# STOCHASTIC DIFFERENTIAL EQUATIONS

BENJAMIN FEHRMAN

ABSTRACT. These notes provide an essentially self-contained introduction to the theory of stochastic differential equations, beginning with the theory of martingales in continuous time. These notes were prepared for lecture courses at the University of Oxford during the Michaelmas Terms of Fall 2019/20. Each course consisted of sixteen fifty-minute lectures.

## CONTENTS

## 1. Preliminaries

In this course, we will assume basic concepts from measure theory including the notions of a sigma algebra, a measure, and a measure space. Furthermore, we will assume the foundational theorems of Lebesgue integration, including the monotone convergence theorem, Fatou's lemma, and the dominated convergence theorem. We will assume a familiarity with martingales in discrete in continuous time, but all necessary concepts will be reviewed in the notes.

1.1. **Probability spaces, random variables, and independence.** A probability space is the basic mathematical object used to describe random phenomena.

**Definition 1.1.** A *probability space* is a measure space $(\Omega, \mathcal{F}, \mathbb{P})$ with total measure one—that is, it is a set $\Omega$, a sigma algebra $\mathcal{F}$ on $\Omega$, and a nonnegative measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ that satisfies $\mathbb{P}(\Omega) = 1$.

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, measurable subsets $A \in \mathcal{F}$ are called *events*. The simplest nontrivial example is the two-point space $\Omega = \{H, T\}$ describing a single flip of a fair coin. The set $\Omega$ is equipped with its power set $\mathcal{F} = \mathcal{P}(\Omega)$ for the sigma algebra, and it is equipped with the probability measure $\mathbb{P}$ defined uniquely by

$$\mathbb{P}[\{H\}] = \mathbb{P}[\{T\}] = \frac{1}{2}.$$

The event $A = \{H\}$ represents an outcome of "heads," and the event $A = \{T\}$ represents the outcome "tails." The *probability* of both events is defined by their measure under $\mathbb{P}$, and so both events have probability $1/2$.

A more expressive probability space, and one that is suitable for many purposes, is the unit interval $\Omega = [0, 1]$ equipped with its Borel sigma algebra and Lebesgue measure. As an exercise, show that on this probability space there exists a sequence of independent, identically distributed random variables $\{X_n\}_{n\in\mathbb{N}}$ satisfying

$$\mathbb{P}[X_n = 1] = \mathbb{P}[X_n = -1] = \frac{1}{2}.$$

The functions $X_n \colon \Omega \to \mathbb{R}$ are *random variables* that represent the outcome of a coin-flip, where we say that the $n$th flip is heads if $X_n = 1$ and is tails if $X_n = -1$. So, what we observe is that the probability space $[0, 1]$ is rich enough to model a countable infinity of independent coin-flips.

**Definition 1.2.** A real-valued *random variable* on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a measurable function $X \colon (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. That is,

$$X^{-1}(B) = \{\omega \in \Omega \colon X(\omega) \in B\} \in \mathcal{F} \ \text{ for every } \ B \in \mathcal{B}(R),$$

where $\mathcal{B}(\mathbb{R})$ is the Borel sigma-algebra of $\mathbb{R}$.

Consider now a probability space that represents flipping a fair coin twice. The set is

$$\Omega = \{(H, H), (H, T), (T, H), (T, T)\},$$

the sigma algebra $\mathcal{F} = \mathcal{P}(\Omega)$ is the power set of $\Omega$, and the measure $\mathbb{P}$ satisfies $\mathbb{P}[\{\omega\}] = 1/4$ for every $\omega = (\omega_1, \omega_2) \in \Omega$. Then consider the two random variables

$$X_1(\omega_1, \omega_2) = \begin{cases} 1 & \text{if } \ w_1 = H, \\ -1 & \text{if } \ w_1 = T, \end{cases} \quad \text{and} \quad X_2(\omega_1, \omega_2) = \begin{cases} 1 & \text{if } \ w_2 = H, \\ -1 & \text{if } \ w_2 = T. \end{cases}$$

The *sigma algebra generated* by the random variables $X_1$ is the smallest sub-sigma algebra $\sigma(X_1) \subseteq \mathcal{F}$ with respect to which $X_1$ is measurable. Precisely,

$$\sigma(X_1) = \{\emptyset, \Omega, \{(H, H), (H, T)\}, \{(T, H), (T, T)\}\}.$$

Similarly, the sigma algebra $\sigma(X_2)$ is

$$\sigma(X_2) = \{\emptyset, \Omega, \{(H, H), (T, H)\}, \{(H, T), (T, T)\}\}.$$

The important thing to notice here is that these sigma algebras encode the dependencies of the random variables $X_1$ and $X_2$. The value of $X_1$ only depends on the first flip, and so too do the the sets in its sigma algebra. Similarly, the values of $X_2$ depend only on the second flip, and the same is true for the sets in its sigma algebra. In this way, we view the sigma algebra of a random variables as carrying the "information" or the "dependencies of" the random variable.

**Definition 1.3.** Let $X\colon \Omega \to \mathbb{R}$ be a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The *sigma algebra generated by* $X$ is the sigma algebra $\sigma(X) \subseteq \mathcal{F}$ defined by

$$\sigma(X) = \{X^{-1}(B)\colon B \in \mathcal{B}(\mathbb{R})\}.$$

That is, $\sigma(X) \subseteq \mathcal{F}$ is the smallest sigma algebra on $\Omega$ with respect to which $X$ is measurable.

Intuitively the random variables $X_1$ and $X_2$ defined above are *independent*. Information concerning the first coin-flip tells us nothing about the outcome of the second. Mathematically we make this notion precise through the notion of independence. Two events are independent if the probability of both of them happening is the product of their probabilities:

$$\mathbb{P}[\{\text{both flips heads}\}] = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}[\{\text{first flip heads}\}] \cdot \mathbb{P}[\{\text{second flip heads}\}].$$

In terms of events, if $A = \{\text{first flip heads}\} = \{(H,H),(H,T)\}$ and $B = \{\text{second flip heads}\} = \{(H,H),(T,H)\}$ then $A \cap B = \{\text{both flips heads}\} = \{(H,H)\}$ and the above equality becomes

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B].$$

This defines what it means for two events to be independent. We say that two random variables are independent if their sigma algebras are independent.

**Definition 1.4.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Two events $A, B \in \Omega$ are *independent* if

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B].$$

Two random variables $X_1, X_2\colon \Omega \to \mathbb{R}$ are *independent* if

$$\mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B] \;\text{ for every }\; A \in \sigma(X_1) \;\text{ and }\; B \in \sigma(X_2).$$

That is, two random variables are independent if their sigma algebras are independent.

Show that if $X_1, X_2\colon \Omega \to \mathbb{R}$ are bounded, independent random variables then

$$\mathbb{E}[X_1X_2] = \mathbb{E}[X_1]\mathbb{E}[X_2],$$

where the expectation $\mathbb{E}$ of a random variable $Y$ is defined as the Lebesgue integral $\int_\Omega Y \,\mathrm{d}\mathbb{P}$. However, show that there exist bounded random variables $X_1, X_2\colon \Omega \to \mathbb{R}$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which satisfy

$$\mathbb{E}[X_1X_2] = \mathbb{E}[X_1]\mathbb{E}[X_2],$$

but which are not independent. The point is that the expectation is rather course information that tells you only about the averages of these random variables. Independence is a much finer condition, which states that all of the information or dependencies of $X_1$ are independent of those of $X_2$.

1.2. **Conditional expectation.** Returning to the probability space describing two flips of a fair coin, the conditional probability answers the following types of question. Given that the first flip is heads what is the probability that both flips are heads? That is, what is the probability that both flips are heads conditioned on the event that the first flip is heads? The answer should be $1/2$, since if we know that the first flip is heads the only randomness lies in the second flip. Mathematically,

$$\mathbb{P}[\{\text{both flips heads}\}|\{\text{given first flip heads}\}] = \frac{\mathbb{P}[\{\text{both flips heads and first flip heads}\}]}{\mathbb{P}[\{\text{first flip heads}\}]},$$

or,

$$\begin{aligned}\mathbb{P}[\{(H,H)\}|\{\text{given } (H,H),(H,T)\}] &= \frac{\mathbb{P}[\{(H,H)\}\cap\{(H,H),(H,T)\}]}{\mathbb{P}[\{(H,H),(H,T)\}]}\\ &= \frac{\mathbb{P}[\{(H,H)\}]}{\mathbb{P}[\{(H,H),(H,T)\}]}\\ &= \frac{1/4}{1/2} = \frac{1}{2}.\end{aligned}$$

In general, for an event $B \in \mathcal{F}$ satisfying $\mathbb{P}[B] > 0$, we define the conditional probability of an event $A \in \mathcal{F}$ with respect to $B$ as

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A\cap B]}{\mathbb{P}[B]}.$$

One could in fact define a new probability space $(\Omega_B, \mathcal{F}_B, \mathbb{P}_B)$ where $\Omega_B = B$, $\mathcal{F}_B = \mathcal{F}\cap B$, and for every $\tilde{A} \in \mathcal{F}_B$ satisfying $\tilde{A} = A\cap B$ for some $A \in \mathcal{F}$, the new probability measure $\mathbb{P}_B$ is defined by

$$\mathbb{P}_B[\tilde{A}] = \frac{\mathbb{P}[A\cap B]}{\mathbb{P}[B]}.$$

In this new probability space, we are living in a world where the event $B$ is guaranteed to occur, and the corresponding probabilities are adjusted accordingly. Observe in particular that, if $A, B$ are independent events, then $\mathbb{P}[A|B] = \mathbb{P}[A]$. That is, the knowledge that the event $B$ will occur yields no information about the likelihood of $A$, as we should expect.

Conditional expectation generalizes the notion of conditional probability to random variables. The sigma algebra $\sigma(X)$ of a random variable $X$ encodes the information or dependencies of the random variable. The conditional expectation of a random variable $Y$ with respect to $X$ is our best guess for $Y$ using only the information or dependencies of $X$.

**Definition 1.5.** Let $X\colon \Omega \to \mathbb{R}$ be a random variable and let $Y\colon \Omega \to \mathbb{R}$ be an integrable random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The *conditional expectation* of $Y$ given $X$ is the unique $\sigma(X)$-measurable random variable $\mathbb{E}[Y|X]$ or $E[Y|\sigma(X)]$ that satisfies, for every $A \in \sigma(X)$,

$$\mathbb{E}[Y\colon A] = \int_A Y \,\mathrm{d}\mathbb{P} = \int_A \mathbb{E}[Y|X]\,\mathrm{d}\mathbb{P} = \mathbb{E}[\mathbb{E}[Y|X]\colon A].$$

We can also take conditional expectations with respect to a sigma-algebra.

**Definition 1.6.** Let $Y\colon \Omega \to \mathbb{R}$ be an integrable random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-sigma algebra. The *conditional expectation* of $Y$ given $\mathcal{G}$ is the unique $\mathcal{G}$-measurable random variable $\mathbb{E}[Y|\mathcal{G}]$ that satisfies, for every $A \in \mathcal{G}$,

$$\mathbb{E}[Y\colon A] = \int_A Y \,\mathrm{d}\mathbb{P} = \int_A \mathbb{E}[Y|\mathcal{G}]\,\mathrm{d}\mathbb{P} = \mathbb{E}[\mathbb{E}[Y|\mathcal{G}]\colon A].$$

In this way, the conditional expectation of $Y$ with respect to a random variable $X$ is the conditional expectation of $Y$ with respect to the sub-sigma algebra $\sigma(X)$.

The existence and uniqueness of the conditional expectation follows from the integrability of $Y$ and the Radon-Nikodym theorem. Precisely, given a sub-sigma algebra $\mathcal{G} \subseteq \mathcal{F}$, the conditional expectation is the Radon-Nikodym derivative of the finite measure $\mu_Y(A) = \int_A Y \,\mathrm{d}\mathbb{P}$ defined on $\mathcal{G}$. Uniqueness is a consequence of the fact that the conditional expectation must be $\mathcal{G}$-measurable, and that every two candidates $\mathbb{E}[Y|\mathcal{G}]$ and $\tilde{\mathbb{E}}[Y|\mathcal{G}]$ satisfy

$$\int_A \left(\mathbb{E}[Y|\mathcal{G}] - \tilde{\mathbb{E}}[Y|\mathcal{G}]\right) \mathrm{d}\mathbb{P} = 0 \text{ for every } A \in \mathcal{G}.$$

The following properties are a consequence of the definition of the conditional expectation.

**Proposition 1.7.** *Let $Y_1, Y_2 \colon \Omega \to \mathbb{R}$ be integrable random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G}_1, \mathcal{G}_2 \subseteq \mathcal{F}$ be sub-sigma-algebras.*

*(a.) Linearity: for every $\alpha, \beta \in \mathbb{R}$,*

$$\mathbb{E}[\alpha Y_1 + \beta Y_2|\mathcal{G}_1] = \alpha\mathbb{E}[Y_1|\mathcal{G}_1] + \beta\mathbb{E}[Y_2|\mathcal{G}_1].$$

*(b.) Order-preserving: if $Y_1 \leq Y_2$ almost surely then, almost surely,*

$$\mathbb{E}[Y_1|\mathcal{G}_1] \leq \mathbb{E}[Y_2|\mathcal{G}_1].$$

*(c.) Tower property: if $\mathcal{G}_1 \subseteq \mathcal{G}_2$ then*

$$\mathbb{E}[\mathbb{E}[Y_1|\mathcal{G}_2]|\mathcal{G}_1] = \mathbb{E}[Y_1|\mathcal{G}_1].$$

Show that a random variable is measurable with respect to the trivial sigma algebra $\{\emptyset, \Omega\}$ if and only if that random variable is constant. Then, using the definition of the conditional expectation, show that for every integrable random variable $Y$ we have that $\mathbb{E}[Y|\{\emptyset, \Omega\}] = \mathbb{E}[Y]$ is constant. Taking this one step further, if $Y$ is independent of a sub-sigma algebra $\mathcal{G}$ in the sense that $\sigma(Y)$ and $\mathcal{G}$ are independent sigma algebras, then the conditional expectation $\mathbb{E}[Y|\mathcal{G}]$ should yield essentially no information. The following proposition shows that this is the case: if $Y$ is independent of $\mathcal{G}$ then $\mathbb{E}[Y|\mathcal{G}] = \mathbb{E}[Y]$ is constant and offers no information beyond what we deduce from the trivial sigma algebra. Conversely, if $Y$ is $\mathcal{G}$-measurable, then $\mathcal{G}$ contains all of the dependencies or information of $Y$ and the conditional expectation $\mathbb{E}[Y|\mathcal{G}] = Y$ describes $Y$ exacty.

**Proposition 1.8.** *Let $Y, Z \colon \Omega \to \mathbb{R}$ be integrable random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let $\mathcal{G} \subseteq \mathcal{F}$, and assume that the product $YZ \colon \Omega \to \mathbb{R}$ is integrable.*

*(a.) Independence: if $Y$ is independent of $\mathcal{G}$ then*

$$\mathbb{E}[Y|\mathcal{G}] = \mathbb{E}[Y].$$

*(b.) Factoring: If $Y$ is $\mathcal{G}$-measurable then*

$$\mathbb{E}[YZ|\mathcal{G}] = Y\mathbb{E}[Z|\mathcal{G}].$$

*In particular, taking $Z = 1$ we have $\mathbb{E}[Y|\mathcal{G}] = Y$.*

The following proposition is a version of Jensen's inequality for the conditional expectation. Observe that both sides of this inequality are random variables, and that the inequality holds for almost every $\omega \in \Omega$.

**Proposition 1.9.** *Let $Y \colon \Omega \to \mathbb{R}$ be an integrable random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-sigma algebra, and let $f \colon \mathbb{R} \to \mathbb{R}$ be a convex function. Then, almost surely,*

$$f(\mathbb{E}[Y|\mathcal{G}]) \leq \mathbb{E}[f(Y)|\mathcal{G}].$$

Finally, the conditional expectation is continuous with respect to $L^1$-convergence and satisfies versions of the monotone convergence theorem, Fatou's lemma, and the dominated convergence theorem. Full details can be found in the Math B8.1 course notes, which are available on the course website.

**Proposition 1.10.** *Let* $\{Y_n\colon \Omega \to \mathbb{R}\}_{n\in\mathbb{N}}$ *and* $Y\colon \Omega \to \mathbb{R}$ *be integrable random variables on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and let* $\mathcal{G} \subseteq \mathcal{F}$ *be a sub-sigma algebra. If* $Y_n \to Y$ *in* $L^1(\Omega)$ *as* $n \to \infty$ *then, as* $n \to \infty$,

$$\mathbb{E}[Y_n|\mathcal{G}] \to \mathbb{E}[Y|\mathcal{G}] \ \text{ in } \ L^1(\Omega).$$

## 2. Martingales

A martingale is a mathematical model for a fair game, or a stochastic process where knowledge of the past does not allow the player to predict the future. Think for example of flipping a coin. The outcome of the first $n$ flips does not reveal any information about the outcome of flip $(n+1)$. More precisely, on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let $\{X_k\}_{k\in\mathbb{N}}$ be independent, identically distributed (i.i.d.) random variables which satisfy

$$\mathbb{P}(X_1 = -1) = \mathbb{P}(X_1 = 1) = \frac{1}{2}.$$

We view 1 as representing a "heads" and $-1$ as representing "tails." For every $n \in \mathbb{N}$ define the sum

$$S_n = X_1 + \ldots + X_n = (\#\text{ "heads"}) - (\#\text{ "tails"}). \tag{2.1}$$

For every $n \in \mathbb{N}$ let $\mathcal{F}_n \subseteq \mathcal{F}$ be the sigma algebra generated by the random variables $X_1, \ldots, X_n$,

$$\mathcal{F}_n = \sigma(X_1, \ldots, X_n). \tag{2.2}$$

The sequence $\{S_n\}_{n\in\mathbb{N}}$ forms a martingale with respect to the filtration $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ in the sense that our best guess for the random variable $S_{n+1}$ given the information $\mathcal{F}_n$ is $S_n$. Or, in terms of the conditional expectation,

$$\mathbb{E}[S_{n+1}|\mathcal{F}_n] = \mathbb{E}[X_{n+1}|\mathcal{F}_n] + \mathbb{E}[S_n|\mathcal{F}_n] = \mathbb{E}[X_{n+1}] + S_n = S_n.$$

A random process is called a submartingale if $S_n$ provides a lower bound for the conditional expectation $\mathbb{E}[S_{n+1}|\mathcal{F}_n]$, and a supermartingale if $S_n$ provides an upper bound for the conditional expectation. That is, roughly speaking, submartingales are "increasing" whereas supermartingales are "decreasing."

The process $\{S_n\}_{n\in\mathbb{N}}$ is known as the *simple random walk*, and it provides our first example of a martingale in discrete time. However, in this course, we will primarily be interested in martingales in continuous time. Our first and most important example is a *Brownian motion*.

**Definition 2.1.** A *Brownian motion* on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a stochastic process $B\colon \Omega \times [0,\infty) \to \mathbb{R}$ that satisfies the following four properties.

(a.) Beginning at zero: almost surely,
$$B_0 = 0.$$
(b.) Normal distribution: for every $s < t \in [0,\infty)$,
$$B_t - B_s \;\text{ is a normally distributed variable with mean zero and variance } t-s.$$
(c.) Independent incrementes: for every $N \in \mathbb{N}$ and $t_0 \le t_1 \le t_2 \le \ldots \le t_n \in [0,\infty)$, the random variables
$$\{B_{t_i} - B_{t_{i-1}}\}_{i=1}^N \;\text{ are mutually independent.}$$
(d.) Continuous sample paths: almost surely the map
$$t \in [0,\infty) \mapsto B_t(\omega) \;\text{ is continuous.}$$

Not every probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is rich enough to support a Brownian motion, and the details of its construction and an overview of its properties can be found in the Math B8.2 notes provided on the course website. What we observe for now is that, if for each $t \in [0,\infty)$ we define $\mathcal{F}_t = \sigma(B_s\colon s \in [0,t])$ to be the sigma algebra generated by the Brownian motion up to time $t$, then we have for every $s \le t \in [0,\infty)$,

$$\mathbb{E}[B_t|\mathcal{F}_s] = \mathbb{E}[B_t - B_s|\mathcal{F}_s] + \mathbb{E}[B_s|\mathcal{F}_s].$$

Since $(B_t - B_s)$ is independent of $\mathcal{F}_s$, and since $B_s$ is $\mathcal{F}_s$-measurable,

$$\mathbb{E}[B_t|\mathcal{F}_s] = \mathbb{E}[B_t - B_s] + B_s = B_s.$$

This is the continuous martingale property, which states that our best guess for a Brownian motion at time $t \geq s$ using only the dependencies or information of the Brownian motion up to time $s$ is $B_s$. Knowledge of the past does not allow you to predict the future.

2.1. **Filtrations, martingales, and stopping times.** Let $(\Omega, \mathcal{F})$ be a measurable space, which is to say that $\Omega$ is a set equipped with a sigma algebra $\mathcal{F}$ of subsets. We will view sigma algebras as carrying information, where in the above the sigma algebra $\mathcal{F}_n$ defined in (2.2) carries the information of the random variables $X_1, \ldots, X_n$. This is to say that, given all of the information of $\mathcal{F}_n$, we can predict the random variables $X_1, \ldots, X_n$ exactly. Or, in terms of the conditional expectation, for every $m \in \{1, \ldots, n\}$,

$$\mathbb{E}[X_m|\mathcal{F}_n] = X_m.$$

A filtration is an increasing family of sub-sigma algebras, where the fact that the sub-sigma algebras are increasing implies that the amount of information carried by the sub-sigma algebras is increasing. We will be primarily interested in filtrations indexed by a continuous parameter $t \in [0, \infty)$ or a discrete parameter $n \in \mathbb{N}$.

**Definition 2.2.** Let $(\Omega, \mathcal{F})$ be a measurable space and let $I \subseteq \mathbb{R}$. A *filtration* on $(\Omega, \mathcal{F})$ indexed by $I$ is an increasing family of sub-sigma algebras $\{\mathcal{F}_t\}_{t \in I}$. That is, for every $s \leq t \in I$ we have $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$.

Filtrations are often generated by a stochastic process, such as in (2.2) where the sub-sigma algebra $\mathcal{F}_n$ was generated by the first $n$ coin-flips. We will be primarily interested in continuous processes indexed by $t \in [0, \infty)$ or discrete processes indexed by $n \in \mathbb{N}$.

**Definition 2.3.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\Theta, \mathcal{G})$ be a measurable space, and let $I \subseteq \mathbb{R}$. A *stochastic process* indexed by $I$ taking values in $\Theta$ is a family of bimeasurable maps $\{X_t \colon \Omega \to \Theta\}_{t \in I}$.

Every stochastic process $\{X_t\}_{t \in [0,\infty)}$ generates a natural filtration

$$\{\mathcal{F}_t = \sigma(X_s \colon s \in [0, t])\}_{t \in [0,\infty)},$$

since it follows by the measurability of the $\{X_t\}_{t \in [0,\infty)}$ and by definition that $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ whenever $s \leq t \in [0, \infty)$. We will often view the parameter $t$ as a "time"-variable, and the sigma algebra $\mathcal{F}_t$ as carrying the information of $\{X_s\}_{s \in [0,\infty)}$ up to time $t$. This was the case in the discrete example (2.2), where the time parameter was discrete, and was the case for Brownian motion.

A martingale is a stochastic process defined on a probability space with respect to a filtration. It is important to understand that a martingale is only ever a martingale with respect to a filtration. If the filtration changes, the martingale need not remain a martingale, and if there is no filtration there is absolutely no martingale. We will be primarily interested in martingales defined over $t \in [0, \infty)$ or $n \in \mathbb{N}$, and taking values some Euclidean space.

**Definition 2.4.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\Theta, \mathcal{G})$ be a measurable space, let $I \subseteq \mathbb{R}$, and let $\{\mathcal{F}_t\}_{t \in I}$ be a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$.

(a.) A $\Theta$-valued stochastic process $(M_t)_{t \in I}$ indexed by $I$ is a *martingale* with respect to $\{\mathcal{F}_t\}_{t \in I}$ if
  (i) For every $t \in I$,
$$\mathbb{E}[|M_t|] < \infty.$$
  (ii) For every $t \in I$, $M_t$ is $\mathcal{F}_t$-measurable.
  (iii) For every $s \leq t \in I$,
$$\mathbb{E}[M_t|\mathcal{F}_s] = M_s.$$

(b.) A $\Theta$-valued stochastic process $(M_t)_{t \in I}$ indexed by $I$ is a *submartingale* with respect to $\{\mathcal{F}_t\}_{t \in I}$ if
  (i) For every $t \in I$,
$$\mathbb{E}[M_t^+] = \mathbb{E}[\max(M_t, 0)] < \infty.$$
  (ii) For every $t \in I$, $M_t$ is $\mathcal{F}_t$-measurable.

(iii) For every $s \leq t \in I$,

$$\mathbb{E}[M_t|\mathcal{F}_s] \geq M_s.$$

(c.) A $\Theta$-valued stochastic process $(M_t)_{t\in I}$ indexed by $I$ is a *supermartingale* with respect to $\{\mathcal{F}_t\}_{t\in I}$ if the stochastic process $(-M_t)_{t\in I}$ is a submartingale. That is, if

(i) For every $t \in I$,

$$\mathbb{E}[M_t^-] = \mathbb{E}[\min(M_t, 0)] > -\infty.$$

(ii) For every $t \in I$, $M_t$ is $\mathcal{F}_t$-measurable.

(iii) For every $s \leq t \in I$,

$$\mathbb{E}[M_t|\mathcal{F}_s] \leq M_s.$$

Observe that a process $(M_t)_{t\in[0,\infty)}$ is a martingale with respect to a filtration $\{\mathcal{F}_t\}_{t\in[0,\infty)}$ if and only if $(M_t)_{t\in[0,\infty)}$ is both a submartingale and a supermartingale with respect to $\{\mathcal{F}_t\}_{t\in[0,\infty)}$. Intuitively, the process $(M_t)_{t\in[0,\infty)}$ is a martingale if our best guess for $M_t$ given the information $\mathcal{F}_s$ is $M_s$. A process is a submartingale if $M_s$ provides a lower bound for our best guess of $M_t$ given $\mathcal{F}_s$, and is a supermartingale if $M_s$ provides an upper bound for our best guess of $M_t$ given $\mathcal{F}_s$. Loosely speaking, we can therefore view submartingales as "increasing" and supermartingales as "decreasing."

An important concept in the study of martingales, and stochastic differential equations, is the notion of a stopping time. Consider the example above, for i.i.d. random variables $\{X_k\}_{k\in\mathbb{N}}$ which satsify

$$\mathbb{P}[X_1 = 1] = \mathbb{P}[X_1 = -1] = \frac{1}{2},$$

for the sigma algebras $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ defined by

$$\mathcal{F}_n = \sigma(X_m \colon m \in \{1, \ldots, n\}),$$

and for random variables $\{S_n\}_{n\in\mathbb{N}}$ defined by $S_0 = 0$ and

$$S_n = X_1 + \ldots + X_n.$$

The random variables $\{S_n\}_{n\in\mathbb{N}}$ define a *random walk* on $\mathbb{Z}$, and for a fixed integer $z \in \mathbb{Z}$ we will be interested in the first time that the walk hits $z$. This is defined by the random variable

$$T_z = \inf\{n \in \mathbb{N}_0 \colon S_n = z\}, \tag{2.3}$$

which is an example of a *hitting time*. The essential property of the variable $T_z$ is that to know if $T_z \leq n$ we only need to know information about the random variables $X_1, \ldots, X_n$. This is to say that

$$\{T_z \leq n\} \in \mathcal{F}_n,$$

which is the defining property of a stopping time.

**Definition 2.5.** Let $(\Omega, \mathcal{F}, \mathbb{P})$, let $I = \mathbb{N}_0$ or $I = [0, \infty)$, and let $\{\mathcal{F}_t\}_{t\in I}$ be a filtration on $(\Omega, \mathcal{F})$. A *stopping time* is a measurable map $T \colon \Omega \to I$ such that, for every $t \in I$,

$$\{T \leq t\} \in \mathcal{F}_t.$$

If we think of a stopping time $T$ as an alarm clock, the condition $\{T \leq t\} \in \mathcal{F}_t$ implies we don't need to look into the future to determine whether or not the alarm has rung. Hitting times like (2.3) above are the most common and most important examples of stopping times that we will encounter. Stopping times come with an associated sigma algebra, which consists of those events that are $\mathcal{F}_t$-measurable conditioned on the event $\{T \leq t\}$.

**Definition 2.6.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $I = \mathbb{N}_0$ or $I = [0, \infty)$, let $\{\mathcal{F}_t\}_{t\in I}$ be a filtration on $(\Omega, \mathcal{F})$, and let $T \colon \Omega \to I$ be a stopping time. We define the sub-sigma algebra $\mathcal{F}_T \subseteq \mathcal{F}$ by

$$\mathcal{F}_T = \{A \in \mathcal{F} \colon (A \cap \{T \leq t\}) \in \mathcal{F}_t \text{ for every } t \in I\}.$$

The following proposition proves that a stopped (sub/super) martingale remains a (sub/super) martingale.

**Proposition 2.7.** *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, let* $I = \mathbb{N}_0$ *or* $I = [0, \infty)$*, let* $\{\mathcal{F}_t\}_{t\in I}$ *be a filtration on* $(\Omega, \mathcal{F})$*, let* $(M_t)_{t\in I}$ *be an* $\mathcal{F}_t$*-(sub/super) martingale, and let* $T \colon \Omega \to I$ *be a stopping time. Then the process*

$$(M^T_t)_{t\in I} = (M_{t\wedge T})_{t\in I},$$

*is a (sub/super) martingale.*

*Proof.* Let $I = \mathbb{N}_0$ and suppose that $(M_n)_{n\in\mathbb{N}_0}$ is a martingale. For every $n \in \mathbb{N}_0$,

$$M^T_n = M_{T\wedge n} = M_n \mathbf{1}_{\{T>n\}} + \sum_{k=1}^{n} M_k \mathbf{1}_{\{T=k\}}. \tag{2.4}$$

Equation (2.4) proves that

$$\mathbb{E}[\left|M^T_n\right|] \leq \sum_{k=1}^{n} \mathbb{E}[|M_k|] < \infty,$$

and that $M^T_n$ is $\mathcal{F}_n$-measurable because every function on the righthand side of (2.4) is $\mathcal{F}_n$-measurable. Finally, by (2.4), for every $n \in \mathbb{N}_0$,

$$\begin{aligned}
\mathbb{E}[M^T_{n+1}|\mathcal{F}_n] &= \mathbb{E}[M_{n+1}\mathbf{1}_{\{T>n+1\}}|\mathcal{F}_n] + \sum_{k=1}^{n+1} \mathbb{E}[M_k \mathbf{1}_{\{T=k\}}|\mathcal{F}_n] \\
&= \mathbb{E}[M_{n+1}\mathbf{1}_{\{T>n\}}|\mathcal{F}_n] + \sum_{k=1}^{n} \mathbb{E}[M_k \mathbf{1}_{\{T=k\}}|\mathcal{F}_n].
\end{aligned}$$

Since $M_k \mathbf{1}_{\{T=k\}}$ and $\mathbf{1}_{\{T>n+1\}}$ are $\mathcal{F}_n$-measurable, the

$$\begin{aligned}
\mathbb{E}[M^T_{n+1}|\mathcal{F}_n] &= \mathbb{E}[M_{n+1}|\mathcal{F}_n]\mathbf{1}_{\{T>n\}} + \sum_{k=1}^{n} M_k \mathbf{1}_{\{T=n\}} \\
&= M_n \mathbf{1}_{T>n} + \sum_{k=1}^{n} M_k \mathbf{1}_{\{T=n\}} \\
&= M_n,
\end{aligned}$$

which completes the proof. $\square$

Stopping terms are sometimes referred to *local times*. We will encounter processes that are not martingales, but which are *locally* a martingale when appropriately stopped. For example, let $\{B_t\}_{t\in[0,\infty)}$ be a standard Brownian motion, and let $T_1$ denote the stopping time $T_1 = \inf\{t \in [0, \infty) \colon B_s = 1\}$. Since $T_1$ is almost surely finite, the process

$$W_t = \begin{cases} B_{T_1 \wedge (\frac{t}{1-t})} & \text{if } t \in [0, 1), \\ 1 & \text{if } b \in [1, \infty), \end{cases}$$

is continuous, but it is not a martingale because

$$\mathbb{E}[W_t] = \begin{cases} 0 & \text{if } t \in [0, 1), \\ 1 & \text{if } t \in [1, \infty). \end{cases}$$

However, Proposition 2.7 proves that, for every $k \in \mathbb{N}$, for the stopping times

$$\tau_{-k} = \inf\{t \in [0, \infty) \colon W_t = -k\},$$

the stopped processes $\{W^{\tau_k}\}_{k\in\mathbb{N}}$ are martingales. The process $W$ is called a *local martingale*.

