tau) \in B\} \cap \{\tau \leq t\} = \bigcup_{s\in\tau(\Omega), s\leq t}\Big(\{X(\tau) \in B\} \cap \{\tau = s\}\Big) \in \mathfrak{F}(t)$$

for all $t \in \mathcal{T}$, i.e. $\{X(\tau) \in B\} \in \mathfrak{F}(\tau)$ and thus $X(\tau)$ is $\mathfrak{F}(\tau)$-measurable. □

In continuous time, as you may have guessed, the question of adaptedness of the stopped process is more delicate and as Exercise 17 shows, the answer is **in general not affirmative**. We shall prove the adaptedness of the stopped process in the case of right continuous processes (which is, once again, not the most general result but sufficient for our purposes). The proof is based on the very useful fact that any stopping time can be approximated from the right by a monotone sequence of stopping times with a finite range.

Proposition 1.25 (Approximation of Stopping Times). *Assume that $\mathcal{T} = [0,\infty)$ and let τ be a stopping time with respect to a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in[0,\infty)}$. Then there exists a sequence $\{\tau_n\}_{n\in\mathbb{N}}$ of stopping times such that*

(i) $\tau_n(\Omega) \subset [0,\infty]$ *is finite for all* $n \in \mathbb{N}$,

(ii) $\tau \leq \tau_{n+1} \leq \tau_n$ *for all* $n \in \mathbb{N}$,

(iii) $\tau < \tau_n$ *for all* $n \in \mathbb{N}$ *on* $\{\tau < \infty\}$,

(iv) $\inf_{n\in\mathbb{N}} \tau_n = \lim_{n\to\infty} \tau_n = \tau$.

Moreover, if σ is another stopping time satisfying $\sigma \leq \tau$, and $\{\sigma_n\}_{n\in\mathbb{N}}$ denotes the corresponding approximating sequence, then this sequence can be chosen such that $\sigma_n \leq \tau_n$ for all $n \in \mathbb{N}$. ◇

Proof. Given $n \in \mathbb{N}$, we define $\tau_n : \Omega \to [0,\infty]$ by

$$\tau_n(\omega) \triangleq \sum_{k=1}^{n2^n} k2^{-n}\mathbb{1}_{\{\tau\in[(k-1)2^{-n},k2^{-n})\}} + \infty\mathbb{1}_{\{\tau\geq n\}} \qquad \text{for all } \omega \in \Omega.$$

Then the sequence $\{\tau_n\}_{n\in\mathbb{N}}$ clearly satisfies the properties (i) to (iv) above. Moreover, each τ_n is a stopping time by Lemma 1.15 (Characterization of Discrete Stopping Times) since

$$\{\tau_n = k2^{-n}\} = \{\tau \geq (k-1)2^{-n}\} \cap \{\tau < k2^{-n}\} \in \mathfrak{F}(k2^{-n})$$

for all $k \in \{1, 2, \ldots, n2^n\}$. Finally, if σ is another stopping time satisfying $\sigma \leq \tau$ and if $\{\sigma_n\}_{n\in\mathbb{N}}$ is defined in the same way as the sequence $\{\tau_n\}_{n\in\mathbb{N}}$ above, then it is obvious that $\sigma_n \leq \tau_n$ for all $n \in \mathbb{N}$. □

It is crucial that the sequence $\{\tau_n\}_{n\in\mathbb{N}}$ approximates the stopping time τ **from the right**. In general, it is not possible to find an approximating sequence from the left, since otherwise we would be able to anticipate the stopping time τ. With this approximating procedure at hand, we can now show by a limiting argument that if we stop adapted right continuous processes, the stopped process is still adapted.

Proposition 1.26 (Measurability of Stopped Processes (Continuous Case)). *Let $X = \{X(t)\}_{t\in[0,\infty)}$ be a right continuous stochastic process which is adapted to a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in[0,\infty)}$. Let moreover τ be a stopping time with respect to $\mathfrak{F}$. Then the stopped process X^τ is adapted to both the stopped filtration $\mathfrak{F}^\tau$ and the original filtration $\mathfrak{F}$. Moreover, if τ takes values in $[0,\infty)$ only, then $X(\tau)$ is $\mathfrak{F}(\tau)$-measurable.* ◇

Proof. Let $\{\tau_n\}_{n\in\mathbb{N}}$ be the approximating sequence of stopping times for τ given by Proposition 1.25 (Approximation of Stopping Times). Then it follows from Lemma 1.24 (Measurability of Stopped Processes (Discrete Case)) that X^{τ_n} is $\mathfrak{F}$-adapted for each $n \in \mathbb{N}$. From this and the right continuity of X, it follows that

$$X^\tau(t) = X(t\wedge\tau) = \lim_{n\to\infty} X(t\wedge\tau_n) = \lim_{n\to\infty} X^{\tau_n}(t)$$

is $\mathfrak{F}(t)$-measurable for all $t\in[0,\infty)$, i.e. X^τ is $\mathfrak{F}$-adapted. Now assume that τ takes values in $[0,\infty)$ and let $t\in[0,\infty)$. Then, since $\{\tau\le t\}\in\mathfrak{F}(t)$ and $X^\tau(t)$ is $\mathfrak{F}(t)$-measurable, it follows that, for all $B\in\mathfrak{S}$, we have

$$\{X(\tau)\in B\}\cap\{\tau\le t\} = \{X^\tau(t) = X(t\wedge\tau)\in B\}\cap\{\tau\le t\}\in\mathfrak{F}(t),$$

and hence $X(\tau)$ is $\mathfrak{F}(\tau)$-measurable. As in the proof of Lemma 1.24 (Measurability of Stopped Processes (Discrete Case)), this implies that $X^\tau(t) = X(t\wedge\tau)$ is $\mathfrak{F}^\tau(t) = \mathfrak{F}(t\wedge\tau)$-measurable for all $t\in[0,\infty)$ and hence X^τ is also adapted to the stopped filtration $\mathfrak{F}^\tau$. □

Chapter

2 Brownian Motion

In this chapter, we study the most important stochastic process in continuous time: **Brownian motion**. Why is it so important? Because it's distribution is in some sense the stochastic process analogue of the standard normal distribution in probability theory in that a version of the central limit theorem holds: Suitably rescaled sequences of random walks converge in distribution to Brownian motion. Moreover, Brownian motion belongs to all classes of stochastic processes we will encounter throughout this course: Gaussian processes, martingales, Lévy processes, and Markov processes.

The plan for this chapter is as follows: We begin with the definition of Brownian motion and give some historical background. Then we take a closer look at **Gaussian processes**, which are stochastic processes whose finite-dimensional distributions are multivariate normal, and provide a characterization of Brownian motion as a Gaussian process. Then we take a closer look at **path properties of Brownian motion**. E.g., as we shall see, despite the fact that all paths of Brownian motion are continuous, almost every path is nowhere differentiable. We shall furthermore see that it is not clear from its definition if Brownian motion exists. For this reason, we then spend the remainder of this chapter with the construction of Brownian motion.

There are different ways to construct Brownian motion. Our construction is based on two theorems due to Kolmogorov: The **Kolmogorov Consistency Theorem**, which provides a convenient way to construct stochastic processes by specifying their finite-dimensional distributions, and the **Kolmogorov-Čentsov Continuity Theorem**, which provides a method for constructing stochastic processes with continuous paths. A combination of the two theorems will not only establish existence of Brownian motion, but for a whole range of other processes as well.

2.1 Definition of Brownian Motion

Throughout this chapter, we fix a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in[0,\infty)}$. We begin with the definition of Brownian motion.

Definition 2.1 (Brownian Motion/Wiener Process). *Let $W = \{W(t)\}_{t\in[0,\infty)}$ be an $\mathfrak{F}$-adapted stochastic process. We say that W is a* ***Brownian motion*** *or a* ***Wiener process*** *with respect to $\mathfrak{F}$ if*

(W1) $W(0) = 0$,

(W2) *for all $s, t \in [0,\infty)$ with $s < t$, the increment $W(t) - W(s)$ is independent of $\mathfrak{F}(s)$,*

(W3) *for all $s, t \in [0,\infty)$ with $s < t$, the increment $W(t) - W(s)$ is normally distributed with mean zero and variance $t - s$,*

(W4) *W is continuous.* ◇

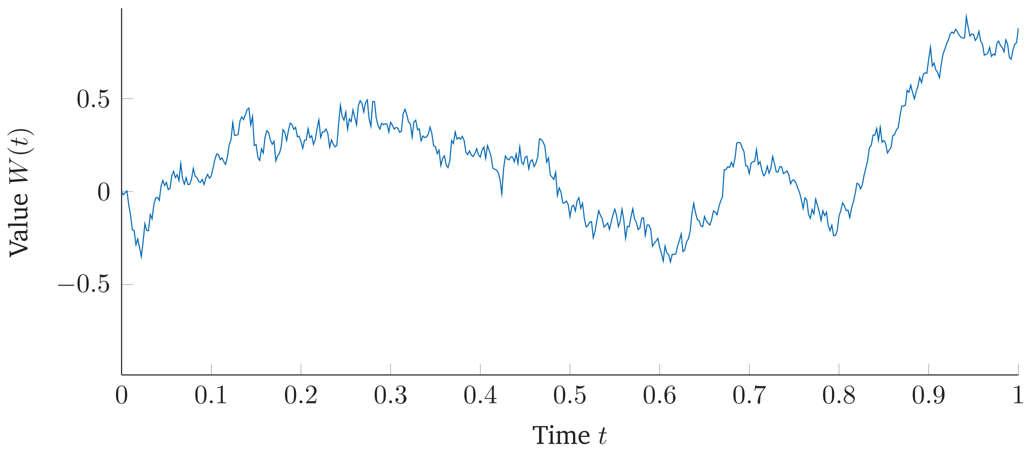

Figure 2.1 Path of a Brownian motion W.

Brownian motion is named after the Scottish botanist **Robert Brown**, who is, among other important contributions, famous for the following experiment: In 1827, Brown observed that microscopic particles within pollen grains suspended in water randomly move around. Up to some scaling, the stochastic process of Definition 2.1 can be thought of as a mathematical model for this kind of motion in one dimension – hence the name Brownian motion. The movement of the pollen particles is now believed to be caused by collisions of the particles with the water molecules, an explanation first given by **Albert Einstein** in 1905. The first existence proof of the stochastic process in Definition 2.1 is due to **Norbert Wiener** in 1923, which is why the process is sometimes also referred to as the Wiener process.

Definition 2.1 does not guarantee existence of Brownian motion and we have to put in a significant amount of effort to arrive at the existence. Especially the question of whether a process satisfying (W1), (W2), and (W3) can have continuous paths turns out to be quite delicate. Before we address this issue, however, let us first look at some properties of Brownian motion in more detail.

2.2 Gaussian Processes

Let us first try to understand the distributional properties of Brownian motion. As highlighted in Chapter 1, the distribution of a stochastic process is uniquely characterized in terms of the **finite-dimensional distributions**. In the case of Brownian motion, the finite-dimensional distributions are easily computed.

For $n \in \mathbb{N}$, let us fix $[t_1, \dots, t_n] \subset [0, \infty)$ and consider the n-dimensional random vector

$$X \triangleq \big(W(t_1), \dots, W(t_n)\big).$$

Writing $t_0 \triangleq 0$ and $Z_j \triangleq W(t_j) - W(t_{j-1})$ for $j = 1, \dots, n$, the k^{th} component X_k of X can be expressed as

$$X_k = W(t_k) = \sum_{j=1}^{k} Z_j \qquad \text{for all } k = 1, \dots, n.$$

In particular, for any $y \in \mathbb{R}^n$ it holds that

$$\langle y, X\rangle = \sum_{k=1}^{n} y_k X_k = \sum_{k=1}^{n} y_k \sum_{j=1}^{k} Z_j = \sum_{j=1}^{n} Z_j \sum_{k=j}^{n} y_k = \sum_{j=1}^{n} \bar{y}_j Z_j,$$

where $\bar{y}_j \triangleq \sum_{k=j}^{n} y_k \in \mathbb{R}$ for $j = 1, \dots, n$. Therefore, since $Z_1, \dots, Z_n$ are independent by (W2) and each Z_j is normally distributed by (W3), it follows that each scalar product $\langle y, X \rangle$, $y \in \mathbb{R}^n$, is normally distributed. But this implies that X has a **multivariate normal distribution** (n-dimensional). In other words: The finite-dimensional distributions of Brownian motion are multivariate normal. Processes with this property are also referred to as **Gaussian processes**.

Definition 2.2 ((Centered) Gaussian Process; Covariance Function)**.** *A real-valued stochastic process $X = \{X(t)\}_{t \in \mathcal{T}}$ is called a **Gaussian process** if for all $[t_1, \ldots, t_n] \subset \mathcal{T}$, $n \in \mathbb{N}$, it holds that*

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

*We say that X is **centered** if*

$$\mathrm{E}\big[X(t)\big] = 0 \qquad \text{for all } t \in \mathcal{T}.$$

Finally, the function

$$\Gamma : \mathcal{T} \times \mathcal{T} \to \mathbb{R}, \qquad \Gamma(s,t) \triangleq \mathrm{Cov}\big[X(s), X(t)\big]$$

*is called the **covariance function** of X.* ◇

Observe that Brownian motion is centered since by (W1) and (W3) we have

$$\mathrm{E}\big[W(t)\big] = \mathrm{E}\big[W(t) - W(0)\big] = 0 \qquad \text{for all } t \in [0, \infty).$$

The distribution of a centered Gaussian process is **uniquely determined** by its covariance function. This is immediate since the finite-dimensional distributions uniquely characterize the distribution of a process and the n-dimensional normal distribution is uniquely determined by its expectation vector and covariance matrix. In particular, we have the following useful characterization of Brownian motion as a Gaussian process.

Theorem 2.3 (Brownian Motion as a Gaussian Process)**.** *A stochastic process $X = \{X(t)\}_{t \in [0,\infty)}$ is a Brownian motion with respect to its natural filtration $\mathfrak{F}^X$ if and only if X is a continuous centered Gaussian process with $X(0) = 0$ and covariance function*

$$\Gamma : [0,\infty) \times [0,\infty) \to \mathbb{R}, \qquad \Gamma(s,t) \triangleq s \wedge t.$$ ◇

Proof. Suppose that X is a Brownian motion. We already know that Brownian motion is a continuous centered Gaussian process with $X(0) = 0$. For $0 \le s \le t$, the fact that Brownian motion starts in zero shows that

$$\begin{aligned}\Gamma(s,t) &= \operatorname{Cov}\big(X(s), X(t)\big)\\ &= \operatorname{Cov}\big(X(s), X(t) - X(s) + X(s)\big)\\ &= \operatorname{Cov}\big(X(s), X(t) - X(s)\big) + \operatorname{Cov}\big(X(s), X(s)\big)\\ &= \operatorname{Cov}\big(X(s), X(t) - X(s)\big) + \operatorname{Cov}\big(X(s) - X(0), X(s) - X(0)\big).\end{aligned}$$

Now $X(s)$ and $X(t) - X(s)$ are independent by (W2) and thus the first covariance is equal to zero. The second covariance is equal to the variance of $X(s) - X(0)$, which by (W3) is equal to s. Thus $\Gamma(s,t) = s = s \wedge t$ as claimed.

On the other hand, since the covariance function uniquely determines the distribution of a centered Gaussian process, it follows that every centered Gaussian process with the above covariance function satisfies (W2) (with respect to its natural filtration) and (W3). Thus, if this process is in addition continuous and satisfies $X(0) = 0$, it must be a Brownian motion. □

The previous theorem is useful since it allows us to show that Brownian motion is stable under various transformations.

Exercise 18. *Let W be a Brownian motion, $a > 0$, and $t_0 \in [0,\infty)$. Show that the following processes are Brownian motions as well:*

$$\big\{-W(t)\big\}_{t\in[0,\infty)}, \quad \big\{\tfrac{1}{\sqrt{a}}W(at)\big\}_{t\in[0,\infty)}, \quad \big\{W(t+t_0) - W(t_0)\big\}_{t\in[0,\infty)}. \qquad \diamond$$

Exercise 19. *Let W be a Brownian motion and $t > 0$. For each $n \in \mathbb{N}$, fix $[t_0^n, \dots, t_n^n] \subset [0,t]$ with $t_0^n = 0$ and $t_n^n = t$ and assume that*

$$\max_{k=1,\dots,n} |t_k^n - t_{k-1}| \to 0 \qquad \text{as } n \to \infty.$$

Show that

$$\lim_{n\to\infty} \mathrm{E}\Big[\Big|t - \textstyle\sum_{k=1}^n \big|W(t_k^n) - W(t_{k-1}^n)\big|^2\Big|^2\Big] = 0. \qquad \diamond$$

Brownian motion is of course not the only Gaussian process. Let us take a look at some more examples.

Example 2.4 (Brownian Bridge)**.** *Let W be a Brownian motion and define a stochastic process $X = \{X(t)\}_{t\in[0,1]}$ by*

$$X(t) = W(t) - tW(1), \qquad t \in [0,1].$$

Then the process X is called a ***Brownian bridge****.* ◇

The Brownian bridge gets it name from its behavior at time $t = 0$ and $t = 1$ as we have

$$X(0) = W(0) - 0W(1) = 0 \quad \text{and} \quad X(1) = W(1) - 1W(1) = 0,$$

i.e. the Brownian bridge starts and ends in zero. The fact that it is a Gaussian process is easily verified.

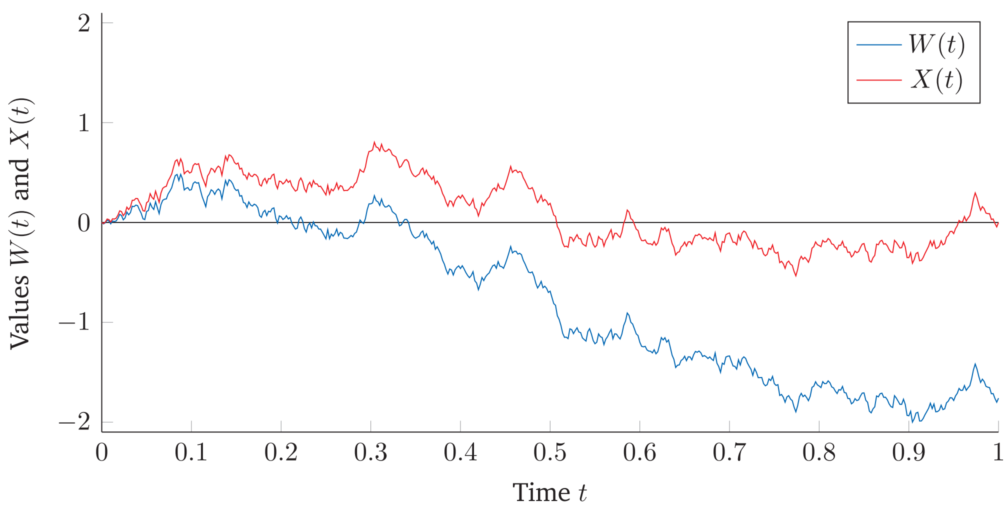

Figure 2.2 Path of a Brownian motion W and corresponding Brownian bridge X.

Exercise 20. *Show that the Brownian bridge $X = \{X(t)\}_{t\in[0,1]}$ constructed from a Brownian motion W is a continuous centered Gaussian process with covariance function*

$$\Gamma : [0,1] \times [0,1] \to \mathbb{R}, \qquad \Gamma(s,t) = s \wedge t - st.$$

Is the Brownian bridge adapted to the natural filtration $\mathfrak{F}^W = \{\mathfrak{F}^W(t)\}_{t\in[0,1]}$ of the underlying Brownian motion? ◇

Another important Gaussian process is the **Ornstein-Uhlenbeck process**.

Example 2.5 (Ornstein-Uhlenbeck Process)**.** *Let $\kappa, \sigma > 0$ be given. The continuous centered Gaussian process $X = \{X(t)\}_{t\in[0,\infty)}$ with covariance function*

$$\Gamma : [0,\infty) \times [0,\infty) \to \mathbb{R}, \qquad \Gamma(s,t) \triangleq \frac{\sigma^2}{2\kappa} e^{-\kappa(s+t)}\left(e^{2\kappa s\wedge t} - 1\right)$$

*is called **Ornstein-Uhlenbeck process**.* ◇

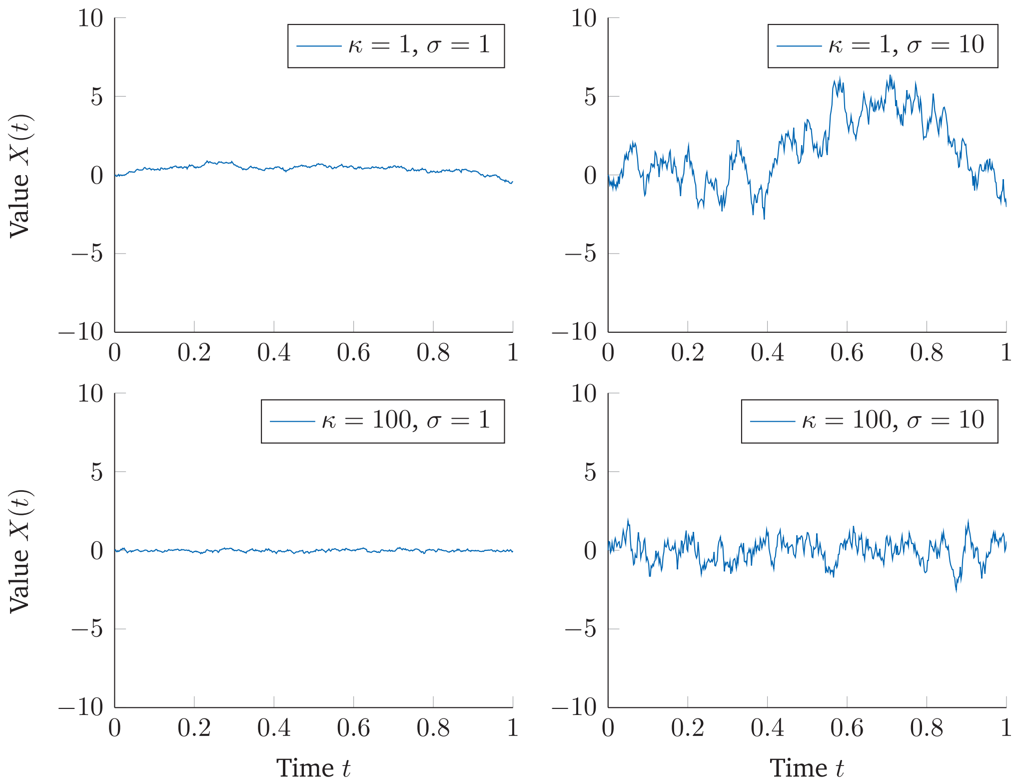

Figure 2.3 Paths of Ornstein-Uhlenbeck processes for various choices of the parameters κ and σ.

As usual, existence of the Ornstein-Uhlenbeck process is a priori not guaranteed. However, as the following exercise shows, the Ornstein-Uhlenbeck process can be constructed in a straightforward manner from a Brownian motion by means of rescaling and a suitable time change.

Exercise 21. *Let* $\kappa, \sigma > 0$ *and let* W *be a Brownian motion. Define a stochastic process* $X = \{X(t)\}_{t\in[0,\infty)}$ *by*

$$X(t) \triangleq e^{-\kappa t} W\big(\tfrac{\sigma^2}{2\kappa}(e^{2\kappa t} - 1)\big) \qquad \text{for all } t \in [0,\infty).$$

Show that X *is an Ornstein-Uhlenbeck process.* ◇

The Ornstein-Uhlenbeck process has the interesting feature that it tends to return to its mean of zero. This feature is sometimes referred to as **mean reversion**. The larger the choice of κ, the faster the Ornstein-Uhlenbeck process returns to its mean. The parameter σ on the other hand controls the volatility of the process. These features are exemplified in Figure 2.3.

Example 2.6 (Fractional Brownian Motion)**.** *Let* $H \in (0,1)$. *The continuous centered Gaussian process* $X = \{X(t)\}_{t\in[0,\infty)}$ *with covariance function*

$$\Gamma : [0,\infty) \times [0,\infty) \to \mathbb{R}, \qquad \Gamma(s,t) \triangleq \tfrac{1}{2}\big(s^{2H} + t^{2H} - |t-s|^{2H}\big)$$

is called ***fractional Brownian motion*** *with* ***Hurst index*** H. ◇

It is not very difficult to verify that a fractional Brownian motion with Hurst index $H = 1/2$ is a Brownian motion. If the Hurst index H is smaller than $1/2$, then the paths of a fractional Brownian motion look rougher than the paths of a Brownian motion. On the other hand, if the Hurst index H is bigger than $1/2$, the paths of the fractional Brownian motion look less rough than those of a Brownian motion. We will return to this feature later, but convince ourselves for now by looking at Figure 2.4.

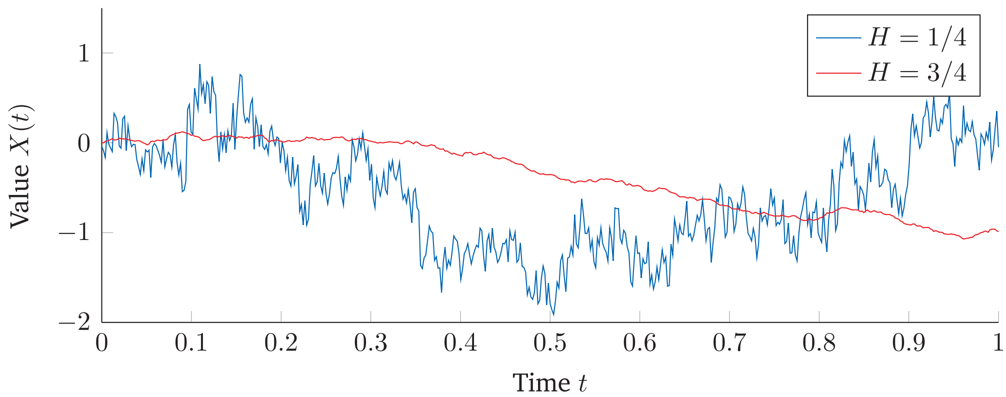

Figure 2.4 Paths of fractional Brownian motions for different choices of H.

2.3 Path Properties of Brownian Motion

Let us now take a closer look at the paths of Brownian motion. Figure 2.1 suggests that the paths are quite rough. As it turns out, almost every path of Brownian motion is nowhere differentiable. More precisely, it can be shown that almost every path is nowhere p-Hölder continuous if $p > 1/2$.

Suppose that (S, d) is a metric space and fix $p \in (0, 1]$. We say that a function $f : \mathcal{T} \to S$ is **p-Hölder continuous at a point** $t \in \mathcal{T}$ if

$$\limsup_{s \to t} \frac{d(f(s), f(t))}{|s - t|^p} < \infty.$$

Moreover, we say that f is **p-Hölder continuous** if there exist constants $\delta > 0$ and $C > 0$ such that

$$d\big(f(s), f(t)\big) \le C|s - t|^p \qquad \text{for all } s, t \in \mathcal{T} \text{ with } |s - t| \le \delta.$$

Clearly, if f is p-Hölder continuous (in t), then it is also q-Hölder continuous (in t) for all $0 < q \le p$. Moreover, p-Hölder continuity implies uniform continuity and if f is differentiable in t, then f must be 1-Hölder continuous in t. Conversely, if there exists no $t \in \mathcal{T}$ such that f is p-Hölder continuous in t for some $p \in (0, 1]$, then f is cannot be differentiable anywhere.