**Definition 2.8.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $I = \mathbb{N}_0$ or $I = [0, \infty)$, a let $\{\mathcal{F}_t\}_{t \in I}$ be a filtration on $(\Omega, \mathcal{F})$. An $\mathcal{F}_t$-adapted process $M$ is called a *local martingale* if there exist a stopping times $\{\tau_k\}_{k \in \mathbb{N}}$ such that

(i) For every $k \in \mathbb{N}$,
$$\tau_k \leq \tau_{k+1} \text{ almost surely.}$$
(ii) Almost surely, as $k \to \infty$,
$$\tau_k \to \infty.$$
(iii) For every $k \in \mathbb{N}$, the stopped process $M^{\tau_k} \mathbf{1}_{\{\tau_k > 0\}}$ is an $\mathcal{F}_t$-martingale.

2.2. **The optional stopping theorem (discrete).** In this section, we will first prove the optional stopping theorem for discrete martingales, which we will use to prove the optional stopping theorem for continuous martingales in Proposition 2.35 below. Let $(S_n)_{n\in\mathbb{N}}$ be a simple random walk, as defined in (2.1). This can be viewed as the outcome of a betting strategy, where for each coin flip we bet £1 that the outcome will be "heads." We thereby gain £1 for every "heads" and lose £1 for every "tails," and the random variables $\{S_n\}_{n\in\mathbb{N}}$ quantify our net profit/loss after $n$ flips.

If we have an infinite bankroll, we can almost surely achieve an arbitrarily large profit. For every $m \in \mathbb{N}$, we simply play until the random walk $\{S_n\}_{n\in\mathbb{N}}$ reaches $m$, which is an event of probability one. In terms of stopping times, if

$$T_m = \inf\{n \in \mathbb{N}_0 \colon S_n = m\},$$

then $T_m$ is almost surely finite and we have

$$m = \mathbb{E}[S_{T_m}] > \mathbb{E}[S_0] = 0.$$

However, if we instead have a finite bankroll £N it is no longer possible to play interminably, since we'll go bankrupt at the first time $n \in \mathbb{N}$ that $S_n = -N$. In this case, the stopping times $\{T_m\}_{m\in\mathbb{N}}$ must be replaced by

$$T_{m,N} = \inf\{n \in \mathbb{N}_0 \colon S_n = m \;\text{ or }\; S_n = -N\}.$$

A simple case of the optional stopping theorem below will prove that, for every $m, N \in \mathbb{N}$,

$$\mathbb{E}[S_{T_{m,N}}] = m\mathbb{P}[S_{T_{m,N}} = m] - N\mathbb{P}[S_{T_{m,N}} = -N] = \mathbb{E}[S_0] = 0,$$

which since

$$\mathbb{P}[S_{T_{m,N}} = m] + \mathbb{P}[S_{T_{m,N}} = -N] = \mathbb{P}[T_{m,N} < \infty] = 1,$$

implies that

$$\mathbb{P}[S_{T_{m,n}} = -N] = \frac{m}{N+m}. \tag{2.5}$$

This is a version of the *gambler's ruin* estimate, which states that a gambler who repeatedly stakes their bankroll to earn any profit, no matter how modest, will eventually lose everything. Indeed, it follows from (2.5) that a gambler with an initial amount $£N$ who repeatedly stakes the entirety of their bankroll, including the profit from their previous bets, to earn a fixed profit $£m$ goes bankrupt with probability

$$\sum_{k=0}^{\infty} \left(\frac{N}{N+km}\right) \cdot \left(\frac{m}{N+(k+1)m}\right) = 1.$$

The optional stopping theorem states that, if the gambler is playing a fair game, there is no strategy (i.e. no stopping time) that gives them an advantage. The proof will use the concept of a predictable process, which is a process for which the past determines the future.

**Definition 2.9.** Let $(\Omega, \mathcal{F})$ be a measurable space and let $\{F_n\}_{n\in\mathbb{N}_0}$ be a filtration on $(\Omega, \mathcal{F})$. An $\mathcal{F}_n$-stochastic process $(H_n)_{n\in\mathbb{N}}$ is called *predictable* if, for every $n \in \mathbb{N}$,

$$H_n \;\text{ is }\; \mathcal{F}_{n-1}\text{-measurable}.$$

In the following proposition, we define the discrete integral of a predictable process with respect to a discrete martingale. The essential point is that the integral of a predictable process with respect to a martingale is again a martingale. A fact we will see again in the context of integrals with respect to continuous martingales.

**Proposition 2.10.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration on $(\Omega, \mathcal{F})$, let $(M_n)_{n\in\mathbb{N}_0}$ be an $\mathcal{F}_n$-martingale, and let $(H_n)_{n\in\mathbb{N}}$ be a bounded $\mathcal{F}_n$-predictable process. Then, defined inductively,*

$$\begin{cases} (H\cdot M)_0 = 0 \\ (H\cdot M)_n = (H\cdot M)_{n-1} + H_n(M_n - M_{n-1}) \;\; \text{if } \; n \in \mathbb{N}, \end{cases}$$

*is a $\mathcal{F}_n$-martingale. If in addition $(H_n)_{n\in\mathbb{N}}$ is nonnegative, then $((H\cdot M)_n)_{n\in\mathbb{N}_0}$ is a (sub/super) martingale if $(M_n)_{n\in\mathbb{N}_0}$ is a (sub/super)-martingale.*

*Proof.* We will carry out the details in the case that $(M_n)_{n\in\mathbb{N}_0}$ is an $\mathcal{F}_n$-martingale. For every $n\in\mathbb{N}$,

$$\mathbb{E}[|(H\cdot M)_n|]\leq \max_{k\in\{1,\ldots,n\}}\|H_k\|_{L^\infty(\Omega)}\sum_{k=1}^n\mathbb{E}[|M_k|]<\infty,$$

and since by definition

$$(H\cdot M)_n=\sum_{k=1}^N H_k(M_n-M_{n-1}),$$

it follows that $(H\cdot M)_n$ is $\mathcal{F}_n$-measurable. Finally, the predictability of $(H_k)_{k\in\mathbb{N}}$ and the martingale property prove that, for every $n\in N_0$,

$$\begin{aligned}\mathbb{E}[(H\cdot M)_{n+1}|\mathcal{F}_n]&=\sum_{k=1}^{n+1}\mathbb{E}[H_k(M_k-M_{k-1})|\mathcal{F}_n]\\&=H_n\mathbb{E}[(M_{n+1}-M_n)|\mathcal{F}_n]+\sum_{k=1}^n H_k(M_k-M_{k-1})\\&=(H\cdot M)_n,\end{aligned}$$

which completes the proof. □

We now prove the *optional stopping theorem* in discrete time for bounded stopping times. We will extend this theorem to continuous time martingales in the sections to follow. The result will rely on the lemma below, which proves that is $S\leq T$ are stopping times then $\mathcal{F}_S\subseteq\mathcal{F}_T$.

**Lemma 2.11.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $I=\mathbb{N}_0$ or $I=[0,\infty)$, let $(\mathcal{F}_t)_{t\in I}$ be a filtration on $(\Omega,\mathcal{F})$, and let $S\leq T$ be $\mathcal{F}_t$-stopping times. Then $\mathcal{F}_S\subseteq\mathcal{F}_T$.*

*Proof.* Let $A\in\mathcal{F}_S$. It is necessary to prove that, for every $t\in I$,

$$A\cap\{T\leq t\}\in\mathcal{F}_t.$$

Since $S\leq T$, it follows that

$$(A\cap\{T\leq t\})=(A\cap\{S\leq t\})\cap((A\cap\{S\leq t\})\setminus\{T>t\}),$$

where the first term on the righthand side is in $\mathcal{F}_t$ by definition of $\mathcal{F}_S$, the second term is in $\mathcal{F}_t$ because $\mathcal{F}_t$ is a sigma algebra and $(A\cap\{S\leq t\})$ and $\{T>t\}=\{T\leq t\}^c$ are in $\mathcal{F}_t$, and the union is in $\mathcal{F}_t$ because $\mathcal{F}_t$ is a sigma algebra. This completes the proof. □

**Theorem 2.12.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $\{\mathcal{F}_n\}_{n\in\mathbb{N}_0}$ be a filtration on $(\Omega,\mathcal{F})$, let $(M_n)_{n\in\mathbb{N}_0}$ be an $\mathcal{F}_n$- (sub/super) martingale, and let $\sigma\leq\tau$ be two bounded stopping times. Then,*

$$\mathbb{E}[M_\tau|\mathcal{F}_\sigma]=M_\sigma.\ \ (\geq/\leq)\tag{2.6}$$

*In particular,*

$$\mathbb{E}[M_\tau]=\mathbb{E}[M_\sigma].\ \ (\geq/\leq)\tag{2.7}$$

*Proof.* We will carry out the details in the case that $(M_n)_{n\in\mathbb{N}_0}$ is a martingale. Let $\sigma\leq\tau$ be two bounded stopping times, and let $m\in\mathbb{N}$ be such that, for almost every $\omega\in\Omega$,

$$\sigma(\omega)\leq\tau(\omega)\leq m.\tag{2.8}$$

We will first prove (2.7). Let $(H_n)_{n\in\mathbb{N}_0}$ denote the process $H_0=0$ and

$$H_n=\mathbf{1}_{\{\sigma<n\leq\tau\}}=\mathbf{1}_{\{n\leq\tau\}}-\mathbf{1}_{\{n\leq\sigma\}}.$$

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

$$\{n \leq \tau\} = (\Omega \setminus \{\tau \leq n-1\}) \in \mathcal{F}_{n-1},$$

and similarly $\{n \leq \tau\} \in \mathcal{F}_{n-1}$. Therefore, for every $n \in \mathbb{N}_0$, the variable $H_n$ is $\mathcal{F}_{n-1}$-measurable, and $(H_n)_{n\in\mathbb{N}_0}$ is a $\mathcal{F}_n$-predictable process. Therefore, by Proposition 2.10, the process defined by $(H \cdot M)_0 = 0$ and,for $n \in \mathbb{N}$ by

$$(H \cdot M)_n = \sum_{k=1}^{n} H_k(M_k - M_{k-1}) = M_{\tau\wedge n} - M_{\sigma\wedge n}, \tag{2.9}$$

is a martingale. In combination (2.8) and (2.9) prove that, for every $n \geq m$,

$$(H \cdot M)_n = M_\tau - M_\sigma.$$

Therefore, since $(H \cdot M)$ is a martingale,

$$0 = \mathbb{E}[(H \cdot M)_0] = \mathbb{E}[(H \cdot M)_m] = \mathbb{E}[M_\tau] - \mathbb{E}[M_\sigma],$$

which proves (2.7). It remains to prove (2.6).

Let $B \in \mathcal{F}_\sigma$. It suffices to prove that

$$\mathbb{E}[M_\tau \colon B] = \mathbb{E}[M_\sigma \colon B].$$

For $m \in \mathbb{N}$ defined in (2.8), define the stopping times $\tau_B, \sigma_B \colon \Omega \to \mathbb{N}_0$ by

$$\tau_B(\omega) = \begin{cases} \tau(\omega) & \text{if } \omega \in B, \\ m & \text{if } \omega \in B^c, \end{cases} \quad \text{and} \quad \sigma_B(\omega) = \begin{cases} \sigma(\omega) & \text{if } \omega \in B, \\ m & \text{if } \omega \in B^c. \end{cases}$$

Indeed, by definition of the sigma algebra $\mathcal{F}_\sigma$ and Lemma 2.11, for every $n \in \{0, 1, 2, \ldots, m-1\}$,

$$\{\tau_B \leq n\} = (\{\tau \leq n\} \cap B) \in \mathcal{F}_n,$$

and for every $n \geq m$,

$$\{\tau_B \leq n\} = \Omega \in \mathcal{F}_n,$$

and the identical argument proves that $\sigma_B$ is a stopping time. Therefore, since $\tau_B$ and $\sigma_B$ are bounded stopping times, it follows from (2.7) that

$$\mathbb{E}[M_{\tau_B}] = \mathbb{E}[M_{\tau_B} \colon B] + \mathbb{E}[M_{\tau_B} \colon B^c] = \mathbb{E}[M_{\sigma_B} \colon B] + \mathbb{E}[M_{\sigma_B} \colon B^c] = \mathbb{E}[M_{\sigma_B}]. \tag{2.10}$$

Since it follows by definition that

$$\mathbb{E}[M_{\tau_B} \colon B] = \mathbb{E}[M_\tau \colon B] \ \text{ and } \ \mathbb{E}[M_{\sigma_B} \colon B] = \mathbb{E}[M_\sigma \colon B], \tag{2.11}$$

and that

$$\mathbb{E}[M_{\tau_B} \colon B^c] = \mathbb{E}[M_m \colon B^c] = \mathbb{E}[M_{\sigma_B} \colon B^c], \tag{2.12}$$

it follows from (2.10), (2.11), and (2.12) that

$$\mathbb{E}[M_\tau \colon B] = \mathbb{E}[M_\sigma \colon B],$$

which completes the proof of (2.6), and therefore the proof. □

The following corollary extends Theorem 2.12 to stopping times that are not necessarily bounded. We will extend these results to continuous martingales in the sections to follow.

**Corollary 2.13.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\{\mathcal{F}_n\}_{n\in\mathbb{N}_0}$ be a filtration on $(\Omega, \mathcal{F})$, let $(M_n)_{\mathbb{N}_0}$ be an $\mathcal{F}_n$- (sub/super) martingale, and let $\sigma \leq \tau$ be two almost surely finite stopping times. Assume that $|M_\tau|$ and $|M_\sigma|$ are integrable, and assume that*

$$\lim_{n\to\infty} \mathbb{E}[|M_n| : \tau > n] = 0. \tag{2.13}$$

*Then,*

$$\mathbb{E}[M_\tau|\mathcal{F}_\sigma] = M_\sigma. \ (\geq / \leq) \tag{2.14}$$

*In particular,*

$$\mathbb{E}[M_\tau] = \mathbb{E}[M_\sigma]. \ (\geq / \leq) \tag{2.15}$$

*Proof.* Let $B \in \mathcal{F}_\sigma$. Theorem 2.12 proves that, for every $n \in \mathbb{N}$, since the stopping times $\tau \wedge n$ and $\sigma \wedge n$ are bounded with $\sigma \wedge n \leq \tau \wedge n$, and since $B \cap \{\sigma \leq n\} \in \mathcal{F}_{\sigma \wedge n}$,

$$\begin{aligned} \mathbb{E}[M_{\tau\wedge n} \colon B, \sigma \leq n] &= \mathbb{E}[M_\tau \colon B, \tau \leq n, \sigma \leq n] + \mathbb{E}[M_n \colon B, \tau > n, \sigma \leq n] \\ &= \mathbb{E}[M_\sigma \colon B, \sigma \leq n] \\ &= \mathbb{E}[M_{\sigma\wedge n} \colon B, \sigma \leq n]. \end{aligned}$$

Since

$$|\mathbb{E}[M_n \colon B, \tau > n, \sigma \leq n]| \leq \mathbb{E}[|M_n| : \tau > n],$$

it follows from (2.13) that

$$\lim_{n\to\infty} |\mathbb{E}[M_n \colon B, \tau > n, \sigma \leq n]| = 0.$$

The dominated convergence theorem, the fact that $\sigma$ and $\tau$ are almost surely finite, and the integrability of $|M_\tau|$ and $|M_\sigma|$ prove that, after passing to the limit $n \to \infty$,

$$\mathbb{E}[M_\tau \colon B] = \mathbb{E}[M_\sigma \colon B].$$

This completes the proof of (2.14), which implies (2.15). This completes the proof. $\square$

2.3. **Uniform integrability.** In this section, will now explain conditions which can be used to upgrade the almost sure convergence or convergence in probability to strong convergence in $L^1(\Omega)$. This will require the notion of uniform integrability, which when combined with almost sure convergence (in fact, convergence in probability) implies convergence in $L^1(\Omega)$. This is a version of Vitalli's convergence theorem below.

**Definition 2.14.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\mathcal{A}$ be a set, and let $\{X_\alpha\}_{\alpha\in\mathcal{A}}$ be a family of random variables. The family $\{X_\alpha\}_{\alpha\in\mathcal{A}}$ is *uniformly integrable* if the following two conditions are satisfied.

(i) $L^1$-boundedness:
$$\sup_{\alpha\in\mathcal{A}} \mathbb{E}[|X_\alpha|] < \infty.$$
(ii) No concentration:
$$\lim_{K\to\infty}\left(\sup_{\alpha\in\mathcal{A}} \mathbb{E}[|X_\alpha| : \{|X_\alpha| \geq K\}]\right) = 0.$$

**Remark 2.15.** The second property of Definition 2.14 guarantees that the family of random variables $\{X_\alpha\}_{\alpha\in\mathcal{A}}$ do not concentrate their mass in the following sense. Let $\rho \in \mathrm{C}^\infty_c(\mathbb{R})$ be a nonnegative, smooth function which satisfies that
$$\int_{\mathbb{R}} \rho(x)\,\mathrm{d}x = 1,$$
and for every $\varepsilon \in (0,1)$ define $\rho^\varepsilon(x) = \varepsilon^{-1}\rho(x/\varepsilon)$. The functions form a Dirac sequence in the sense that, for every $f \in \mathrm{C}^\infty(\mathbb{R})$,
$$\lim_{\varepsilon\to 0}\left(\int_{\mathbb{R}} f(x)\rho^\varepsilon(x)\,\mathrm{d}x\right) = \int_{\mathbb{R}} f(x)\delta_0(x)\,\mathrm{d}x = f(0),$$
where $\delta_0$ is the Dirac delta distribution at zero. That is, as $\varepsilon \to 0$, as distributions,
$$\rho^\varepsilon \rightharpoonup \delta_0. \tag{2.16}$$
Since the distribution $\delta_0$ is not an $L^1$-function, we conclude that the family $\{\rho^\varepsilon\}_{\varepsilon\in(0,1)}$ is not precompact in $L^1(\mathbb{R})$ despite the fact that, for every $\varepsilon \in (0,1)$,
$$\|\rho^\varepsilon\|_{L^1(\mathbb{R})} = 1,$$
and despite the fact that $\rho^\varepsilon \to 0$ almost everywhere on $\mathbb{R}$, as $\varepsilon \to 0$. This lack of compactness is due to the fact that the functions $\{\rho^\varepsilon\}_{\varepsilon\in(0,1)}$ concentrate the entirety of their mass at the origin, as $\varepsilon \to 0$. In particular, they do not satisfy condition (ii) of Definition 2.14.

**Remark 2.16.** Definition 2.14 is the probabilistic version of uniform integrability, which is equivalent to standard definition of uniform integrability. The reason we use this definition is that, as we will see below, it is very convenient to verify when dealing with families of conditional expectations. Definition 2.14 also relies on the fact that a probability space $\Omega$ is a finite measure space with $\mathbb{P}(\Omega) < \infty$. On an infinite measure space, like $\mathbb{R}^d$, it is also necessary to impose a tightness condition that precludes mass from escaping to infinity. For example, on $\mathbb{R}$ and for the standard convolution kernel $\rho$ above the family of translates $\{\rho(\cdot - n)\}_{n\in\mathbb{Z}}$ converges pointwise to zero and satisfies the conditions of Definition 2.14, but this family does not converge in $L^1$.

The following theorem is Vitalli's convergence theorem specialized to the case of a probability space.

**Theorem 2.17.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $I = \mathbb{N}_0$ or let $I = [0,\infty)$. Let $\{f_t\}_{t\in I}$ be family of random variables which satisfy the following two properties.*

*(i) Uniform integrability: the family $\{f_t\}_{t\in I}$ is uniformly integrable.*

*(ii) Convergence in probability: there exists a random variable $f$ such that, for every $\varepsilon \in (0,1)$,*

$$\lim_{t\to\infty} \mathbb{P}[|f_t - f| > \varepsilon] = 0.$$

*Then $f \in L^1(\Omega)$ and, as $t \to \infty$,*

$$\lim_{t\to\infty} \mathbb{E}[|f_t - f|] = 0.$$

*That is, as $t \to \infty$, the $\{f_t\}_{t\in I}$ converge to $f$ in $L^1(\Omega)$.*

2.4. **The optional stopping theorem (continuous).** In this section, we will prove a version of the optional stopping theorem for uniformly integrable martingales in continuous time. As in the case of martingales in discrete time, it is necessary to impose some conditions on the stopping time. For instance, if $(B_t)_{t\in[0,\infty)}$ is a standard Brownian motion then

$$T = \inf\{s \in [0,\infty) \colon B_s \geq 1\},$$

is an almost surely finite stopping time. However, in this case the conclusion of the optional stopping theorem fails, since

$$\mathbb{E}[B_T] = 1 \neq 0 = \mathbb{E}[B_0].$$

We will therefore first prove the optional stopping theorem for bounded stopping times. We first state a useful lemma which states that a family consisting of a collection of conditional expectations of a fixed random variable is uniformly integrable.

**Lemma 2.18.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\mathcal{A}$ be a set, let $\{\mathcal{G}_\alpha\}_{\alpha\in\mathcal{A}}$, and let $X \in L^1(\Omega)$. Then the family*

$$\{\mathbb{E}[X|\mathcal{G}_\alpha]\}_{\alpha\in\mathcal{A}},$$

*is uniformly integrable.*

*Proof.* We will first prove property (i) of Definition 2.14. Since Jensen's inequality proves that, for every $\alpha \in \mathcal{A}$,

$$P\,|\mathbb{E}[X|\mathcal{G}_\alpha]| \leq \mathbb{E}[|X|\,|\mathcal{G}_\alpha], \tag{2.17}$$

it follows from $X \in L^1(\Omega)$ that, for every $\alpha \in \mathcal{A}$,

$$\mathbb{E}[|\mathbb{E}[X|\mathcal{G}_\alpha]|] \leq \mathbb{E}[\mathbb{E}[|X|\,|\mathcal{G}_\alpha]] = \mathbb{E}[|X|] < \infty.$$

Therefore,

$$\sup_{\alpha\in\mathcal{A}} \mathbb{E}[|E[X|\mathcal{G}_\alpha]|] < \infty. \tag{2.18}$$

It remains to prove property (ii). Since it follows from (2.17) that, for every $\alpha \in \mathcal{A}$ and $K \in (0,\infty)$,

$$\{|\mathbb{E}[X|\mathcal{G}_\alpha]| \geq K\} \subseteq \{\mathbb{E}[|X|\,|\mathcal{G}_\alpha] \geq K\},$$

it follows by properties of the conditional expectation and (2.17) that, for every $\alpha \in \mathcal{A}$ and $K \in (0,\infty)$,

$$\begin{aligned}\mathbb{E}[|\mathbb{E}[X|\mathcal{G}_\alpha]| : \{|\mathbb{E}[X|\mathcal{G}_\alpha]| \geq K\}] &\leq \mathbb{E}[\mathbb{E}[|X|\,|\mathcal{G}_\alpha] \colon \{\mathbb{E}[|X|\,|\mathcal{G}_\alpha] \geq K\}] \\ &= \mathbb{E}[|X| : \{\mathbb{E}[|X|\,|\mathcal{G}_\alpha] \geq K\}].\end{aligned} \tag{2.19}$$

By Chebyshev's inequality and (2.18), for every $\alpha \in \mathcal{A}$ and $K \in (0,\infty)$,

$$\mathbb{P}[\mathbb{E}[|X|\,|\mathcal{G}_\alpha] \geq K] \leq \frac{1}{K}\mathbb{E}[\mathbb{E}[|X|\,|\mathcal{G}_\alpha]] \leq \frac{1}{K}\mathbb{E}[|X|]. \tag{2.20}$$

We therefore conclude by continuity of the Lebesgue integral, $X \in L^1(\Omega)$, (2.19), and (2.20) that

$$\lim_{K\to\infty}\left(\sup_{\alpha\in\mathcal{A}} \mathbb{E}[|\mathbb{E}[X|\mathcal{G}_\alpha]| : \{|\mathbb{E}[X|\mathcal{G}_\alpha]| \geq K\}]\right) = 0,$$

which completes the proof. □

**Theorem 2.19.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\{\mathcal{F}_t\}_{t\in[0,\infty)}$ be a filtration on $(\Omega, \mathcal{F})$, let $(M_t)_{t\in[0,\infty)}$ be an $\mathcal{F}_t$-martingale (sub/super), and let $\sigma \leq \tau$ be two bounded bounded $\mathcal{F}_t$-stopping times. Then,*

$$\mathbb{E}[M_\tau|\mathcal{F}_\sigma] = M_\sigma \;\; (\geq / \leq).$$

*In particular,*

$$\mathbb{E}[M_\tau] = \mathbb{E}[M_\sigma] \;\; (\geq / \leq).$$

*Proof.* We will present the proof in the case that $(M_t)_{t\in[0,\infty)}$ is a martingale. Let $K\in\mathbb{N}$ almost surely satisfy $\tau\leq M$ and $\sigma\leq M$. For each $k\in\mathbb{N}$ let $\tau_k,\sigma_k\colon\Omega\to[0,\infty)$ be defined by

$$\tau_k=\sum_{n=0}^{\infty}\left(\frac{n+1}{2^k}\right)\mathbf{1}_{\left\{\frac{n}{2^k}<\tau\leq\frac{n+1}{2^k}\right\}}\ \text{ and }\ \sigma_k=\sum_{n=0}^{\infty}\left(\frac{n+1}{2^k}\right)\mathbf{1}_{\left\{\frac{n}{2^k}<\sigma\leq\frac{n+1}{2^k}\right\}}.$$

The $\{\tau_k\}_{k\in\mathbb{N}}$ and $\{\sigma_k\}_{k\in\mathbb{N}}$ are decreasing sequences of stopping times which satisfy $\tau\leq\tau_k$ and $\sigma\leq\sigma_k$ for every $k\in\mathbb{N}$ and, almost surely as $k\to\infty$,

$$\tau_k\to\tau\ \text{ and }\ \sigma_k\to\sigma. \tag{2.21}$$

The boundedness of $\tau$ and $\sigma$ prove that $\tau_k$ and $\sigma_k$ take only finitely many values and almost surely satisfy $\tau_k\leq K+1$ and $\sigma_k\leq K+1$. A repetition of the proof of Theorem 2.12 applied to the stopping time $\tau_k$ and the constant stopping time $K+1$, and the stopping time $\sigma_k$ and constant stopping time $K+1$, proves that, for every $k\in\mathbb{N}$,

$$\mathbb{E}[M_{K+1}|\mathcal{F}_{\tau_k}]=S_{\tau_k}\ \text{ and }\ \mathbb{E}[M_{K+1}|\mathcal{F}_{\sigma_k}]=S_{\sigma_k}. \tag{2.22}$$

Lemma 2.18 and (2.22) prove that that the families of random variables $\{M_{\tau_k}\}_{k\in\mathbb{N}}$ and $\{M_{\sigma_k}\}_{k\in\mathbb{N}}$ are uniformly integrable, and (2.21) proves that, almost surely as $k\to\infty$,

$$M_{\tau_k}\to M_\tau\ \text{ and }\ M_{\sigma_k}\to M_\sigma.$$

Vitalli's convergence theorem therefore proves that, as $k\to\infty$,

$$M_{\tau_k}\to M_\tau\ \text{ and }\ M_{\sigma_k}\to M_\sigma\ \text{ in }\ L^1(\Omega). \tag{2.23}$$

Since a repetition of the proof of Theorem 2.12 applied to the stopping time $\tau_k$ and $\sigma_k$ proves that, for each $k\in\mathbb{N}$,

$$\mathbb{E}[M_{\tau_k}]=\mathbb{E}[M_{\sigma_k}],$$

the convergence (2.23) proves that, for any two bounded stopping times $\sigma\leq\tau$,

$$\mathbb{E}[M_\tau]=\mathbb{E}[M_\sigma]. \tag{2.24}$$

Now let $B\in\mathcal{F}_\sigma$ and define the stopping times $\tau_B=\tau\mathbf{1}_B+K\mathbf{1}_{B^c}$ and $\sigma_B=\tau\mathbf{1}_B+K\mathbf{1}_{B^c}$. It follows from Lemma 2.11, which proves that $\mathcal{F}_\sigma\subseteq\mathcal{F}_\tau$, that $\tau_B$ and $\tau_\sigma$ are bounded stopping times. Therefore, the definitions and (2.24) prove that

$$\mathbb{E}[M_{\tau_B}]=\mathbb{E}[M_\tau\colon B]+\mathbb{E}[M_K\colon B^c]=\mathbb{E}[M_\sigma\colon B]+\mathbb{E}[M_K\colon B^c]=\mathbb{E}[M_{\sigma_B}].$$

We therefore conclude that, for every $B\in\mathcal{F}_\sigma$,

$$\mathbb{E}[M_\tau\colon B]=\mathbb{E}[M_\sigma\colon B].$$

Since $M_\sigma$ is $\mathcal{F}_\sigma$-measurable, this proves that $\mathbb{E}[M_\tau|\mathcal{F}_\sigma]=M_\sigma$ and completes the proof. □

In this course, we will most often be dealing with stopping times that are not bounded. The following corollary will therefore be very useful in our applications of the optional stopping theorem.

**Corollary 2.20.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $\{\mathcal{F}_t\}_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, let $(M_t)_{t\in[0,\infty)}$ be an $\mathcal{F}_t$-martingale (sub/super), and let $\sigma\leq\tau$ be two finite $\mathcal{F}_t$-stopping times. Assume that $M_\tau$ and $M_\sigma$ are integrable and that*

$$\lim_{n\to\infty}\mathbb{E}[|M_n|:\tau>n]=0. \tag{2.25}$$

*Then,*

$$\mathbb{E}[M_\tau|\mathcal{F}_\sigma]=M_\sigma\ \ (\geq/\leq).$$

*In particular,*

$$\mathbb{E}[M_\tau]=\mathbb{E}[M_\sigma]\ \ (\geq/\leq).$$

*Proof.* Let $B \in \mathcal{F}_\sigma$. For every $n \in \mathbb{N}$, since $\sigma \wedge n \leq \tau \wedge n$ are stopping times and since $B \cap \{\sigma \leq n\} \in \mathcal{F}_{\sigma \wedge n}$, Theorem 2.19 proves that

$$\begin{aligned}\mathbb{E}[M_{\tau\wedge n} \colon B, \sigma \leq n] &= \mathbb{E}[M_\tau \colon B, \tau \leq n, \sigma \leq n] + \mathbb{E}[M_N \colon B, \tau > n, \sigma \leq n] \\ &= \mathbb{E}[M_\sigma \colon B, \sigma \leq n] \\ &= \mathbb{E}[M_{\sigma\wedge n} \colon B, \sigma \leq n].\end{aligned}$$

The integrability of $M_\tau$ and $M_\sigma$, the finiteness of the stopping times, and the dominated convergence theorem prove that

$$\lim_{N\to\infty} \mathbb{E}[M_\tau \colon B, \tau \leq n, \sigma \leq n] = \mathbb{E}[M_\tau \colon B] \;\text{ and }\; \lim_{N\to\infty} \mathbb{E}[M_\sigma \colon B, \sigma \leq n] = \mathbb{E}[M_\sigma \colon B].$$

Assumption (2.25) proves that

$$\lim_{N\to\infty} |\mathbb{E}[M_n \colon B, \sigma \leq n, \tau > n]| \leq \lim_{N\to\infty} \mathbb{E}[|M_n| : \tau > n] = 0.$$

Therefore, for every $B \in \mathcal{F}_\sigma$,

$$\mathbb{E}[M_\tau \colon B] = \mathbb{E}[M_\sigma \colon B],$$

which completes the proof. □

Finally, we will see in Proposition 2.35 below that there is also a version of the optional stopping theorem for uniformly integrable martingales that completely avoids imposing integrability conditions on the stopping time itself. However, in this course, the martingales we encounter will not usually be uniformly integrable. For this reason, and because the stopping times we encounter will not usually be bounded, it is good to keep Corollary (2.20) in mind.

2.5. **Doob's martingale inequality.** In this section, we will establish fundamental inequalities for submartingales. In a sense that will be made more precise by the Doob-Meyer decomposition to follow in the next section, recall that a submartingale $(X_t)_{t\in[0,\infty)}$ is "increasing" in the sense that, for every $s \leq t \in [0,\infty)$,

$$X_s \leq \mathbb{E}[X_t|\mathcal{F}_s].$$

This implies in particular that, for every $s \leq t \in [0,\infty)$,

$$\mathbb{E}[X_s] \leq \mathbb{E}[X_t].$$

Doob's martingale inequality proves that the running maximum $X_T^*$ of a nonnegative submartingale, definedy for every $T \in [0,\infty)$ by

$$X_T^* = \max_{0\leq t\leq T} |X_t|\,, \tag{2.26}$$

can be estimated in $L^p$-norms by $X_T$. We will first present the results for discrete martingales. We will extend these results to continuous martingales in the next section.

**Definition 2.21.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $X\colon \Omega \to \mathbb{R}$ be a random variable. For every $p \in [1,\infty)$,

$$\|X\|_p = \left(\int_\Omega |X|^p \,\mathrm{d}\mathbb{P}\right)^{\frac{1}{p}},$$

and

$$\|X\|_\infty = \text{ess sup}_{\omega\in\Omega} |X(\omega)|\,.$$

**Definition 2.22.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $X\colon \Omega \to \mathbb{R}$ be a random variable. For every $B \in \mathcal{F}$,

$$\mathbb{E}[X\colon B] = \int_B X \,\mathrm{d}\mathbb{P}.$$

**Proposition 2.23.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\{\mathcal{F}_t\}_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(X_t)_{t\in[0,\infty)}$ be a nonnegative $\mathcal{F}_t$-submartingale. Then, for every $T \in [0,\infty)$ and $\lambda \in [0,\infty)$,*

$$\lambda\mathbb{P}[X_T^* \geq \lambda] \leq \mathbb{E}[X_T\colon \{X_T^* \geq \lambda\}].$$

*Proof.* Let $T \in [0,\infty)$ and $\lambda \in [0,\infty)$. Define the stopping time $\tau$ by

$$\tau = (\inf\{s \in [0,\infty)\colon X_s \geq \lambda\}) \wedge T.$$

Since the constant function $T$ is also a stopping time, and since by definition $\tau \leq T$, the optional stopping theorem, the fact that $(X_t)_{t\in[0,\infty)}$ is a submartingale, and the definition of $\tau$ prove that

$$\begin{aligned}\mathbb{E}[X_T] &\geq \mathbb{E}[X_\tau]\\ &= \mathbb{E}[X_T\colon \{X_T^* < \lambda\}] + \mathbb{E}[X_\tau\colon \{X_T^* \geq \lambda\}]\\ &\geq \mathbb{E}[X_T\colon \{X_T^* < \lambda\}] + \lambda\mathbb{P}[X_T^* \geq \lambda].\end{aligned}$$

Therefore, by linearity of the expectation,

$$\mathbb{E}[X_T\colon \{X_T^* \geq \lambda\}] \geq \lambda\mathbb{P}[X_T^* \geq \lambda],$$

which completes the proof. □

In the following proposition, we prove for every $p \in (1,\infty)$ that the $L^p$-norm of the running maximum $X_T^*$ is controlled from above by the $L^p$-norm of $X_T$. The case $p = 2$ is the *Doob-Kolmogorov inequality.*

**Proposition 2.24.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\{\mathcal{F}_t\}_{t\in[0,\infty)}$ be a filtration on $(\Omega, \mathcal{F})$, and let $(X_t)_{t\in[0,\infty)}$ be a nonnegative $\mathcal{F}_t$-submartingale. Then, for every $T \in [0,\infty)$ and $\lambda \in [0,\infty)$,*

$$\lambda \mathbb{P}[X_T^* \geq \lambda] \leq \mathbb{E}[X_T \colon \{X_T^* \geq \lambda\}] \leq \mathbb{E}[X_T], \tag{2.27}$$

*and, for every $p \in (1,\infty)$,*

$$\|X_T^*\|_p \leq \left(\frac{p}{1-p}\right) \|X_T\|_p \,. \tag{2.28}$$

*Proof.* Estimate (2.29) is an immediate consequence of Proposition 2.23. It remains to prove (2.30). Let $p \in (1,\infty)$ and $T \in [0,\infty)$. It follows from (2.29) that

$$\begin{aligned}
\mathbb{E}[|X_T^*|^p] &= p\int_0^\infty \lambda^{p-1}\mathbb{P}[X_T^* \geq \lambda]\,\mathrm{d}\lambda \\
&\leq p\int_0^\infty \lambda^{p-2}\mathbb{E}[X_T \colon \{X_T^* \geq \lambda\}]\,\mathrm{d}\lambda \\
&= p\int_\Omega\int_0^\infty \lambda^{p-2}X_T\mathbf{1}_{\{X_T^*\geq\lambda\}}\,\mathrm{d}\lambda\,\mathrm{d}\mathbb{P} \\
&= \left(\frac{p}{p-1}\right)\int_\Omega (X_T^*)^{p-1}X_T\,\mathrm{d}\mathbb{P}.
\end{aligned}$$

Hölder's inequality with exponents $p$ and $\frac{p}{p-1}$ then proves that

$$\mathbb{E}[|X_T^*|^p] \leq \left(\frac{p}{p-1}\right)\mathbb{E}[|X_T^*|^p]^{\frac{p-1}{p}}\,\mathbb{E}[|X_T|^p]^{\frac{1}{p}}.$$

In terms of $L^p$-norms, this implies that

$$\|X_T^*\|_p^p \leq \left(\frac{p}{p-1}\right)\|X_T^*\|^{p-1}\,\|X_T\|_p\,,$$

which proves that, after dividing by $\|X_T^*\|^{p-1}$,

$$\|X_T^*\|_p \leq \left(\frac{p}{p-1}\right)\|X_T\|_p\,.$$

This completes the proof. □

The following corollary is an immediate consequence of Jensen's inequality and Propositions 2.23 and 2.24.

**Corollary 2.25.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\{\mathcal{F}_t\}_{t\in[0,\infty)}$ be a filtration on $(\Omega, \mathcal{F})$, and let $(X_t)_{t\in[0,\infty)}$ be a $\mathcal{F}_t$-martingale. Then, for every $N \in \mathbb{N}_0$ and $\lambda \in [0,\infty)$,*

$$\lambda \mathbb{P}[X_T^* \geq \lambda] \leq \mathbb{E}[|X_T| : \{X_T^* \geq \lambda\}] \leq \mathbb{E}[|X_T|], \tag{2.29}$$

*and, for every $p \in (1,\infty)$,*

$$\|X_T^*\|_p \leq \left(\frac{p}{1-p}\right) \|X_T\|_p \,. \tag{2.30}$$

*Proof.* Jensen's inequality proves that $(|X_t|)_{t\in[0,\infty)}$ is a submartingale, since the absolute value function is convex. The proof then follows from Propositions 2.23 and 2.24. □

2.6. **Doob's martingale convergence theorem.** In this section, we will prove the martingale convergence theorem. The convergence is obtained by proving that, in the long run, an integrable martingale cannot oscillate to infinity. This is most easily seen on the level of $L^2$-bounded martingales. Precisely, suppose that $(M_n)_{n\in\mathbb{N}_0}$ is an $L^2$-bounded martingale on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ in the sense that

$$\sup_{n\in\mathbb{N}_0} \mathbb{E}[M_n^2] < \infty. \tag{2.31}$$

Since martingale increments in some ways behave like independent random variables, in the sense that, for every $n \in \mathbb{N}_0$,

$$\begin{aligned}
\mathbb{E}[M_n^2] &= \mathbb{E}\left[\left(\sum_{k=1}^n (M_k - M_{k-1})\right)^2\right] \\
&= \mathbb{E}\left[\sum_{k=0}^n (M_k - M_{k-1})^2\right] + 2\mathbb{E}\left[\sum_{j<k=1}^n (M_k - M_{k-1})(M_j - M_{j-1})\right] \\
&= \mathbb{E}\left[\sum_{k=0}^n (M_k - M_{k-1})^2\right] + 2\mathbb{E}\left[\sum_{j<k=1}^n \mathbb{E}[(M_k - M_{k-1})(M_j - M_{j-1})|\mathcal{F}_{k-1}]\right] \\
&= \mathbb{E}\left[\sum_{k=0}^n (M_k - M_{k-1})^2\right] + 2\mathbb{E}\left[\sum_{j<k=1}^n (\mathbb{E}[M_k|\mathcal{F}_{k-1}] - M_{k-1})(M_j - M_{j-1})\right] \\
&= \mathbb{E}\left[\sum_{k=0}^n (M_k - M_{k-1})^2\right].
\end{aligned}$$

Therefore, after passing to the limit $n \to \infty$, it follows from (2.31) that

$$\mathbb{E}\left[\sum_{k=0}^\infty (M_k - M_{k-1})^2\right] < \infty. \tag{2.32}$$

Since the same computation proves that, for every $n \leq m \in \mathbb{N}_0$,

$$\mathbb{E}[(M_n - M_m)^2] = \mathbb{E}\left[\sum_{k=m+1}^n (M_k - M_{k-1})^2\right],$$

it follows from (2.32) that $\{M_n\}_{n\in\mathbb{N}_0}$ is a Cauchy sequence in $L^2(\Omega)$. Therefore, there exists $M_\infty \in L^2(\Omega)$ such that, as $n \to \infty$,

$$M_n \to M_\infty \text{ strongly in } L^2(\Omega).$$

The estimate (2.32) implies almost surely that the sequence $\{M_n\}_{n\in\mathbb{N}_0}$ does not oscillate to infinity in the sense that the differences $|M_k - M_{k-1}|^2$ decay in an summable fashion.

We will make the notion of oscillation precise by defining an *upcrossing*. Let $a < b \in \mathbb{R}$, and suppose that $(M_t)_{t\in[0,\infty)}$ is a continuous (sub/super)martingale. We then define the following infinite sequence of stopping times $\{\tau_k\}_{k\in\mathbb{N}}$ by

$$\begin{aligned}
\tau_0 &= 0\\
\tau_1 &= \inf\{t\in[0,\infty)\colon M_t\leq a\}\\
\tau_2 &= \inf\{t\in[0,\infty)\colon t\geq \tau_1 \text{ and } M_t\geq b\}\\
\tau_3 &= \inf\{t\in[0,\infty)\colon t\geq \tau_2 \text{ and } M_t\leq a\}\\
\tau_4 &= \inf\{t\in[0,\infty)\colon t\geq \tau_3 \text{ and } M_t\geq b\}\\
&\ \vdots\\
\tau_{2k-1} &= \inf\{t\in[0,\infty)\colon t\geq \tau_{2k-2} \text{ and } M_t\leq a\}\\
\tau_{2k} &= \inf\{t\in[0,\infty)\colon t\geq \tau_{2k-1} \text{ and } M_t\geq b\}\\
&\ \vdots
\end{aligned}$$

We are interested in the number of instances over the interval $[0,T]$ that the martingale moves from below the value $a$ to above the value $b$. Such an event is called an *upcrossing* from $a$ to $b$. For a (sub/super) martingale $(M_t)_{t\in[0,\infty)}$ and for $T\in(0,\infty)$, we define $U(a,b;M,T)$ to be the number of upcrossings on $[0,T]$,

$$U(a,b;M,T)=\sup\{k\in\mathbb{N}\colon \tau_{2k}\leq T\}. \tag{2.33}$$

The following proposition, which is *Doob's upcrossing inequality*, estimates the expectation of $U(a,b;M,T)$ for a super martingale $(M_t)_{t\in[0,\infty)}$.

**Proposition 2.26.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(M_t)_{t\in[0,\infty)}$ be a continuous super martingale. Then, for every $T\in(0,\infty)$ and $a<b\in\mathbb{R}$,*

$$\mathbb{E}[U(a,b;M,T)]\leq \frac{1}{b-a}\mathbb{E}[(X_T-a)^-],$$

*where $(M_T-a)^-=-\min((M_T-a),0)$.*

*Proof.* Let $T\in(0,\infty)$ and let $K\in\mathbb{N}$. Then,

$$M_{\tau_{2K}\wedge T}=\sum_{k=1}^K\left(M_{\tau_{2k}\wedge T}-M_{\tau_{2k-1}\wedge T}\right)+\sum_{k=1}^K\left(M_{\tau_{2k-1}\wedge T}-M_{\tau_{2k-2}\wedge T}\right)+M_0.$$

We first observe that, by continuity of the process $(M_t)_{t\in[0,\infty)}$,

$$\begin{aligned}
&\sum_{k=1}^K\left(M_{\tau_{2k}\wedge T}-M_{\tau_{2k-1}\wedge T}\right)\\
&=\sum_{k=1}^{K\wedge U(a,b;M,T)}\left(M_{\tau_{2k}\wedge T}-M_{\tau_{2k-1}\wedge T}\right)+\sum_{k=U(a,b;M,T)+1}^{K}\left(M_{\tau_{2k}\wedge T}-M_{\tau_{2k-1}\wedge T}\right)\\
&=(b-a)(U(a,b;M,T)\wedge K)+(M_T-M_{\tau_{2U(a,b;M,T)+1}})\mathbf{1}_{\{\tau_{2U(a,b;M,T)+1}\leq T\}}\\
&=(b-a)(U(a,b;M,T)\wedge K)+(M_T-a)\mathbf{1}_{\{\tau_{2U(a,b;M,T)+1}\leq T\}}
\end{aligned}$$

Therefore, after rearranging the sums and observing the cancellation of the terms involving $M_0$,

$$\begin{aligned}
&(b-a)(U(a,b;M,T)\wedge K)+(M_T-M_{\tau_{2U(a,b;M,T)+1}})\mathbf{1}_{\{\tau_{2U(a,b;M,T)+1}\leq T\}}\\
&=\sum_{k=1}^K M_{\tau_{2k}\wedge T}-M_{\tau_{2k-1}\wedge T}.
\end{aligned}$$

Since the super martingale property and the optional stopping theorem applied to the bounded stopping times $\tau_i \wedge T$ prove that

$$\mathbb{E}\left[\sum_{k=1}^{K}\left(M_{\tau_{2k}\wedge T} - M_{\tau_{2k-1}\wedge T}\right)\right] = \sum_{k=2}^{K}\left(\mathbb{E}[M_{\tau_{2k}\wedge T}] - \mathbb{E}[M_{\tau_{2k-1}\wedge T}]\right) \leq 0,$$

it follows that

$$(b-a)\mathbb{E}[(U(a,b;M,T)\wedge K)] + \mathbb{E}[(M_T - a)\mathbf{1}_{\{\tau_{2U(a,b;M,T)+1}\leq T\}}] \leq 0. \tag{2.34}$$

We therefore conclude from (2.34) that, for every $K \in \mathbb{N}$,

$$\mathbb{E}[U(a,b;M,T)\wedge K] \leq \frac{1}{(b-a)}\mathbb{E}[-(M_T-a)\mathbf{1}_{\{\tau_{2U(a,b;M,T)+1}\leq T\}}] \leq \frac{1}{(b-a)}\mathbb{E}[(M_T-a)^-].$$

Since by continuity of $(M_t)_{t\in[0,\infty)}$, which implies the uniform continuity of $(M_t)_{t\in[0,\infty)}$ on $[0,T]$, it follows that, for almost every $\omega \in \Omega$,

$$U(a,b;M,T)(\omega) < \infty,$$

the monotone convergence theorem proves that, after passing to the limit $K \to \infty$,

$$\mathbb{E}[U(a,b;M,T)] \leq \frac{1}{b-a}\mathbb{E}[(M_T-a)^-],$$

which completes the proof. □

Upcrossings are convenient for supermartingales, and downcrossings are convenient for submartingales. That is, given a (sub/super) martingale $(M_t)_{t\in[0,\infty)}$ define the stopping times

$$\begin{aligned}
\tau_0 &= 0\\
\tau_1 &= \inf\{t\in[0,\infty)\colon M_t \geq b\}\\
\tau_2 &= \inf\{t\in[0,\infty)\colon t\geq\tau_1 \text{ and } M_t \leq a\}\\
\tau_3 &= \inf\{t\in[0,\infty)\colon t\geq\tau_2 \text{ and } M_t \geq b\}\\
\tau_4 &= \inf\{t\in[0,\infty)\colon t\geq\tau_3 \text{ and } M_t \leq a\}\\
&\vdots\\
\tau_{2k-1} &= \inf\{t\in[0,\infty)\colon t\geq\tau_{2k-2} \text{ and } M_t \geq b\}\\
\tau_{2k} &= \inf\{t\in[0,\infty)\colon t\geq\tau_{2k-1} \text{ and } M_t \leq a\}\\
&\vdots
\end{aligned}$$

and let $D(b,a;M,T) = \inf\{k\in\mathbb{N}_0\colon \tau_{2k} \leq T\}$. This is the total number of downcrossings of $(M_t)_{t\in[0,\infty)}$ from $b$ to $a$. The following is *Doob's downcrossing inequality.*

**Proposition 2.27.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(M_t)_{t\in[0,\infty)}$ be a continuous submartingale. Then, for every $T\in(0,\infty)$ and $a<b\in\mathbb{R}$,*

$$\mathbb{E}[D(b,a;M,T)] \leq \frac{1}{b-a}\mathbb{E}[(M_T-b)^+],$$

*where $(M_T-a)^+- = \max((M_T-a),0)$.*

*Proof.* Since $(M_t)_{t\in[0,\infty)}$ is a submartingales, its negative $(-M_t)_{t\in[0,\infty)}$ is a supermartingale. Furthermore, a down crossing of $(M_t)_{t\in[0,\infty)}$ from $b$ to $a$ is an upcrossing of $(-M_t)_{t\in[0,\infty)}$ from $-b$ to

$-a$. That is, $D(b,a;M,T) = U(-b,-a;-M,T)$. Therefore, by Proposition 2.26,

$$\begin{aligned} E[D(b,a;M,T)] &= \mathbb{E}[U(-b,-a;M,T)] \\ &\leq \frac{1}{((-a)) - (-b))}\mathbb{E}[(-M_T - (-b))^-] \\ &= \frac{1}{(b-a)}\mathbb{E}[(M_t - b)^+], \end{aligned}$$

which completes the proof. □

Given a (sub/super) martingale $(M_t)_{t\in[0,\infty)}$ and $a < b \in \mathbb{R}$, we define the total number of upcrossings from $a$ to $b$ to be

$$U(a,b;M) = \lim_{T\to\infty} U(a,b;M,T),$$

and the total number of downcrossings to be

$$D(b,a;M) = \lim_{T\to\infty} D(b,a;M,T).$$

This limit always exists, as the limit of an increasing sequence, but it may be infinite. We will now prove the first version of the martingale convergence theorem, which states that an $L^1$-bounded submartingale converges almost surely as $t \to \infty$.

**Proposition 2.28.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(M_t)_{t\in[0,\infty)}$ be a continuous submartingale. Assume that $M_0 \in L^1(\Omega)$ and that*

$$\sup_{t\in[0,\infty)} \mathbb{E}[M_t^+] < \infty. \tag{2.35}$$

*Then, there exists $M_\infty \in L^1(\Omega)$ such that, for almost every $\omega \in \Omega$, as $t \to \infty$,*

$$M_t(\omega) \to M_\infty(\omega). \tag{2.36}$$

**Remark 2.29.** Note carefully that (2.40) implies only almost sure convergence, it does not imply convergence in $L^1(\Omega)$. Convergence in $L^1(\Omega)$ requires the additional assumption of uniform integrability, as explained below.

*Proof.* It follows from Proposition 2.27 and (2.39) that there exists $c \in (0,\infty)$ such that, for every $T \in (0,\infty)$ and $a < b \in \mathbb{R}$,

$$\mathbb{E}[D(b,a;M,T)] \leq \frac{1}{(b-a)}\mathbb{E}[(X_T - b)^+] \leq \frac{1}{(b-a)}\left(|b| + \mathbb{E}[X_T^+]\right) \leq \frac{c+|a|}{b-a}.$$

Therefore, for every $a < b \in \mathbb{R}$, after passing to the limit $T \to \infty$ using the monotone convergence theorem,

$$\mathbb{E}[D(b,a;M)] \leq \frac{c+|a|}{b-a}.$$

We therefore conclude that, for every $a < b \in \mathbb{R}$, there exists a subset $\Omega_{a,b} \subseteq \Omega$ of full probability such that

$$D(b,a;M)(\omega) < \infty \;\text{ for every }\; \omega \in \Omega.$$

We then define the subset $\Omega' \subseteq \Omega$ of full probability

$$\Omega' = \cap_{a<b\in\mathbb{Q}}\Omega_{a,b},$$

and claim that, for every $\omega \in \Omega'$,

$$\lim_{t\to\infty} M_t(\omega) \;\text{ exists.} \tag{2.37}$$

Since the limit (2.37) exists if and only if

$$\limsup_{t\to\infty} M_t(\omega) = \liminf_{t\to\infty} M_t(\omega),$$

suppose by contradiction that there exist $\omega \in \Omega'$ and $\overline{a}, \overline{b} \in [-\infty, \infty]$ such that

$$\overline{a} = \liminf_{t\to\infty} M_t(\omega) < \limsup_{t\to\infty} M_t(\omega) = \overline{b}. \tag{2.38}$$

By density of the rationals, there exist $a < b \in \mathbb{Q}$ such that $\overline{a} < a < b < \overline{b}$, and by definition of the lim inf and lim sup it follows from (2.38) that

$$D(b, a; M)(\omega) = \infty,$$

which contradicts the assumption that $\omega \in \Omega' \subseteq \Omega_{a,b}$.

We therefore define the limit $M_\infty$ for every $\omega \in \Omega'$ by

$$M_\infty(\omega) = \lim_{t\to\infty} M_t(\omega),$$

and conclude using Fatou's lemma, (2.39), and the submartingale property that

$$\begin{aligned}
\mathbb{E}[|M_\infty|] = \mathbb{E}[(\lim_{t\to\infty} |M_t|)] &\leq \liminf_{t\to\infty} (\mathbb{E}[|M_t|]) \\
&= \liminf_{t\to\infty} \big(2\mathbb{E}[M_t^+] - \mathbb{E}[M_t]\big) \\
&\leq \liminf_{t\to\infty} \big(2\mathbb{E}[M_t^+] - \mathbb{E}[M_0]\big) \\
&\leq \liminf_{t\to\infty} \big(2\mathbb{E}[M_t^+] + \mathbb{E}[|M_0|]\big) < \infty,
\end{aligned}$$

which proves that $M_\infty \in L^1(\Omega)$. This completes the proof. □

**Proposition 2.30.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega, \mathcal{F})$, and let $(M_t)_{t\in[0,\infty)}$ be a continuous super martingale. Assume that $M_0 \in L^1(\Omega)$ and that*

$$\sup_{t\in[0,\infty)} \mathbb{E}[M_t^-] < \infty. \tag{2.39}$$

*Then, there exists $M_\infty \in L^1(\Omega)$ such that, for almost every $\omega \in \Omega$, as $t \to \infty$,*

$$M_t(\omega) \to M_\infty(\omega). \tag{2.40}$$

*Proof.* The proof is identical to the proof of Proposition 2.28, with the exception that Proposition 2.26 is used in place of Proposition 2.27. Also, in this case, to prove the integrability of $M_\infty$ we observe using the supermartingale property that

$$\begin{aligned}
\mathbb{E}[|M_\infty|] = \mathbb{E}[(\lim_{t\to\infty} |M_t|)] &\leq \liminf_{t\to\infty} (\mathbb{E}[|M_t|]) \\
&= \liminf_{t\to\infty} \big(2\mathbb{E}[M_t^-] + \mathbb{E}[M_t]\big) \\
&\leq \liminf_{t\to\infty} \big(2\mathbb{E}[M_t^-] + \mathbb{E}[M_0]\big) \\
&\leq \liminf_{t\to\infty} \big(2\mathbb{E}[M_t^-] + \mathbb{E}[|M_0|]\big) < \infty,
\end{aligned}$$

which completes the proof. □

**Remark 2.31.** In particular, it follows from Proposition 2.30 that nonnegative super martingales with an integrable initial condition always converge in almost sure sense to an integrable function. Again, I emphasize that this is only almost sure convergence. In Theorem 2.32 below, we will upgrade this convergence to strong convergence in $L^1(\Omega)$ for uniformly integrable martingales.

The following theorem is the *Doob's martingale convergence theorem* for a uniformly integrable martingale. It follows from either Proposition 2.28 or Proposition 2.30 that if $(M_t)_{t\in[0,\infty)}$ is uniformly integrable then, in an almost sure sense, $M_\infty(\omega) = \lim_{t\to\infty} M_t(\omega)$ exists and $M_\infty \in L^1(\Omega)$. The following theorem proves that not only is this true, but if $(M_t)_{t\in[0,\infty)}$ is uniformly integrable then $\lim_{t\to\infty} \mathbb{E}[|M_t - M_\infty|] = 0$ and $M_t = \mathbb{E}[M_\infty|\mathcal{F}_t]$ for every $t \in [0, \infty)$.

**Theorem 2.32.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega, \mathcal{F})$, and let $(M_t)_{t\in[0,\infty)}$ be a continuous $\mathcal{F}_t$-martingale. Then, the following three statements are equivalent.*

*(i) $(M_t)_{t\in[0,\infty)}$ is* closed*: there exists $M_\infty \in L^1(\Omega)$ such that, for every $t \in [0, \infty)$,*

$$M_t = \mathbb{E}[M_\infty|\mathcal{F}_t].$$

*(ii) The family $(M_t)_{t\in[0,\infty)}$ is uniformly integrable.*

*(iii) The family $(M_t)_{t\in[0,\infty)}$ converges, as $t \to \infty$, almost surely and in $L^1(\Omega)$.*

*Proof.* We will first prove that (i) implies (ii). Since $M_\infty \in L^1(\Omega)$, and since $M_t = \mathbb{E}[M_\infty|\mathcal{F}_t]$ for every $t \in [0, \infty)$, it follows from Lemma 2.18 that $(M_t)_{t\in[0,\infty)}$ is uniformly integrable. We will now prove that (ii) implies (iii). Since the uniform integrability implies that $(M_t)_{t\in[0,\infty)}$ is bounded in $L^1(\Omega)$, Proposition 2.28 proves that there exists $M_\infty \in L^1(\Omega)$ such that, as $t \to \infty$,

$$M_t \to M_\infty \text{ almost surely.}$$

Since almost sure convergence implies convergence in probability, the uniform integrability and Theorem 2.17 prove that

$$\lim_{t\to\infty} \mathbb{E}[|M_t - M_\infty|] = 0.$$

Finally, we will prove that $(iii)$ implies $(i)$. By assumption, there exists $M_\infty \in L^1(\Omega)$ such that $(M_t)_{t\in[0,\infty)}$ converges, as $t \to \infty$, to $M_\infty$ almost surely and strongly in $L^1(\Omega)$. Since the martingale property implies that, for every $s \le t \in [0, \infty)$,

$$M_s = \mathbb{E}[M_t|\mathcal{F}_s], \tag{2.41}$$

and since the conditional expectation is stable with respect to convergence in $L^1(\Omega)$, after passing to the limit $t \to \infty$ in (2.41) we have that, for every $s \in [0, \infty)$,

$$M_s = \mathbb{E}[M_\infty|\mathcal{F}_s].$$

This completes the proof. □

The following corollary applies Theorem 2.32 to martingales bounded in $L^p$, for $p \in (1, \infty)$. It is a consequence of the following lemma, which proves that a family of random variables that is bounded in $L^p(\Omega)$, for some $p \in (1, \infty)$, is necessarily uniformly integrable.

**Lemma 2.33.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{A}$ be a set, and let $\{X_\alpha\}_{\alpha\in\mathcal{A}}$ be a family of random variables. Assume that, for $p \in (1, \infty)$,*

$$\sup_{\alpha\in\mathcal{A}} \|X_\alpha\|_p < \infty. \tag{2.42}$$

*Then the family $\{X_\alpha\}_{\alpha\in\mathcal{A}}$ is uniformly integrable.*

*Proof.* We will first prove condition (i) of Definition 2.14. It follows from Hölder's inequality and (2.42) that

$$\sup_{\alpha\in\mathcal{A}} \|X_\alpha\|_{L^1(\Omega)} \le \sup_{\alpha\in\mathcal{A}} \|X_\alpha\|_p < \infty.$$

We will no prove condition (ii). For every $\alpha \in \mathcal{A}$ and $K \in (0, \infty)$, Hölder's inequality, Chebyshev' inequality, and $p \in (1, \infty)$ prove that

$$\mathbb{E}[|X_\alpha| : \{|X_\alpha| \ge K\}] \le \|X\|_p \mathbb{P}[|X_\alpha| \ge K]^{\frac{p-1}{p}} \le \|X\|_p \left(\frac{1}{K}\|X\|_p\right)^{p-1} = K^{-(p-1)} \|X\|_p^p.$$

Therefore, it follows from (2.42) and $p \in (1, \infty)$ that

$$\lim_{K\to\infty} \left(\sup_{\alpha\in\mathcal{A}} \mathbb{E}[|X_\alpha| : \{|X_\alpha| \ge K\}]\right) \le \lim_{K\to\infty} \left(K^{-(p-1)} \sup_{\alpha\in\mathcal{A}} \|X_\alpha\|_p^p\right) = 0.$$

This completes the proof. □

**Theorem 2.34.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega, \mathcal{F})$, and let $(M_t)_{t\in[0,\infty)}$ be a continuous $\mathcal{F}_t$-martingale. Assume that there exists $p \in (1,\infty)$ such that*

$$\sup_{t\in[0,\infty)} \|M_t\|_p < \infty. \tag{2.43}$$

*Then there exists $M_\infty \in L^p(\Omega)$ such that*

$$\lim_{t\to\infty} \|M_t - M_\infty\|_p = 0.$$

*Proof.* It follows from Lemma 2.33 that $(M_t)_{t\in[0,\infty)}$ is uniformly integrable. Therefore, by Theorem 2.32, there exists $M_\infty \in L^1(\Omega)$ such that, as $t \to \infty$,

$$M_t \to M_\infty \text{ almost surely and strongly in } L^1(\Omega). \tag{2.44}$$

It remains to prove that $M_\infty \in L^p(\Omega)$ and that the convergence takes place in $L^p(\Omega)$. For the first point, Fatou's lemma proves that

$$\mathbb{E}[|M_\infty|^p] = \mathbb{E}[(\lim_{t\to\infty} |M_t|^p)] \leq \liminf_{t\to\infty} \mathbb{E}[|M_t|^p] = \liminf_{t\to\infty} \|M_t\|_p^p < \infty.$$

Then, since it follows from Doob's martingale inequality that

$$\left\| \sup_{t\in[0,\infty)} |M_t|^p \right\|_p \leq \frac{p}{p-1} \limsup_{t\to\infty} \|M_t\|_p < \infty,$$

and since we have for every $t \in [0,\infty)$ that

$$|M_\infty - M_t| \leq \left( |M_\infty| + \sup_{t\in[0,\infty)} |M_t| \right) \in L^p(\Omega),$$

it follows from (2.44) and the dominated convergence theorem that

$$\lim_{t\to\infty} \|M_t - M_\infty\|_p = 0.$$

This completes the proof. □

In the final proposition of this section, we extend the optional stopping theorem to uniformly integrable martingales. Suppose that $(M_t)_{t\in[0,\infty)}$ is a continuous, uniformly integrable martingale. Then, by Theorem 2.32, there exists $M_\infty \in L^1(\Omega)$ such that, as $t \to \infty$,

$$M_t \to M_\infty \text{ almost surely and strongly in } L^1(\Omega).$$

Therefore, given a possibly infinite stopping time $T \colon \Omega \to [0,\infty]$, we define

$$M_T = M_T \mathbf{1}_{\{T<\infty\}} + M_\infty \mathbf{1}_{\{T=\infty\}}.$$

The following proposition proves that, when dealing with a uniformly integrable martingale, the optional stopping theorem applies even to stopping times that are infinite with positive probability. This stands in stark contrast to the example of the simple random walk above, which is not uniformly integrable, and for which the optional stopping theorem fails for a stopping time that is almost surely finite.