Exercise 22. *Let $q \in (0, 1]$ and $f : [0, \infty) \to [0, \infty)$ be given by $f(x) \triangleq x^q$ for all $x \in [0, \infty)$. For which $p \in (0, 1]$ is this function p-Hölder continuous?* ◇

Exercise 23. *Construct a function which is uniformly continuous, but not p-Hölder continuous for any $p \in (0, 1]$.* ◇

Theorem 2.7 (Nowhere Differentiability of Brownian Paths)**.** *Let W be a Brownian motion and $p > 1/2$. Then for $\mathbb{P}$-almost every $\omega \in \Omega$, the path $W(\cdot, \omega)$ is continuous, but nowhere p-Hölder continuous. In particular, almost every path of W is nowhere differentiable.* ◇

Proof. Choose $r \in \mathbb{N}$ such that $(p - 1/2)r > 1$, which is possible since $p > 1/2$. For each $t \in [0, 1]$, we define

$$H(t) \triangleq \big\{f : [0, \infty) \to \mathbb{R} : f \text{ is } p\text{-Hölder continuous in } t\big\} \subset \mathbb{R}^{[0,\infty)}.$$

We now proceed to construct a set $A \subset \mathbb{R}^{[0,\infty)}$ with $A \in \mathfrak{B}(\mathbb{R})^{[0,\infty)}$ such that $H(t) \subset A$ for all $t \in [0,1]$ and $\mathbb{P}[W \in A] = 0$. From this, it follows that almost every path is nowhere p-Hölder continuous on $[0,1]$. Using the scaling invariance of Brownian motion, we then extend the result to $[0,\infty)$.

Step 1: Construction of A. Given $K, n \in \mathbb{N}$, and $k \in \{1,\dots,n\}$, define

$$A(K,n,k) \triangleq \bigcap_{j=1}^{r} \left\{ f : [0,\infty) \to \mathbb{R} : \left| f\left(\tfrac{k+j}{n}\right) - f\left(\tfrac{k+j-1}{n}\right) \right| \le K n^{-p} \right\}.$$

With this, we then proceed to set

$$A(K,n) \triangleq \bigcup_{k=1}^{n} A(K,n,k)$$

followed by

$$A(K) \triangleq \liminf_{n\to\infty} A(K,n) = \bigcup_{n=1}^{\infty} \bigcap_{N=n}^{\infty} A(K,N),$$

and finally

$$A \triangleq \bigcup_{K\in\mathbb{N}} A(K).$$

Since the set A is constructed from countably many unions and intersections of sets $A(K,n,k)$ and we clearly have $A(K,n,k) \in \mathfrak{B}(\mathbb{R})^{[0,\infty)}$, we see that have $A \in \mathfrak{B}(\mathbb{R})^{[0,\infty)}$, i.e. A is a measurable set. We claim that

$$\bigcup_{t\in[0,1]} H(t) \subset A.$$

For this, let us fix $t \in [0,1]$ as well as $f \in H(t)$. To see that $f \in A$, it suffices to show that there exists K sufficiently large, such that for eventually all $n \in \mathbb{N}$ there exists some $k \in \{1,\dots,n\}$ with $f \in A(K,n,k)$. For this, we first observe that we can find $\delta = \delta(f,t) > 0$ and $C = C(f,t) > 0$ such that

$$|f(t) - f(s)| \le C|t-s|^p \qquad \text{for all } s \in [t-\delta, t+\delta] \cap [0,\infty).$$

Now choose $n \in \mathbb{N}$ sufficiently large such that $(r+1)/n < \delta$. Next, let $k \in \{1,\dots,n\}$ such that $(k-1)/n < t \le k/n$ (and $k=1$ if $t=0$). Then

$$\left|\tfrac{k+j}{n} - t\right| = \tfrac{k+j}{n} - t = \left(\tfrac{k}{n} - t\right) + \tfrac{j}{n} \le \tfrac{1}{n} + \tfrac{j}{n} \le \tfrac{r+1}{n} < \delta \quad \text{for all } j = 0,\dots,r.$$

But then we must have

$$\begin{aligned} \left| f\left(\tfrac{k+j}{n}\right) - f\left(\tfrac{k+j-1}{n}\right) \right| &\le \left| f\left(\tfrac{k+j}{n}\right) - f(t) \right| + \left| f(t) - f\left(\tfrac{k+j-1}{n}\right) \right| \\ &\le 2C\left(\tfrac{r+1}{n}\right)^p = 2C(r+1)^p n^{-p}, \qquad j = 1,\dots,r. \end{aligned}$$

Put differently, choosing $K \in \mathbb{N}$ such that $K \geq 2C(r+1)^p$, we have $f \in A(K,n,k)$ for all $n \in \mathbb{N}$ sufficiently large and some $k \in \{1,\dots,n\}$ and therefore $f \in A$ as claimed.

Step 2: We show that $\mathbb{P}[W \in A] = 0$. By definition of A, it suffices to show that $\mathbb{P}[W \in A(K)] = 0$ for all $K \in \mathbb{N}$. To see this, let $n \in \mathbb{N}$ and choose $k \in \{1,\dots,n\}$. Using the independence and stationarity of the increments of Brownian motion, it follows that

$$\begin{aligned}
\mathbb{P}\big[W \in A(K,n,k)\big] &= \mathbb{P}\Big[\big|W\big(\tfrac{k+j}{n}\big) - W\big(\tfrac{k+j-1}{n}\big)\big| \leq Kn^{-p} \text{ for all } j = 1,\dots,r\Big] \\
&= \textstyle\prod_{j=1}^{r} \mathbb{P}\big[\big|W\big(\tfrac{k+j}{n}\big) - W\big(\tfrac{k+j-1}{n}\big)\big| \leq Kn^{-p}\big] \\
&= \mathbb{P}\big[\big|W\big(\tfrac{1}{n}\big)\big| \leq Kn^{-p}\big]^r \\
&= \mathbb{P}\big[|W(1)| \leq Kn^{-p+1/2}\big]^r.
\end{aligned}$$

Since the density of the standard normal distribution is bounded from above by 1, we can estimate the probability of the event $\{|W(1)| \leq Kn^{-p+1/2}\}$ by the Lebesgue measure of the set $[-Kn^{-p+1/2}, Kn^{-p+1/2}]$, i.e.

$$\mathbb{P}\big[W \in A(K,n,k)\big] \leq (2K)^r n^{(-p+1/2)r}.$$

Using Fatou's lemma and the σ-subadditivity of $\mathbb{P}$, we therefore find that

$$\begin{aligned}
\mathbb{P}\big[W \in A(K)\big] &\leq \liminf_{n\to\infty} \mathbb{P}\big[W \in A(K,n)\big] \\
&\leq \limsup_{n\to\infty} \textstyle\sum_{k=1}^{n} \mathbb{P}\big[W \in A(K,n,k)\big] \\
&\leq \limsup_{n\to\infty} n(2K)^r n^{(-p+1/2)r} \\
&= (2K)^r \limsup_{n\to\infty} n^{1-(p-1/2)r} = 0
\end{aligned}$$

since we have chosen r such that $(p-1/2)r > 1$. Thus $\mathbb{P}[W \in A] = 0$.

Step 3: Extension to $[0,\infty)$. Thus far, we have seen that $\mathbb{P}$-almost every path is nowhere p-Hölder continuous on $[0,1]$. Now assume by contradiction that there exists an event $F \in \mathfrak{A}$ with $\mathbb{P}[F] > 0$ such that for each $\omega \in F$, there exists some $t(\omega) \in [0,\infty)$, such that $W(\cdot,\omega)$ is p-Hölder continuous in $t(\omega)$. Since $[0,\infty) = \bigcup_{m\in\mathbb{N}}[m-1,m)$, we may without loss of generality assume that there exists $M \in \mathbb{N}$ such that every $t(\omega) \leq M$ for all $\omega \in F$. Now define a stochastic process $X = \{X(t)\}_{t\in[0,\infty)}$ by

$$X(t) \triangleq \tfrac{1}{\sqrt{M}} W(Mt) \qquad \text{for all } t \in [0,\infty).$$

By Exercise 18, X is a Brownian motion and thus almost surely nowhere p-Hölder continuous on $[0,1]$. On the other hand, since for each $\omega \in F$ the path $W(\cdot,\omega)$ is p-Hölder continuous in $t(\omega)$, $X(\cdot,\omega)$ must be p-Hölder continuous in $t(\omega)/M \leq 1$. This is a contradiction, hence the proof is finished. □

Let us emphasize the following: Almost every path of Brownian motion is **nowhere differentiable**, which is a lot stronger than the statement that almost every path is **not differentiable**. Indeed, for the latter statement it suffices to find one $t \in [0,\infty)$ at which $W(\cdot,\omega)$ is not differentiable. The statement of the previous theorem tells us however that $t \mapsto W(t,\omega)$ is not differentiable in any $t \in [0,\infty)$. Intuitively, this means that the paths of Brownian motion must have a lot of kinks.

Differentiability is not the only property which fails for paths of Brownian motion. For example, it can also be shown that almost every path is **nowhere monotone**.

Exercise 24. *Let W be a Brownian motion. Show that*

$$\mathbb{P}\Big[\Big\{\omega \in \Omega : \textit{there exist } t_0, t_1 \in [0,\infty) \textit{ with } t_0 < t_1 \textit{ such that} \\ W(s,\omega) \leq W(t,\omega) \textit{ for all } t_0 \leq s \leq t \leq t_1\Big\}\Big] = 0. \quad \diamond$$

2.4 The Kolmogorov Consistency Theorem

We now turn to the issue of **existence of Brownian motion**. As already mentioned, there are several ways to establish the existence result. The approach taken in this course is a rather systematic one in that the results establish existence for a whole range of stochastic processes at once, including all Gaussian processes introduced in Section 2.2.

We follow a two step procedure. First, we derive a theorem which allows to construct stochastic processes by specifying **finite-dimensional distributions** satisfying a certain consistency condition. This will allow us to construct **raw versions** of stochastic processes satisfying the correct distributional properties, i.e. in case of Brownian motion the theorem can be used to construct a raw Brownian motion satisfying the independent increments property (W2) and the stationary normal increments property (W3), but not necessarily the continuity property (W4). The theorem facilitating this raw construction is called **Kolmogorov's Consistency Theorem** and is the main subject of this section.

In the second step, we show that we can **modify** the raw process on suitable nullsets to ensure the **continuous paths** property. We shall see that this is always possible provided that the increments of the raw process satisfy a certain integrability condition (which is a distributional property). This result is called the **Kolmogorov-Čentsov Continuity Theorem** and will be the derived in the subsequent section.

In order to establish the consistency theorem, we first need to following **regularity result** concerning probability measures on $\mathbb{R}^d$, stating that the probability of a Borel set can be approximated from below by probabilities of compact sets and from above by probabilities of open sets. A measure which satisfies these two properties is often referred to as being regular.

Lemma 2.8 (Regularity of Probability Measures on $\mathbb{R}^d$). *Let μ be a probability measure on $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$ and let $B \in \mathfrak{B}(\mathbb{R}^d)$ be a Borel set. Then*

$$\mu(B) = \sup\big\{\mu(K) : K \subset B \text{ compact}\big\} = \inf\big\{\mu(O) : O \supset B \text{ open}\big\}, \tag{2.1}$$

i.e. μ is a regular measure. ◇

Proof. The proof is based on a monotone class argument and divided into three steps.

Step 1: The good set. We denote by $\mathfrak{G}$ the set of Borel sets for which the claim of the lemma holds, i.e.

$$\mathfrak{G} \triangleq \big\{G \in \mathfrak{B}(\mathbb{R}^d) : G \text{ satisfies Equation (2.1)}\big\}.$$

Our aim is of course to show that $\mathfrak{G} = \mathfrak{B}(\mathbb{R}^d)$. For this, we show that $\mathfrak{G}$ is a λ-system which contains a π-system generating $\mathfrak{B}(\mathbb{R}^d)$. The result then follows from Dynkin's π-λ lemma (which states, as we recall, that the σ-field generated by a π-system coincides with the λ-system generated by this

π-system). The fact that $\mathfrak{G}$ contains a π-system generating $\mathfrak{B}(\mathbb{R}^d)$ is easily seen since $\mathfrak{G}$ clearly contains the system of sets of the form

$$\times_{i=1}^{d}(-\infty, a_i] \qquad \text{for all } a_1, \dots, a_d \in \mathbb{R}.$$

Indeed, $\times_{i=1}^{d}(-\infty, a_i]$ can be approximated from below by the compact sets $\{K_n\}_{n\in\mathbb{N}}$ and from above by the open sets $\{O_n\}_{n\in\mathbb{N}}$ given by

$$K_n \triangleq \times_{i=1}^{d}[-n, a_i] \qquad \text{and} \qquad O_n \triangleq \times_{i=1}^{d}(-\infty, a_i + 1/n), \qquad n \in \mathbb{N}.$$

By the continuity of μ from below and above it then follows that

$$\mu\big(\times_{i=1}^{d}(-\infty, a_i]\big) = \lim_{n\to\infty} \mu(K_n) = \lim_{n\to\infty} \mu(O_n),$$

i.e. $\times_{i=1}^{d}(-\infty, a_i] \in \mathfrak{G}$. To see that $\mathfrak{G}$ is a λ-system, we have to show that $\mathbb{R}^d \in \mathfrak{G}$ (which follows by a similar argument as above), and that $\mathfrak{G}$ is stable under taking complements and countable unions of disjoint sets.