**Proposition 2.35.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega, \mathcal{F})$, and let $(M_t)_{t\in[0,\infty)}$ be a continuous $\mathcal{F}_t$-martingale. Assume that $(M_t)_{t\in[0,\infty)}$ is uniformly integrable. Then for every pair of $\mathcal{F}_t$-stopping times $S \leq T \colon \Omega \to [0,\infty]$,*

$$\mathbb{E}[M_T | \mathcal{F}_S] = M_S.$$

*In particular,*

$$\mathbb{E}[M_T] = \mathbb{E}[M_S].$$

*Proof.* For every bounded stopping time $T$, Theorem 2.19 proves that

$$\mathbb{E}[M_N|\mathcal{F}_T] = M_T.$$

The martingale convergence theorem then proves that, after passing to the limit $N \to \infty$,

$$\mathbb{E}[M_\infty|\mathcal{F}_T] = M_T.$$

Lemma 2.18 proves that the family of random variables $M_T$, for $T$ a bounded stopping time, is uniformly integrable. A repetition of the proof of Theorem 2.19 then proves that, for every stopping time $T\colon \Omega \to [0, \infty]$,

$$\mathbb{E}[M_\infty|\mathcal{F}_T] = M_T,$$

which proves using Lemma 2.18 that the family of random variables $M_T$, for $T$ a stopping time, is uniformly integrable. Then, for every pair of stopping times $S \leq T$, the tower property of conditional expectation and $\mathcal{F}_S \subseteq \mathcal{F}_T$ prove that

$$\mathbb{E}[M_T\colon \mathcal{F}_S] = \mathbb{E}[\mathbb{E}[M_\infty|\mathcal{F}_T]|\mathcal{F}_S] = \mathbb{E}[M_\infty|\mathcal{F}_S] = M_S,$$

which completes the proof. $\square$

## 3. Itô's Formula

In this section, we will develop the theory of stochastic integration on the way to proving Itô's formula, which is the fundamental theorem of the calculus of stochastic processes. We first observe that the integral with respect to a process of bounded variation is almost surely well-defined. Observe in particular that a nondecreasing or nonincreasing process is necessarily of finite variation.

**Definition 3.1.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $a < b \in \mathbb{R}$. A stochastic process $(A_t)_{t\in[0,\infty)}$ has *bounded variation* on $[a, b]$ if

$$V(a, b; A) = \sup_{\Delta\subseteq[a,b]} \sum_{k=1}^{n(\Delta)} \left|A_{t_k} - A_{t_{k-1}}\right| < \infty \text{ almost surely,}$$

where the sup is over all partitions $\Delta = \{a = t_0 < t_1 < \ldots < t_{n(\delta)-1} < t_{n(\Delta)} = b\}$ of $[a, b]$.

Given a process of finite variation $(A_t)_{t\in[0,\infty)}$ and a continuous stochastic $(X_t)_{t\in[0,\infty)}$, we can define the integral of $(X_t)_{t\in[0,\infty)}$ with respect to $(A_t)_{t\in[0,\infty)}$ as the limit

$$\int_a^b X_t \, \mathrm{d}A_t = \lim_{|\Delta|\to 0} \sum_{k=1}^{n(\Delta)} X_{t_{k-1}}(A_{t_k} - A_{t_{k-1}}),$$

where the facts that $(A_t)_{t\in[0,\infty)}$ has finite variation and that $(X_t)_{t\in[0,\infty)}$ is continuous prove that the above limit is almost surely well-defined. We aim to extend the above definition to martingales, and later to semimartingales.

**Definition 3.2.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega, \mathcal{F})$. A stochastic process $(X_t)_{t\in[0,\infty)}$ is a *continuous semimartingale* if there exists a continuous process of finite variation $(A_t)_{t\in[0,\infty)}$ and a continuous local martingale $(M_t)_{t\in[0,\infty)}$ such that

$$X_t = M_t + A_t.$$

Given a martingale $(M_t)_{t\in[0,\infty)}$ and a continuous process $(X_t)_{t\in[0,\infty)}$, we will ultimately define the stochastic integral of $(X_t)_{t\in[0,\infty)}$ with respect to $(M_t)_{t\in[0,\infty)}$ to be the limit

$$\int_a^b X_t \, \mathrm{d}M_t = \lim_{|\Delta|\to 0, \Delta\subseteq[a,b]} \sum_{k=1}^{n(\Delta)} X_{t_{k-1}}(M_{t_k} - M_{t_{k-1}}).$$

However, unlike the finite variation case, the above limit is not well-defined in a pointwise sense. The above limit is well-defined because of certain stochastic cancellations that rely on the martingale property. These will imply in particular that the stochastic integral is itself a martingale beginning from zero, and therefore that, for every $a \le b \in [0, \infty)$,

$$\mathbb{E}\left[\int_a^b X_t \, \mathrm{d}M_t\right] = 0.$$

We have seen that a martingale is of finite variation if and only if it is constant. Specifically, for a Brownian motion $(B_t)_{t\in[0,\infty)}$, the limit along a sequence of partitions $\{\Delta_k \subseteq [0, t]\}_{k\in\mathbb{N}}$ satisfying $|\Delta_k| \to 0$ as $k \to \infty$,

$$\lim_{k\to\infty} \sum_{k=1}^{n(\Delta_k)} (B_{t_k} - B_{t_{k-1}})^2 = b - a = \langle B\rangle_b - \langle B\rangle_a,$$

exists almost surely. The continuous increasing process $(\langle B\rangle_t = t)_{t\in[0,\infty)}$ is the *quadratic variation* of Brownian motion. We will define the quadratic variation of a general local martingale $(M_t)_{t\in[0,\infty)}$ in the section to follow.

We will use the quadratic variation and a stochastic integration to prove Itô's formula, which is the fundamental theorem of the calculus of semimartingales. Let

$$(M_t = (M_t^1, \ldots, M_t^d))_{t\in[0,\infty)},$$

be a $d$-dimensional continuous martingale beginning from zero, and let $f \in \mathrm{C}^2(\mathbb{R}^d)$ be a twice-differentiable function. We are essentially interested in establishing a fundamental theorem of calculus for the process $f((M_t))_{t\in[0,\infty)}$. A first guess, based on the classical fundamental theorem of calculus, would be that

$$f(M_t) = f(M_0) + \int_0^t f'(M_s)\,\mathrm{d}M_s.$$

However, we can see immediately that such a formula cannot hold true unless $M = 0$ is constant. Indeed, if we choose $f(x) = |x|^2$, since stochastic integral is a martingale, for every $t \in [0,\infty)$,

$$\mathbb{E}[|M_t|^2] = \mathbb{E}[|M_0|] + \mathbb{E}\left[\int_0^t f'(M_s)\,\mathrm{d}M_s\right] = \mathbb{E}[|M_0|^2].$$

This implies that

$$\mathbb{E}[(M_0 + M_t)\cdot(M_t - M_0)] = 0,$$

from which it follows by the continuity of $(M_t)_{t\in[0,\infty)}$ that $M$ is constant. We therefore seek a higher order approximation of the process.

A second-order Taylor expansion of $f$ about $M_0$ proves for every $t \in [0,\infty)$ that

$$f(M_t) = f(M_0) + \nabla f(M_0)\cdot(M_t - M_0) + \frac{1}{2}\langle \nabla^2 f(M_0)(M_t - M_0), (M_t - M_0)\rangle + \mathcal{O}(|M_t - M_0|^3),$$

where

$$\nabla F(M_0)\cdot(M_t - M_0) = \sum_{i=1}^d \frac{\partial f}{\partial x_i}(M_0)(M_t^i - M_0^i),$$

and where

$$\langle \nabla^2 f(M_0)(M_t - M_0), (M_t - M_0)\rangle = \sum_{i,j=1}^d \frac{\partial^2 f}{\partial x_i \partial x_j}(M_0)(M_t^i - M_0^i)(M_t^j - M_0^j).$$

Therefore, on the interval $[0,1]$, for a partition

$$\Delta = \{0 = t_0 < t_1 < \ldots < t_{n(\Delta)-1} < t_{n(\Delta)} = 1\},$$

we have

(3.1)

$$\begin{aligned} f(M_1) - f(M_0) &= \sum_{k=1}^{n(\Delta)} (f(M_k) - f(M_{k-1})) \\ &= \sum_{k=1}^{n(\Delta)} \nabla f(M_{t_{k-1}})\cdot(M_{t_k} - M_{t_{k-1}}) \\ &\quad + \frac{1}{2}\sum_{k=1}^{n(\Delta)} \langle \nabla^2 f(M_{t_{k-1}})(M_{t_k} - M_{t_{k-1}}), (M_{t_k} - M_{t_{k-1}})\rangle + \sum_{k=1}^{n(\Delta)} \mathcal{O}(\left|M_{t_k} - M_{t_{k-1}}\right|^3). \end{aligned}$$

As the mesh $|\Delta| \to 0$, we expect by stochastic integration that

$$\sum_{k=1}^{n(\Delta)} \nabla f(M_{t_{k-1}})\cdot(M_{t_k} - M_{t_{k-1}}) \to \sum_{i=1}^d \int_0^1 \frac{\partial f}{\partial x_i}(M_s)\,\mathrm{d}M_s^i.$$

Motivated by the quadratic variation, we expect that, as $|\Delta| \to 0$,

$$\frac{1}{2}\sum_{k=1}^{n(\Delta)}\langle \nabla^2 f(M_{t_{k-1}})(M_{t_k}-M_{t_{k-1}}),(M_{t_k}-M_{t_{k-1}})\rangle \to \frac{1}{2}\sum_{i,j=1}^{d}\int_0^1 \frac{\partial^2 f}{\partial x_i \partial x_j}(M_s)\,\mathrm{d}\langle M^i, M^j\rangle_s,$$

where will define $(\langle M^i, M^j\rangle_t)_{t\in[0,\infty)}$ to be the quadratic covariation of the one-dimensional martingales $(M^i_t)_{t\in[0,\infty)}$ and $(M^j_t)_{t\in[0,\infty)}$, for each $i,j\in\{1,\ldots,d\}$. Finally, since the quadratic variation is bounded, we expect by continuity that, as $|\Delta|\to 0$,

$$\sum_{k=1}^{n(\Delta)}\mathcal{O}(\left|M_{t_k}-M_{t_{k-1}}\right|^3)\to 0.$$

In combination this will prove *Itô's formula*, which states that, for every $t\in[0,\infty)$,

$$f(M_t)-f(M_0)=\sum_{i=1}^{d}\int_0^t \frac{\partial f}{\partial x_i}(M_s)\,\mathrm{d}M^i_s+\frac{1}{2}\sum_{i,j=1}^{d}\int_0^t \frac{\partial^2 f}{\partial x_i\partial x_j}(M_s)\,\mathrm{d}\langle M^i,M^j\rangle_s.$$

In particular, if $(B_t)_{t\in[0,\infty)}$ is a one-dimensional Brownian motion,

$$f(B_t)=f(0)+\int_0^t f(B_x)\,\mathrm{d}B_s+\frac{1}{2}\int_0^t f''(B_s)\,\mathrm{d}s,$$

which implies in particular that

$$\int_0^t B_s\,\mathrm{d}B_s = B_t^2 - t.$$

Notice as well that Itô's formula implies that the process $(f(B_t)_t)_{t\in[0,\infty)}$ is a semimartingale, because the stochastic integral is martingale and the deterministic integral is a of bounded variation. This fact is true in general, we will prove below that if $(Z_t)_{t\in[0,\infty)}$ is a semimartingale then $(f(Z_t))_{t\in[0,\infty)}$ is also a semimartingale.

Indeed, the integration theory will be developed within the class of semimartingales. The theory will therefore apply to the case of Brownian motion, but it will not in general cover the case of integration with respect to *fractional Brownian motion*. We define a real-valued centered Gaussian process to be a real valued process $(X_t)_{t\in[0,\infty)}$ on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ with finite dimensional distributions that are normally distributed and mean zero. A centered Gaussian process $(X_t)_{t\in[0,\infty)}$ is a fractional Brownian motion with Hurst parameter $h\in(0,1)$ if $\mathbb{P}[X_0=0]=1$ and if, for every $s,t\in[0,\infty)$,

$$\mathbb{E}[X_tX_s]=\frac{1}{2}\left(t^{2h}+s^{2h}-|t-s|^{2h}\right).$$

A fractional Brownian motion satisfies the following four properties.

(a) For every $s\le t\in[0,\infty)$, the increment $X_t-X_s$ has mean zero and variance $|t-s|^{2h}$.
(b) For every $p\in(0,\infty)$ there exists $c_p\in(0,\infty)$ such that, for every $s\le t\in[0,\infty)$,

$$\mathbb{E}[|X_t-X_s|^p]=c_p\,|t-s|^{hp}.$$

(c) For every $\gamma\in(0,h)$, Fractional Brownian motion has a $\gamma$-Hölder continuous modification.
(d) Factional Brownian motion is not a semimartingale if $h\neq\frac{1}{2}$.

*Proof.* We will first prove (a). Let $h\in(0,1)$. It follows by assumption that, since the finite dimensional distributions are normally distributed with mean zero, for every $s\le t\in[0,\infty)$,

$$\mathbb{E}[X_t-X_s]=0.$$

For the variance, for every $s \leq t \in [0,\infty)$,

$$\begin{aligned}\mathbb{E}[|X_t - X_s|^2] &= \mathbb{E}[X_t^2] + \mathbb{E}[X_s^2] - 2\mathbb{E}[X_sX_t] \\ &= t^{2h} + s^{2h} - (t^{2h} + s^{2h} - |t-s|^{2h}) \\ &= |t-s|^{2h}.\end{aligned}$$

We will now prove (b). Let $h \in (0,1)$. Let $s \leq t \in [0,\infty)$. Since $(X_t - X_s)$ is normally distributed with mean zero and variance $|t-s|^{2h}$, it follows for every $p \in (0,\infty)$ that

$$\begin{aligned}\mathbb{E}[|X_t - X_s|^p] &= \int_{\mathbb{R}} |x|^p \left(2\pi |t-s|^{2h}\right)^{-\frac{1}{2}} \exp\left(-\frac{|x|^2}{2|t-s|^{2h}}\right) \mathrm{d}x \\ &= |t-s|^{hp} \int_{\mathbb{R}} |x|^p (2\pi)^{-\frac{1}{2}} \exp\left(-\frac{|x|^2}{2}\right) \mathrm{d}x,\end{aligned}$$

where the final equality follows from the change of variables $\tilde{x} = {x}/{|t-s|^h}$. Therefore, for every $p \in (0,\infty)$, after defining

$$c_p = \int_{\mathbb{R}} |x|^p (2\pi)^{-\frac{1}{2}} \exp\left(-\frac{|x|^2}{2}\right) \mathrm{d}x,$$

it follows for every $p \in (0,\infty)$ that

$$\mathbb{E}[|X_t - X_s|^p] = c_p |t-s|^{hp}.$$

We will now prove (c). Let $h \in (0,1)$. The Kolmogorov continuity criterion states that, since for every $p \in (0,\infty)$,

$$\mathbb{E}[|X_t - X_s|^p] = c_p |t-s|^{hp},$$

fractional Brownian motion has a Hölder continuous modification with Hölder exponent $\gamma \in (0, h - \frac{1}{p})$ for every $p \in (0,\infty)$. That is, fractional Brownian motion is almost surely Hölder continuous for any Hölder exponent $\gamma \in (0,h)$. We will now prove (d). We will first identify the scaling properties of fractional Brownian motion. Let $h \in (0,1)$ and $\lambda \in (0,\infty)$. We aim to identify $\alpha \in \mathbb{R}$, such that the process $(\lambda X_{\lambda^\alpha s})_{s\in[0,\infty)}$ is a fractional Brownian motion. Since this process is a centered Gaussian process with mean zero, it is sufficient to identify $\alpha \in \mathbb{R}$ such that the covariance satisfies

$$\mathbb{E}[\lambda X_{\lambda^\alpha t} \lambda X_{\lambda^\alpha s}] = t^{2h} + s^{2h} - |t-s|^{2h}.$$

For this, notice that

$$\mathbb{E}[\lambda X_{\lambda^\alpha t} \lambda X_{\lambda^\alpha s}] = \lambda^2 \mathbb{E}[X_{\lambda^\alpha t} X_{\lambda^\alpha s}] = \lambda^{2-2\alpha h}\left(t^{2h} + s^{2h} - |t-s|^{2h}\right).$$

We therefore require that $\lambda^{2-2\alpha h} = 1$, which implies that $\alpha = -{1}/{h}$. We conclude that, for every $\alpha \in (0,\infty)$, if $(X_t)_{t\in[0,\infty)}$ is a fractional Brownian motion with Hurst parameter $h \in (0,\infty)$ then $(\lambda X_{\lambda^{-1/h} t})_{t\in[0,\infty)}$ is also a fractional Brownian motion with the same Hurst parameter.

We will now identify $p \in [1,\infty)$ depending on $h \in (0,1)$ such that fractional Brownian motion with Hurst parameter $h \in (0,1)$ has finite $p$-variation. By this we mean that, if $\{\Delta_n = \{0 = t_0 < t_1 < \ldots < t_{k_{n-1}} < t_{k_n} = 1\}\}_{n\in\mathbb{N}}$ is a sequence of partitions of $[0,1]$ such that $|\Delta_n| \to 0$ as $n \to \infty$, then

$$\lim_{k\to\infty} \sum_{\Delta_n} \left|B_{t_k} - B_{t_{k-1}}\right|^p \text{ exists and is finite almost surely.}$$

For simplicity, we consider the sequence of partitions

$$\{\Delta_n = 0 < {1}/{2^n} < {2}/{2^n} < \ldots < {2^n - 1}/{2^n} < 1\}_{n\in\mathbb{N}},$$

where, for every $n \in \mathbb{N}$, the scaling properties of fractional Brownian motion prove that in law, for every $p \in [1, \infty)$,

$$\sum_{k=1}^{2^n} \left| B_{\frac{k}{2^n}} - B_{\frac{k-1}{2^n}} \right|^p = \sum_{k=1}^{2^n} 2^{-nph} \left| 2^{nh} B_{\frac{k}{2^n}} - B_{\frac{k-1}{2^n}} \right|^p = 2^{-nph} \sum_{k=1}^{2^n} \left| B_k - B_{k-1} \right|. \tag{3.2}$$

The sequence $\{B_k - B_{k-1}\}_{k \in \mathbb{N}}$ is a stationary and ergodic sequence. This means that the random variables $\{B_k - B_{k-1}\}_{k \in \mathbb{N}}$ are identically distributed, so that in particular we have, for every $n \in \mathbb{N}$, since $B_0 = 0$,

$$\mathbb{E}\left[ \sum_{k=1}^{2^n} \left| B_{\frac{k}{2^n}} - B_{\frac{k-1}{2^n}} \right|^p \right] = 2^{-nph} \cdot 2^n \mathbb{E}[|B_1|] = 2^{n(1-ph)} \mathbb{E}[|B_1|]. \tag{3.3}$$

(For the problem sheet, observing that the above equality implies that, if the $p$-variation is finite and nonzero, then the above inequality implies that $p = {}^1\!/\!_h$ is sufficient.) The ergodicity states that, for the measure $W^h$ induced the space of continuous paths beginning from zero $\mathrm{C}_0([0, \infty))$ by fractional Brownian motion, the only measurable subsets $A \subseteq \mathrm{C}_0([0, \infty))$ left invariant by the shift operators $\{\tau_k\}_{k \in \mathbb{N}_0}$ defined by

$$\tau_k(\sigma)(t) = \sigma(k + t) - \sigma(k) \ \text{ for every } \ t \in [0, \infty),$$

have measure zero or measure one. The ergodic theorem and the estimate (3.3) prove that, almost surely and in $L^1(\Omega)$, as $n \to \infty$,

$$2^{-n} \sum_{k=1}^{2^n} |B_k - B_{k-1}| \to \mathbb{E}[|X_1|].$$

Therefore, returning to (3.2), it follows that in probability

$$\lim_{n \to \infty} \sum_{k=1}^{2^n} \left| B_{\frac{k}{2^n}} - B_{\frac{k-1}{2^n}} \right|^p = \begin{cases} 0 & \text{if } \ ph > 1, \\ \mathbb{E}[|X_1|] & \text{if } \ p = \dfrac{1}{h}, \\ \infty & \text{if } \ ph < 1. \end{cases}$$

In particular, because semi-martingales have finite quadratic variation, it follows that fractional Brownian motion is a semi-martingale if and only if $h = \frac{1}{2}$. That is, if and only if the fractional Brownian motion is a Brownian motion. $\square$

3.1. **Quadratic variation.** In this section, we will define the quadratic variation $(\langle M\rangle_t)_{t\in[0,\infty)}$ of a continuous local martinagle $(M_t)_{t\in[0,\infty)}$ as a continuous increasing process that is defined by the limit, along a sequence of partitions locally finite partitions

$$\Delta_n = \{0 = t_{0_n} < t_{1_n} < \ldots < t_{k_n} < t_{(k+1)_n} < \ldots\},$$

of $[0,\infty)$ satisfying $|\Delta_n| \to 0$ as $n\to\infty$ by

$$\langle M\rangle_t = \lim_{n\to\infty}\sum_{k=1}^{\infty}(M_{t_{k_n}\wedge t} - M_{t_{(k-1)_n}\wedge t})^2.$$

The construction is based on approximating this limit by the corresponding discrete sums. Let $\Delta \subseteq [0,\infty)$ be a locally finite partition

$$\Delta = \{0 = t_0 < t_1 < t_2 < \ldots < t_k < t_{k+1} < \ldots\},$$

and define the process $(T_t^\Delta(M))_{t\in[0,\infty)}$ by

$$T_t^\Delta(M) = \sum_{k=1}^{\infty}(M_{t_k\wedge t} - M_{t_{k-1}\wedge t})^2.$$

That is, if $t_n \le t < t_{n+1}$ then

$$T_t^\Delta(M) = \sum_{k=1}^{n}(M_{t_k}^2 - M_{t_{k-1}}^2)^2 + (M_t - M_{t_n})^2.$$

The following proposition proves that, for every partition $\Delta$, the process $(M_t^2 - T_t^\Delta(M))_{t\in[0,\infty)}$ is a martingale.

**Proposition 3.3.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(M_t)_{t\in[0,\infty)}$ be a bounded $\mathcal{F}_t$-martingale. Then, for every locally finite partition $\Delta\subseteq[0,\infty)$,*

$$(M_t^2 - T_t^\Delta(M))_{t\in[0,\infty)} \text{ is a martingale.}$$

*Proof.* Let $\Delta \subseteq [0,\infty)$ be a locally finite partition

$$\Delta = \{0 = t_0 < t_1 < t_2 < \ldots < t_k < t_{k+1} < \ldots\}.$$

Since $(M_t)_{t\in[0,\infty)}$ is bounded, it follows by definition that, for every $t\in[0,\infty)$,

$$\mathbb{E}[\left|M_t^2 - T_t^\Delta(M)\right|] < \infty.$$

Let $s \le t \in [0,\infty)$. Fix $n\in\mathbb{N}$ such that $t_{n-1} \le s < t_n$. Since properties of the conditional expectation and the martingale property prove that, for every $k > n$,

$$\mathbb{E}[(M_{t_k} - M_{t_{k-1}})^2|\mathcal{F}_s] = \mathbb{E}[M_{t_k}^2 - M_{t_{k-1}}^2|\mathcal{F}_s],$$

and since

$$\begin{aligned}\mathbb{E}[(M_{t_n} - M_{t_{n-1}})^2|\mathcal{F}_s] &= \mathbb{E}[M_{t_n}^2|\mathcal{F}_s] - 2M_sM_{t_{n-1}} + M_{t_{n-1}}^2 \\ &= \mathbb{E}[M_{t_n}^2|\mathcal{F}_s] + (M_s - M_{t_{n-1}})^2 - M_s^2\end{aligned}$$

it follows that

$$\begin{aligned}\mathbb{E}[T_t^\Delta(M)|\mathcal{F}_s] = &\sum_{k=n+1}^{\infty}\mathbb{E}[(M_{t_k\wedge t} - M_{t_{k-1})^2\wedge t}|\mathcal{F}_s] + \mathbb{E}[(M_{t_n} - M_{t_{n-1}})^2|\mathcal{F}_s] + \sum_{k=1}^{n-1}(M_{t_k} - M_{t_{k-1}})^2 \\ &= \mathbb{E}[M_t^2|\mathcal{F}_s] - \mathbb{E}[M_{t_n}^2|\mathcal{F}_s] + \mathbb{E}[M_{t_n}^2|\mathcal{F}_s] - M_s^2 + T_s^\Delta(M).\end{aligned}$$

Therefore, we conclude that

$$\mathbb{E}[M_t^2 - T_s^\Delta(M)|\mathcal{F}_s] = M_s^2 - T_s^\Delta(M),$$

which completes the proof. □

Observe that the processes $\{T_t^{\Delta}(M)\}_{\Delta\subseteq[0,\infty)}$ are not necessarily increasing. However, whenever $k\leq j$ and $t_k,t_j\in\Delta$, it follows by definition that $T_{t_k}^{\Delta}(M)\leq T_{t_j}^{\Delta}(M)$. This observation and the boundedness of the martingale allow us to pass strongly to the limit $|\Delta|\rightarrow 0$ in the following proposition, and thereby construct a continuous, increasing process $(\langle M\rangle)_{t\in[0,\infty)}$ such that the process $(M_t^2-\langle M\rangle_t)_{t\in[0,\infty)}$ is a martingale.

**Theorem 3.4.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(M_t)_{t\in[0,\infty)}$ be a bounded continuous $\mathcal{F}_t$-martingale. Then there exists a unique nondecreasing process $(\langle M\rangle_t)_{t\in[0,\infty)}$ which vanishes at zero such that*

$$(M_t^2-\langle M\rangle_t)_{t\in[0,\infty)}\ \text{ is a martingale.}$$

*Proof.* To prove uniqueness, observe that if $(A_t)_{t\in[0,\infty)}$ and $(B_t)_{t\in[0,\infty)}$ are two increasing processes vanishing zero such that

$$(M_t^2-A_t)_{t\in[0,\infty)}\ \text{ and }\ (M_t^2-B_t)_{t\in[0,\infty)}\ \text{ are martingales,}$$

then the difference $(A_t-B_t)_{t\in[0,\infty)}$ is a martingale of finite variation. Hence, the difference is constant and identically equal to zero. To prove existence, choose a nested sequence of partitions $\Delta_1\subseteq\Delta_2\subseteq\Delta_3\subseteq\ldots$ satisfying $|\Delta_k|\rightarrow 0$ and $k\rightarrow\infty$. Prove that the sequence $\{T_t^{\Delta_k}(M)\}_{k\in\mathbb{N}}$ converges strongly in $L^2(\Omega)$, and that the limit is satisfies the desired properties. (See Math B8.2 notes for full details, which will shortly be added here.) ☐

This theorem has the following important extension to continuous local martingales.

**Theorem 3.5.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(M_t)_{t\in[0,\infty)}$ be a continuous local $\mathcal{F}_t$-martingale. Then there exists a unique continuous increasing process $(\langle M\rangle_t)_{t\in[0,\infty)}$ such that $(M_t^2-\langle M\rangle_t)_{t\in[0,\infty)}$ is a continuous local martingale. Furthermore, for every $t\in[0,\infty)$, for any sequence of partitions $\{\Delta_k\}_{k\in\mathbb{N}}$ of $[0,t]$ satisfying $|\Delta_k|\rightarrow 0$ as $k\rightarrow\infty$,*

$$\lim_{k\rightarrow\infty}\left(\sup_{s\in[0,T]}\left|T_s^{\Delta_k}(M)-\langle M\rangle_s\right|\right)=0\ \text{ in probability.}$$

We can now define the quadratic covariation or bracket process of two local martingales $(M_t)_{t\in[0,\infty)}$ and $(N_t)_{t\in[0,\infty)}$. Based upon the computations above, for $t\in[0,\infty)$ and for a sequence of partitions $\{\Delta_k\}_{k\in\mathbb{N}}$ of $[0,T]$ satisfying $|\Delta_k|\rightarrow 0$ as $k\rightarrow\infty$, we expect to define

$$\langle M,N\rangle_t=\lim_{k\rightarrow\infty}\sum_{k=1}^{n(\Delta_k)}(M_{t_k}-M_{t_{k-1}})(N_{t_k}-N_{t_{k-1}}).$$

where by convention we define $\langle M,M\rangle_t=\langle M\rangle_t$. Indeed, by polarization, the work is already done.