Step 2: To see that $\mathfrak{G}$ is stable under taking complements, let $G \in \mathfrak{G}$. Then there exists a sequence of compact sets $\{K_n\}_{n\in\mathbb{N}}$ and a sequence of open sets $\{O_n\}_{n\in\mathbb{N}}$ such that $K_n \subset G \subset O_n$ for all $n \in \mathbb{N}$ and

$$\mu(G) = \lim_{n\to\infty} \mu(K_n) = \lim_{n\to\infty} \mu(O_n).$$

For each $n \in \mathbb{N}$, it follows that G^c is contained in the open set K_n^c, and hence

$$\mu(G^c) = 1 - \mu(G) = 1 - \lim_{n\to\infty} \mu(K_n) = \lim_{n\to\infty} \mu(K_n^c).$$

On the other hand, for each $n, m \in \mathbb{N}$, the set $F_n^m \triangleq O_n^c \cap [-m, m]^d$ is a compact subset of G^c and

$$\mu(O_n^c) = \lim_{m\to\infty} \mu(F_n^m) \qquad \text{for all } n \in \mathbb{N}$$

by continuity of μ from below. Now given $j \in \mathbb{N}$, there exists $N(j) \in \mathbb{N}$ with

$$\mu(G) \geq \mu(O_n) - \tfrac{1}{2j} \qquad \text{for all } n \geq N(j)$$

and for each $n \in \mathbb{N}$ we can find $M(n) \in \mathbb{N}$ such that

$$\mu(O_n^c) \leq \mu(F_n^m) + \tfrac{1}{2j} \qquad \text{for all } m \geq M(n).$$

Now define a sequence of sets $\{F_j\}_{j\in\mathbb{N}}$ by

$$F_j \triangleq F_{N(j)}^{M(N(j))}, \qquad j \in \mathbb{N}.$$

Then it follows that $F_j \subset G^c$ is compact and

$$\mu(F_j) \geq \mu(O_{N(j)}^c) - \tfrac{1}{2j} = 1 - \mu(O_{N(j)}) - \tfrac{1}{2j} \geq 1 - \mu(G) - \tfrac{1}{j} = \mu(G^c) - \tfrac{1}{j}$$

for all $j \in \mathbb{N}$, i.e. $\mu(G^c) = \lim_{j\to\infty} \mu(F_j)$ and thus $G^c \in \mathfrak{G}$.

Step 3: To finish the proof, it remains to show that whenever $\{G_n\}_{n\in\mathbb{N}}$ is a sequence of disjoint sets in $\mathfrak{G}$, then $G \triangleq \bigcup_{n\in\mathbb{N}} G_n \in \mathfrak{G}$. For this, let us fix $\varepsilon > 0$ and for each $n \in \mathbb{N}$ we select $K_n \subset G_n$ compact and $O_n \supset G_n$ open with

$$\mu(O_n) - \varepsilon 2^{-n} \leq \mu(G_n) \leq \mu(K_n) + \tfrac{1}{2}\varepsilon 2^{-n}.$$

Now the set $O \triangleq \bigcup_{n\in\mathbb{N}} O_n$ is open, contains G, and

$$\begin{aligned}\mu(G) \leq \mu(O) \leq \textstyle\sum_{n=1}^{\infty} \mu(O_n) \leq \sum_{n=1}^{\infty} \mu(G_n) + \varepsilon 2^{-n} \\ = \varepsilon + \textstyle\sum_{n=1}^{\infty} \mu(G_n) = \varepsilon + \mu(G),\end{aligned} \tag{2.2}$$

where the last equality is a consequence of the σ-additivity of μ since the sequence $\{G_n\}_{n\in\mathbb{N}}$ is disjoint. From $\sum_{n=1}^{\infty} \mu(G_n) = \mu(G) < \infty$, it furthermore follows that there exists $m \in \mathbb{N}$ such that

$$\textstyle\sum_{n=m+1}^{\infty} \mu(G_n) \leq \tfrac{1}{2}\varepsilon.$$

Now define $K \triangleq \bigcup_{n=1}^{m} K_n$ and observe that K is a compact subset of G and

$$\begin{aligned}\mu(K) \leq \mu(G) = \textstyle\sum_{n=1}^{m} \mu(G_n) + \sum_{n=m+1}^{\infty} \mu(G_n) \leq \sum_{n=1}^{m} \mu(G_n) + \tfrac{1}{2}\varepsilon \\ \leq \textstyle\sum_{n=1}^{m} \mu(K_n) + \tfrac{1}{2}\varepsilon 2^{-n} + \tfrac{1}{2}\varepsilon \leq \varepsilon + \sum_{n=1}^{m} \mu(K_n) = \varepsilon + \mu(K),\end{aligned} \tag{2.3}$$

where we have used that $K_1, \dots, K_m$ are disjoint since $K_n \subset G_n$ for all $n \in \mathbb{N}$ and the sequence $\{G_n\}_{n\in\mathbb{N}}$ is disjoint. Since $\varepsilon > 0$ was chosen arbitrarily, it follows from Equation (2.2) and Equation (2.3) that $G \in \mathfrak{G}$. □

Let us now turn to Kolmogorov's consistency theorem. For this, let us denote $S = \mathbb{R}^d$ and $\mathfrak{S} = \mathfrak{B}(\mathbb{R}^d)$. Now for each $n \in \mathbb{N}$ and each finite subset $\mathcal{F} \triangleq [t_1, \dots, t_n]$ of $\mathcal{T}$, let $\mathbb{P}_{\mathcal{F}}$ be a probability measure on $(S^{\mathcal{F}}, \mathfrak{S}^{\mathcal{F}})$. The question we ask ourselves now is whether there exists a stochastic process

$X = \{X(t)\}_{t\in\mathcal{T}}$ such that $\{\mathbb{P}_{\mathcal{F}} : \mathcal{F} = [t_1,\ldots,t_n] \subset \mathcal{T}, n \in \mathbb{N}\}$ are the finite-dimensional distributions of X. In general, this need not be true as we have not yet specified any **conditions on the probability measures** $\mathbb{P}_{\mathcal{F}}$.

To see what kind of conditions are necessary, let us assume for a second that $\{\mathbb{P}_{\mathcal{F}} : \mathcal{F} = [t_1,\ldots,t_n] \subset \mathcal{T}, n \in \mathbb{N}\}$ really are the finite-dimensional distributions of some stochastic process $X = \{X(t)\}_{t\in\mathcal{T}}$. Let us furthermore fix two finite subsets $\mathcal{F} = [t_1,\ldots,t_n]$ and $\mathcal{G} = [s_1,\ldots,s_m]$ of $\mathcal{T}$ such that $\mathcal{F} \subset \mathcal{G}$. We denote by

$$\pi^{\mathcal{G}}_{\mathcal{F}} : S^{\mathcal{G}} \to S^{\mathcal{F}}, \qquad \{x(t)\}_{t\in\mathcal{G}} \mapsto \{x(t)\}_{t\in\mathcal{F}}$$

the **coordinate projection** from $S^{\mathcal{G}}$ to $S^{\mathcal{F}}$. Let us fix $B \in \mathfrak{S}^{\mathcal{F}}$ of the form

$$B = \bigtimes\nolimits_{t\in\mathcal{F}} B_t \qquad \text{where } B_t \in \mathfrak{S} \text{ for all } t \in \mathcal{F}.$$

Now set $\bar{B} \triangleq \{\pi^{\mathcal{G}}_{\mathcal{F}} \in B\}$ and observe that $\bar{B}$ is given explicitly by

$$\bar{B} = \bigtimes\nolimits_{t\in\mathcal{G}} \bar{B}_t \qquad \text{where } \bar{B}_t = \begin{cases} B_t & \text{if } t \in \mathcal{F}, \\ S & \text{if } t \in \mathcal{G}\setminus\mathcal{F}. \end{cases}$$

Now since $\{\mathbb{P}_{\mathcal{F}} : \mathcal{F} = [t_1,\ldots,t_n] \subset \mathcal{T}, n \in \mathbb{N}\}$ are the finite dimensional distributions of X, it follows that

$$\mathbb{P}\big[\{X(t)\}_{t\in\mathcal{F}} \in B\big] = \mathbb{P}_{\mathcal{F}}[B].$$

On the other hand, since $\bar{B}_t = S$ whenever $t \in \mathcal{G}\setminus\mathcal{F}$ and $\{X(t) \in S\} = \Omega$ for all $t \in \mathcal{T}$, we also have

$$\begin{aligned} \mathbb{P}\big[\{X(t)\}_{t\in\mathcal{F}} \in B\big] &= \mathbb{P}\big[\{X(t)\}_{t\in\mathcal{F}} \in B, X(t) \in \bar{B}_t \text{ for all } t \in \mathcal{G}\setminus\mathcal{F}\big] \\ &= \mathbb{P}\big[\{X(t)\}_{t\in\mathcal{G}} \in \bar{B}\big] = \mathbb{P}_{\mathcal{G}}[\bar{B}] = \mathbb{P}_{\mathcal{G}}\big[\pi^{\mathcal{G}}_{\mathcal{F}} \in B\big], \end{aligned}$$

i.e. for $\{\mathbb{P}_{\mathcal{F}} : \mathcal{F} = [t_1,\ldots,t_n] \subset \mathcal{T}, n \in \mathbb{N}\}$ to be the finite dimensional distributions of a stochastic process X we must necessarily have

$$\mathbb{P}_{\mathcal{F}}[B] = \mathbb{P}_{\mathcal{G}}\big[\pi^{\mathcal{G}}_{\mathcal{F}} \in B\big] \qquad \text{for all } B \in \mathfrak{S}^{\mathcal{F}} \tag{2.4}$$

for any choice of finite sets $\mathcal{F}$ and $\mathcal{G}$ with $\mathcal{F} \subset \mathcal{G} \subset \mathcal{T}$. Equation (2.4) is called **Kolmogorov's Consistency Condition**, and Kolmogorov's consistency theorem below shows that this consistency condition is not just necessary for the existence of the process X, but also sufficient.

Let us think for a second about the strategy to construct a stochastic process $X = \{X(t)\}_{t\in\mathcal{T}}$ given the finite-dimensional distributions $\{\mathbb{P}_\mathcal{F} : \mathcal{F} = [t_1, \ldots, t_n] \subset \mathcal{T}, n \in \mathbb{N}\}$. Observe that we have some **additional freedom** in our construction: We can choose the probability space $(\Omega, \mathfrak{A}, \mathbb{P})$ in whatever way we want and we shall use this to our advantage. If we choose $\Omega = S^\mathcal{T}$ to be the space of functions from $\mathcal{T}$ to S, let the σ-field $\mathfrak{A} = \mathfrak{S}^\mathcal{T}$ be the product σ-field on Ω, we can choose $X = \{X(t)\}_{t\in\mathcal{T}}$ to be the **coordinate process** or **canonical process** given by

$$X(t, \omega) \triangleq \omega(t) \qquad \text{for all } t \in \mathcal{T} \text{ and } \omega \in \Omega.$$

Considering the process X as a path-valued random variable, i.e.

$$X : \Omega \to S^\mathcal{T}, \qquad \omega \mapsto X(\cdot, \omega),$$

we see that X is the **identity mapping** from Ω into itself. Thus choosing the probability measure $\mathbb{P}$ on $(\Omega, \mathfrak{A})$ is the same as choosing the distribution of the process X. Since the distribution of a process is determined uniquely by its finite-dimensional distributions, we should be able to construct $\mathbb{P}$ from the given set of probability measures $\{\mathbb{P}_\mathcal{F} : \mathcal{F} = [t_1, \ldots, t_n] \subset \mathcal{T}, n \in \mathbb{N}\}$.

Theorem 2.9 (Kolmogorov's Consistency Theorem). *Let $S = \mathbb{R}^d$ and $\mathfrak{S} = \mathfrak{B}(\mathbb{R}^d)$. For each $\mathcal{F} = [t_1, \ldots, t_n] \subset \mathcal{T}$, $n \in \mathbb{N}$, assume that $\mathbb{P}_\mathcal{F}$ is a probability measure on $(S^\mathcal{F}, \mathfrak{S}^\mathcal{F})$ and assume that whenever $\mathcal{G} = [s_1, \ldots, s_m] \subset \mathcal{T}$, $m \in \mathbb{N}$, is such that $\mathcal{F} \subset \mathcal{G}$, then Kolmogorov's consistency condition*

$$\mathbb{P}_\mathcal{F}[B] = \mathbb{P}_\mathcal{G}\left[\pi^\mathcal{G}_\mathcal{F} \in B\right] \qquad \textit{for all } B \in \mathfrak{S}^\mathcal{F}$$

holds. Then there exists a probability space $(\Omega, \mathfrak{A}, \mathbb{P})$ as well as a stochastic process $X = \{X(t)\}_{t\in\mathcal{T}}$ on $(\Omega, \mathfrak{A}, \mathbb{P})$ such that the finite-dimensional distributions of X are given by $\{\mathbb{P}_\mathcal{F} : \mathcal{F} = [t_1, \ldots, t_n] \subset \mathcal{T}, n \in \mathbb{N}\}$. ◇

Proof. As outlined above, we choose

$$\Omega \triangleq S^\mathcal{T} \qquad \text{and} \qquad \mathfrak{A} \triangleq \mathfrak{S}^\mathcal{T},$$

and let $X = \{X(t)\}_{t\in\mathcal{T}}$ be the canonical process given by

$$X(t, \omega) \triangleq \omega(t) \qquad \text{for all } t \in \mathcal{T} \text{ and } \omega \in \Omega.$$

Then specifying a probability measure $\mathbb{P}$ on $(\Omega, \mathfrak{A})$ is equivalent to specifying the distribution of X. To construct the measure $\mathbb{P}$, we construct a premeasure $\mathbb{P}^\circ$ from $\{\mathbb{P}_\mathcal{F} : \mathcal{F} = [t_1, \ldots, t_n] \subset \mathcal{T}, n \in \mathbb{N}\}$ and then use Carathéodory's extension theorem.

Step 1: Construction of $\mathbb{P}^\circ$. Given $\mathcal{F} = [t_1, \ldots, t_n] \subset \mathcal{T}$, we introduce the short hand notation $\pi_\mathcal{F} \triangleq \pi^\mathcal{T}_\mathcal{F}$. Now consider the system of sets

$$\begin{aligned}\mathfrak{R} &\triangleq \bigcup_{n\in\mathbb{N}} \bigcup_{\mathcal{F}=[t_1,\ldots,t_n]\subset\mathcal{T}} \sigma(\pi_\mathcal{F}) \\ &= \Big\{\{\pi_\mathcal{F} \in B\} : \mathcal{F} = [t_1, \ldots, t_n] \subset \mathcal{T}, n \in \mathbb{N}, B \in \mathfrak{S}^\mathcal{F}\Big\}.\end{aligned}$$

It is clear that $\sigma(\mathfrak{R}) = \mathfrak{A}$ since $\mathfrak{A} = \mathfrak{S}^\mathcal{T}$ is the smallest σ-field on $S^\mathcal{T}$ containing all finite-dimensional cylinder sets. Moreover, $\mathfrak{R}$ is a ring of sets, i.e. it is nonempty, closed under unions, and closed under taking set differences. Indeed, $\mathfrak{R}$ is non-empty since each $\sigma(\pi_\mathcal{F})$ is nonempty. Let now $\{\pi_{\mathcal{F}_1} \in B_1\}, \{\pi_{\mathcal{F}_2} \in B_2\} \in \mathfrak{R}$ for $\mathcal{F}_1, \mathcal{F}_2 \subset \mathcal{T}$ finite and $B_1 \in \mathfrak{S}^{\mathcal{F}_1}$ and $B_2 \in \mathfrak{S}^{\mathcal{F}_2}$. Setting $\mathcal{F} \triangleq \mathcal{F}_1 \cup \mathcal{F}_2$, we observe that

$$\{\pi_{\mathcal{F}_i} \in B_i\} = \big\{\pi_\mathcal{F} \in \{\pi^\mathcal{F}_{\mathcal{F}_i} \in B_i\}\big\} \qquad i = 1, 2.$$

But then

$$\{\pi_{\mathcal{F}_1} \in B_1\} \cup \{\pi_{\mathcal{F}_2} \in B_2\} = \Big\{\pi_\mathcal{F} \in \big\{\pi^\mathcal{F}_{\mathcal{F}_1} \in B_1\big\} \cup \big\{\pi^\mathcal{F}_{\mathcal{F}_2} \in B_2\big\}\Big\} \in \mathfrak{R}$$

and similarly

$$\{\pi_{\mathcal{F}_1} \in B_1\} \setminus \{\pi_{\mathcal{F}_2} \in B_2\} = \Big\{\pi_\mathcal{F} \in \big\{\pi^\mathcal{F}_{\mathcal{F}_1} \in B_1\big\} \setminus \big\{\pi^\mathcal{F}_{\mathcal{F}_2} \in B_2\big\}\Big\} \in \mathfrak{R}$$

Thus $\mathfrak{R}$ really is a ring. With this, we now define

$$\mathbb{P}^\circ : \mathfrak{R} \to [0, 1], \qquad \mathbb{P}^\circ\big[\pi_\mathcal{F} \in B\big] \triangleq \mathbb{P}_\mathcal{F}[B].$$

We have to be careful here to ensure that $\mathbb{P}^\circ$ is well-defined as

$$\{\pi_\mathcal{F} \in B\} = \big\{\pi_\mathcal{G} \in \{\pi^\mathcal{G}_\mathcal{F} \in B\}\big\} \quad \text{if } \mathcal{G} \subset \mathcal{T} \text{ is finite and contains } \mathcal{F}.$$

However, by Kolmogorov's consistency condition, $\mathbb{P}^\circ$ is well-defined. If we can show that $\mathbb{P}^\circ$ is a premeasure, it follows from Carathéodories extension theorem that $\mathbb{P}^\circ$ extends uniquely to a probability measure $\mathbb{P}$ on $(\Omega, \mathfrak{A})$. Since

$$\mathbb{P}\big[\{X(t)\}_{t\in\mathcal{F}} \in B\big] = \mathbb{P}\big[\big\{\omega \in \Omega : \{\omega(t)\}_{t\in\mathcal{F}} \in B\big\}\big] = \mathbb{P}^\circ\big[\pi_\mathcal{F} \in B\big] = \mathbb{P}_\mathcal{F}[B]$$

for all $B \in \mathfrak{S}^{\mathcal{F}}$ and $\mathcal{F} \subset \mathcal{T}$ finite, showing that $\mathbb{P}^\circ$ is a premeasure thus finishes the proof.

Step 2: We show that $\mathbb{P}^\circ$ is a content. For this, we first observe that $\mathbb{P}^\circ[\emptyset] = 0$. Indeed, for any $t \in \mathcal{T}$ we have $\emptyset = \{\pi_{\{t\}} \in \emptyset\}$ and thus

$$\mathbb{P}^\circ[\emptyset] = \mathbb{P}^\circ\big[\pi_{\{t\}} \in \emptyset\big] = \mathbb{P}_{\{t\}}[\emptyset] = 0.$$

Moreover, $\mathbb{P}^\circ$ is finitely additive. Indeed, let $\mathcal{F}_1, \dots, \mathcal{F}_n$ be finite subsets of $\mathcal{T}$ and $B_1, \dots, B_n$ such that $B_i \in \mathfrak{S}^{\mathcal{F}_i}$, $i = 1, \dots, n$, and such that $\{\pi_{\mathcal{F}_i} \in B_i\}$, $i = 1, \dots, n$, are disjoint. Now set $\mathcal{F} = \bigcup_{i=1}^n \mathcal{F}_i$ and observe that, as above,

$$\{\pi_{\mathcal{F}_i} \in B_i\} = \big\{\pi_{\mathcal{F}} \in \{\pi^{\mathcal{F}}_{\mathcal{F}_i} \in B_i\}\big\} \qquad i = 1, \dots, n.$$

With $\{\pi_{\mathcal{F}_i} \in B_i\}$, $i = 1, \dots, n$, being disjoint, it hence follows that the sets $\{\pi^{\mathcal{F}}_{\mathcal{F}_i} \in B_i\}$, $i = 1, \dots, n$, are disjoint as well. But then

$$\begin{aligned}
\mathbb{P}^\circ\big[\textstyle\bigcup_{i=1}^n \{\pi_{\mathcal{F}_i} \in B_i\}\big] &= \mathbb{P}^\circ\big[\textstyle\bigcup_{i=1}^n \big\{\pi_{\mathcal{F}} \in \{\pi^{\mathcal{F}}_{\mathcal{F}_i} \in B_i\}\big\}\big] \\
&= \mathbb{P}^\circ\big[\pi_{\mathcal{F}} \in \textstyle\bigcup_{i=1}^n \{\pi^{\mathcal{F}}_{\mathcal{F}_i} \in B_i\}\big] \\
&= \mathbb{P}_{\mathcal{F}}\big[\textstyle\bigcup_{i=1}^n \{\pi^{\mathcal{F}}_{\mathcal{F}_i} \in B_i\}\big] \\
&= \textstyle\sum_{i=1}^n \mathbb{P}_{\mathcal{F}}\big[\pi^{\mathcal{F}}_{\mathcal{F}_i} \in B_i\big] \\
&= \textstyle\sum_{i=1}^n \mathbb{P}_{\mathcal{F}_i}\big[B_i\big] \\
&= \textstyle\sum_{i=1}^n \mathbb{P}^\circ\big[\pi_{\mathcal{F}_i} \in B_i\big].
\end{aligned}$$

We have hence argued that $\mathbb{P}^\circ$ is a content.

Step 3: We are left with showing that $\mathbb{P}^\circ$ is σ-additive. Since we already know that $\mathbb{P}^\circ$ is a content and $\mathbb{P}^\circ[A] \leq 1$ for all $A \in \mathfrak{R}$, this is equivalent to showing that $\mathbb{P}^\circ$ is continuous from above in $\emptyset$, i.e. whenever $\{A_n\}_{n\in\mathbb{N}}$ is a sequence in $\mathfrak{R}$ with $A_n \supset A_{n+1}$ for all $n \in \mathbb{N}$ and $A \triangleq \bigcap_{n\in\mathbb{N}} A_n = \emptyset$, then $\lim_{n\to\infty} \mathbb{P}^\circ[A_n] = 0$. Evidently this is equivalent to showing that if $\delta \triangleq \inf_{n\in\mathbb{N}} \mathbb{P}^\circ[A_n] > 0$, then $A \neq \emptyset$.

Since $A_n \in \mathfrak{R}$, there exists $\mathcal{F}_n \subset \mathcal{T}$ finite and $B_n \in \mathfrak{S}^{\mathcal{F}_n}$ such that $A_n = \{\pi_{\mathcal{F}_n} \in B_n\}$. Clearly, we may choose $\mathcal{F}_n$ in such a way that $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for all $n \in \mathbb{N}$. Applying Lemma 2.8 to the probability measure $\mathbb{P}_{\mathcal{F}_n}$, there exists a compact set $K_n \subset B_n$ such that

$$\mathbb{P}^\circ\big[A_n \setminus \{\pi_{\mathcal{F}_n} \in K_n\}\big] = \mathbb{P}^\circ\big[\pi_{\mathcal{F}_n} \in B_n \setminus K_n\big] = \mathbb{P}_{\mathcal{F}_n}\big[B_n \setminus K_n\big] \leq \delta 2^{-n}.$$

Now define

$$L_n \triangleq \bigcap_{k=1}^n \{\pi_{\mathcal{F}_k} \in K_k\} \text{ for all } n \in \mathbb{N} \qquad \text{and} \qquad L \triangleq \bigcap_{n\in\mathbb{N}} L_n.$$

Then, for each $n \in \mathbb{N}$, we observe that

$$L_{n+1} \subset L_n \subset \{\pi_{\mathcal{F}_n} \in K_n\} \subset \{\pi_{\mathcal{F}_n} \in B_n\} = A_n.$$

Now suppose that $x \in A_n \setminus L_n$ for some $n \in \mathbb{N}$. Then $x \in A_k$ for all $k \leq n$ since the sequence $\{A_n\}_{n\in\mathbb{N}}$ is nonincreasing, and $x \notin L_n$, i.e. $x \notin \{\pi_{\mathcal{F}_k} \in K_k\}$ for some $k \leq n$. But then $x \in A_k \setminus \{\pi_{\mathcal{F}_k} \in K_k\}$ for some $k \leq n$ and thus $A_n \setminus L_n \subset \bigcup_{k=1}^n A_k \setminus \{\pi_{\mathcal{F}_k} \in K_k\}$. From this and $\delta \leq \mathbb{P}[A_n]$ it follows that

$$\delta - \mathbb{P}^\circ[L_n] \leq \mathbb{P}^\circ\big[A_n \setminus L_n\big] \leq \textstyle\sum_{k=1}^n \mathbb{P}^\circ\big[A_k \setminus \{\pi_{\mathcal{F}_k} \in K_k\}\big] \leq \delta \sum_{k=1}^n 2^{-k} < \delta$$

and hence $\mathbb{P}^\circ[L_n] > 0$ for all $n \in \mathbb{N}$. In particular, this implies that $L_n \neq \emptyset$ for all $n \in \mathbb{N}$. Now, for each $n \in \mathbb{N}$, we fix some arbitrary $x_n \in L_n$ and observe that this implies

$$x_n \in \{\pi_{\mathcal{F}_k} \in K_k\}, \qquad \text{i.e.} \qquad \pi_{\mathcal{F}_k}(x_n) \in K_k \qquad \text{for all } k \leq n.$$

Put differently, this shows that $\pi_{\mathcal{F}_1}(x_n) \in K_1$ for all $n \in \mathbb{N}$, $\pi_{\mathcal{F}_2}(x_n) \in K_2$ for all $n \geq 2$, and so on. Since K_1 is compact, there exists a subsequence $\{n_j^1\}_{j\in\mathbb{N}} \subset \mathbb{N}$ such that

$$y_1 \triangleq \lim_{j\to\infty} \pi_{\mathcal{F}_1}(x_{n_j^1}) \in K_1 \text{ exists.}$$

Iteratively, for each $k \geq 2$, we find $\{n_j^k\}_{j\in\mathbb{N}} \subset \{n_j^{k-1}\}_{j\in\mathbb{N}}$ with

$$y_k \triangleq \lim_{j\to\infty} \pi_{\mathcal{F}_k}(x_{n_j^k}) \in K_k \text{ exists.}$$

Since $\{n_j^m\}_{j\in\mathbb{N}} \subset \{n_j^k\}_{j\in\mathbb{N}}$ and the projection $\pi_{\mathcal{F}_k}^{\mathcal{F}_m}$ is continuous for each $k, m \in \mathbb{N}$ with $k \leq m$, we observe moreover that

$$\begin{aligned} \pi_{\mathcal{F}_k}^{\mathcal{F}_m}(y_m) &= \pi_{\mathcal{F}_k}^{\mathcal{F}_m}\big(\lim_{j\to\infty} \pi_{\mathcal{F}_m}(x_{n_j^m})\big) \\ &= \lim_{j\to\infty} \pi_{\mathcal{F}_k}^{\mathcal{F}_m}\big(\pi_{\mathcal{F}_m}(x_{n_j^m})\big) = \lim_{j\to\infty} \pi_{\mathcal{F}_k}(x_{n_j^m}) = y_k. \end{aligned}$$

But this can only be the case if there exists $y \in S^{\mathcal{T}}$ such that

$$y_n = \pi_{\mathcal{F}_n}(y) \qquad \text{for all } n \in \mathbb{N},$$

and since $y_n \in K_n$ for each $n \in \mathbb{N}$, this shows that $y \in \{\pi_{\mathcal{F}_n} \in K_n\}$ for all $n \in \mathbb{N}$. By the definition of L, this yields $y \in L \subset A$ and thus A is nonempty. Hence $\mathbb{P}^\circ$ is indeed a premeasure and we conclude. □

It is not very difficult to see that Theorem 2.9 remains valid if we replace $S = \mathbb{R}^d$ and $\mathfrak{S} = \mathfrak{B}(\mathbb{R}^d)$ by any **discrete space** I together with its power set. The main reason for this is that any discrete space may be identified with $\mathbb{Z} \subset \mathbb{R}$. This is important since it allows us to construct I-valued stochastic processes. As a matter of fact, Theorem 2.9 is true for S being any complete, separable metric space and $\mathfrak{S} = \mathfrak{B}(S)$ the Borel σ-field on S.