**Theorem 3.6.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(M_t)_{t\in[0,\infty)}$ and $(N_t)_{t\in[0,\infty)}$ be a continuous local $\mathcal{F}_t$-martingales. Then there exists a unique finite variation process $(\langle M\rangle_t)_{t\in[0,\infty)}$ such that*

$$(M_tN_t-\langle M,N\rangle_t)_{t\in[0,\infty)}\ \text{ is a continuous local martingale.}$$

*Furthermore, for every $t\in[0,\infty)$, for any sequence of partitions $\{\Delta_n\}_{n\in\mathbb{N}}$ of $[0,t]$ satisfying $|\Delta_n|\rightarrow 0$ as $n\rightarrow\infty$,*

$$\lim_{n\rightarrow\infty}\left(\sup_{s\in[0,T]}\left|T_s^{\Delta_n}(M,N)-\langle M\rangle_s\right|\right)=0\ \text{ in probability,}$$

*where*

$$T_s^{\Delta_n}(M,N)=\sum_{k=1}^{\infty}(M_{t_k\wedge s}-M_{t_{k-1}\wedge s})(N_{t_k\wedge s}-N_{t_{k-1}\wedge s}).$$

*Proof.* Since $(M_t+N_t)_{t\in[0,\infty)}$ and $(M_t-N_t)_{t\in[0,\infty)}$ are continuous local martingales, we define by polarization the bracket

$$\langle M,N\rangle_t=\frac{1}{4}\left(\langle M+N,M+N\rangle_t-\langle M-N,M-N\rangle_t\right).$$

The formula

$$M_tN_t=\frac{1}{4}\left((M+N)_t^2-(M-N)_t^2\right),$$

proves that $(\langle M,N\rangle_t)_{t\in[0,\infty)}$ satisfies the desired properties. □

In the final proposition of this section, we prove an important estimate for the integral of measurable processes with respect to the quadratic covariation. This is the *Kunita-Watanabe inequality.*

**Proposition 3.7.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(M_t)_{t\in[0,\infty)}$ and $(N_t)_{t\in[0,\infty)}$ be a continuous local $\mathcal{F}_t$-martingales. Let $(H_s)_{s\in[0,\infty)}$ and $(K_s)_{s\in[0,\infty)}$ be measurable stochastic processes. Then, for every $t\in[0,\infty)$,*

$$\int_0^t |H_s|\,|K_s|\,\mathrm{d}\,|\langle M,N\rangle|_s\leq\left(\int_0^t H_s^2\,\mathrm{d}\langle M\rangle_s\right)^{\frac{1}{2}}\left(\int_0^t K_s^2\,\mathrm{d}\langle N\rangle_s\right)^{\frac{1}{2}}\quad\textit{almost surely.}$$

*Furthermore, for every $p,q\in[1,\infty)$ satisfying $1/p+1/q=1$,*

$$\left\|\int_0^t |H_s|\,|K_s|\,\mathrm{d}\,|\langle M,N\rangle|_s\right\|_1\leq\left\|\left(\int_0^t H_s^2\,\mathrm{d}\langle M\rangle_s\right)^{\frac{1}{2}}\right\|_p\left\|\left(\int_0^t K_s^2\,\mathrm{d}\langle N\rangle_s\right)^{\frac{1}{2}}\right\|_q.$$

*Proof.* Let $t\in[0,\infty)$. By density, it suffices to prove that statement for functions of the form

$$K=\sum_{k=1}^n K_k\mathbf{1}_{[t_{k-1},t_k)}\ \text{ and }\ H=\sum_{k=1}^n H_k\mathbf{1}_{[t_{k-1},t_k)},$$

where $\{0=t_0<t_1<\ldots<t_{n-1}<t_n=t\}$ is a partition of $[0,T]$, and where $\{K_k\}_{k\in\{1,\ldots,n\}}$ and $\{H_k\}_{k\in\{1,\ldots,n\}}$ are bounded random variables. It then follows by definition that

$$\int_0^t |H_s|\,|K_s|\,\mathrm{d}\,|\langle M,N\rangle|_s=\sum_{k=1}^n |H_k|\,|K_k|\,|\langle M,N\rangle|_{t_{k-1}}^{t_k},$$

where $|\langle M,N\rangle|_{t_{k-1}}^{t_k}=|\langle M,N\rangle|_{t_k}-|\langle M,N\rangle|_{t_{k-1}}$. Since it follows for every $r\in\mathbb{Q}$ that, for each $k\in\{1,\ldots,n\}$, almost surely,

$$\langle M\rangle_{t_{k-1}}^{t_k}-2r\langle M,N\rangle_{t_{k-1}}^{t_k}+r^2\langle N\rangle_{t_{k-1}}^{t_k}=\langle M+rN,M+rN\rangle_{t_{k-1}}^{t_k}\geq 0,$$

and therefore by continuity this inequality holds for every $r\in\mathbb{R}$. Minimizing the lefthand side of $r\in\mathbb{R}$ proves that, for every $k\in\{1,\ldots,n\}$, almost surely,

$$\langle M,N\rangle_{t_{k-1}}^{t_k}\leq\left(\langle M\rangle_{t_{k-1}}^{t_k}\right)^{\frac{1}{2}}\left(\langle N\rangle_{t_{k-1}}^{t_k}\right)^{\frac{1}{2}}.$$

It then follows by Hölder's inequality that

$$\begin{aligned}\int_0^t |H_s|\,|K_s|\,\mathrm{d}\,|\langle M,N\rangle|_s&\leq\left(\sum_{k=1}^N H_k^2\langle M\rangle_{t_{k-1}}^{t_k}\right)^{\frac{1}{2}}\left(\sum_{k=1}^n K_k^2\langle N\rangle_{t_{k-1}}^{t_k}\right)^{\frac{1}{2}}\\&=\left(\int_0^t H_s^2\,\mathrm{d}\langle M\rangle_s\right)^{\frac{1}{2}}\left(\int_0^t K_s^2\,\mathrm{d}\langle N\rangle_s\right)^{\frac{1}{2}}.\end{aligned}$$

This proves the first statement. The final statement is then a consequence of Hölder's inequality. □

3.2. **Stochastic integration.** In this section, we will develop the stochastic integral with respect to a continuous martingale. We henceforth fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $(\mathcal{F}_t)_{t\in[0,\infty)}$ on $(\Omega, \mathcal{F})$. The construction is based on duality. Indeed, we will consider the class of integrators

$$\mathbb{H} = \left\{ (M_t)_{t\in[0,\infty)} \colon (M_t)_{t\in[0,\infty)} \text{ is a martingale with } \sup_{t\in[0,\infty)} \mathbb{E}[M_t^2] < \infty \right\}.$$

For every $(M_t)_{t\in[0,\infty)} \in \mathbb{H}$, it follows from Theorem 2.34 that there exists $M_\infty \in L^2(\Omega)$ such that, as $t \to \infty$,

$$M_t \to M_\infty \text{ almost surely and in } L^2(\Omega).$$

Furthermore, it follows from Doob's inequality that

$$\mathbb{E}\left[\sup_{t\in[0,\infty)} M_t^2\right] \leq 4\mathbb{E}[M_\infty^2].$$

We therefore define an inner product and norm on $\mathbb{H}$ for $M, N \in \mathbb{H}$ by

$$\langle\langle M, N\rangle\rangle = \mathbb{E}[M_\infty N_\infty] \text{ and } \|M\|_{\mathbb{H}}^2 = \mathbb{E}[M_\infty^2],$$

where, since $(M_t - \langle M\rangle_t)_{t\in[0,\infty)}$ is a martingale,

$$\mathbb{E}[M_\infty^2] = \mathbb{E}[\langle M\rangle_\infty].$$

The following proposition characterizes the space $\mathbb{H}$ as a complete Hilbert space isomorphic to an $L^2$-space with respect to the sigma algebra $\mathcal{F}_\infty = \cup_{t\in[0,\infty)}\mathcal{F}_t$.

**Proposition 3.8.** *The space $(\mathbb{H}, \langle\langle\cdot,\cdot\rangle\rangle)$ is a complete Hilbert space isomorphic to $L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$.*

*Proof.* Let $(M^n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathbb{H}$. It follows by definition of $\|\cdot\|_{\mathbb{H}}$ that $(M_\infty^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$. Therefore, there exists $M_\infty \in L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$ such that, as $n \to \infty$,

$$M_\infty^n \to M_\infty \text{ strongly in } L^2(\Omega).$$

We define the $L^2$-bounded martingale, for $t \in [0, \infty)$,

$$M_t = \mathbb{E}[M_\infty|\mathcal{F}_t],$$

and conclude by definition of $\|\cdot\|_{\mathbb{H}}$ that $(M_t)_{t\in[0,\infty)} \in \mathbb{H}$ and that, as $n \to \infty$,

$$(M_t^n)_{t\in[0,\infty)} \to (M_t)_{t\in[0,\infty)} \text{ strongly in } \mathbb{H}.$$

These arguments prove that the maps

$$\begin{aligned} (M_t)_{t\in[0,\infty)} \in \mathbb{H} \quad &\mapsto \quad M_\infty \in L^2(\Omega, \mathcal{F}_\infty, \mathbb{P}), \\ M \in L^2(\Omega, \mathcal{F}_\infty, \mathbb{P}) \quad &\mapsto \quad (\mathbb{E}[M|\mathcal{F}_t])_{t\in[0,\infty)} \in \mathbb{H}, \end{aligned}$$

define a Hilbert space isomorphism between $\mathbb{H}$ and $L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$. This completes the proof. ☐

We will integrate with respect to continuous $L^2$-bounded martingales, and therefore define the subspace $\mathcal{H} \subseteq \mathbb{H}$ of continuous $L^2$-bounded martingales

$$\mathcal{H} = \{(M_t)_{t\in[0,\infty)} \in \mathbb{H} \colon t \in [0,\infty) \mapsto M_t \text{ is almost surely continuous.}\}.$$

The following proposition proves that $(\mathcal{H}, \langle\langle\cdot,\cdot\rangle\rangle)$ remains is a complete Hilbert subspace of $\mathbb{H}$.

**Proposition 3.9.** *The space $(\mathcal{H}, \langle\langle\cdot,\cdot\rangle\rangle)$ is a complete Hilbert space.*

*Proof.* It is necessary to prove that $\mathcal{H}$ is complete. Suppose that $(M^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{H}$. Then, there exists $M_\infty \in L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$ such that, as $n \to \infty$,

$$M^n_\infty \to M_\infty \text{ strongly in } L^2(\Omega, \mathcal{F}_\infty, \mathbb{P}).$$

Define the $L^2$-bounded martingale, for $t \in [0,\infty)$,

$$M_t = \mathbb{E}[M|\mathcal{F}_t].$$

Since Jensen's inequality proves for every $n \in \mathbb{N}$ that $|M^n - M|$ is a submartingale, Doob's inequality proves for every $n \in \mathbb{N}$ that

$$\mathbb{E}\left[\sup_{t\in[0,\infty)} |M^n_t - M_t|^2\right] \leq 4\mathbb{E}[|M^n_\infty - M_\infty|^2].$$

Therefore, there exists a subsequence $\{n_k\}_{k\in\mathbb{N}}$ such that, as $k \to \infty$,

$$\sup_{t\in[0,\infty)} |M^{n_k}_t - M_t|^2 \to 0 \text{ almost surely.}$$

This proves that $t \in [0,\infty) \mapsto M_t$ is almost surely the uniform limit of continuous functions, which proves that $(M_t)_{t\in[0,\infty)}$ is almost surely continuous. This implies that $(M_t)_{t\in[0,\infty)} \in \mathcal{H}$, which completes the proof that $\mathcal{H}$ is complete. □

The following two propositions provide a useful characterizations of $\mathcal{H}$.

**Proposition 3.10.** *Let $(M_t)_{t\in[0,\infty)}$ be a continuous local martingale. Then, $(M_t)_{t\in[0,\infty)} \in \mathcal{H}$ if and only if the following two conditions are satisfied.*

*(i) We have that*

$$M_0 \in L^2(\Omega).$$

*(ii) We have that*

$$\mathbb{E}[\langle M\rangle_\infty] < \infty.$$

*Furthermore, in this case, $(M^2_t - \langle M\rangle_t)_{t\in[0,\infty)}$ is uniformly integrable.*

*Proof.* If $(M_t)_{t\in[0,\infty)} \in \mathcal{H}$ then by $L^2$-boundedness we have that $M_0 \in L^2(\Omega)$ and that $(M^2_t - \langle M\rangle_t)_{t\in[0,\infty)}$ is a martingale. And, by the $L^2$-boundedness and martingale converge theorem,

$$\mathbb{E}[\langle M\rangle_\infty] = \mathbb{E}[M^2_\infty] < \sup_{t\in[0,\infty)} \mathbb{E}[M^2_t] < \infty.$$

For the converse, because $(M_t)_{t\in[0,\infty)}$ is a continuous local martingale, there exist an increasing sequence of stopping times $\{\tau_n\}_{n\in\mathbb{N}}$ satisfying almost surely that $\tau_n \to \infty$ as $n \to \infty$. For every $n \in \mathbb{N}$, since by assumption $M^{\tau_n} = (M_{t\wedge\tau_n}\mathbf{1}_{\{\tau_n>0\}})_{t\in[0,\infty)}$ is a martingale, and since

$$\langle M^{\tau_n}\rangle = \langle M\rangle^{\tau_n}\mathbf{1}_{\{\tau_n>0\}},$$

it follows from $M_0 \in L^2(\Omega)$ and $\mathbb{E}[\langle M\rangle_\infty] < \infty$ that

$$((M^{\tau_n})^2_t - \langle M^{\tau_n}\rangle_t)_{t\in[0,\infty)} \text{ is a martingale.}$$

The martingale property and the fact that the quadratic variation is nondecreasing prove that, for every $n \in \mathbb{N}$ and $t \in [0,\infty)$,

$$\mathbb{E}[M^2_{\tau_n\wedge t}] = \mathbb{E}[M^2_0] + \mathbb{E}[\langle M^{\tau_n}\rangle_t] \leq \mathbb{E}[M^2_0] + \mathbb{E}[\langle M\rangle_t] \leq \mathbb{E}[M^2_0] + \mathbb{E}[\langle M\rangle_\infty].$$

We then pass to the limit using Fatou's lemma. Indeed, since the stopping times diverge to infinity almost surely, we have that, as $n \to \infty$, for every $t \in [0,\infty)$,

$$M_{\tau_n\wedge t} \to M_t \text{ almost surely,}$$

and therefore, for every $t \in [0, \infty)$,

$$\mathbb{E}[M_t^2] = \mathbb{E}[M_0^2] + \mathbb{E}[\lim_{n\to\infty} M_{\tau_n \wedge t}^2] \leq \mathbb{E}[M_0^2] + \liminf_{n\to\infty} \mathbb{E}[M_{\tau_n\wedge t}^2] < \mathbb{E}[M_0^2] + \mathbb{E}[\langle M\rangle_\infty],$$

which proves that $(M_t)_{t\in[0,\infty)}$ is a uniformly bounded $L^2(\Omega)$ and therefore uniformly integrable, from which it follows that $(M_t)_{t\in[0,\infty)}$ is a continuous martingale. This completes the proof that $(M_t)_{t\in[0,\infty)} \in \mathcal{H}$.

We will now prove that if $(M_t)_{t\in[0,\infty)} \in \mathcal{H}$ then $(M_t^2 - \langle M\rangle_t)_{t\in[0,\infty)}$ is uniformly integrable. The martingale convergence theorem proves that there exists $M_\infty \in L^2(\Omega)$ such that, as $t \to \infty$,

$$M_t \to M_\infty \text{ almost surely and strongly in } L^2(\Omega).$$

This implies by Hölder's inequality that

$$M_t^2 \to M_\infty^2 \text{ almost surely and strongly in } L^1(\Omega),$$

using the equality $\left|M_t^2 - M_\infty^2\right| = |(M_t - M_\infty)|\,|(M_t + M_\infty)|$. Therefore, since the bracket process is an increasing function, the dominated convergence theorem with dominating function $\langle M_\infty\rangle_\infty$ proves that, as $t \to \infty$,

$$\langle M\rangle_t \to \langle M\rangle_\infty \text{ almost surely and strongly in } L^1(\Omega).$$

In combination, this implies that, as $t \to \infty$,

$$M_t^2 - \langle M\rangle_t \to M_\infty^2 - \langle M_\infty\rangle_\infty \text{ almost surely and strongly in } L^1(\Omega).$$

The equivalent properties of the martingale convergence theorem then prove that $(M_t^2-\langle M\rangle_t)_{t\in[0,\infty)}$ is uniformly integrable. □

**Proposition 3.11.** *Let $(M_t)_{t\in[0,\infty)}$ be a continuous local martingale. Then, $(M_t)_{t\in[0,\infty)}$ converges on the set $\{\langle M\rangle_\infty < \infty\}$.*

*Proof.* For every $n \in \mathbb{N}$ define the stopping time

$$\sigma_n = \inf\{t \in [0,\infty) \colon \langle M\rangle_t \geq n\}.$$

For every $n \in \mathbb{N}$, consider the stopped local martingale $(M_{\sigma_n\wedge t})_{t\in[0,\infty)}$. It follows by definition of the $\{\sigma_n\}_{n\in\mathbb{N}}$ that, for every $n \in \mathbb{N}$,

$$\mathbb{E}[\langle M^{\sigma_n}\rangle_\infty] \leq n.$$

The above proposition proves that there exists $M_{\infty,n} \in L^2(\Omega)$ such that, as $t \to \infty$,

$$M_{\sigma_n\wedge t} \to M_{\infty,n} \text{ almost surely and strongly in } L^2(\Omega).$$

Observe that for every $n \leq m \in \mathbb{N}$, it follows that

$$M_{\infty,n} = M_{\infty,m} \text{ on the set } \{\langle M\rangle_\infty < n\}. \tag{3.4}$$

We therefore define

$$M_\infty(\omega) = \begin{cases} M_{\infty,n}(\omega) & \text{if } \omega \in \{\langle M\rangle_\infty < n\} \text{ for some } n \in \mathbb{N}, \\ 0 & \text{if } \omega \in \{\langle M\rangle_\infty = \infty\}. \end{cases}$$

It then follows from (3.4) that, for almost every $\omega \in \{\langle M\rangle_\infty < \infty\}$, as $t \to \infty$,

$$M_t(\omega) \to M_\infty(\omega),$$

which completes the proof. □

Let $(M_t)_{t\in[0,\infty)} \in \mathcal{H}$, and suppose that

$$\Delta = \{0 = t_0 < t_1 < t_2 < t_3 < \ldots\},$$

is a locally finite partition of $[0,\infty)$. Here locally finite means that for every compact set $K \subseteq [0,\infty)$ the intersection $K \cap \Delta$ is finite. For uniformly bounded random variables $\{K_i\}_{i\in\mathbb{N}}$ satisfying that $K_i$ is $\mathcal{F}_{t_i}$-measurable for every $i \in \mathbb{N}$, consider the simple process, for $t \in [0,\infty)$,

$$K_t = \sum_{i=0}^{\infty} K_i \mathbf{1}_{(t_i,t_{i+1}]}(t).$$

The stochastic integral of the process $(K_t)_{t\in[0,\infty)}$ with respect to $(M_t)_{t\in[0,\infty)}$ should mimic the discrete stochastic integral, and were therefore expect that

$$\int_0^\infty K_s \,\mathrm{d}M_s = \sum_{i=0}^{\infty} K_i(M_{t_{i+1}} - M_{t_i}).$$

Since the random variables $\{K_i\}_{i\in\mathbb{N}}$ are uniformly bounded and since $K_i$ is $\mathcal{F}_{t_i}$-measurable for every $i \in \mathbb{N}$, it follows from Proposition 2.10 that the process

$$\left(\int_0^t K_s \,\mathrm{d}M_s = \sum_{i=0}^{\infty} K_i(M_{t_{i+1}\wedge t} - M_{t_i\wedge t})\right)_{t\in[0,\infty)},$$

is a martingale. Furthermore, by the martingale property and the predictability and uniform boundedness of $(K_t)_{t\in[0,\infty)}$,

$$\mathbb{E}\left[\left(\int_0^\infty K_s \,\mathrm{d}M_s\right)^2\right] = \mathbb{E}\left[\sum_{i=1}^{\infty} K_i^2(M_{t_{i+1}} - M_{t_i})^2\right] \leq \left\|\sup_{t\in[0,\infty)} K_t^2\right\|_\infty \mathbb{E}\left[M_\infty^2\right].$$

We therefore conclude that the stochastic integral of a uniformly bounded, predictable simple process with respect to an element of $\mathcal{H}$ is again an element of $\mathcal{H}$. In the following proposition, we observe that the stochastic integral of a discrete process defines an isometry by which we come to extend the integral to more general processes.

**Proposition 3.12.** *Let $(M_t)_{t\in[0,\infty)}$ and $(N_t)_{t\in[0,\infty)}$ be elements of $\mathcal{H}$, let*

$$\Delta = \{0 = t_0 < t_1 < t_2 < t_3 < \ldots\},$$

*be a locally finite partition of $[0,\infty)$, and for uniformly bounded random variables $\{K_i\}_{i\in\mathbb{N}}$ satisfying that $K_i$ is $\mathcal{F}_{t_i}$-measurable for every $i \in \mathbb{N}$ let $(K_t)_{t\in[0,\infty)}$ be defined by*

$$K_t = \sum_{i=0}^{\infty} K_i \mathbf{1}_{(t_i,t_{i+1}]}.$$

*Then,*

$$\langle\langle \int_0^\cdot K_s \,\mathrm{d}M_s, N_\cdot\rangle\rangle = \mathbb{E}\left[\int_0^\infty K_s \,\mathrm{d}\langle M,N\rangle_s\right].$$

*Proof.* Since

$$\lim_{t\to\infty}\int_0^t K_s \,\mathrm{d}M_s = \int_0^\infty K_s \,\mathrm{d}M_s \text{ strongly in } L^2(\Omega),$$

it follows by definition and the martingale property that

$$\begin{aligned}\langle\langle \int_0^{\cdot} K_s M_s, N_\cdot\rangle\rangle &= \mathbb{E}\left[\left(\int_0^\infty K_s \,\mathrm{d}M_s\right) N_\infty\right] \\ &= \mathbb{E}\left[\left(\sum_{i=0}^\infty K_i(M_{t_{i+1}} - t_i)\right)\left(\sum_{j=1}^\infty (N_{t_{j+1}} - N_{t_j})\right)\right] \\ &= \mathbb{E}\left[\sum_{i=0}^\infty K_i(M_{t_{i+1}} - M_{t_i})(N_{t_{i+1}} - N_{t_i})\right] \\ &= \mathbb{E}\left[\sum_{i=0}^\infty K_i(M_{t_{i+1}}N_{t_{i+1}} - M_{t_i}N_{t_i})\right].\end{aligned}$$

We will now analyze the final term on the righthand side of this equality. The martingale property and the predictability of $(K_t)_{t\in[0,\infty)}$ prove that, for every $i \in \mathbb{N}_0$,

$$\mathbb{E}[K_i(M_{t_{i+1}}N_{t_{i+1}} - M_{t_i}N_{t_i})] = \mathbb{E}[K_i(\mathbb{E}[M_{t_{i+1}}N_{t_{i+1}}|\mathcal{F}_{t_i}] - M_{t_i}N_{t_i})].$$

Since the process $(M_tN_t - \langle M,N\rangle_t)_{t\in[0,\infty)}$ is a martingale,

$$\mathbb{E}[M_{t_{i+1}}N_{t_{i+1}}|\mathcal{F}_{t_i}] = \mathbb{E}[\langle M,N\rangle_{t_{i+1}}|\mathcal{F}_{t_i}] + M_{t_i}N_{t_i} - \langle M,N\rangle_{t_i}.$$

Therefore, by the martingale property and predictability of $(K_t)_{t\in[0,\infty)}$, for every $i \in \mathbb{N}_0$,

$$\begin{aligned}\mathbb{E}[K_i(M_{t_{i+1}}N_{t_{i+1}} - M_{t_i}N_{t_i})] &= \mathbb{E}[K_i(\mathbb{E}[\langle M,N\rangle_{t_{i+1}}|\mathcal{F}_{t_i}] - \langle M,N\rangle_{t_i})] \\ &= \mathbb{E}[K_i(\langle M,N\rangle_{t_{i+1}} - \langle M,N\rangle_{t_i})].\end{aligned}$$

Therefore, by definition of the deterministic integral with respect to the quadratic covariation,

$$\begin{aligned}\langle\langle \int_0^{\cdot} K_s M_s, N_\cdot\rangle\rangle &= \mathbb{E}\left[\sum_{i=0}^\infty K_i(\langle M,N\rangle_{t_{i+1}} - \langle M,N\rangle_{t_i})\right] \\ &= \mathbb{E}\left[\int_0^\infty K_s \,\mathrm{d}\langle M,N\rangle_s\right].\end{aligned}$$

This completes the proof. □

Observe in particular that, by choosing

$$\left(N_t = \int_0^t K_s \,\mathrm{d}M_s\right)_{t\in[0,\infty)} \in \mathcal{H},$$

the above computation proves that

$$\mathbb{E}\left[\left(\int_0^\infty K_s \,\mathrm{d}M_s\right)^2\right] = \mathbb{E}\left[\int_0^\infty K_s^2 \,\mathrm{d}\langle M\rangle_s\right].$$

This equality is the Itô isometry, and it suggests the class of processes for which we can define the stochastic integral with respect to an element $(M_t)_{t\in[0,\infty)} \in \mathcal{H}$.

**Definition 3.13.** For every $(M_t)_{t\in[0,\infty)} \in \mathcal{H}$, we define the space

$$L^2(M) = \left\{(K_t)_{t\in[0,\infty)} \colon (K_t)_{t\in[0,\infty)} \text{ is predictable and satisfies } \mathbb{E}\left[\int_0^\infty K_s^2 \,\mathrm{d}\langle M\rangle_s\right] < \infty\right\}.$$

We define, for every $(K_t)_{t\in[0,\infty)}$,

$$\|K\|_{L^2(M)} = \mathbb{E}\left[\int_0^\infty K_s^2 \,\mathrm{d}\langle M\rangle_s\right].$$

In the following theorem, for every $(M_t)_{t\in[0,\infty)} \in \mathcal{H}$ and $(K_t)_{t\in[0,\infty)} \in L^2(M)$, we construct the stochastic integral

$$\left(\int_0^t K_s \,\mathrm{d}M_s\right)_{t\in[0,\infty)},$$

as an element of $\mathcal{H}$. The proof is a consequence of the Riesz representation theorem, and it is for this reason that we proved that $\mathcal{H}$ is a complete Hilbert space.

**Proposition 3.14.** *For every $(M_t)_{t\in[0,\infty)} \in \mathcal{H}$ and $(K_t)_{t\in[0,\infty)} \in L^2(M)$, there exists a unique element $((K \bullet M)_t)_{t\in[0,\infty)} \in \mathcal{H}$ which satisfies for every $(N_t)_{t\in[0,\infty)}$ that*

$$\langle\langle K \bullet M, N\rangle\rangle = \mathbb{E}\left[\int_0^\infty K_s \,\mathrm{d}\langle M, N\rangle_s\right].$$

*We will henceforth write, for every $t \in [0,\infty)$,*

$$(K \bullet M)_t = \int_0^t K_s \,\mathrm{d}M_s.$$

*Proof.* The Kunita-Watanabe inequality proves that, for every $(N_t)_{t\in[0,\infty)} \in \mathcal{H}$,

$$\begin{aligned}
\left|\mathbb{E}\left[\int_0^\infty K_s \,\mathrm{d}\langle M, N\rangle_s\right]\right| &\leq \mathbb{E}\left[\int_0^\infty |K_s| \,\mathrm{d}\left|\langle M, N\rangle\right|_s\right] \\
&\leq \mathbb{E}\left[\left(\int_0^\infty K_s^2 \,\mathrm{d}\langle M\rangle_s\right)^{\frac{1}{2}}\right]\mathbb{E}\left[\left(\int_0^\infty \mathrm{d}\langle N\rangle_s\right)^{\frac{1}{2}}\right] \\
&= \|K\|_{L^2(M)}\, \mathbb{E}[\langle N\rangle_\infty]^{\frac{1}{2}}. \\
&= \|K\|_{L^2(M)}\, \|N\|_{\mathbb{H}}\,.
\end{aligned}$$

We therefore conclude that the map

$$(N_t)_{t\in[0,\infty)} \in \mathcal{H} \mapsto \mathbb{E}\left[\int_0^\infty K_s \,\mathrm{d}\langle M, N\rangle_s\right] \in \mathbb{R}$$

is a continuous linear functional on $\mathcal{H}$. The Riesz representation theorem therefore proves that there exists a unique elements $((K \bullet M)_t)_{t\in[0,\infty)} \in \mathcal{H}$ which satisfies that, for every $(N_t)_{t\in[0,\infty)} \in \mathcal{H}$,

$$\langle\langle K \bullet M, N\rangle\rangle = \mathbb{E}\left[\int_0^\infty K_s \,\mathrm{d}\langle M, N\rangle_s\right].$$

This completes the proof. ☐

Observe in particular that, for $(M_t)_{t\in[0,\infty)} \in \mathcal{H}$ and $K \in L^2(M)$, by choosing $(N_t = (K \bullet M)_t)_{t\in[0,\infty)}$ we have

$$\langle\langle K \bullet M, N\rangle\rangle = \mathbb{E}\left[\left(\int_0^\infty K_s \,\mathrm{d}M_s\right)^2\right] = \mathbb{E}\left[\int_0^\infty K_s^2 \,\mathrm{d}\langle M\rangle_s\right],$$

which is the general Itô isometry. This isometry implies that if $(K^n)_{n\in\mathbb{N}} \subseteq L^2(M)$ and $K \in L^2(M)$ satisfy that

$$\lim_{n\to\infty} \|K - K^n\|_{L^2(M)} = \lim_{n\to\infty} \mathbb{E}\left[\int_0^\infty |K_s - K_s^n|^2 \,\mathrm{d}\langle M\rangle_s\right] = 0,$$

then it follows that

$$\lim_{n\to\infty} \mathbb{E}\left[\left(\int_0^\infty K_s^n \,\mathrm{d}\langle M\rangle_s - \int_0^\infty K_s \,\mathrm{d}\langle M\rangle_s\right)^2\right] = 0.$$

This fact can be used to justify the more standard construction of the stochastic integral based upon Riemann sum approximations. Precisely, for a sequence of locally finite partitions

$$\{\Delta^n = \{0 = t_0^n < t_1^n < t_2^n < \ldots\}\}_{n\in\mathbb{N}}$$

which satisfy that $|\Delta^n| \to 0$ as $n \to \infty$, define suitable simple approximations

$$K_t^n = \sum_{i=0}^{\infty} K_i^n \mathbf{1}_{(t_i^n, t_{i+1}^n]}(t),$$

which satisfy

$$\lim_{n\to\infty} \|K^n - K\|_{L^2(M)} = 0.$$

Since for simple processes we know that, for every $t \in [0,\infty)$,

$$\int_0^t K_s^n \,\mathrm{d}M_s = \sum_{i=0}^{\infty} K_i^n (M_{t_{i+1}^n \wedge t} - M_{t_i^n \wedge t})$$

it follows that, as $n \to \infty$,

$$\left(\sum_{i=0}^{\infty} K_i^n (M_{t_{i+1}^n \wedge t} - M_{t_i^n \wedge t})\right)_{t\in[0,\infty)} \to \left(\int_0^t K_s \,\mathrm{d}M_s\right)_{t\in[0,\infty)} \text{ strongly in } \mathcal{H}.$$

We will now extend the definition of the stochastic integral to the class of continuous local martingales. If $(M_t)_{t\in[0,\infty)}$ is a continuous local martingale, then there exists a nondecreasing sequence of stopping times $\{tau_n\}_{n\in\mathbb{N}}$ such that $\tau_n \to \infty$ almost surely as $n \to \infty$, and such that for every $n \in \mathbb{N}$ we have

$$(M_t^{\tau_n})_{t\in[0,\infty)} = (M_{\tau_n\wedge t}\mathbf{1}_{\{\tau_n>0\}})_{t\in[0,\infty)} \in \mathcal{H}.$$

Since it holds that

$$\langle M^{\tau_n}\rangle_t = \langle M\mathbf{1}_{\{\tau_n>0\}}\rangle_{\tau_n\wedge t},$$

we define the space $L^2_{\text{loc}}(M)$ to be the space of predictable processes $(K_t)_{t\in[0,\infty)}$ for which there exists a nondecreasing sequence of stopping times $\{\sigma_n\}_{n\in\mathbb{N}}$ satisfying almost surely that $\sigma_n \to \infty$ as $n \to \infty$ such that

$$\mathbb{E}\left[\int_0^{\sigma_n} K_s \,\mathrm{d}\langle M\rangle_s\right] = \mathbb{E}\left[\int_0^{\infty} K_s \,\mathrm{d}\langle M^{\sigma_n}\rangle_s\right] < \infty.$$

The following proposition is then a consequence of localization.

**Proposition 3.15.** *Let $(M_t)_{t\in[0,\infty)}$ be a continuous local martingale and let $(K_t)_{t\in[0,\infty)} \in L^2_{loc}(M)$. Then, there exists a unique continuous local martingale $((K \bullet M)_t)_{t\in[0,\infty)}$ such that, for every continuous local martingale $(N_t)_{t\in[0,\infty)}$ there exists an nondecreasing sequence of stopping times $\{\tau_n\}_{n\in\mathbb{N}}$ satisfying almost surely $\tau_n \to \infty$ as $n \to \infty$ such that, for every $n \in \mathbb{N}$,*

$$(M_t^{\tau_n})_{t\in[0,\infty)}, (N_t^{\tau_n})_{t\in[0,\infty)} \in \mathcal{H} \text{ and } (K_t)_{t\in[0,\infty)} \in L^2(M^{\tau_n}),$$

*such that, for every $n \in \mathbb{N}$,*

$$\left(\int_0^t K_s \,\mathrm{d}\langle M^{\tau_n}\rangle_s\right)_{t\in[0,\infty)} = ((K \bullet M)_t^{\tau_n})_{t\in[0,\infty)} \in \mathcal{H},$$

*and such that, for every $n \in \mathbb{N}$,*

$$\langle\langle (K \bullet M)^{\tau_n}, N\rangle\rangle = \mathbb{E}\left[\int_0^{\infty} K_s \,\mathrm{d}\langle M^{\tau_n}, N^{\tau_n}\rangle_s\right].$$

In the final definition of this section, we define the stochastic integral with respect to a continuous semimartingale. The integral with

**Definition 3.16.** Let $(X_t = A_t + M_t)_{t\in[0,\infty)}$ be a continuous semimartingale, where $(A_t)_{t\in[0,\infty)}$ is a continuous process of bounded variation, and where $(M_t)_{t\in[0,\infty)}$ is a continuous local martingale. For every bounded process $(K_t)_{t\in[0,\infty)} \in L^2_{\text{loc}}(M)$, we define the continuous semimartingale

$$\left(\int_0^t K_s \,\mathrm{d}X_s\right)_{t\in[0,\infty)} = \left(\int_0^t K_s \,\mathrm{d}A_s + \int_0^t K_s \,\mathrm{d}M_s\right)_{t\in[0,\infty)}.$$

3.3. **The integration-by-parts formula.** In this section, we will prove an integration-by-parts formula that plays the role of the Leibniz rule in stochastic calculus. We will use this formula in the proof of Itô's formula below. The proof is a straightforward consequence of the definition of the quadratic variation and the existence of the stochastic integral.