The main difficulty in applying Kolmogorov's consistency theorem is checking if a given set $\{\mathbb{P}_\mathcal{F} : \mathcal{F} = [t_1, \ldots, t_n] \subset \mathcal{T}, n \in \mathbb{N}\}$ of finite-dimensional distributions is consistent. By induction, it suffices to check that

$$\mathbb{P}_\mathcal{G}\big[\pi^\mathcal{G}_\mathcal{F} \in B\big] = \mathbb{P}_\mathcal{F}[B] \qquad \text{for all } B \in \mathfrak{B}(\mathbb{R}^d)^\mathcal{F}$$

whenever $\mathcal{F} \subset \mathcal{G}$ are finite subsets of $\mathcal{T}$ and $\mathcal{G}$ has **exactly one more element** than $\mathcal{F}$. Moreover, since both $\mathbb{P}_\mathcal{G}\big[\pi^\mathcal{G}_\mathcal{F} \in \cdot\big]$ and $\mathbb{P}_\mathcal{F}$ are probability measures on $\mathfrak{B}(\mathbb{R}^d)^\mathcal{F}$, it suffices to check the consistency for all B contained in some π-system generating $\mathfrak{B}(\mathbb{R}^d)^\mathcal{F}$.

Let us now put Kolmogorov's consistency theorem to use and construct stochastic processes. The first example proves the existence of sequences of independent random variables — a result which is often silently assumed in introductory courses on probability theory, but seldomly proved.

Corollary 2.10 (Existence of Independent Sequences)**.** *For each $n \in \mathbb{N}$, let μ_n be a probability measure on $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$. Then there exists a sequence of independent random variables $\{X_n\}_{n\in\mathbb{N}}$ on some probability space $(\Omega, \mathfrak{A}, \mathbb{P})$ such that X_n has distribution μ_n for each $n \in \mathbb{N}$.* ◇

Proof. We may treat $\{X_n\}_{n\in\mathbb{N}}$ as a stochastic process with finite-dimensional distributions $\{\mathbb{P}_\mathcal{F} : \mathcal{F} = [t_1, \ldots, t_n] \subset \mathbb{N}, n \in \mathbb{N}\}$ given by $\mathbb{P}_\mathcal{F} \triangleq \bigotimes_{t\in\mathcal{F}} \mu_t$ for all $\mathcal{F} = [t_1, \ldots, t_n] \subset \mathbb{N}$, $n \in \mathbb{N}$. By Theorem 2.9, we only have to check consistency. For this, let $\mathcal{F}$ and $\mathcal{G}$ be finite subsets of $\mathbb{N}$ with $\mathcal{F} \subset \mathcal{G}$, and for each $t \in \mathcal{F}$ fix some $B_t \in \mathfrak{B}(\mathbb{R}^d)$. Then $B \triangleq \bigtimes_{t\in\mathcal{F}} B_t \in \mathfrak{B}(\mathbb{R}^d)^\mathcal{F}$ and

$$\{\pi^\mathcal{G}_\mathcal{F} \in B\} = \bigtimes_{t\in\mathcal{G}} B_t, \qquad \text{where } B_t \triangleq \mathbb{R}^d \text{ whenever } t \notin \mathcal{F}.$$

But then, since $\mu_n(\mathbb{R}^d) = 1$ for all $n \in \mathbb{N}$, it follows that

$$\mathbb{P}_\mathcal{F}[B] = \prod_{t\in\mathcal{F}} \mu_t(B_t) = \prod_{t\in\mathcal{G}} \mu_t(B_t) = \mathbb{P}_\mathcal{G}\big[\bigtimes_{t\in\mathcal{G}} B_t\big] = \mathbb{P}_\mathcal{G}\big[\pi^\mathcal{G}_\mathcal{F} \in B\big],$$

implying that $\{\mathbb{P}_\mathcal{F} : \mathcal{F} = [t_1, \ldots, t_n] \subset \mathbb{N}, n \in \mathbb{N}\}$ is consistent. □

Exercise 25. *Let $(\Omega, \mathfrak{A}, \mathbb{P})$ be a finite probability space, i.e. $\Omega = \{\omega_1, \dots, \omega_n\}$ for some $n \in \mathbb{N}$. Show that there exists no nontrivial sequence $\{Z_n\}_{n\in\mathbb{N}}$ of independent random variables on this probability space.* ◇

Exercise 26. *Let $\Omega = [0, 1]$, $\mathfrak{A} = \mathfrak{B}([0, 1])$, and $\mathbb{P}$ be the Lebesgue measure on $[0, 1]$. Construct a sequence $\{Z_n\}_{n\in\mathbb{N}}$ of independent Bernoulli distributed random variables with parameter $p = 1/2$ on $(\Omega, \mathfrak{A}, \mathbb{P})$.* ◇

The next example is the construction of a **raw Brownian motion**, i.e. a stochastic process $X = \{X(t)\}_{t\in[0,\infty)}$ which has the same distribution as a Brownian motion (but need not necessarily be continuous).

Corollary 2.11 (Existence of Raw Brownian Motion)**.** *There exists a stochastic process $X = \{X(t)\}_{t\in[0,\infty)}$ on some probability space $(\Omega, \mathfrak{A}, \mathbb{P})$ which satisfies the properties (W1), (W2) with respect to its natural filtration $\mathfrak{F}^X$, and (W3) of Definition 2.1.* ◇

Proof. We already know that the finite-dimensional distributions of a process satisfying (W1), (W2), and (W3) are multivariate normal. More precisely, we look for a stochastic process $X = \{X(t)\}_{t\in[0,\infty)}$ such that whenever $\mathcal{F} = [t_1, \dots, t_n] \subset [0, \infty)$, $n \in \mathbb{N}$, then $(X(t_1), \dots, X(t_n))$ is multivariate normally distributed with mean zero and covariance matrix $\Sigma^{\mathcal{F}} = (\Sigma^{\mathcal{F}}_{i,j})_{i,j=1,\dots,n}$ given by

$$\Sigma^{\mathcal{F}}_{i,j} = t_i \wedge t_j, \qquad i, j = 1, \dots, n.$$

Denote by $\mathbb{P}_{\mathcal{F}}$ the probability measure on $(\mathbb{R}^{\mathcal{F}}, \mathfrak{B}(\mathbb{R})^{\mathcal{F}})$ corresponding to this distribution, the existence of which is guaranteed since $\Sigma^{\mathcal{F}}$ is positive semidefinite: If $a = (a_1, \dots, a_n) \in \mathbb{R}^{\mathcal{F}}$, then

$$\textstyle\sum_{i,j=1}^n a_i a_j \Sigma^{\mathcal{F}}_{i,j} = \sum_{i,j=1}^n a_i a_j (t_i \wedge t_j) = \int_0^\infty \big(\sum_{i=1}^n a_i \mathbb{1}_{[0,t_i]}\big)^2 \mathrm{d}x \geq 0.$$

It is straightforward to check that this yields a set of finite-dimensional distributions satisfying Kolmogorov's consistency condition, and hence Theorem 2.9 yields the existence of a stochastic process $X = \{X(t)\}_{t\in[0,\infty)}$ with these finite-dimensional distributions. Observe that the distribution of $X(0)$ degenerates, i.e. it has a Dirac distribution at zero, meaning that $X(0) = 0$ almost surely. Clearly, it makes no difference for the distribution of X if we set replace $X(0)$ by constant zero, and hence (W1) is satisfied. The properties (W2) with respect to $\mathfrak{F}^X$ and (W3) are argued for exactly as in the proof of Theorem 2.3 (Brownian Motion as a Gaussian Process). □

A closer inspection of the proof of Corollary 2.11 shows that we make use of the specific covariance structure of Brownian motion only to guarantee that $\Sigma^{\mathcal{F}}$ is positive semidefinite, which in turn is only needed for the multivariate normal distribution to exist in the first place. The same argument can hence be used to construct raw Gaussian processes.

Exercise 27. *Show the following equivalence: There exists a centered Gaussian process $X = \{X(t)\}_{t\in\mathcal{T}}$ with covariance function $\Gamma : \mathcal{T} \times \mathcal{T} \to \mathbb{R}$ if and only if for all $\mathcal{F} = [t_1, \ldots, t_n] \subset \mathcal{T}$, $n \in \mathbb{N}$, the matrix $\Sigma^{\mathcal{F}} = (\Sigma^{\mathcal{F}}_{i,j})_{i,j=1,\ldots,n}$ with*

$$\Sigma^{\mathcal{F}}_{i,j} \triangleq \Gamma(t_i, t_j), \qquad i, j = 1, \ldots, n,$$

is positive semidefinite. ◇

Exercise 28. *Show that there exist raw versions (i.e. not necessarily continuous) of the following Gaussian processes: Brownian bridge, Ornstein-Uhlenbeck process, and fractional Brownian motion.* ◇

A word of **warning** is in order here: The processes constructed in Corollary 2.11, Exercise 27, and Exercise 28 are in general **not continuous**, so we have to put in some extra effort to construct continuous modifications of these processes without changing their distributions.

A natural idea to get a continuous modification of raw Brownian motion would be to argue that the probability measure obtained from Kolmogorov's consistency theorem is concentrated on the subset of continuous functions. However, as the following exercise makes evident, this set is **not measurable** and hence it does not even make sense to assign any probability to it. Thus, we have to come up with a different idea.

Exercise 29. *Let us denote by $C \subset \mathbb{R}^{[0,\infty)}$ the set of continuous functions from $[0,\infty)$ to $\mathbb{R}$. Show that $C \notin \mathfrak{B}(\mathbb{R})^{[0,\infty)}$.* ◇

2.5 The Kolmogorov-Čentsov Continuity Theorem

So how can we construct Brownian motion W from raw Brownian motion X? Clearly, if we modify each $X(t)$, $t \in [0,\infty)$, on a nullset, the finite-dimensional distributions of X remain unchanged. The idea is hence to make these modifications in a way to end up with continuous paths. It is quite easy to imagine that this will not work for any arbitrary stochastic process X, and hence we need to identify a condition under which this is possible. This condition is provided by the **Kolmogorov-Čentsov Continuity Theorem**, and involves a condition on the moments of the increments of X. We start our endeavors with a definition.

Definition 2.12 (Modification; Indistinguishability). *Let $X = \{X(t)\}_{t\in\mathcal{T}}$ and $Y = \{Y(t)\}_{t\in\mathcal{T}}$ be two stochastic processes defined on the same probability space. We say that Y is a **modification** of X if*

$$X(t) = Y(t) \text{ almost surely for each } t \in \mathcal{T}.$$

*Moreover, we say that X and Y are **indistinguishable** if*

$$X(t) = Y(t) \text{ for all } t \in \mathcal{T} \text{ almost surely}$$

in which case we also write $X = Y$ a.s. for brevity. ◇

Y being a modification of X means that for each $t \in \mathcal{T}$, there exists a nullset $N_t \in \mathfrak{A}$ such that $X(t,\omega) = Y(t,\omega)$ for all $\omega \notin N_t$. It is crucial to understand that the nullset N_t depends on t. If X and Y are indistinguishable, there exist a universal nullset $N \in \mathfrak{A}$ such that $X(t,\omega) = Y(t,\omega)$ for all $t \in \mathcal{T}$ and all $\omega \notin N$, i.e. almost all paths of X and Y coincide. In particular, the notion of indistinguishability is stronger than the notion of modification, i.e. if X and Y are indistinguishable, then Y is also a modification of X. Moreover, in both cases, X and Y have the same finite-dimensional distributions as

$$\big(X(t_1), \dots, X(t_n)\big) = \big(Y(t_1), \dots, Y(t_n)\big) \text{ a.s.} \quad \text{for all } t_1, \dots, t_n \in \mathcal{T}, n \in \mathbb{N}.$$

If $\mathcal{T}$ is countable, then $\bigcup_{t\in\mathcal{T}} N_t$ is still a nullset and the two notions coincide. If $\mathcal{T}$ is uncountable this need not be the case.

Exercise 30. *Show that indistinguishability is stronger than the notion of modification by constructing processes $X = \{X(t)\}_{t\in[0,1]}$ and $Y = \{Y(t)\}_{t\in[0,1]}$ such that Y is a modification of X, but X and Y are not indistinguishable.* ◇

If, on the other hand, X and Y are, say, right continuous, then the paths of X and Y are determined by the values at countably many time points. In this situation, it is not surprising that modifications are indistinguishable.

Exercise 31. *Let $Y = \{Y(t)\}_{t\in[0,\infty)}$ be a modification of $X = \{X(t)\}_{t\in[0,\infty)}$. Show that if X and Y are either both left continuous or both right continuous, then they are indistinguishable.* ◇

It will turn out that Brownian motion can be constructed as a modification of raw Brownian motion. To arrive at this result, we first need some preliminary notation and basic results on increments on **dyadic rational numbers**.

Definition 2.13 (Dyadic Rationals; Dyadic Neighbors; Dyadic Modulus). *For each $n \in \mathbb{N}$, we define the* ***dyadic rationals*** *via*

$$\mathbb{D}_n \triangleq \big\{k2^{-n} \in [0,1) : k = 0,\dots,2^n - 1\big\} \qquad \text{and} \qquad \mathbb{D}_\infty \triangleq \textstyle\bigcup_{n\in\mathbb{N}} \mathbb{D}_n.$$

We say that $s,t \in \mathbb{D}_n$ are ***neighbors in $\mathbb{D}_n$*** *if $|s-t| \le 2^{-n}$. If (S,d) is a metric space and $f : \mathbb{D}_\infty \to S$ is an arbitrary function, then we call*

$$\varpi_f : (0,\infty) \to [0,\infty], \qquad \varpi_f(\delta) \triangleq \sup_{|s-t|\le\delta} d\big(f(s), f(t)\big)$$

the ***dyadic modulus*** *of f.* ◇

Note that for any $t \in \mathbb{D}_n$, there are at most three neighbors in $\mathbb{D}_n$: $t - 2^{-n}$, t, and $t + 2^{-n}$. Moreover, $\mathbb{D}_n$ consists of exactly 2^n points. Finally, for $p \in (0,1]$ given, a function $f : \mathbb{D}_\infty \to S$ is p-Hölder continuous whenever

$$\varpi_f(2^{-n}) \le C2^{-np} \qquad \text{for eventually all } n \in \mathbb{N}.$$

Indeed, choose n large enough such that the above estimate holds and choose $s,t \in \mathbb{D}_\infty$ with $0 < |s-t| \le 2^{-n}$. Then there exists $m \in \mathbb{N}$ with $m \ge n$ such that

$$2^{-(m+1)} \le |s-t| \le 2^{-m}.$$

But then it follows that

$$d\big(f(s), f(t)\big) \le \varpi_f\big(2^{-m}\big) \le C2^{-mp} = 2^pC2^{-(m+1)p} \le 2^pC|s-t|^p.$$

But this means that f is p-Hölder continuous.

Lemma 2.14 (Estimate on the Dyadic Modulus). *If (S,d) is a metric space and $f : \mathbb{D}_\infty \to S$, then*

$$\varpi_f\big(2^{-n}\big) \le 3\textstyle\sum_{k=n}^{\infty} \varpi_f^k \qquad \text{for all } n \in \mathbb{N},$$

where, for each $k \in \mathbb{N}$, the constant ϖ_f^k is defined as

$$\varpi_f^k \triangleq \sup\big\{d\big(f(s), f(t)\big) : s,t \in \mathbb{D}_k \text{ are neighbors in } \mathbb{D}_k\big\}.$$ ◇

Proof. For each $n \in \mathbb{N}$, we define $u_n : \mathbb{D}_\infty \to \mathbb{D}_n$ by

$$u_n(t) \triangleq \max\{k2^{-n} : k \in \mathbb{N}_0, k2^{-n} \leq t\}.$$

Clearly, $u_n(t)$ is simply the left neighbor in $\mathbb{D}_n$ of t. In particular, it is easily seen that $u_{n+1}(t)$ and $u_n(t)$ are neighbors in $\mathbb{D}_{n+1}$ and $u_n(t) \leq u_{n+1}(t) \leq t$. Moreover, it is clear that $u_n(t) = t$ if $t \in \mathbb{D}_n$, implying that for every $t \in \mathbb{D}_\infty$ there exists some $N \in \mathbb{N}$ (depending on t) such that $u_n(t) = t$ for all $n \geq N$. From this and the triangle inequality, it follows that

$$\begin{aligned} d\big(f(t), f(u_n(t))\big) = d\big(f(u_N(t)), f(u_n(t))\big) &\leq \textstyle\sum_{k=n}^{N-1} d\big(f(u_k(t)), f(u_{k+1}(t))\big) \\ &\leq \textstyle\sum_{k=n}^{\infty} d\big(f(u_k(t)), f(u_{k+1}(t))\big) \\ &\leq \textstyle\sum_{k=n}^{\infty} \varpi_f^k \end{aligned}$$

for all $n \in \mathbb{N}$. Since the last estimate does not depend on t any longer, this shows that the estimate holds uniformly in $t \in \mathbb{D}_\infty$. Now suppose that $s, t \in \mathbb{D}_\infty$ are such that $|s-t| \leq 2^{-n}$. Then $u_n(s)$ and $u_n(t)$ must be neighbors in $\mathbb{D}_n$ and hence $d(f(u_n(s)), f(u_n(t))) \leq \varpi_f^n$. But then

$$\begin{aligned} d\big(f(s), f(t)\big) &\leq d\big(f(s), f(u_n(s))\big) + d\big(f(u_n(s)), f(u_n(t))\big) + d\big(f(u_n(t)), f(t)\big) \\ &\leq 2\textstyle\sum_{k=n}^{\infty} \varpi_f^k + \varpi_f^n \leq 3\textstyle\sum_{k=n}^{\infty} \varpi_f^k. \end{aligned}$$

Again, since the last estimate no longer depends on the particular choice of s and t, this implies the result. □

We now put this estimate to use by showing that every stochastic process defined on $\mathcal{T} = \mathbb{D}_\infty$ satisfying a suitable moment condition must necessarily have Hölder continuous paths.

Lemma 2.15 (Hölder Continuity from Moments). *Let $X = \{X(t)\}_{t \in \mathbb{D}_\infty}$ be a stochastic process taking values in a metric space (S, d). Suppose that there exist constants $\alpha, \beta, \gamma > 0$ such that*

$$\mathrm{E}\big[d\big(X(s), X(t)\big)^\alpha\big] \leq \gamma |s-t|^{1+\beta} \qquad \textit{for all } s, t \in \mathbb{D}_\infty.$$

Then, for every $p \in (0, \beta/\alpha)$ with $p \leq 1$, there exists an event $N \in \mathfrak{A}$ with $\mathbb{P}[N] = 0$ such that $t \mapsto X(t, \omega)$ is p-Hölder continuous on $\mathbb{D}_\infty$ for all $\omega \in N^c$. ◇

Proof. Step 1: Let $p \in (0, \beta/\alpha)$ with $p \leq 1$ and $\theta > 0$ such that $p + \theta < \beta/\alpha$. For each $s, t \in \mathbb{D}_\infty$ with $s \neq t$, the Markov inequality shows that

$$\begin{aligned}\mathbb{P}\Big[\tfrac{d(X(s),X(t))}{|s-t|^p} \geq |s-t|^\theta\Big] &= \mathbb{P}\big[d\big(X(s),X(t)\big)^\alpha \geq |s-t|^{\alpha(p+\theta)}\big] \\ &\leq |s-t|^{-\alpha(p+\theta)}\mathrm{E}\big[d\big(X(s),X(t)\big)^\alpha\big] \\ &\leq \gamma|s-t|^{1+\beta-\alpha(p+\theta)}.\end{aligned}$$

Since $p+\theta < \beta/\alpha$, it follows that $\kappa \triangleq \beta - \alpha(p+\theta) > 0$. Now for each $n \in \mathbb{N}$, define an event

$$\begin{aligned}A_n &\triangleq \Big\{\tfrac{d(X(s),X(t))}{|s-t|^p} \geq 2^{-n\theta} \text{ for some distinct } s,t \in \mathbb{D}_n \text{ neighbors in } \mathbb{D}_n\Big\} \\ &= \textstyle\bigcup_{s\in\mathbb{D}_n}\bigcup_{t\in\mathbb{D}_n,|s-t|=2^{-n}}\Big\{\tfrac{d(X(s),X(t)}{|s-t|^p} \geq 2^{-n\theta} = |s-t|^\theta\Big\}.\end{aligned}$$

Recall that there are at most 2^n points $s \in \mathbb{D}_n$, and there are at most two points $t \in \mathbb{D}_n$ with $|s-t| = 2^{-n}$. This implies that

$$\mathbb{P}[A_n] \leq 2 \cdot 2^n\gamma(2^{-n})^{1+\beta-\alpha(p+\theta)} = \gamma 2^{1+n-n(1+\kappa)} = \gamma 2^{1-n\kappa} \quad \text{for all } n \in \mathbb{N}.$$

In particular, $\sum_{n=1}^\infty \mathbb{P}[A_n] < \infty$ since $\kappa > 0$, and thus the Borel-Cantelli lemma implies that

$$N \triangleq \limsup_{n\to\infty} A_n = \textstyle\bigcap_{n=1}^\infty\bigcup_{m=n}^\infty A_m \in \mathfrak{A} \qquad \text{is a set of probability zero.}$$

Step 2: p-Hölder continuity outside of N. Let us fix $\omega \in N^c$ and write $f \triangleq X(\cdot,\omega)$ for brevity. Since $\omega \notin N$, there exists some $N \in \mathbb{N}$ such that $\omega \notin A_n$ for all $n \geq N$, i.e. for all $n \in \mathbb{N}$ with $n \geq N$ it holds that

$$\tfrac{d(f(s),f(t))}{|s-t|^p} < 2^{-n\theta} \qquad \text{whenever } s,t \in \mathbb{D}_n \text{ are neighbors in } \mathbb{D}_n \text{ and } s \neq t.$$

This, however, simply means that $\varpi_f^n \leq 2^{-n(p+\theta)}$ for all $n \geq N$. But then Lemma 2.14 (Estimate on the Dyadic Modulus) shows that

$$\varpi_f\big(2^{-n}\big) \leq 3\textstyle\sum_{k=n}^\infty \varpi_f^k \leq 3\sum_{k=n}^\infty 2^{-k(p+\theta)} = 2^{-n(p+\theta)}\frac{3}{1-2^{-(p+\theta)}} \leq C2^{-np}$$

for all $n \geq N$, where $C \triangleq \frac{3}{1-2^{-(p+\theta)}}$. But then f is p-Hölder continuous. □

The next step is to show that any Hölder continuous function f on $\mathbb{D}_\infty$ can be extended to a Hölder continuous function on $[0,1]$.

Lemma 2.16 (Hölder Continuous Extension). *Suppose that (S, d) is a complete metric space and let $n \in \mathbb{N}$. Assume that there exists $p \in (0,1]$ such that $f : 2^n\mathbb{D}_\infty \to S$ is p-Hölder continuous. Then there exists a uniquely determined p-Hölder continuous extension $g : [0, 2^n] \to S$ of f.* ◇

Proof. Step 1: Construction of g. Let $t \in [0, 2^n]$ and choose a sequence $\{s_k\}_{k\in\mathbb{N}} \subset 2^n\mathbb{D}_\infty$ converging to t. We claim that $\{f(s_k)\}_{k\in\mathbb{N}}$ is a Cauchy sequence in S. Indeed, let $\varepsilon > 0$ be given. Since f is p-Hölder continuous on $2^n\mathbb{D}_\infty$, there exist $\delta, C > 0$ such that

$$d\big(f(r), f(s)\big) \leq C|r-s|^p \qquad \text{for all } r, s \in 2^n\mathbb{D}_\infty \text{ with } |r-s| \leq \delta.$$

Since the sequence $\{s_k\}_{k\in\mathbb{N}}$ converges, it must be Cauchy, and hence there exists $K \in \mathbb{N}$ such that $|s_k - s_j| < \min\{(\varepsilon/C)^{1/p}, \delta\}$ for all $k, j \geq K$. But then

$$d\big(f(s_k), f(s_j)\big) \leq C|s_k - s_j|^p < \varepsilon \qquad \text{for all } k, j \geq K,$$

showing that $\{f(s_k)\}_{k\in\mathbb{N}}$ is Cauchy. Since (S, d) is complete, this implies that the sequence converges. Now set

$$g(t) \triangleq \lim_{k\to\infty} f(s_k).$$

Observe that the right hand side is equal to $f(t)$ if $t \in 2^n\mathbb{D}_\infty$ by the Hölder continuity of f, and hence g is an extension of f.

Step 2: We show that $g(t)$ does not depend on the particular choice of the sequence $\{s_k\}_{k\in\mathbb{N}}$. Indeed, let $\{\bar{s}_k\}_{k\in\mathbb{N}} \subset 2^n\mathbb{D}_\infty$ be another sequence converging to t. Then $\{r_k\}_{k\in\mathbb{N}}$ defined by

$$r_{2k-1} \triangleq s_k \qquad \text{and} \qquad r_{2k} \triangleq \bar{s}_k \qquad \text{for all } k \in \mathbb{N}$$

also converges to t and contains $\{s_k\}_{k\in\mathbb{N}}$ and $\{\bar{s}_k\}_{k\in\mathbb{N}}$ as subsequences. Now as in the first step, we see that $\{f(r_k)\}_{k\in\mathbb{N}}$ is Cauchy and hence converges. But then any subsequence converges to the same point and thus

$$g(t) = \lim_{k\to\infty} f(s_k) = \lim_{k\to\infty} f(r_k) = \lim_{k\to\infty} f(\bar{s}_k).$$

Step 3: Uniqueness of g. Let $\bar{g}$ be another p-Hölder continuous extension of f. Then g and $\bar{g}$ must coincide on $2^n\mathbb{D}_\infty$ since they both coincide with f

on this set. If $t \in [0, 2^n] \setminus 2^n\mathbb{D}_\infty$, there exists a sequence $\{s_k\}_{k\in\mathbb{N}} \subset 2^n\mathbb{D}_\infty$ converging to t. But then, since both g and $\bar{g}$ are continuous, it follows that

$$g(t) = \lim_{k\to\infty} g(s_k) = \lim_{k\to\infty} f(s_k) = \lim_{k\to\infty} \bar{g}(s_k) = \bar{g}(t),$$

i.e. $g = \bar{g}$ everywhere on $[0, 2^n]$.

Step 4: We are left with showing that g is p-Hölder continuous on $[0, 2^n]$. Since f is p-Hölder continuous, there exist $\delta, C > 0$ such that

$$d\big(f(s), f(t)\big) \leq C|s-t|^p \qquad \text{for all } s,t \in 2^n\mathbb{D}_\infty \text{ with } |s-t| \leq \delta.$$

Now let $s, t \in [0, 2^n]$ with $|s-t| \leq \delta/2$ and choose sequences $\{s_k\}_{k\in\mathbb{N}}$ and $\{t_k\}_{k\in\mathbb{N}}$ in $2^n\mathbb{D}_\infty$ converging to s and t, respectively, and $|s_k - t_k| \leq \delta$ for all $k \in \mathbb{N}$. Then

$$\begin{aligned} d\big(g(s), g(t)\big) &\leq d\big(g(s), f(s_k)\big) + d\big(f(s_k), f(t_k)\big) + d\big(f(t_k), g(t)\big) \\ &\leq d\big(g(s), f(s_k)\big) + d\big(f(t_k), g(t)\big) + C|s_k - t_k|^p \end{aligned}$$

for all $k \in \mathbb{N}$. Taking the limit $k \to \infty$ and using that s and t where chosen arbitrarily thus shows that

$$d\big(g(s), g(t)\big) \leq C|s-t|^p \qquad \text{for all } s,t \in [0, 2^n] \text{ with } |s-t| \leq \delta/2,$$

i.e. g is p-Hölder continuous. □

Putting the previous lemmas together allows us to prove the Kolmogorov-Čentsov continuity theorem.

Theorem 2.17 (Kolmogorov-Čentsov Continuity Theorem)**.** *Let (S, d) be a complete metric space and $X = \{X(t)\}_{t\in[0,\infty)}$ be an S-valued stochastic process. Assume that there exist constants $\alpha, \beta, \gamma > 0$ such that*

$$\mathrm{E}\Big[d\big(X(s), X(t)\big)^\alpha\Big] \leq \gamma|s-t|^{1+\beta} \qquad \textit{for all } s,t \in [0,\infty).$$

Then, for every $p \in (0, \beta/\alpha)$ with $p \leq 1$, there exists a p-Hölder continuous modification $Y = \{Y(t)\}_{t\in[0,\infty)}$ of X. ◇

Proof. We first construct Y and then show that Y is indeed a modification of X. We keep $p \in (0, \beta/\alpha)$ with $p \leq 1$ fixed.