**Proposition 3.17.** *Let $(M_t)_{t\in[0,\infty)}$ and $(N_t)_{t\in[0,\infty)}$ be continuous semimartingales. Then the process $(M_tN_t)_{t\in[0,\infty)}$ is a continuous semimartingale which satisfies, for every $t \in [0,\infty)$,*

$$M_tN_t = M_0N_0 + \int_0^t N_s \,\mathrm{d}M_s + \int_0^t M_s \,\mathrm{d}N_s + \langle M, N\rangle_t.$$

*Or, in differential notation,*

$$\mathrm{d}(M_tN_t) = N_t \,\mathrm{d}M_t + M_t \,\mathrm{d}N_t + \mathrm{d}\langle M, N\rangle_t.$$

*Proof.* Let $\Delta \subseteq [0,\infty)$ be a locally finite partition

$$\Delta = \{0 = t_0 < t_1 < t_2 < \ldots\}.$$

For every $i \in \mathbb{N}_0$, observe that

$$M_{t_{i+1}}N_{t_{i+1}} - M_{t_i}N_{t_i} = M_{t_i}(N_{t_{i+1}} - N_{t_i}) + N_{t_i}(M_{t_{i+1}} - M_{t_i}) + (M_{t_{i+1}} - M_{t_i})(N_{t_{i+1}} - N_{t_i}).$$

Therefore, for every $t \in [0,\infty)$,

$$\begin{aligned} M_tN_t &= M_0N_0 + \sum_{i=0}^{\infty}(M_{t_{i+1}\wedge t}N_{t_{i+1}\wedge t} - N_{t_i\wedge t}M_{t_i\wedge t}) \\ &= M_0N_0 + \sum_{i=0}^{\infty}\left(M_{t_i}(N_{t_{i+1}} - N_{t_i}) + N_{t_i}(M_{t_{i+1}} - M_{t_i}) + (M_{t_{i+1}} - M_{t_i})(N_{t_{i+1}} - N_{t_i})\right). \end{aligned}$$

Since the fact that $(M_t)_{t\in[0,\infty)}$ and $(N_t)_{t\in[0,\infty)}$ are continuous implies that they are respectively in $L^2_{\mathrm{loc}}(N)$ and $L^2_{\mathrm{loc}}(M)$, it follows after passing to the limit $|\Delta| \to 0$ that, for every $t \in [0,\infty)$,

$$M_tN_t = M_0N_0 + \int_0^t M_s \,\mathrm{d}N_s + \int_0^t N_s \,\mathrm{d}M_s + \langle M, N\rangle_t,$$

which completes the proof. □

3.4. **Proof of Itô's formula.** We are now prepared to prove Itô's formula, which is essentially the fundamental theorem of the calculus of semimartingales. There are many approaches to prove this statement, much like for the construction of the stochastic integral, including a direct argument based on Taylor expansions and discrete approximations along a sequence of partitions whose mesh approaches zero. The approach we take is based on the integration-by-parts formula and the Stone-Weierstrass theorem, which states that polynomials are dense in the space of continuous functions on a compact set of Euclidean space.

**Theorem 3.18.** *Let $(X_t)_{t\in[0,\infty)}$ be a continuous, $d$-dimensional semimartingale and let $f \in \mathrm{C}^2(\mathbb{R}^d)$. Then $(f(X_t))_{t\in[0,\infty)}$ is a semimartingale which satisifes*

$$f(X_t) = f(X_0) + \sum_{i=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(X_s)\,\mathrm{d}X_s^i + \frac{1}{2}\sum_{i,j=1}^{d}\int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)\,\mathrm{d}\langle X^i, X^j\rangle_s. \tag{3.5}$$

*Proof.* We may assume without loss of generality that the semimartingale $(X_t)_{t\in[0,\infty)}$ takes values in a compact subset of $\mathbb{R}^d$. For, if not, we can introduce the stopping times $\{T_R\}_{R\in(0,\infty)}$ defined by

$$T_R = \inf\{t \in [0,\infty) \colon X_t \notin B_R\},$$

prove the theorem for the stopped process $(X_t^{T_R})_{t\in[0,\infty)}$, and pass to the limit $R \to \infty$ using the continuity of $(X_t)_{t\in[0,\infty)}$.

We observe that for every $i \in \{1,\ldots,d\}$ equation (3.5) is satisfied by the function $f_i(x) = x_i$, since

$$X_t^i = X_0^i + \int_0^t \mathrm{d}X_s^i.$$

Similarly, formula (3.5) is satisfied by the constant function $f = 1$ and if functions $f, g \in \mathrm{C}^2(\mathbb{R}^d)$ and $c \in \mathbb{R}$ then by linearity if $f$ and $g$ satisfy (3.5) so too do $cf$ and $f+g$. It remains to prove that the product $fg$ satisfy (3.5).

Assume that $f, g \in \mathrm{C}^2(\mathbb{R}^d)$ satisfy (3.5). Then, the processes $(f(X_t))_{t\in[0,\infty)}$ and $(g(X_t))_{t\in[0,\infty)}$ are semimartingales, and the integration-by-parts formula proves that

$$\mathrm{d}(f(X_t)g(X_t)) = f(X_t)\,\mathrm{d}g(X_t) + g(X_t)\,\mathrm{d}f(X_t) + \mathrm{d}\langle f(X_t), g(X_t)\rangle_t, \tag{3.6}$$

where

$$\mathrm{d}g(X_t) = \sum_{i=1}^{d} \frac{\partial g}{\partial x_i}(X_t)\,\mathrm{d}X_t^i + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2 g}{\partial x_i \partial x_j}(X_t)\,\mathrm{d}\langle X^i, X^j\rangle_t, \tag{3.7}$$

where

$$\mathrm{d}f(X_t) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(X_t)\,\mathrm{d}X_t^i + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_t)\,\mathrm{d}\langle X^i, X^j\rangle_t, \tag{3.8}$$

and where

$$\mathrm{d}\langle f(X), g(X)\rangle_t = \sum_{i,j=1}^{d} \frac{\partial g}{\partial x_i}(X_t)\frac{\partial f}{\partial x_j}(X_t)\,\mathrm{d}\langle X^i, X^j\rangle_t \tag{3.9}$$

Returning to (3.6), it follows from the identities

$$\frac{\partial (fg)}{\partial x_i} = \frac{\partial f}{\partial x_i}g + f\frac{\partial g}{\partial x_i} \quad\text{and}\quad \frac{\partial (fg)}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j}g + \frac{\partial f}{\partial x_i}\frac{\partial g}{\partial x_j} + \frac{\partial f}{\partial x_j}\frac{\partial g}{\partial x_i} + f\frac{\partial^2 g}{\partial x_i \partial x_j},$$

from the symmetry of the quadratic covariation, and from (3.7), (3.8), and (3.9) that

$$\mathrm{d}(f(X_t)g(X_t)) = \sum_{i=1}^{d} \frac{\partial(fg)}{\partial x_i}(X_t)\,\mathrm{d}X_s^i + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2(fg)}{\partial x_i \partial x_j}(X_t)\,\mathrm{d}\langle X^i, X^j\rangle_t,$$

which completes the proof that the product $(fg)$ satisfies (3.5). We therefore conclude that every polynomial function satisfies (3.5).

Let $f \in \mathrm{C}^2(\mathbb{R}^d)$ be arbitrary and fix $R \in (0,\infty)$ such that $X_t \in B_R$ for every $t \in [0,\infty)$. Then, by the Stone-Weierstrass theorem, there exists a sequence of polynomials $\{p_n\}_{n\in\mathbb{N}}$ such that, in the $\mathrm{C}^2$-topology,

$$\lim_{n\to\infty}\left(\sup_{x\in B_R}\left[|f(x)-p_n(x)| + |\nabla f(x) - \nabla p_n(x)| + \left|\nabla^2 f(x) - \nabla^2 p_n(x)\right|\right]\right) = 0.$$

Since each of the polynomials $\{p_n\}_{n\in\mathbb{N}}$ satisfies (3.5), after passing to the limit $n\to\infty$ the uniform $\mathrm{C}^2$-convergence of the $\{p_n\}_{n\in\mathbb{N}}$ to $f$ and the Itô isometry prove that $f$ satisfies (3.5). This completes the proof. $\square$

In the case of a standard $d$-dimensional Brownian motion $(B_t)_{t\in[0,\infty)}$, it follows by choosing $f(x) = |x|^2$ in Itô's formula that

$$\begin{aligned} |B_t|^2 &= 2\sum_{i=1}^{d}\int_0^t B_s^i\,\mathrm{d}B_s^i + \sum_{i,j=1}^{d}\int_0^t \mathrm{d}\langle B^i, B^j\rangle_t \\ &= 2\sum_{i=1}^{d}\int_0^t B_s^i\,\mathrm{d}B_s^i + \sum_{i=1}^{d}\int_0^t \mathrm{d}t \qquad\qquad . \\ &= 2\sum_{i=1}^{d}\int_0^t B_s^i\,\mathrm{d}B_s^i + dt \end{aligned}$$

Hence, the square of the Euclidean norm of a $d$-dimensional Brownian motion satisfies

$$\sum_{i=1}^{d}\int_0^t B_s^i\,\mathrm{d}B_s^i = \frac{1}{2}\left(|B_t|^2 - dt\right).$$

In particular, in one-dimension,

$$\int_0^t B_s\,\mathrm{d}B_s = \frac{1}{2}(B_t^2 - t),$$

which reproves the statement that $(B_t^2 - t)_{t\in[0,\infty)}$ is a martingale.

## 4. Applications of stochastic calculus

In this section, we will explain some important applications of stochastic calculus. The first is the stochastic analogue of the exponential function, which we will subsequently use to prove Levy's characterization of Brownian motion and the martingale representation theorem. We will use Levy's theorem to prove the Dambis-Dubins-Schwarz theorem. We will use Itô's formula to provide a proof of the Burkholder-Davis-Gundy inequality.

4.1. **The stochastic exponential.** The exponential function $Y_t = \exp(t)$ is the unique solution to the ordinary differential equation

$$\begin{cases} \mathrm{d}Y_t = Y_t \,\mathrm{d}t & \text{in } (0,\infty), \\ Y_0 = 1, \end{cases}$$

and in this way plays an essential role in the theory of ordinary and partial differential equations. For instance, we can reduce the nonlinear partial differential equation

$$\begin{cases} \partial_t u = \Delta u + |\nabla u|^2 & \text{in } \mathbb{R}^d \times (0,\infty), \\ u = u_0 & \text{on } \mathbb{R}^d \times \{0\}, \end{cases}$$

by making the exponential transformation

$$v(x,t) = \exp(u(x,t)),$$

whereby the chain rule implies that $v$ solves the linear heat equation

$$\begin{cases} \partial_t v = \Delta v & \text{in } \mathbb{R}^d \times (0,\infty), \\ v = \exp(u_0) & \text{on } \mathbb{R}^d \times \{0\}, \end{cases}$$

which can be solved explicitly in terms of the heat kernel. We will see below, for the simpler example of a geometric Brownian motion, that the stochastic exponential defined below can be similarly used to simplify terms appearing stochastic differential equations.

**Proposition 4.1.** *Let $(X_t)_{t\in[0,\infty)}$ be a continuous semimartingale that vanishes at zero. Then the continuous semimartingale $(\mathcal{E}(X)_t)_{t\in[0,\infty)}$ defined by*

$$\mathcal{E}(X)_t = \exp\left(X_t - \frac{\langle X\rangle_t}{2}\right),$$

*is the unique continuous semimartingale solution of the stochastic differential equation*

$$\begin{cases} \mathrm{d}Z_t = Z_t \,\mathrm{d}X_t & \text{in } (0,\infty), \\ Z_0 = 1. \end{cases} \tag{4.1}$$

*The process $(\mathcal{E}(X)_t)_{t\in[0,\infty)}$ is the stochastic exponential of $(X_t)_{t\in[0,\infty)}$.*

*Proof.* We first observe formally that if $(Z_t)_{t\in[0,\infty)}$ solves (4.1) then Itô's formula would prove that $(\log(Z_t))_{t\in[0,\infty)}$ solves

$$\begin{aligned} \mathrm{d}(\log(Z_t)) &= \frac{1}{Z_t}\,\mathrm{d}Z_t - \frac{1}{2Z_t^2}\,\mathrm{d}\langle Z\rangle_t \\ &= \mathrm{d}X_t - \frac{1}{2}\,\mathrm{d}\langle X\rangle_t. \end{aligned}$$

This would imply that

$$\log(Z_t) = X_t - \frac{\langle X\rangle_t}{2},$$

from which it follows that

$$Z_t = \exp\left(X_t - \frac{\langle X\rangle_t}{2}\right). \tag{4.2}$$

This argument, however, is not justified because we do not know a priori that a semimartingale solution to (4.1) exists, and because the logarithm is not a $\mathrm{C}^2$-function on the whole of $\mathbb{R}$. The latter of these points is not so serious, however, because we could instead analyze approximating processes that are stopped before hitting a neighborhood of the origin. Nevertheless, equation (4.2) provides a candidate for the solution, and we define the stochastic exponential $(\mathcal{E}(X)_t)_{t\in[0,\infty)}$ by

$$\mathcal{E}(X)_t = \exp\left(X_t - \frac{\langle X\rangle_t}{2}\right).$$

It then follows by definition that $\mathcal{E}(X)_0 = 0$ and by using Itô's formula that

$$\begin{aligned}\mathrm{d}\mathcal{E}(X)_t &= \mathcal{E}(X)_t\,\mathrm{d}X_t - \frac{1}{2}\mathcal{E}(X)_t\,\mathrm{d}\langle X\rangle_x + \frac{1}{2}\mathcal{E}(X)_t\,\mathrm{d}\langle X\rangle_t \\ &= \mathcal{E}(X)_t\,\mathrm{d}X_t,\end{aligned}$$

which completes the proof that $(\mathcal{E}(X)_t)_{t\in[0,\infty)}$ solves (4.1).

In order to prove uniqueness, observe using Itô's formula that the inverse process $(\mathcal{E}(X)_t^{-1})_{t\in[0,\infty)}$ defined by

$$\mathcal{E}(X)_t^{-1} = \exp\left(-X_t + \frac{\langle X\rangle_t}{2}\right),$$

solves the equation

$$\mathrm{d}\mathcal{E}(X)_t^{-1} = -\mathcal{E}(X)_t^{-1}\,\mathrm{d}X_t + \mathcal{E}(X)_t^{-1}\,\mathrm{d}\langle X\rangle_t.$$

If $(Z_t)_{t\in[0,\infty)}$ is a continuous semimartingale solution of (4.1), then the integration-by-parts formula proves that

$$\begin{aligned}\mathrm{d}(Z_t\mathcal{E}(X)_t^{-1}) &= Z_t\,\mathrm{d}\mathcal{E}(X)_t^{-1} + \mathcal{E}(X)_t^{-1}\,\mathrm{d}Z_t + \mathrm{d}\langle\mathcal{E}(X)^{-1}, Z_t\rangle_t \\ &= -Z_t\mathcal{E}(X)_t^{-1}\,\mathrm{d}X_t + Z_t\mathcal{E}(X)_t^{-1}\,\mathrm{d}X_t + Z_t\mathcal{E}(X)_t^{-1}\,\mathrm{d}\langle X\rangle_t - Z_t\mathcal{E}(X)_t^{-1}\,\mathrm{d}\langle X\rangle_t \\ &= 0.\end{aligned}$$

Therefore, for every $t\in[0,\infty)$,

$$Z_t\mathcal{E}(X)_t^{-1} = Z_0\mathcal{E}(X)_t^{-1} = 1 \;\text{ and }\; Z_t = \mathcal{E}(X)_t,$$

which completes the proof. ☐

A *geometric Brownian motion* is a stochastic process $(S_t)_{t\in[0,\infty)}$ that satisfies the stochastic differential equation, for a standard one dimensional Brownian motion $(B_t)_{t\in[0,\infty)}$, for some $\mu,\sigma\in\mathbb{R}$,

$$\begin{cases}\mathrm{d}S_t = \mu S_t\,\mathrm{d}t + \sigma S_t B_t & \text{in } (0,\infty), \\ S_0 = 1.\end{cases} \tag{4.3}$$

Versions of such processes play an important role in mathematical finance, where $(S_t)_{t\in[0,\infty)}$ represents the value of a stock or bond, where the parameter $\mu$ is the interest rate, and where the parameter $\sigma$ quantifies the volatility of a financial market. In the following proposition, we use the stochastic exponential to solve equation (4.3). Observe, in particular, that the solution is positive.

**Proposition 4.2.** *Let $(B_t)_{t\in[0,\infty)}$ be a standard one-dimensional Brownian motion and let $\mu,\sigma\in\mathbb{R}$. Then the stochastic process $(S_t)_{t\in[0,\infty)}$ defined by*

$$S_t = \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t\right), \tag{4.4}$$

*is the unique continuous semimartingale solution of* (4.3).

*Proof.* The fact that $(S_t)_{t\in[0,\infty)}$ is a solution of (4.3) is an immediate consequence of Itô's formula. Now, suppose that $(S_t)_{t\in[0,\infty)}$ is an arbitrary continuous semimartingale solution of (4.3), and let $(\mathcal{E}(\sigma B)_t)_{t\in[0,\infty)}$ be the unique solution of the equation

$$\begin{cases}\mathrm{d}\mathcal{E}(\sigma B)_t = \sigma\mathcal{E}(\sigma B)_t\,\mathrm{d}B_t & \text{in } (0,\infty), \\ \mathcal{E}(\sigma B)_0 = 1,\end{cases}$$

defined in Proposition 4.1 by

$$\mathcal{E}(\sigma B)_t = \exp\left(\sigma B_t - \frac{\sigma^2 t}{2}\right),$$

since $\langle B \rangle_t = t$. It was shown in Proposition 4.1 that the inverse process $(\mathcal{E}(\sigma B)_t^{-1})_{t\in[0,\infty)}$ satisfies

$$\mathcal{E}(\sigma B)_t^{-1} = -\sigma \mathcal{E}(\sigma B)_t^{-1}\,\mathrm{d}B_t + \sigma^2 \mathcal{E}(\sigma B)_t\,\mathrm{d}t.$$

Therefore, by the integration-by-parts formula,

$$\begin{aligned}
\mathrm{d}(S_t\mathcal{E}(\sigma B)_t^{-1}) &= S_t\,\mathrm{d}\mathcal{E}(\sigma B)_t^{-1} + \mathcal{E}(\sigma B)_t^{-1}\,\mathrm{d}S_t + \mathrm{d}\langle S, \mathcal{E}(\sigma B)^{-1}\rangle_t \\
&= -\sigma S_t\mathcal{E}(\sigma B)_t^{-1}\,\mathrm{d}B_t + \sigma S_t\mathcal{E}(\sigma B)_t^{-1}\sigma\,\mathrm{d}B_t + \mu S_t\mathcal{E}(\sigma B)_t^{-1}\,\mathrm{d}t \\
&\quad + \sigma^2 S_t\mathcal{E}(\sigma B)_t^{-1}\,\mathrm{d}t - \sigma^2 S_t\mathcal{E}(\sigma B)_t^{-1}\,\mathrm{d}t \\
&= \mu S_t\mathcal{E}(\sigma B)_t^{-1}\,\mathrm{d}t.
\end{aligned}$$

We therefore conclude that

$$S_t\mathcal{E}(\sigma B)_t^{-1} = \exp(\mu t),$$

and therefore that

$$S_t = \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t\right).$$

This completes the proof. $\square$

4.2. **Levy's characterization of Brownian motion.** In this section, we will use stochastic calculus to prove that Brownian motion is uniquely characterized in the class of continuous local martingales by its quadratic variation. Observe in particular that for general martingales $(M_t)_{t\in[0,\infty)}$ the quadratic variation process $(\langle M\rangle_t)_{t\in[0,\infty)}$ is a random process, in the sense that for every $t\in(0,\infty)$ the random variable $\langle M\rangle_t$ is not deterministic, which is not the case for a standard Brownian motion.

**Theorem 4.3.** *Let $(M_t)_{t\in[0,\infty)}$ be a continuous local martingale vanishing at zero. Then $(M_t)_{t\in[0,\infty)}$ is a standard Brownian motion if and only if, for every $t\in[0,\infty)$,*

$$\langle M\rangle_t = t.$$

*Proof.* If $(M_t)_{t\in[0,\infty)} = (B_t)_{t\in[0,\infty)}$ is a Brownian motion, then for every $T\in(0,\infty)$ and $N\in\mathbb{N}$,

$$\sum_{k=0}^{\infty}\left(B_{\frac{k+1}{N}\wedge T} - B_{\frac{k}{N}\wedge T}\right)^2,$$

is a sum of independent random variables with mean $1/N$. Hence, after rewriting this sum,

$$\sum_{k=0}^{\infty}\left(B_{\frac{k+1}{N}\wedge T} - B_{\frac{k}{N}\wedge T}\right)^2 = \frac{1}{N}\left(\sum_{k=0}^{\infty}N\left(B_{\frac{k+1}{N}\wedge T} - B_{\frac{k}{N}\wedge T}\right)^2\right),$$

the law of large numbers proves that, as $N\to\infty$,

$$\frac{1}{N}\left(\sum_{k=0}^{\infty}N\left(B_{\frac{k+1}{N}\wedge T} - B_{\frac{k}{N}\wedge T}\right)^2\right)\to T \text{ almost surely,}$$

from which we conclude that $\langle B\rangle_T = T$, for every $T\in(0,\infty)$.

Now suppose that $(M_t)_{t\in[0,\infty)}$ is a continuous local martingale satisfying $\langle M\rangle_t = t$ for every $t\in[0,\infty)$. It then follows that, for every $T\in(0,\infty)$, the stopped martingale $(M^T_t)_{t\in[0,\infty)} = (M_{t\wedge T})_{t\in[0,\infty)}$ satisfies $\langle M^T\rangle_\infty = \langle M\rangle_T = T$. We therefore conclude that, for every $T\in(0,\infty)$, the martingale $(M^T_t)_{t\in[0,\infty)}$ is $L^2$-bounded and hence uniformly integrable. This implies that $(M^T_t)_{t\in[0,\infty)}$ is a martingale, for every $T\in(0,\infty)$, and hence that $(M_t)_{t\in[0,\infty)}$ is a martingale. We will not prove that $(M_t)_{t\in[0,\infty)}$ is a Brownian motion.

Since it follows by definition that $M_0 = 0$ and that $(M_t)_{t\in[0,\infty)}$ is almost surely continuous, it remains only to prove that $M_t - M_s$ is normally distributed with mean zero and variance $t-s$, for every $s\le t$, and that $M_t - M_s$ is independent of $\mathcal{F}_s$, for every $s\le t$. Let $s\le t$. For the proof, we will essentially study the characteristic function, for each $\alpha\in\mathbb{R}$,

$$\mathbb{E}[\exp(i\alpha(M_t - M_s))],$$

which is related to the complex exponential

$$\left(\exp\left(i\alpha M_r + \frac{\alpha^2}{2}\langle M\rangle_r\right)\right)_{t\in[0,\infty)} = \left(\exp\left(i\alpha M_r + \frac{\alpha^2 r}{2}\right)\right)_{r\in[0,\infty)}.$$

Observe that this is the stochastic exponential of the complex martingale $(i\alpha M_r)_{r\in[0,\infty)}$.

We say that a complex-valued process is a complex martingale if both its real and complex parts are martingales. In the case of the complex exponential, for every $r\in[0,\infty)$,

$$\exp\left(i\alpha M_r + \frac{\alpha^2 r}{2}\right) = \exp\left(\frac{\alpha^2 r}{2}\right)\left(\cos(\alpha M_r) + i\sin(\alpha M_r)\right).$$

Since the argument for both the real and complex parts are the same, we will prove that

$$\left(\exp\left(\frac{\alpha^2 r}{2}\right)\sin(\alpha M_r)\right)_{r\in[0,\infty)}$$

is a martingale. Define the process of bounded variation

$$(A_r)_{r\in[0,\infty)} = (\sin(\alpha M_r))_{r\in[0,\infty)},$$

and define the process $(Z_r)_{r\in[0,\infty)}$ by

$$Z_r = \sin(\alpha M_r).$$

. The integration-by-parts formula proves that

$$\mathrm{d}\left(\exp\left(\frac{\alpha^2 r}{2}\right)\sin(\alpha M_r)\right) = A_r\,\mathrm{d}Z_r + Z_r\,\mathrm{d}A_r + \mathrm{d}\langle A, Z\rangle_r,$$

where Itô's formula and $\langle M\rangle_r = r$ prove that

$$\mathrm{d}Z_r = \alpha\cos(\alpha M_r)\,\mathrm{d}M_r - \frac{\alpha^2}{2}\sin(\alpha M_r)\,\mathrm{d}r,$$

where ordinary calculus proves that

$$\mathrm{d}A_r = \frac{\alpha^2}{2}\exp\left(\frac{\alpha^2 r}{2}\right),$$

and where $\mathrm{d}\langle A, Z\rangle_r = 0$ since $(A_r)_{r\in[0,\infty)}$ has bounded variation. Hence,

$$\mathrm{d}\left(\exp\left(\frac{\alpha^2 r}{2}\right)\sin(\alpha M_r)\right) = \alpha\exp\left(\frac{\alpha^2 r}{2}\right)\cos(\alpha M_r)\,\mathrm{d}M_r,$$

which implies using the boundedness of the cos function that the lefthand side is a martingale. And, therefore, we conclude that the complex stochastic exponential is an exponential martingale.

The martingale property proves that

$$\mathbb{E}\left[\exp\left(i\alpha M_t + \frac{\alpha^2 t}{2}\right)|\mathcal{F}_s\right] = \exp\left(i\alpha M_s + \frac{\alpha^2 s}{2}\right).$$

Since the deterministic functions are $\mathcal{F}_s$-measurable, and since $M_s$ is $\mathcal{F}_s$-measurable, this implies that

$$\mathbb{E}\left[\exp(i\alpha(M_t - M_s))|\mathcal{F}_s\right] = \exp\left(-\frac{\alpha^2}{2}(t-s)\right).$$

This implies that $(M_t - M_s)$ is normally distributed and independent of $\mathcal{F}_s$, which completes the proof. (The full proof will be added after the next problem session.) $\square$

4.3. **The Dambis-Dubins-Schwarz theorem.** In the case of a standard Brownian motion, the oscillations of the process effectively act as a clock in the sense that the quadratic variation $\langle B\rangle_t = t$ for every $t \in [0,\infty)$. In this section, we will prove that every continuous martingale $(M_t)_{t\in[0,\infty)}$ that "oscillates enough" in the sense that $\langle M\rangle_\infty = \infty$ almost surely is merely a time-changed Brownian motion. That is, on the event $\{\langle M\rangle_t > t\}$ the martingale is behaving like a Brownian motion that has been "sped-up", and on the event $\{\langle M\rangle_t < t\}$ the martingale is behaving like a Brownian motion that has been "slowed-down." We make the notion of a time change precise in the next definition.

**Definition 4.4.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$. A *time change* is a nondecreasing sequence of right-continuous $\mathcal{F}_t$-stopping times $(\sigma_t)_{t\in[0,\infty)}$ which satisfy $\sigma_s \to \infty$ as $s \to \infty$ almost surely.

Observe in particular that, given a uniformly integrable martingale $(M_t)_{t\in[0,\infty)}$ and a time change $(\sigma_t)_{t\in[0,\infty)}$, we can define the time-changed process $(M_{\sigma_t})_{t\in[0,\infty)}$. The optional stopping theorem proves that, for every $s \le t \in [0,\infty)$,

$$\mathbb{E}[M_{\sigma_t}|\mathcal{F}_{\sigma_s}] = M_{\sigma_s}.$$

That is, the time-changed process $(M_{\sigma_t})_{t\in[0,\infty)}$ is a martingale with respect to $(\mathcal{F}_{\sigma_t})_{t\in[0,\infty)}$, which is the filtration generated by the time change $(\sigma_t)_{t\in[0,\infty)}$.

We will now show that every continuous local martingale $(M_t)_{t\in[0,\infty)}$ that satisfies $\langle M\rangle_\infty = \infty$ almost surely is a time-changed Brownian motion. The time change is determined by the quadratic variation process $(\langle M\rangle_t)_{t\in[0,\infty)}$. We will "slow down" the martingale if its quadratic variation is larger than $t$, and we will "speed up" the martingale if the quadratic variation is smaller than $t$. The theorem follows immediately after the next two lemmas.

**Lemma 4.5.** *Let $(M_t)_{t\in[0,\infty)}$ be a continuous local martingale vanishing at zero. Then $M_t = 0$ for every $t \in [0,\infty)$ if and only if $\langle M\rangle_t = 0$ for every $t \in [0,\infty)$.*

*Proof.* If $M_t = 0$ for every $t \in [0,\infty)$ then it follows by definition that $\langle M\rangle_t = 0$ for every $t \in [0,\infty)$. To prove that, suppose that $\langle M\rangle_t = 0$ for every $t \in [0,\infty)$. Then, since $(M_t^2 - \langle M\rangle_t)_{t\in[0,\infty)}$ is a local martingale, it follows that there exist a sequence of stopping times $\{\tau_n\}_{n\in\mathbb{N}}$ that satisfy $\tau_n \to \infty$ almost surely as $n \to \infty$ such that

$$\mathbb{E}[M^2_{t\wedge\tau_n}] = 0.$$

Therefore, by continuity, it follows almost surely that $M_t^2 = 0$ for every $t \in [0,\infty)$. This completes the proof. □

**Lemma 4.6.** *Let $(M_t)_{t\in[0,\infty)}$ be a continuous local martingale vanishing at zero. Then, the intervals of constancy of $(M_t)_{t\in[0,\infty)}$ and $(\langle M\rangle_t)_{t\in[0,\infty)}$ coincide almost surely. That is, almost surely, $M_t(\omega) = M_b(\omega)$ for every $t \in [a,b]$ if and only if $\langle M\rangle_t(\omega) = \langle M\rangle_b(\omega)$ for every $t \in [a,b]$.*

*Proof.* It follows by definition that $(\langle M\rangle_t)_{t\in[0,\infty)}$ is constant on an interval $[a,b]$ if $(M_t)_{t\in[0,\infty)}$ is constant on $[a,b]$. Let $a \in \mathbb{Q}\cap[0,\infty)$ and let $T_a$ denote the stopping time

$$T_a = \inf\{s \in (a,\infty) \colon \langle M\rangle_s > \langle M\rangle_a\}.$$

A repetition of the above argument proves almost surely that $M_t = M_a$ for every $t \in [0,T_a]$. It then follows by density of the rationals and continuity that if $\langle M\rangle_t$ is constant on an interval $[a,b]$ then almost surely $M_t$ is constant on $[a,b]$. This completes the proof. □

**Theorem 4.7.** *Let $(M_t)_{t\in[0,\infty)}$ be a continuous local martingale vanishing at zero that satisfies $\langle M\rangle_\infty = \infty$ almost surely. Define the stopping times $\{\sigma_t\}_{t\in[0,\infty)}$ by*

$$\sigma_t = \inf\{s \in [0,\infty) \colon \langle M\rangle_s > t\} = \sup\{s \in [0,\infty) \colon \langle M\rangle_s = t\}.$$

*Then,* $(M_{\sigma_t})_{t\in[0,\infty)}$ *is a standard Brownian motion* $(B_t)_{t\in[0,\infty)}$ *and, for every* $t \in [0,\infty)$*,*

$$M_t = B_{\langle M\rangle_t}.$$

*Proof.* First, we observe that the stopping times $(\sigma_t)_{t\in[0,\infty)}$ are indeed stopping times. This relies on the right continuity of the filtration $(\mathcal{F}_t)_{t\in[0,\infty)}$, which means that, for every $t \in [0,\infty)$,

$$\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s.$$

Let $M_\sigma$ denote the time-changed process $(M_{\sigma_t})_{t\in[0,\infty)}$. Then, $M_\sigma$ is a $\mathcal{F}_{\sigma_t}$-martingale. Indeed, for every $t \in [0,\infty)$, the stopped process $(M_{\sigma_t\wedge s})_{s\in[0,\infty)}$ is $L^2$-bounded and hence uniformly integrable since $\langle M^{\sigma_t}\rangle_\infty = \langle M\rangle_{\sigma_t} = t$ almost surely. Therefore, the optional stopping theorem applies and, for every $s \leq t \in [0,\infty)$,

$$\mathbb{E}[M_{\sigma_t}|\mathcal{F}_{\sigma_s}] = M_{\sigma_s}.$$

Since $t \in [0,\infty)$ was arbitrary, this completes the proof that $(M_{\sigma_t})_{t\in[0,\infty)}$ is a $\mathcal{F}_{\sigma_t}$-martingale.

Let $M^\sigma$ denote the $\mathcal{F}_{\sigma_t}$-martingale $(M_{\sigma_t})_{t\in[0,\infty)}$. We aim to apply Levy's characterization of Brownian motion, and for this it is necessary to prove continuity and to prove that the quadratic variation satisfies $\langle M^\sigma\rangle_t = t$. For the quadratic variation, since the assumption $\langle M\rangle_\infty = \infty$ implies that $\sigma_t < \infty$ almost surely, we have that by definition of the time change, for every $t \in [0,\infty)$,

$$\langle M^\sigma\rangle_t = \langle M^{\sigma_t}\rangle_\infty = \langle M\rangle_{\sigma_t} = t.$$

To prove continuity, for every $t \in [0,\infty)$, the definition of the time change proves that

$$\lim_{s\to t^-} M_{\sigma_s} = M_{\underline{t}},$$

where by continuity of $(\langle M\rangle_t)_{t\in[0,\infty)}$,

$$\underline{t} = \sup\{s \in [0,\infty)\colon \langle M\rangle_s < t\} = \min\{s \in [0,\infty)\colon \langle M\rangle_s = t\}.$$

Similarly,

$$\lim_{s\to t^-} M_{\sigma_s} = M_{\overline{t}},$$

where

$$\overline{t} = \inf\{s \in [0,\infty)\colon \langle M\rangle_s > t\} = \max\{s \in [0,\infty)\colon \langle M\rangle_s = t\} = \sigma_t.$$

Since the intervals of constancy of the quadratic variation and the continuous local martingale almost surely coincide, and since the quadratic variation is constant on the interval $[\underline{t},\overline{t}] = [\underline{t},\sigma_t]$, we conclude that

$$M_{\underline{t}} = M_{\overline{t}} = M_{\sigma_t},$$

and therefore that

$$\lim_{s\to t} M_{\sigma_s} = M_{\sigma_t}.$$

Levy's characterization of Brownian motion implies that $(M_{\sigma_t})_{t\in[0,\infty)}$ is a standard $\mathcal{F}_{\sigma_t}$-Brownian motion $(B_t)_{t\in[0,\infty)}$ where by definition $B_t = M_{\sigma_t}$.

It remains to prove that, for every $t \in [0,\infty)$,

$$M_t = B_{\langle M\rangle_t}.$$

By definition, for every $t \in [0,\infty)$,

$$B_{\langle M\rangle_t} = B_{\sigma_{\langle M\rangle_t}},$$

where by definition, for every $t \in [0,\infty)$,

$$\sigma_{\langle M\rangle_t} = \inf\{s \in [0,\infty)\colon \langle M\rangle_s > \langle M\rangle_t\} = \sup\{s \in [0,\infty)\colon \langle M\rangle_s = \langle M\rangle_t\}.$$

Therefore, for every $t \in [0,\infty)$, it follows that $\sigma_{\langle M\rangle_t} \geq t$ and that the quadratic variation is constant on the interval $[t,\sigma_{\langle M\rangle_t}]$. Since the intervals of constancy of the quadratic variation and the martingale almost surely coincide, we conclude by definition of $(B_t)_{t\in[0,\infty)}$ that

$$M_t = M_{\sigma_{\langle M\rangle_t}} = B_{\langle M\rangle_t},$$

which completes the proof. □

4.4. **The Burkholder-Davis-Gundy inequality.** We have already observed that if $(M_t)_{t\in[0,\infty)}$ is an $L^2$-bounded continuous martingale, then the quadratic variation process $(\langle M\rangle_t)_{t\in[0,\infty)}$ exists and we have, for every $t\in[0,\infty)$,

$$\mathbb{E}[M_t^2]=\mathbb{E}[\langle M\rangle_t].$$

If we define the running maximum

$$M_t^*=\sup_{s\in[0,t]}|M_s|,$$

it is then immediate that

$$\mathbb{E}[\langle M\rangle_t]=\mathbb{E}[M_t^2]\leq\mathbb{E}[(M_t^*)^2].$$

Alternately, Doob's inequality proves for every $t\in[0,\infty)$ that, since $(M_t^2)_{t\in[0,\infty)}$ is a submartingale,

$$\mathbb{E}[(M_t^*)^2]\leq 4\mathbb{E}[M_t^2]=4\mathbb{E}[\langle M\rangle_t].$$

In combination, therefore, we have for every $t\in[0,\infty)$ that

$$\frac{1}{4}\mathbb{E}[(M_t^*)^2]\leq\mathbb{E}[\langle M\rangle_t]\leq\mathbb{E}[(M_t^*)^2],$$

which implies that the $L^2$-norm of the running maximum and the $L^1$-norm of the quadratic variation are equivalent norms. The Burkholder-Davis-Gundy inequalities prove that this is the case for every $L^p$-norm, for every $p\in(0,\infty)$, in the sense that the $L^{2p}$-norm of the running maximum and the $L^p$-norm of the quadratic variation are equivalent.

**Theorem 4.8.** *For every $p\in(0,\infty)$ there exist constants $c_p,C_p\in(0,\infty)$ such that, for every continuous martingale $(M_t)_{t\in[0,\infty)}$ vanishing at zero, for every $t\in[0,\infty)$,*

$$c_p\mathbb{E}[(M_t^*)^{2p}]\leq\mathbb{E}[\langle M\rangle_t^p]\leq C_p\mathbb{E}[(M_t^*)^{2p}]. \tag{4.5}$$

*Proof.* Let $(M_t)_{t\in[0,\infty)}$ be a continuous martingale vanishing at zero. For every $n\in\mathbb{N}$ let $T_n$ be the stopping time defined by

$$T_n=\inf\{t\in[0,\infty)\colon\ |M_t|\geq n\ \text{ or }\ \langle M\rangle_t\geq n\}.$$

Observe that if we prove that the theorem holds for the stopped martingales $(M_t^{T_n})_{t\in[0,\infty)}$, for constants that are independent of $n\in\mathbb{N}$, then we can apply the monotone convergence theorem to deduce that the theorem holds for $(M_t)_{t\in[0,\infty)}$. The monotone convergence theorem applies because, for every $t\in[0,\infty)$, the functions $(M^{T_n})_t^*$ and $\langle M^{T_n}\rangle_t$ are nondecreasing functions of $n\in\mathbb{N}$. We can therefore assume without loss of generality that $(M_t)_{t\in[0,\infty)}$ is a bounded martingale with bounded quadratic variation. This assumption guarantees that, in the applications of Itô's formula to follow, all of the integrals appearing are true martingales, as opposed to only local martingales.

Henceforth, let $(M_t)_{t\in[0,\infty)}$ be a bounded continuous martingale vanishing at zero with bounded quadratic variation. We have already proven (4.5) in the case $p=1$. Suppose now that $p\in(1,\infty)$. Then, since the function $f(x)=|x|^{2p}$ is twice continuously differentiable, it follows from Itô's formula and the fact that $(M_t)_{t\in[0\infty)}$ vanishes at zero that

$$|M_t|^{2p}=2p\int_0^t M_s^{2p-1}\mathrm{sgn}(M_s)\,\mathrm{d}M_s+p(2p-1)\int_0^t|M_s|^{2(p-1)}\ \mathrm{d}\langle M\rangle_s.$$

Hence, after taking the expectation, using the fact that the first term on the righthand side is a martingale,

$$\mathbb{E}\left[|M_t|^{2p}\right]=p(2p-1)\mathbb{E}\left[\int_0^t|M_s|^{2(p-1)}\ \mathrm{d}\langle M\rangle_s\right]\leq p(2p-1)\mathbb{E}\left[(M_t^*)^{2(p-1)}\langle M\rangle_t\right].$$

Hence, by Hölder's inequality with exponents $p$ and $p/p-1$,

$$\mathbb{E}\left[|M_t|^{2p}\right]\leq p(2p-1)\mathbb{E}\left[(M_t^*)^{2p}\right]^{\frac{p-1}{p}}\mathbb{E}\left[\langle M\rangle_t^p\right]^{\frac{1}{p}}.$$

Since Jensen's inequality implies that $|M_t|^{2p}$ is a submartingale, it follows by Doob's inequality that

$$\mathbb{E}[(M_t^*)^{2p}] \leq \left(\frac{2p}{2p-1}\right)^{2p} \mathbb{E}[|M_t|^{2p}],$$

and therefore we have that

$$\mathbb{E}[(M_t^*)^{2p}] \leq \left(\frac{2p}{2p-1}\right)^{2p} p(2p-1)\mathbb{E}[(M_t^*)^{2p}]^{\frac{p-1}{p}}\mathbb{E}[\langle M\rangle_t^p]^{\frac{1}{p}}.$$

After dividing through by $\mathbb{E}[(M_t^*)^{2p}]^{\frac{p-1}{p}}$,

$$\mathbb{E}\left[(M_t^*)^{2p}\right]^{\frac{1}{p}} \leq \left(\frac{2p}{2p-1}\right)^{2p} p(2p-1)\mathbb{E}\left[\langle M\rangle_t^p\right]^{\frac{1}{p}},$$

from which we conclude that

$$\mathbb{E}\left[|M_t|^{2p}\right] \leq \left(\frac{2p}{2p-1}\right)^{3p} (p(2p-1))^p\, \mathbb{E}\left[\langle M\rangle_t^p\right].$$

This completes the proof of the leftmost inequality in (4.5) with

$$c_p = \left(\frac{2p}{2p-1}\right)^{-3p} (p(2p-1))^{-p}.$$

For the rightmost inequality in (4.5), define the process

$$N_t = \int_0^t \langle M\rangle_s^{\frac{p-1}{2}}\,\mathrm{d}M_s.$$

The reason for considering this process is that its quadratic variation is defined by

$$\langle N\rangle_t = \int_0^t \langle M\rangle_s^{p-1}\,\mathrm{d}\langle M\rangle_s = \frac{1}{p}\langle M\rangle_t^p.$$

Therefore, owing to the fact that $(N_t - \langle N\rangle_t)_{t\in[0,\infty)}$ is a martingale vanishing at zero,

$$\frac{1}{p}\mathbb{E}[\langle M\rangle_t^p] = \mathbb{E}[\langle N\rangle_t] = \mathbb{E}[N_t^2].$$

However, we observe using the integration-by-parts formula that, since $(\langle M\rangle_t^{\frac{p-1}{2}})_{t\in[0,\infty)}$ is a process of bounded variation owing to the facts that $p\in(1,\infty)$ and that $(\langle M\rangle_t)_{t\in[0,\infty)}$ is a bounded process of bounded variation,

$$\begin{aligned} M_t\langle M\rangle_t^{\frac{p-1}{2}} &= \int_0^t \langle M\rangle_s^{\frac{p-1}{2}}\,\mathrm{d}M_s + \int_0^t M_s\,\mathrm{d}\left(\langle M\rangle_s^{\frac{p-1}{2}}\right) \\ &= N_t + \int_0^t M_s\,\mathrm{d}\left(\langle M\rangle_s^{\frac{p-1}{2}}\right). \end{aligned}$$

We therefore conclude that

$$|N_t| \leq |M_t|\,\langle M\rangle_t^{\frac{p-1}{2}} + \int_0^t |M_s|\,\mathrm{d}\left(\langle M\rangle_s^{\frac{p-1}{2}}\right) \leq 2M_t^*\langle M\rangle_t^{\frac{p-1}{2}}.$$

Hence, again applying Hölder's inequality with exponents $p$ and $p/p-1$,

$$\frac{1}{p}\mathbb{E}[\langle M\rangle_t^p] = \mathbb{E}[N_t^2] \leq 4\mathbb{E}[(M_t^*)^2\langle M\rangle_t^{p-1}] \leq 4\mathbb{E}[(M_t^*)^{2p}]^{\frac{1}{p}}\mathbb{E}[\langle M\rangle_t^p]^{\frac{p-1}{p}}.$$

We conclude that

$$\mathbb{E}[\langle M\rangle_t^p]^{\frac{1}{p}} \leq 4p\mathbb{E}[(M_t^*)^{2p}]^{\frac{1}{p}},$$

and therefore that

$$\mathbb{E}[\langle M\rangle_t^p] \leq (4p)^p \mathbb{E}[(M_t^*)^{2p}].$$

This completes the proof of the rightmost inequality of (4.5) with

$$C_p = (4p)^p,$$

and therefore the proof of (4.5) in the case $p \in (1, \infty)$. We have by now completed the proof for every $p \in [1, \infty)$. It remains to consider the case $p \in (0, 1)$.

Let $p \in (0, 1)$. For every $\alpha \in (0, 1)$ let $(N_t^\alpha)_{t\in[0,\infty)}$ be defined by

$$N_t^\alpha = \int_0^t (\langle M\rangle_s + \alpha)^{\frac{p-1}{2}} \,\mathrm{d}M_s.$$

The quadratic variation process $(\langle N^\alpha\rangle_t)_{t\in[0,\infty)}$ satisfies

$$\langle N^\alpha\rangle_t = \int_0^t (\langle M\rangle_s + \alpha)^{p-1} \,\mathrm{d}\langle M\rangle_s = \frac{1}{p} (\langle M\rangle_t + \alpha)^p, \tag{4.6}$$

where the final equality uses the fact that $\mathrm{d}(\langle M\rangle_t + \alpha) = \mathrm{d}\langle M\rangle_t$. Furthermore, since it follows by definition that

$$\mathrm{d}N_t^\alpha = (\langle M\rangle_t + \alpha)^{\frac{p-1}{2}} \,\mathrm{d}M_t,$$

it follows that

$$\mathrm{d}M_t = (\langle M\rangle_t + \alpha)^{\frac{1-p}{2}} \,\mathrm{d}N_t^\alpha.$$

The integration-by-parts formula proves that, since the process

$$\left((\langle M\rangle_t + \alpha)^{\frac{1-p}{2}}\right)_{t\in[0,\infty)}$$

is a process of bounded variation,

$$\begin{aligned} N_t^\alpha (\langle M\rangle_t + \alpha)^{\frac{1-p}{2}} &= \int_0^t (\langle M\rangle_s + \alpha)^{\frac{1-p}{2}} \,\mathrm{d}N_s^\alpha + \int_0^t N_s^\alpha \,\mathrm{d}\left((\langle M\rangle_s + \alpha)^{\frac{1-p}{2}}\right) \\ &= M_t + \int_0^t N_s^\alpha \,\mathrm{d}\left((\langle M\rangle_s + \alpha)^{\frac{1-p}{2}}\right). \end{aligned}$$

Therefore,

$$|M_t| \leq |N_t^\alpha| (\langle M\rangle_t + \alpha)^{\frac{1-p}{2}} + \int_0^t |N_t^\alpha| \,\mathrm{d}\left((\langle M\rangle_s + \alpha)^{\frac{1-p}{2}}\right) \leq 2(N^\alpha)_t^* (\langle M\rangle_t + \alpha)^{\frac{1-p}{2}}. \tag{4.7}$$

Let $t \in [0, \infty)$. It follows from (4.7) that, for every $s \in [0, t]$,

$$|M_s| \leq 2(N^\alpha)_s^* (\langle M\rangle_s + \alpha)^{\frac{1-p}{2}} \leq 2(N^\alpha)_t^* (\langle M\rangle_t + \alpha)^{\frac{1-p}{2}},$$

from which it follows that

$$|M_t^*| \leq 2(N^\alpha)_t^* (\langle M\rangle_t + \alpha)^{\frac{1-p}{2}}.$$

It then follows from Hölder's inequality with exponents $1/p$ and $1/1-p$ that

$$\mathbb{E}[(M_t^*)^{2p}] \leq 2^{2p}\mathbb{E}\left[(N_t^{\alpha,*})^{2p}(\langle M\rangle_t + \alpha)^{p(1-p)}\right] \leq 2^{2p}\mathbb{E}\left[(N_t^{\alpha,*})^2\right]^p \mathbb{E}\left[(\langle M\rangle_t + \alpha)^p\right]^{1-p}. \tag{4.8}$$

It then follows from Doob's inequality and (4.6) that

$$\mathbb{E}\left[(N_t^{\alpha,*})^2\right] \leq 4\mathbb{E}[(N_t^\alpha)^2] = 4\mathbb{E}[\langle N^\alpha\rangle_t] = \frac{4}{p}\mathbb{E}\left[(\langle M\rangle_t + \alpha)^p\right].$$

Returning to (4.8), we have that

$$\mathbb{E}[(M_t^*)^{2p}] \leq \frac{2^{2(p+1)}}{p}\mathbb{E}\left[(\langle M\rangle_t + \alpha)^p\right].$$

The monotone convergence theorem then proves that, after passing to the limit $\alpha\rightarrow 0$,

$$\mathbb{E}[(M_t^*)^{2p}]\leq \frac{2^{2(p+1)}}{p}\mathbb{E}\left[\langle M\rangle_t^p\right],$$

which completes the leftmost inequality of (4.5) in the case $p\in(0,1)$ with

$$c_p=p2^{-2(p+1)}.$$

It remains to prove the rightmost inequality of (4.5) in the case $p\in(0,1)$. Let $\alpha\in(0,\infty)$. Since we have that

$$\langle M\rangle_t^p=\left(\langle M\rangle_t\,(M_t^*+\alpha)^{2(p-1)}\right)^p(M_t^*+\alpha)^{2p(1-p)},$$

it follows from Hölder's inequality with exponents $1/p$ and $1/1-p$ that

$$\mathbb{E}[\langle M\rangle_t^p]\leq \mathbb{E}\left[\langle M\rangle_t\,(M_t^*+\alpha)^{2(p-1)}\right]^p\mathbb{E}\left[(M_t^*+\alpha)^{2p}\right]^{1-p}.$$

Let $(N_t^\alpha)_{t\in[0,\infty)}$ be the process

$$N_t^\alpha=\int_0^t(M_s^*+\alpha)^{p-1}\,\mathrm{d}M_s,$$

for which

$$\langle N^\alpha\rangle_t=\int_0^t(M_s^*+\alpha)^{2(p-1)}\,\mathrm{d}\langle M\rangle_s\geq (M_t^*+\alpha)^{2(p-1)}\langle M\rangle_t. \tag{4.9}$$

The integration-by-parts formula proves that, since as a nondecreasing process $(M_t^*)_{t\in[0,\infty)}$ is a bounded process of bounded variation,

$$\begin{aligned}M_t(M_t^*+\alpha)^{p-1}&=\int_0^t(M_s^*+\alpha)^{p-1}\,\mathrm{d}M_s+\int_0^t M_s\,\mathrm{d}\left((M_s^*+\alpha)^{p-1}\right)\\&=N_t^\alpha+(p-1)\int_0^t M_s(M_s^*+\alpha)^{p-2}\,\mathrm{d}M_s^*.\end{aligned}$$

We therefore conclude that, since $p\in(0,1)$,

$$\begin{aligned}|N_t^\alpha|&\leq |M_t|\,(M_t^*+\alpha)^{p-1}+(1-p)\int_0^t|M_s|\,(M_s^*+\alpha)^{p-2}\,\mathrm{d}M_s^*\\&\leq M_t^*(M_t^*+\alpha)^{p-1}+(1-p)\int_0^t(M_s^*+\alpha)^{p-1}\,\mathrm{d}M_s^*.\\&\leq M_t^*(M_t^*+\alpha)^{p-1}+\left(\frac{1}{p}-1\right)(M_t^*+\alpha)^p.\end{aligned} \tag{4.10}$$

Since it follows from (4.9) and Hölder's inequality with exponents $1/p$ and $1/1-p$ that

$$\begin{aligned}\mathbb{E}[\langle M\rangle_t^p]&\leq \mathbb{E}[\langle N^\alpha\rangle_t^p(M_t^*+\alpha)^{2p(1-p)}]\\&\leq \mathbb{E}[\langle N^\alpha\rangle_t]^p\mathbb{E}[(M_t^*+\alpha)^{2p}]^{1-p}\\&=\mathbb{E}[(N_t^\alpha)^2]^p\mathbb{E}[(M_t^*+\alpha)^{2p}]^{1-p},\end{aligned}$$

and since it follows from (4.10) that

$$\mathbb{E}\left[(N_t^\alpha)^2\right]\leq \mathbb{E}\left[(M_t^*)^2(M_t^*+\alpha)^{2(p-1)}+\left(\frac{1}{p}-1\right)^2(M_t^*+\alpha)^{2p}\right],$$

we have that

$$\mathbb{E}[\langle M\rangle_t^p]\leq \mathbb{E}\left[(M_t^*)^2(M_t^*+\alpha)^{2(p-1)}+\left(\frac{1}{p}-1\right)^2(M_t^*+\alpha)^{2p}\right]^p\mathbb{E}[(M_t^*+\alpha)^p]^{1-p}.$$

By the monotone convergence theorem, after passing to the limit $\alpha \to 0$, we conclude that

$$\mathbb{E}[\langle M\rangle_t^p] \leq p^{-2p}\mathbb{E}\left[(M_t^*)^{2p}\right],$$

which completes the proof of the rightmost inequality of (4.5) in the case $p \in (0,1)$ with

$$C_p = p^{-2p}.$$

This completes the proof. □

In the following three corollaries, we will show that the Burkholder-Davis-Gundy inequality applies equally to stopped continuous martingales, continuous local martingales, and to integrals of bounded predictable processes.

**Corollary 4.9.** *For every $p \in (0,\infty)$ there exist constants $c_p, C_p \in (0,\infty)$ such that, for every continuous martingale $(M_t)_{t\in[0,\infty)}$ vanishing at zero, for every almost surely finite stopping time $T$,*

$$c_p\mathbb{E}[(M_T^*)^{2p}] \leq \mathbb{E}[\langle M\rangle_T^p] \leq C_p\mathbb{E}[(M_T^*)^{2p}].$$

*Proof.* Let $p \in (0,\infty)$, let $(M_t)_{t\in[0,\infty)}$ be a continuous martingale vanishing at zero, and let $T$ be an almost surely finite stopping time. Then, since the stopped process $(M_t^T)_{t\in[0,\infty)}$ is a martingale, it follows from Theorem 4.8 that there exist universal constants $c_p, C_p \in (0,\infty)$ such that, for every $t \in [0,\infty)$,

$$c_p\mathbb{E}[(M_{t\wedge T}^*)^{2p}] \leq \mathbb{E}[\langle M\rangle_{t\wedge T}^p] \leq C_p\mathbb{E}[(M_{t\wedge T}^*)^{2p}].$$

Since the stopping time is almost surely finite, the monotone convergence theorem proves that, after passing to the limit $t \to \infty$,

$$c_p\mathbb{E}[(M_T^*)^{2p}] \leq \mathbb{E}[\langle M\rangle_T^p] \leq C_p\mathbb{E}[(M_T^*)^{2p}],$$

which completes the proof. □

**Corollary 4.10.** *For every $p \in (0,\infty)$ there exist constants $c_p, C_p \in (0,\infty)$ such that, for every continuous local martingale $(M_t)_{t\in[0,\infty)}$ vanishing at zero, for every almost surely finite stopping time $T$,*

$$c_p\mathbb{E}[(M_T^*)^{2p}] \leq \mathbb{E}[\langle M\rangle_T^p] \leq C_p\mathbb{E}[(M_T^*)^{2p}].$$

*Proof.* Let $p \in (0,\infty)$, let $(M_t)_{t\in[0,\infty)}$ be a continuous local martingale vanishing at zero, and let $T$ be an almost surely finite stopping time. Then there exist stopping times $\{\tau_n\}_{n\to\infty}$ such that $\tau_n \to \infty$ almost surely as $n \to \infty$, and such that

$$(M_t^{\tau_n})_{t\in[0,\infty)} = (M_{t\wedge\tau_n}\mathbf{1}_{\{\tau_n>0\}})_{t\in[0,\infty)} \text{ is a martingale.}$$

Therefore, by the previous Corollary, there exist universal constants $c_p, C_p \in (0,\infty)$ such that

$$c_p\mathbb{E}[(M_{\tau_n\wedge T}^*)^{2p}\mathbf{1}_{\{\tau_n>0\}}] \leq \mathbb{E}[\langle M\rangle_{\tau_n\wedge T}^p\mathbf{1}_{\{\tau_n>0\}}] \leq C_p\mathbb{E}[(M_{\tau_n\wedge T}^*)^{2p}\mathbf{1}_{\{\tau_n>0\}}].$$

Since the stopping time is almost surely finite, the monotone convergence theorem proves that, after passing to the limit $n \to \infty$,

$$c_p\mathbb{E}[(M_T^*)^{2p}] \leq \mathbb{E}[\langle M\rangle_T^p] \leq C_p\mathbb{E}[(M_T^*)^{2p}],$$

which completes the proof. □

**Corollary 4.11.** *For every $p \in (0,\infty)$ there exist constants $c_p, C_p \in (0,\infty)$ such that, for every continuous local martingale $(M_t)_{t\in[0,\infty)}$ vanishing at zero, for every bounded predictable process $(H_t)_{t\in[0,\infty)}$, for every almost surely finite stopping time $T$,*

$$c_p\mathbb{E}\left[\left(\sup_{t\in[0,T]}\left(\int_0^t H_s\,\mathrm{d}M_s\right)^2\right)^p\right] \leq \mathbb{E}\left[\left(\int_0^T H_s^2\,\mathrm{d}\langle M\rangle_s\right)^p\right] \leq C_p\mathbb{E}\left[\left(\sup_{t\in[0,T]}\left(\int_0^t H_s\,\mathrm{d}M_s\right)^2\right)^p\right].$$

*Proof.* The statement is an immediate consequence of the previous corollary, since the boundedness of the process $(H_t)_{t\in[0,\infty)}$ guarantees that

$$\left(\int_0^t H_s\,\mathrm{d}M_s\right)_{t\in[0,\infty)}\quad\text{is a local martingale}$$

with quadratic variation process

$$\left(\langle\int_0^\cdot H_s\,\mathrm{d}M_s\rangle_t\right)_{t\in[0,\infty)}=\left(\int_0^t H_s^2\,\mathrm{d}\langle M\rangle_s\right)_{t\in[0,\infty)}.$$

$\square$

4.5. **The martingale representation theorem.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probabilty space, and let $X\colon \Omega \to \mathbb{R}$ be a random variable. Then, it follows that for every random variable $Y\colon \Omega \to \mathbb{R}$ that is measurable with respect to $\sigma(X)$, there exists a measurable function $f_Y\colon \mathbb{R} \to \mathbb{R}$ such that

$$Y = f_Y(X). \tag{4.11}$$

This fact emphasizes the view that $\sigma(X)$ carries all of the information of $X$, in the sense that every random variable measurable with respect to $\sigma(X)$ is actually a function of $X$.

The martingale representation theorem is the analogue of this fact for Brownian motion. Suppose that $(B_t)_{t\in[0,\infty)}$ is a standard Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to a filtration $(\mathcal{F}_t)_{t\in[0,\infty)}$. We define the sigma algebra

$$\mathcal{F}_\infty = \sigma(\cup_{s\in[0,\infty)}\mathcal{F}_s),$$

which carries the information of the entire Brownian path. We therefore expect that every random variable $M \in L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$, that is every $L^2$-random variable that is measurable with respect to $\mathcal{F}_\infty$, is a "function of Brownian motion." We will see that $M = I(B_t)$ is a functional of the Brownian paths, and that the functional is a stochastic integral. That is, for every $M \in L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$ there exists a unique process $(H_s)_{s\in[0,\infty)} \in L^2(B)$ such that

$$M = \mathbb{E}[M] + \int_0^\infty M_s \,\mathrm{d}B_s. \tag{4.12}$$

In particular, if we define the $L^2$-bounded martingale $(M_t)_{t\in[0,\infty)}$ by

$$M_t = \mathbb{E}[M|\mathcal{F}_t] = \mathbb{E}[M] + \int_0^t H_s \,\mathrm{d}B_s,$$

it follows that $(M_t)_{t\in[0,\infty)}$ is a solution to the stochastic differential equation

$$\begin{cases} \mathrm{d}M_t = H_t \,\mathrm{d}B_t \quad \text{in } (0,\infty), \\ M_0 = \mathbb{E}[M] \end{cases}$$

Much like the proof of (4.11), the proof of (4.12) is done by proving the density of certain simple functions, which in this case are provided by the stochastic exponential.

**Remark 4.12.** Some care if necessary when applying the above intuition. It is not true, in general, that Brownian motion can be replaced by an arbitrary martingale $(X_t)_{t\in[0,\infty)}$. In particular, the proof presented below uses the fact that Brownian motion has a deterministic quadratic variation process $\langle B\rangle_t = t$.

Define the set of simple functions

$$\mathcal{S} = \left\{ f = \sum_{i=0}^{n-1} \lambda_i \mathbf{1}_{(t_i,t_{i+1}]} \colon \lambda_i \in \mathbb{R} \text{ and } 0 = t_0 < t_1 < \ldots < t_n < \infty \right\}.$$

Then, for every $f \in \mathcal{S}$, let $(M_t^f)_{t\in[0,\infty)}$ denote the martingale

$$M_t^f = \int_0^t f_s \,\mathrm{d}B_s = \sum_{i=0}^{n-1} \lambda_i (B_{t_{i+1}\wedge t} - B_{t_i\wedge t}),$$

and the stochastic exponential $(\mathcal{E}(M^f)_t)_{t\in[0,\infty)}$ defined by

$$\mathcal{E}(M^f)_t = \exp\left( \sum_{i=0}^{n-1} \lambda_i (B_{t_{i+1}\wedge t} - B_{t_i\wedge t}) - \sum_{i=0}^{n-1} \frac{\lambda_i^2((t_{i+1}\wedge t) - (t_i \wedge t))}{2} \right).$$

solves

$$\begin{cases} \mathrm{d}\mathcal{E}(M^f)_t = \mathcal{E}(M^f)_t \,\mathrm{d}M^f_t = \mathcal{E}(M^f)_t f_t \,\mathrm{d}B_t \quad \text{in} \;\; (0,\infty), \\ \mathcal{E}(M^f)_0 = 1. \end{cases}$$

We will write $\mathcal{E}(M^f) = \mathcal{E}(M^f)_\infty$, and observe that this is to say

$$\mathcal{E}(M^f) = 1 + \int_0^\infty \mathcal{E}(M^f)_t f_t \,\mathrm{d}B_t,$$

which takes the form (4.12) for $H_t = \mathcal{E}(M^f)_t f_t$. In order to prove (4.12), it will therefore suffice to prove that the linear span of $\{\mathcal{E}(M^f)\}_{f\in\mathcal{S}}$ is dense in $L^2(\Omega,\mathcal{F},\mathbb{P})$.