Step 1: Construction of Y. Fix $n \in \mathbb{N}$ and consider the stochastic process $X^n = \{X^n(t)\}_{t \in \mathbb{D}_\infty}$ given by

$$X^n(t) \triangleq X(t2^n) \qquad \text{for all } t \in \mathbb{D}_\infty.$$

It is clear that for all $s, t \in \mathbb{D}_\infty$

$$\mathrm{E}\big[d\big(X^n(s), X^n(t)\big)^\alpha\big] = \mathrm{E}\big[d\big(X(s2^n), X(t2^n)\big)^\alpha\big] \leq \gamma 2^{(1+\beta)n}|s-t|^{1+\beta}$$

and hence Lemma 2.15 (Hölder Continuity From Moments) shows that there exists an event $N_n \in \mathfrak{A}$ with $\mathbb{P}[N_n] = 0$ such that $X^n(\cdot, \omega)$ is p-Hölder continuous on $\mathbb{D}_\infty$ for all $\omega \in N_n^c$. This however means that $X(\cdot, \omega)$ is p-Hölder continuous on $2^n\mathbb{D}_\infty$ for all $\omega \in N_n^c$ since, for the p-Hölder continuity constants $C, \delta > 0$ of $X^n(\cdot, \omega)$,

$$\begin{aligned} d\big(X(s,\omega), X(t,\omega)\big) &= d\big(X^n(s2^{-n},\omega), X^n(t2^{-n},\omega)\big) \\ &\leq C\big|s2^{-n} - t2^{-n}\big|^p = C2^{-np}\big|s-t\big|^p \end{aligned}$$

for all $s, t \in 2^n\mathbb{D}_\infty$ with $|s-t| \leq \delta 2^n$. By Lemma 2.16, for each such ω, there exists a unique p-Hölder continuous extension $Y_n(\cdot, \omega) : [0, 2^n] \to S$ of $X(\cdot, \omega) : 2^n\mathbb{D}_\infty \to S$. Now set $N \triangleq \bigcup_{n \in \mathbb{N}} N_n \in \mathfrak{A}$ and observe that $\mathbb{P}[N] = 0$. Due to the uniqueness of the extension and since $2^m\mathbb{D}_\infty \subset 2^n\mathbb{D}_\infty$ whenever $m, n \in \mathbb{N}$ with $m \leq n$, we observe that

$$Y_m(t,\omega) = Y_n(t,\omega) \text{ for all } t \in [0, 2^m] \text{ and } \omega \in N^c \text{ if } m, n \in \mathbb{N} \text{ with } m \leq n.$$

Denoting by x some arbitrary element of S, we define $Y = \{Y(t)\}_{t \in [0,\infty)}$ by setting $Y(t,\omega) = x$ for all $t \in [0,\infty)$ if $\omega \in N$ and

$$\begin{aligned} Y(t,\omega) &\triangleq x && \text{whenever } t \in [0,\infty) \text{ and } \omega \in N \text{ and} \\ Y(t,\omega) &\triangleq Y_n(t,\omega) && \text{whenever } t \in [0, 2^n] \text{ and } \omega \in N^c. \end{aligned}$$

Step 2: We are left with showing that Y is a stochastic process and a modification of X. By construction of Y we have

$$Y(s) = X(s)\mathbb{1}_{N^c} + x\mathbb{1}_N \qquad \text{for all } s \in \textstyle\bigcup_{n \in \mathbb{N}} 2^n\mathbb{D}_\infty.$$

In particular, for every such s, we have $Y(s) = X(s)$ a.s. and $Y(s)$ is $\mathfrak{A}$-measurable since $X(s)$ is $\mathfrak{A}$-measurable and $N \in \mathfrak{A}$. For any $t \in [0, \infty)$, there exists $\{s_k\}_{k\in\mathbb{N}} \subset \bigcup_{n\in\mathbb{N}} 2^n \mathbb{D}_\infty$ converging to t. Continuity of Y implies

$$Y(t, \omega) = \lim_{k\to\infty} Y(s_k, \omega) \qquad \text{for all } \omega \in \Omega$$

and thus $Y(t)$ is $\mathfrak{A}$-measurable for all $t \in [0, \infty)$ and therefore Y is a stochastic process. It remains to show that $Y(t) = X(t)$ almost surely. For this, we first observe that the Markov inequality and the moment estimate show that

$$\begin{aligned}\mathbb{P}\big[d\big(Y(s_k), X(t)\big) \geq \varepsilon\big] &= \mathbb{P}\big[d\big(X(s_k), X(t)\big) \geq \varepsilon\big] \\ &= \mathbb{P}\big[d\big(X(s_k), X(t)\big)^\alpha \geq \varepsilon^\alpha\big] \\ &\leq \varepsilon^{-\alpha}\mathrm{E}\big[d\big(X(s_k), X(t)\big)^\alpha\big] \leq \varepsilon^{-\alpha}\gamma|s_k - t|^{1+\beta}\end{aligned}$$

for all $\varepsilon > 0$. Since the right hand side converges to zero, it follows that $Y(s_k)$ converges in probability to $X(t)$. But $Y(s_k, \omega)$ converges to $Y(t, \omega)$ for all $\omega \in \Omega$, hence also in probability. Since limits with respect to convergence in probability are almost surely unique, we must have $Y(t) = X(t)$ a.s. and thus Y is a modification of X and the proof is complete. □

Observe that if X is $\mathfrak{F}$-adapted, it is in general **not true** that the modification Y is also $\mathfrak{F}$-adapted since we do not necessarily have $N \in \mathfrak{F}(t)$ for all $t \in [0, \infty)$. If we want to ensure that the continuous modification Y is adapted, we need an additional assumption on the filtration $\mathfrak{F}$.

Definition 2.18 (Complete Filtration). *We say that a filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$ is **complete** if it contains all $\mathbb{P}$-nullsets in $\mathfrak{A}$, i.e.*

$$N \in \mathfrak{F}(t) \text{ for all } t \in \mathcal{T} \text{ whenever } N \in \mathfrak{A} \text{ satisfies } \mathbb{P}[N] = 0.$$ ◇

If the filtration $\mathfrak{F}$ is complete and X is $\mathfrak{F}$-adapted, it is straightforward to see that any modification Y of X is again $\mathfrak{F}$-adapted. We leave the prove as an exercise.

Exercise 32. *Let $Y = \{Y(t)\}_{t\in\mathcal{T}}$ be a modification of $X = \{X(t)\}_{t\in\mathcal{T}}$ and assume that X is adapted to a complete filtration $\mathfrak{F} = \{\mathfrak{F}(t)\}_{t\in\mathcal{T}}$. Show that Y is $\mathfrak{F}$-adapted as well.* ◇

2.6 Existence of Brownian Motion

At this point, existence of Brownian motion requires little more effort than piecing together the results of the previous two sections.

Theorem 2.19 (Existence of Brownian Motion)**.** *Brownian motion exists. Moreover, almost every path of Brownian motion is p-Hölder continuous for all $p \in (0, 1/2)$.* ◇

Proof. Step 1: Existence. Let $X = \{X(t)\}_{t\in[0,\infty)}$ denote the raw Brownian motion process constructed in Corollary 2.11 and fix $s, t \in [0, \infty)$ with $s < t$. Then $X(t)-X(s)$ has normal distribution with mean zero and variance $t-s$. Thus, if Z denotes a standard normal random variable,

$$\mathrm{E}\big[\big|X(t) - X(s)\big|^n\big] = (t - s)^{n/2}\mathrm{E}\big[|Z|^n\big] \qquad \text{for all } n \in \mathbb{N}.$$

For $n \geq 3$, setting $\alpha \triangleq n$, $\beta \triangleq n/2 - 1 > 0$, and $\gamma \triangleq \mathrm{E}[|Z|^n] < \infty$, this can equivalently be written as

$$\mathrm{E}\big[\big|X(t) - X(s)\big|^\alpha\big] \leq \gamma|t - s|^{1+\beta},$$

and thus Theorem 2.17 (Kolmogorov-Čentsov) implies that X admits a continuous modification $W = \{W(t)\}_{t\in[0,\infty)}$. Since $\{W(0) \neq 0\} \in \mathfrak{A}$ is a nullset, we may without loss of generality assume that W satisfies (W1) (set W to zero on $\{W(0) \neq 0\}$ otherwise). Since X and W share the same finite-dimensional distributions, it follows moreover that W satisfies (W2) (with respect to its natural filtration) and (W3). (W4) is clear. Thus W is a Brownian motion.

Step 2: Hölder continuity. Let W be an arbitrary Brownian motion. Using the same arguments as in the first step, it follows that W admits a modification $\bar{W} = \{\bar{W}(t)\}_{t\in[0,\infty)}$ which is p-Hölder continuous for any p satisfying

$$p < \tfrac{\beta}{\alpha} = \tfrac{n/2-1}{n} = \tfrac{1}{2} - \tfrac{1}{n} \qquad \text{for some } n \in \mathbb{N},$$

i.e. for all $p < 1/2$. But since both W and $\bar{W}$ are continuous it follows from Exercise 31 that W and $\bar{W}$ are indistinguishable, i.e. up to a nullset, the paths of W and $\bar{W}$ coincide. □

Comparing with the examples in Section 2.2, Theorem 2.19 furthermore implies that the Brownian bridge and the Ornstein-Uhlenbeck process exist since they can be constructed directly from Brownian motion. The existence of fractional Brownian must be argued for separately, but of course the arguments resemble those of Brownian motion.

Exercise 33. *Show that fractional Brownian motion with Hurst index* $H \in (0,1)$ *exists. Moreover, show that almost every path of fractional Brownian motion is* p*-Hölder continuous for all* $p \in (0, H)$*.* ◇

Exercise 34. *Let* $X = \{X(t)\}_{t\in[0,\infty)}$ *be an Ornstein-Uhlenbeck process and* $Y = \{Y(t)\}_{t\in[0,1]}$ *be a Brownian bridge. For which* $p \in (0,1]$ *are* X *and* Y p*-Hölder continuous?* ◇

Another interesting consequence of the existence of Brownian motion is the existence of functions which are nowhere differentiable. If this does not impress you, try to construct such a function by yourself!

Corollary 2.20 (Existence of Nowhere Differentiable Functions)**.** *There exists a function* $f : [0,\infty) \to \mathbb{R}$ *which is nowhere differentiable.* ◇

Proof. This is an immediate consequence of Theorem 2.19 (Existence of Brownian motion) and Theorem 2.7 (Nowhere Differentiability of Brownian Paths). □

Another consequence of the Kolmogorov-Čentsov continuity theorem is the following time inversion result, with which we conclude this chapter on Brownian motion.

Corollary 2.21 (Time Inversion for Brownian Motion)**.** *Given a Brownian motion* $W = \{W(t)\}_{t\in[0,\infty)}$*, define a stochastic process* $X = \{X(t)\}_{t\in[0,\infty)}$ *by*

$$X(0) \triangleq 0 \qquad \text{and} \qquad X(t) \triangleq tW\big(1/t\big) \qquad \text{for all } t \in (0,\infty).$$

Then there exists a Brownian motion Y *(with respect to its natural filtration* $\mathfrak{F}^Y$*) which is indistinguishable from* X*.* ◇

Proof. It is immediately seen that X is a centered Gaussian process since W is a centered Gaussian process. The covariance function Γ of X is given by $\Gamma(s,t) = 0 = s \wedge t$ if either s or t is zero and otherwise

$$\Gamma(s,t) = \operatorname{Cov}\big[sW(1/s), tW(1/t)\big] = st\min\{1/s, 1/t\} = s \wedge t.$$

Thus X is a raw Brownian motion and each path of X is continuous on $(0,\infty)$, thus left continuous on $[0,\infty)$. But then the continuous modification Y of X obtained from Theorem 2.17 (Kolmogorov-Čentsov) is a Brownian motion with respect to $\mathfrak{F}^Y$ and indistinguishable from X by Exercise 31. □

Index

VOCABULARY

English	German
σ-field	σ-Algebra
σ-field of the τ-past	σ-Algebra der τ-Vergangenheit
adapted	adaptiert
Brownian motion	Brownsche Bewegung
Brownian bridge	Brownsche Brücke
canonical process	kanonischer Prozess
centered	zentriert
complete	vollständig
continuous extension	stetige Fortsetzung
continuous time	zeitstetig
coordinate process	Koordinatenprozess

English	German
coordinate projection	Koordinatenprojektion
countable	abzählbar
counting process	Zählprozess
discrete time	zeitdiskret
distribution	Verteilung
dyadic modulus	dyadisches (Stetigkeits-)Modul
dyadic rationals	dyadische Zahlen
filtered probability space	filtrierter Wahrscheinlichkeitsraum
filtration	Filtrierung
finite-dimensional cylinder set	endlich-dimensionale Zylindermenge
finite-dimensional distributions	endlich-dimensionale Verteilungen
fractional Brownian motion	fraktionale Brownsche Bewegung
Gaussian process	Gaußprozess
hitting time	Treffzeit
identity mapping	Identitätsabbildung
indistinguishable	ununterscheidbar
Kolmogorov's consistency condition	Kolmogorov's Konsistenzbedingung
Markov chain	Markov-Kette

English	German
mean reversion	Mittelwertrückkehr
measurable space	messbarer Raum bzw. Messraum
modification	Modification
optional time	optionale Zeit
probability space	Wahrscheinlichkeitsraum
path	Pfad
premeasure	Prämaß
random variable	Zufallsvariable
raw process	roher Prozess
renewal process	Erneuerungsprozess
stationary	stationär
stochastic process	stochastischer Prozess
stopping time	Stoppzeit
time index set	Zeitindexmenge
total variation	Totalvariation