**Proposition 4.13.** *The linear span of* $\{\mathcal{E}(M^f)_\infty\}_{f\in\mathcal{S}}$ *is dense in* $L^2(\Omega,\mathcal{F}_\infty,\mathbb{P})$.

*Proof.* Let $Y \in L^2(\Omega,\mathcal{F}_\infty,\mathbb{P})$. It suffices to prove that if

$$\mathbb{E}[Y\mathcal{E}(M^f)] = 0 \;\text{ for every }\; f\in\mathcal{S} \;\text{ if and only if }\; Y=0.$$

It is clear that the expectation vanishes if $Y = 0$. To prove the converse, suppose that $Y \in L^2(\Omega,\mathcal{F}_\infty,\mathbb{P})$ satisfies that, for every $f\in\mathcal{S}$,

$$\mathbb{E}[Y\mathcal{E}(M^f)] = 0.$$

This implies that for every $\lambda_0,\ldots,\lambda_{n-1}\in\mathbb{R}$ and $0=t_0<t_1<\ldots<t_n<\infty$,

$$\mathbb{E}\left[\exp\left(\sum_{i=0}^{n-1}\lambda_i(B_{t_{i+1}}-B_{t_i}) - \sum_{i=0}^{n-1}\frac{\lambda_i^2(t_{i+1}-t_i)}{2}\right)Y\right] = 0,$$

and therefore that

$$\mathbb{E}\left[\exp\left(\sum_{i=0}^{n-1}\lambda_i(B_{t_{i+1}}-B_{t_i})\right)Y\right] = 0. \tag{4.13}$$

Fix a sequence $0=t_0<t_1<\ldots<t_n<\infty$. Since $Y\in L^2(\Omega,\mathcal{F}_\infty,\mathbb{P})$, it follows that the function

$$(z_0,\ldots,z_{n-1})\in\mathbb{C}^n \mapsto \mathbb{E}\left[\exp\left(\sum_{i=0}^{n-1}z_i(B_{t_{i+1}}-B_{t_i})\right)Y\right] \;\text{ is holomorphic.} \tag{4.14}$$

Since the holomorphic function (4.14) vanishes on the connected open set $\mathbb{R}^n\subseteq\mathbb{C}^n$ by (4.13), the function (4.14) is identically zero. In particular, for every $\lambda_0,\ldots,\lambda_{n-1}\in\mathbb{R}$,

$$\mathbb{E}\left[\exp\left(i\sum_{i=0}^{n-1}\lambda_i(B_{t_{i+1}}-B_{t_i})\right)Y\right] = 0. \tag{4.15}$$

Define the measure $\mathrm{d}\mathbb{Q} = Y\,\mathrm{d}\mathbb{P}$ and define the measurable map $F\colon\Omega\to\mathbb{R}^n$ by

$$\begin{aligned} F(\omega) &= \left(B_{t_1}(\omega)-B_{t_0}(\omega), B_{t_2}(\omega)-B_{t_1}(\omega),\ldots,B_{t_n}(\omega)-B_{t_{n-1}}(\omega)\right) \\ &= \left(B_{t_1}(\omega), B_{t_2}(\omega)-B_{t_1}(\omega),\ldots,B_{t_n}(\omega)-B_{t_{n-1}}(\omega)\right). \end{aligned}$$

Let $F_*\mathbb{Q}$ be the pushforward measure on $\mathbb{R}^n$ defined for every Borel measurable set $A\subseteq\mathbb{R}^n$ by

$$F_*\mathbb{Q}(A) = \mathbb{Q}(F^{-1}(A)) = \mathbb{Q}\left(\left(B_{t_1}, B_{t_2}-B_{t_1},\ldots,B_{t_n}-B_{t_{n-1}}\right)\in A\right). \tag{4.16}$$

It follows by definition and (4.15) that, for every $\lambda=(\lambda_0,\ldots,\lambda_{n-1})\in\mathbb{R}^n$,

$$\mathbb{E}\left[\exp\left(i\sum_{i=0}^{n-1}\lambda_i(B_{t_{i+1}}-B_{t_i})\right)Y\right] = \int_{\mathbb{R}^n}\exp(i\langle\lambda,x\rangle)F_*\mathbb{Q}(\,\mathrm{d}x) = 0.$$

We conclude that the Fourier transform of $F_*\mathbb{Q}$ vanishes, and therefore that $F_*\mathbb{Q}$ is the zero measure on $\mathbb{R}^n$. Returning to (4.16), it follows from the definition of $\mathbb{Q}$ that, for every set $B \in \sigma(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}}) = \sigma(B_{t_1}, B_{t_2}, \ldots, B_{t_n})$,

$$\mathbb{E}[Y\mathbf{1}_B] = 0.$$

Since the collection

$$\{B \in \mathcal{F}_\infty \colon B \in \sigma(B_{t_1}, B_{t_2}, \ldots, B_{t_n}) \text{ for some } 0 = t_0 < t_1 < \ldots < t_n < \infty\}$$

is a $\pi$-system on which the measure $\mathrm{d}\mathbb{Q} = Y\,\mathrm{d}\mathbb{P}$ vanishes, and since

$$\mathcal{F}_\infty = \sigma\left(\{B \in \mathcal{F}_\infty \colon B \in \sigma(B_{t_1}, B_{t_2}, \ldots, B_{t_n}) \text{ for some } 0 = t_0 < t_1 < \ldots < t_n < \infty\}\right),$$

we conclude that $\mathrm{d}\mathbb{Q} = Y\,\mathrm{d}\mathbb{P}$ is the zero measure on $\mathcal{F}_\infty$. This implies that $Y = 0$, since $Y$ is $\mathcal{F}_\infty$-measurable, which completes the proof. $\square$

**Theorem 4.14.** *Let $M \in L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$. Then there exists a unique process $(H_s)_{s\in[0,\infty)} \in L^2(B)$ such that*

$$M = \mathbb{E}[M] + \int_0^\infty H_t\,\mathrm{d}B_t.$$

*Proof.* We observe that if $(H_s)_{s\in[0,\infty)} \in L^2(B)$ is a predictable process, and if $X$ is the random variable defined for some $c \in \mathbb{R}$ by

$$X = c + \int_0^\infty H_t\,\mathrm{d}B_t,$$

then it follows that

$$c = \mathbb{E}[X] \text{ since } \mathbb{E}\left[\int_0^\infty H_t\,\mathrm{d}B_t\right] = 0.$$

We therefore consider the linear subspace

$$\mathcal{I} = \left\{X = \mathbb{E}[X] + \int_0^\infty H_t\,\mathrm{d}B_t \colon (H_s)_{s\in[0,\infty)} \in L^2(B)\right\} \subseteq L^2(\Omega, \mathcal{F}_\infty, \mathbb{P}),$$

where the rightmost inclusion follows from the fact that $(H_s)_{s\in[0,\infty)} \in L^2(B)$ and from the definition of $\mathcal{F}_\infty$. The fact that $\mathcal{I}$ is a linear subspace follows from the linearity of the expectation, the linearity of the stochastic integral, and the fact that $L^2(B)$ is a vector space.

We aim to prove that $\mathcal{I} = L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$. For this, it is sufficient to prove (i) that $\mathcal{I}$ is closed and (ii) that $\mathcal{I}$ contains a dense subset. For (i), suppose that $\{X_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{I}$ corresponding to a sequence $\{(H^n_s)_{s\in[0,\infty)}\}_{n\in\mathbb{N}}$ in $L^2(B)$. It follows from the Itô isometry, Young's inequality, and Hölder's inequality that, for every $n, m \in \mathbb{N}$,

$$\begin{aligned}\mathbb{E}\left[\int_0^\infty |H^n_s - H^m_s|^2\,\mathrm{d}s\right] &= \mathbb{E}\left[\left(\int_0^\infty (H^n_s - H^m_s)\,\mathrm{d}B_s\right)^2\right] \\ &\leq 2\left(\mathbb{E}\left[|X_n - X_m|^2\right] + |\mathbb{E}[X_n] - \mathbb{E}[X_m]|^2\right). \\ &\leq 4\mathbb{E}\left[|X_n - X_m|^2\right].\end{aligned}$$

We therefore conclude that $\{(H^n_s)\}_{s\in[0,\infty)}$ is a Cauchy sequence in $L^2(B)$.

It follows that there exists $X_\infty \in L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$ and $(H^\infty_s)_{s\in[0,\infty)}$ such that, as $n \to \infty$,

$$X_n \to X_\infty \text{ strongly in } L^2(\Omega, \mathcal{F}_\infty, \mathbb{P}),$$

such that, as $n \to \infty$,

$$\mathbb{E}[X_n] \to \mathbb{E}[X_\infty],$$

and such that, as $n \to \infty$,

$$\int_0^\infty H_s^n \,\mathrm{d}B_s \to \int_0^\infty H_s^\infty \,\mathrm{d}B_s \text{ strongly in } L^2(\Omega, \mathcal{F}_\infty, \mathbb{P}),$$

where this final convergence is again a consequence of the Itô isometry. After passing to the limit $n \to \infty$ in the equation

$$X_n = \mathbb{E}[X_n] + \int_0^\infty H_s^n \,\mathrm{d}B_s,$$

it follows that

$$X_\infty = \mathbb{E}[X_\infty] + \int_0^\infty H_s^\infty \,\mathrm{d}B_s \text{ in } L^2(\Omega, \mathcal{F}_\infty, \mathbb{P}).$$

Therefore, we have $X_\infty \in \mathcal{I}$ and we conclude that $\mathcal{I}$ is closed.

To prove (ii), since for every $f \in \mathcal{S}$ properties of the stochastic exponential prove that

$$\mathcal{E}(M^f) = 1 + \int_0^\infty \mathcal{E}(M^f)_s \,\mathrm{d}B_s = \mathbb{E}[\mathcal{E}(M^f)] + \int_0^\infty \mathcal{E}(M^f)_s \,\mathrm{d}B_s,$$

we conclude that $\{\mathcal{E}(M^f)\}_{f\in\mathcal{S}} \subseteq \mathcal{I}$. Proposition 4.13 then proves that $\mathcal{I}$ contains a dense subset. This completes the proof of (ii), and therefore the proof that $\mathcal{I} = L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$. That is, for every $M \in L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$, there exists $(H_s)_{s\in[0,\infty)} \in L^2(B)$ such that

$$M = \mathbb{E}[M] + \int_0^\infty H_s \,\mathrm{d}B_s.$$

This completes the proof. (The proof of uniqueness will be added after the third problem session.) $\square$

## 5. Local Times

In this section, we will develop the notion of the *local time* associated to a one-dimensional semimartingale $(X_t)_{t\in[0,\infty)}$. Roughly speaking, for $a \in \mathbb{R}$, the local time $(L_t^a)_{t\in[0,\infty)}$ or sometimes written $(L_t^a(X))_{t\in[0,\infty)}$ measures the time $(X_t)_{t\in[0,\infty)}$ spends at the point $a$ up to time $t$. However, time in this case is measured with respect to the quadratic variation process $(\langle X\rangle_t)_{t\in[0,\infty)}$. As motivated by the Dambis-Dubins-Schwarz theorem, the quadratic variation process acts as an internal clock for the semimartingale $(X_t)_{t\in[0,\infty)}$.

Let $A \subseteq \mathbb{R}^d$. The random time $t_A$ the process $(X_t)_{t\in[0,\infty)}$ spends occupying $A$ up to time $t$ is defined by the integral

$$t_A = \int_0^t \mathbf{1}_A(X_s)\,\mathrm{d}\langle X\rangle_s.$$

In particular, $t_{\mathbb{R}} = \langle X\rangle_t$ is simply the total time elapsed as measured by the quadratic variation, since the process resides always in the whole space $\mathbb{R}$. However, we expect that this occupation time can be similarly expressed in terms of local times. That is,

$$t_A = \int_A L_t^a\,\mathrm{d}a,$$

or, more generally, for any nonnegative Borel measurable function $\Phi\colon \mathbb{R} \to \mathbb{R}$,

$$\int_0^t \Phi(X_s)\,\mathrm{d}\langle X\rangle_s = \int_{\mathbb{R}} \Phi(a) L_t^a\,\mathrm{d}a.$$

This is the occupation formula, which we prove below. The construction of the local time will rely on properties of convex functions, and an important extension of Itô's formula from $\mathrm{C}^2$-functions to functions that can be written as the differenceof two convex functions.

5.1. **Convex functions.** In this section, we will prove that a convex function $f\colon \mathbb{R}\rightarrow\mathbb{R}$ is always twice-differnetiable, at least in a distributional sense.

**Definition 5.1.** A function $f\colon \mathbb{R}\rightarrow\mathbb{R}$ is convex if, for every $a,b\in\mathbb{R}$ and every $\lambda\in[0,1]$,

$$f(\lambda a+(1-\lambda)b)\leq \lambda f(a)+(1-\lambda)f(b).$$

We observe that the following proposition implies, in particular, that finite convex functions are always locally bounded and continuous.

**Proposition 5.2.** *Let $f\colon\mathbb{R}\rightarrow\mathbb{R}$ be convex. Then the left derivative $f'_-\colon\mathbb{R}\rightarrow\mathbb{R}$ defined by*

$$f'_-(x)=\lim_{h\rightarrow 0}\frac{f(x)-f(x-h)}{h}\;\; \textit{exists and is nondecreasing and left continuous.}$$

*Similarly, the right derivative $f'_+(x)\colon\mathbb{R}\rightarrow\mathbb{R}$ defined by*

$$f'_+(x)=\lim_{h\rightarrow 0}\frac{f(x+h)-f(x)}{h}\;\; \textit{exists and is nondecreasing and right continuous.}$$

*Furthermore, for all $\phi\in \mathrm{C}^\infty_c(\mathbb{R})$,*

$$\int_{\mathbb{R}} f(x)\phi'(x)\,\mathrm{d}x=-\int_{\mathbb{R}} f'_-(x)\phi(x)\,\mathrm{d}x=-\int_{\mathbb{R}} f'_+(x)\phi(x)\,\mathrm{d}x.$$

*Proof.* We will prove the statement for the left derivative, since the case of the right derivative is virtually identical. Let $h_1<h_2\in(0,\infty)$. Then, by convexity,

$$f(x-h_1)\leq\frac{h_1}{h_2}f(x-h_2)+\left(1-\frac{h_1}{h_2}\right)f(x).$$

Hence, after reordering this inequality,

$$\frac{f(x)-f(x-h_2)}{h_2}\leq\frac{f(x)-f(x-h_1)}{h_1},$$

which implies that, as $h\rightarrow 0^+$,

$$\frac{f(x)-f(x-h)}{h}\;\;\text{is nondecreasing.}$$

Furthermore, for every $h\in(0,\infty)$, it follows by convexity that

$$f(x)\leq\frac{1}{1+h}f(x-h)+\frac{h}{1+h}f(x+1).$$

Hence, after reordering this inequality,

$$\frac{f(x)-f(x-h)}{h}\leq f(x+1)-f(x),$$

which implies that

$$\left\{\frac{f(x)-f(x-h)}{h}\right\}_{h\in(0,\infty)}\;\;\text{is bounded from above.}$$

We therefore conclude that, as an increasing sequence that is bounded from above,

$$f'_-(x)=\lim_{h\rightarrow 0}\frac{f(x)-f(x-h)}{h}\;\;\text{exists.}$$

A repetition of the above arguments prove that, for every $x<y\in\mathbb{R}$ and $h\in(0,|x-y|)$,

$$\frac{f(x)-f(x-h)}{h}\leq\frac{f(y-h)-f(x)}{y-x-h}\leq\frac{f(y)-f(y-h)}{h},$$

which proves after passing to the limit $h \to 0$ that $f'_-(x) \leq {}^{f(y)-f(x)}/_{y-x} \leq f'_-(y)$. This completes the proof that $f'_-$ is nondecreasing. Finally, since the sequence is decreasing in both $h, h' \in (0, \infty)$,

$$\begin{aligned}\lim_{h\to 0}\left(f'_-(x-h)\right) &= \lim_{h\to 0}\left(\lim_{h'\to 0}\frac{f(x-h)-f(x-h-h')}{h'}\right)\\ &= \lim_{h'\to 0}\left(\lim_{h\to 0}\frac{f(x-h)-f(x-h-h')}{h'}\right)\\ &= \lim_{h'\to 0}\frac{f(x)-f(x-h')}{h'}\\ &= f'_-(x),\end{aligned}$$

which completes the proof of left continuity.

Finally, suppose that $\phi \in \mathrm{C}_c^\infty(\mathbb{R})$. Then, since $f$ and $f'_-$ are locally bounded, the dominated convergence theorem and a change of variables prove that

$$\begin{aligned}\int_{\mathbb{R}} f(x)\phi'(x)\,\mathrm{d}x &= \lim_{h\to 0}\int_{\mathbb{R}} f(x)\left(\frac{\phi(x+h)-\phi(x)}{h}\right)\mathrm{d}x\\ &= -\left(\lim_{h\to 0}\int_{\mathbb{R}}\left(\frac{f(x)-f(x-h)}{h}\right)\phi(x)\,\mathrm{d}x\right)\\ &= -\int_{\mathbb{R}} f'_-(x)\phi(x)\,\mathrm{d}x,\end{aligned}$$

which completes the proof. □

In the final proposition of this section, we prove that the second derivative of a convex function exists in a distributional sense. Precisely, the second derivative is the Riemann-Stietjes measure associated to $f'_-$. We emphasize that the choice to work with the left derivative as opposed to the right derivative is not essential, but it is necessary to choose one or the other.

**Proposition 5.3.** *Let $f\colon \mathbb{R} \to \mathbb{R}$ be a convex function. Let $\mu_f$ denote the nonnegative Riemann-Stieljes measure associated to the nondecreasing left derivative $f'_-$ defined for every $a \leq b \in \mathbb{R}$ by*

$$\mu_f[(a,b)] = f'_-(b) - f'_-(a+) = f'_-(b) - \left(\lim_{h\to 0} f'_-(a+h)\right).$$

*Then, for all $\phi \in \mathrm{C}_c^\infty(\mathbb{R})$,*

$$\int_{\mathbb{R}} f(x)\phi''(x)\,\mathrm{d}x = -\int_{\mathbb{R}} f'_-(x)\phi'(x)\,\mathrm{d}x = \int_{\mathbb{R}}\phi''(x)\mu_f(\,\mathrm{d}x).$$

*Proof.* The first equality was proven above. It remains only to prove that

$$-\int_{\mathbb{R}} f'_-(x)\phi'(x)\,\mathrm{d}x = \int_{\mathbb{R}}\phi''(x)\mu_f(\,\mathrm{d}x),$$

which, owing to the fact that $\phi \in \mathrm{C}_c^\infty(\mathbb{R})$, is the integration-by-parts formula for Riemann-Stieltjes integrals. □

5.2. **The Meyer-Tanaka formula.** We will now define the local time of a continuous semimartingale $(X_t)_{t\in[0,\infty)}$, which will lead to an important generalization of Itô's formula. In particular, the next proposition proves that a convex function of a semimartingale is again a semimartingale, much like Itô's formula proved that a $\mathrm{C}^2$-function of a semimartingale is again a semimartingale.

**Proposition 5.4.** *Let $(X_t)_{t\in[0,\infty)}$ be a continuous semimartingale, and let $f\colon\mathbb{R}\to\mathbb{R}$ be convex. Then, there exists a nondecreasing process $(A^f_t)_{t\in[0,\infty)}$ such that*

$$f(X_t)=f(X_0)+\int_0^t f'_-(X_s)\,\mathrm{d}X_s+A^f_t.$$

*Proof.* By localization using stopping times, we can assume without loss of generality that $(X_t=M_t+A_t)_{t\in[0,\infty)}$, for a local martingale $(M_t)_{t\in[0,\infty)}$ and a process of bounded variation $(A_t)_{t\in[0,\infty)}$, is bounded in the sense that there exists $N\in\mathbb{N}$ such that, almost surely,

$$\sup_{t\in[0,\infty)}\left(|X_t|+\int_0^t|\,\mathrm{d}A_t|+\langle M\rangle_t\right)<N.$$

Let $\rho\in\mathrm{C}^\infty_c(\mathbb{R})$ be a smooth function supported on the set $(-\infty,0]\subseteq\mathbb{R}$. For every $\varepsilon\in(0,1]$ define $\rho^\varepsilon\colon\mathbb{R}\to\mathbb{R}$ by

$$\rho^\varepsilon(x)=\varepsilon^{-1}\rho\left(\frac{x}{\varepsilon}\right).$$

For every $\varepsilon\in(0,1]$ define $f^\varepsilon\colon\mathbb{R}\to\mathbb{R}$ to be the convolution

$$f^\varepsilon(x)=(f*\rho^\varepsilon)(x)=\int_{\mathbb{R}}f(y)\rho^\varepsilon(y-x)\,\mathrm{d}y=\int_{\mathbb{R}}f(y+x)\rho^\varepsilon(y)\,\mathrm{d}y.$$

Since the convexity of $f$ implies that $f$ is locally bounded and continuous, the functions $\{f^\varepsilon\}_{\varepsilon\in(0,1]}$ are smooth. Therefore, Itô's formula proves that, for every $t\in[0,\infty)$,

$$f^\varepsilon(X_t)=f^\varepsilon(X_0)+\int_0^t(f^\varepsilon)'(X_s)\,\mathrm{d}X_s+\frac{1}{2}\int_0^t(f^\varepsilon)''(X_s)\,\mathrm{d}\langle X\rangle_s.$$

For the left derivative $f'_-$ of $f$ and for the Riemann-Stieljes measure $\mu_f$ of $f'_-$, we have that

$$(f^\varepsilon)'(x)=\int_{\mathbb{R}}f'_-(x+y)\rho^\varepsilon(y)\,\mathrm{d}y,$$

and

$$(f^\varepsilon)''(x)=\int_{\mathbb{R}}\rho^\varepsilon(y-x)\mu_f(\,\mathrm{d}x).$$

In particular, we observe that the convolution preserves the convexity in the sense that the nonegativity of the measure $\mu_f$ implies that, for every $x\in\mathbb{R}$,

$$(f^\varepsilon)''(x)=\int_{\mathbb{R}}\rho^\varepsilon(y-x)\mu_f(\,\mathrm{d}x)\geq 0,$$

and therefore that the process $(A^{f^\varepsilon}_t)_{t\in[0,\infty)}$ defined by

$$A^{f^\varepsilon}_t=\frac{1}{2}\int_0^t(f^\varepsilon)''(X_s)\,\mathrm{d}\langle X\rangle_s\ \text{ is nondecreasing.}$$

Since $f$ is continuous and hence uniformly continuous on compact sets, it follows using properties of the convolution that

$$\lim_{\varepsilon\to0}\left(\sup_{\{|x|\leq N\}}|f^\varepsilon(x)-f(x)|\right)=0.$$

Similarly, the triangle inequality, the Itô isometry prove, the construction of the convolution kernel, the left continuity of $f'_-$, the local boundedness of $f'_-$, the dominated convergence theorem, and Doob's inequality prove that, for every $t \in [0,\infty)$,

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\left(\mathbb{E}\left[\sup_{s\in[0,t]}\left(\int_0^s (f^\varepsilon)'(X_r) - f'_-(X_r)\,\mathrm{d}X_r\right)^2\right]^{\frac{1}{2}}\right) \\
&= \lim_{\varepsilon\to 0}\left(4\mathbb{E}\left[\int_0^t \left((f^\varepsilon)'(X_s) - f'_-(X_s)\right)^2 \mathrm{d}\langle X\rangle_s\right]^{\frac{1}{2}} + \mathbb{E}\left[\left(\int_0^t \left((f^\varepsilon)'(X_s) - f'_-(X_s)\right)^2 |\,\mathrm{d}A_s|\right)^2\right]^{\frac{1}{2}}\right) \\
&= 0.
\end{aligned}$$

After passing to a subsequence $\{\varepsilon_k\}_{k\in\mathbb{N}}$, we conclude almost surely that, for every $t \in [0,\infty)$,

$$\lim_{k\to\infty}\left(\sup_{t\in[0,\infty)}\left(\left|\int_0^t (f^{\varepsilon_k})'(X_s)\,\mathrm{d}X_s - \int_0^t f'_-(X_s)\,\mathrm{d}X_s\right|\right)\right) = 0.$$

Since by definition, for every $k \in N$,

$$A_t^{f^{\varepsilon_k}} = f_t^{\varepsilon_k}(X_t) - f^{\varepsilon_k}(X_0) - \int_0^t (f^{\varepsilon_k})'(X_s)\,\mathrm{d}X_s,$$

we conclude that, as a limit of nondecreasing processes the process $(A_t^f)_{t\in[0,\infty)}$ defined by

$$A_t^f = \lim_{k\to\infty} A_t^{f^{\varepsilon_k}} \text{ exists almost surely,}$$

and almost surely satisfies, for every $t \in [0,\infty)$,

$$f(X_t) = f(X_0) + \int_0^t f'_-(X_s)\,\mathrm{d}X_s + A_t^f,$$

which completes the proof. □

Roughly speaking, for $a \in \mathbb{R}$ we expect the local time $(L_t^a)_{t\in[0,\infty)}$ of a continuous semimartingale $(X_t)_{t\in[0,\infty)}$ to be defined by

$$L_t^a = \int_0^t \delta_0(X_s - a)\,\mathrm{d}\langle X\rangle_s,$$

where $\delta_0$ denotes the Dirac distribution at zero. Since for a convex $f$ the process $(A_t^f)_{t\in[0,\infty)}$ constructed above is in spirit given by

$$A_t^f = \frac{1}{2}\int_0^t f''(X_s)\,\mathrm{d}\langle X\rangle_s,$$

we expect that the local time will be defined by the process $(A_t^f)_{t\in[0,\infty)}$ for $f(x) = |x-a|$, since in this case

$$f'_-(x) = \mathbf{1}_{\{x>a\}} - \mathbf{1}_{\{x\le a\}} \text{ and } f''(x) = \delta_0(x-a) \text{ as distributions.}$$

This is the content of the next proposition. We write sgn for the left continuous version of the sign function

$$\operatorname{sgn}(x) = \mathbf{1}_{\{x>0\}} - \mathbf{1}_{\{x\le 0\}}.$$

Again, this choice is made because we work with the left derivative as opposed to the right derivative. The following is the *Tanaka formula.*

**Proposition 5.5.** *Let $(X_t)_{t\in[0,\infty)}$ be a continuous semimartingale. For every $a\in\mathbb{R}$ there exists an increasing process $(L^a_t)_{t\in[0,\infty)}$ called the local time of $(X_t)_{t\in[0,\infty)}$ in $a$ such that, for every $t\in[0,\infty)$,*

$$(X_t-a)_+=(X_0-a)_++\int_0^t \mathbf{1}_{\{X_s>a\}}\,\mathrm{d}X_s+\frac{1}{2}L^a_t,$$
$$(X_t-a)_-=(X_0-a)_-+\int_0^t \mathbf{1}_{\{X_s\leq a\}}\,\mathrm{d}X_s+\frac{1}{2}L^a_t,$$
$$|X_t-a|=|X_0-a|+\int_0^t sgn(X_s-a)\,\mathrm{d}X_s+L^a_t.$$

*Proof.* By convexity of the functions $f(x)_+=(x-a)_+$ and $f(x)_-=(x-a)_-$, there exist increasing processes $(A^+_t)_{t\in[0,\infty)}$ and $(A^-_t)_{t\in[0,\infty)}$ such that

$$(X_t-a)_+=(X_0-a)_++\int_0^t \mathbf{1}_{\{X_s>a\}}\,\mathrm{d}X_s+A^+_t,$$
$$(X_t-a)_-=(X_0-a)_-+\int_0^t \mathbf{1}_{\{X_s\leq a\}}\,\mathrm{d}X_s+A^-_t.$$

Subtracting these equalities proves that

$$X_t=X_0+\int_0^t \mathrm{d}X_s+(A^+_t-A^-_t),$$

and therefore by continuity that, almost surely,

$$A^+_t=A^-_t \;\text{ for every }\; t\in[0,\infty).$$

We therefore define $(L^a_t)_{t\in[0,\infty)}=(2A^+_t)_{t\in[0,\infty)}$, and conclude by subtracting the previous equalities that

$$|X_t-a|=|X_0-a|+\int_0^t \mathrm{sgn}(X_s)\,\mathrm{d}X_s+L^a_t,$$

which completes the proof. □

**Proposition 5.6.** *Let $(X_t)_{t\in[0,\infty)}$ be a continuous semimartingale and let $(L^a_t)_{a\in\mathbb{R},t\in[0,\infty)}$ denote the local times of $(X_t)_{t\in[0,\infty)}$. There exists a $\mathcal{B}(\mathbb{R}\times[0,\infty))\otimes\mathcal{F}$-measurable process*

$$(a,t,\omega)\to\tilde{L}(a,t,\omega)$$

*such that, for every $a\in\mathbb{R}$, the processes $(t,\omega)\mapsto\tilde{L}(a,t,\omega)$ and $(t,\omega)\mapsto L^a_t(\omega)$ are indistinguishable.*

*Proof.* The proof is an immediate consequence of Fubini's theorem for stochsatic integrals. The proof will be added after the fourth problem session. □

**Remark 5.7.** Given a continuous semimartingale $(X_t)_{t\in[0,\infty)}$ we by indistinguishability we can and will henceforth assume without loss of generality that the process $(a,t,\omega)\mapsto L^a_t(\omega)$ is $\mathcal{B}(\mathbb{R}\times[0,\infty))\otimes\mathcal{F}$-measurable.

Again, the intuition is that $(L^a_t)_{t\in[0,\infty)}$ measures the time $(X_t)_{t\in[0,\infty)}$ spends at $a$. The following proposition proves that the random Riemann-Stieltjes measure associated to $(L^a_t)_{t\in[0,\infty)}$ is supported on the random set $\{X_t=a\}$. However, it is clear that, in general, the support of the measure will be smaller than the set $\{X_t=a\}$. This can be seen, for instance, by considering the constant process $X_t=0$. In this case, for $a=0$, we have that $\{X_t=0\}=[0,\infty)$. But, the local time $L^0_t=0$ vanishes identically as well, due to the fact "no time elapses" because the quadratic variation $\langle X\rangle_t=0$ vanishes identically.

**Proposition 5.8.** *Let $(X_t)_{t\in[0,\infty)}$ be a continuous local martingale, let $a\in\mathbb{R}$, and let $(L^a_t)_{t\in[0,\infty)}$ be the local time of $(X_t)_{t\in[0,\infty)}$ in $a$. Then the random Riemann-Stieltjes measure $\mathrm{d}L^a_t$ on $[0,\infty)$ is almost surely supported on the random set $\{X_t=a\}$.*

*Proof.* The proof is based on the equality $(X_t - a)^2 = |X_t - a| \cdot |X_t - a|$. On the lefthand side, Itô's formula proves that

$$\begin{aligned}(X_t - a)^2 &= (X_0 - a)^2 + 2\int_0^2 (X_s - a)\,\mathrm{d}X_s + \int_0^t \mathrm{d}\langle X\rangle_s \\ &= (X_0 - a)^2 + 2\int_0^t (X_s - a)^2\,\mathrm{d}X_s + \langle X\rangle_t.\end{aligned}$$

On the righthand side, the integration-by-parts formula and Tanaka's formula prove that

$$\begin{aligned}(X_t - a)^2 &= (X_0 - a)^2 + 2\int_0^t |X_s - a|\ \mathrm{d}\left(|X_s - a|\right) + \langle |X_t - a|\rangle_t, \\ &= (X_0 - a)^2 + 2\int_0^t \mathrm{sgn}(X_s - a)\,|X_s - a|\ \mathrm{d}X_s + \int_0^t |X_s - a|\ \mathrm{d}L_s^a + \int_0^t \mathrm{d}\langle X\rangle_s \\ &= (X_0 - a)^2 + 2\int_0^t (X_s - a)\,\mathrm{d}X_s + \int_0^t |X_s - a|\ \mathrm{d}L_s^a + \langle X\rangle_t.\end{aligned}$$

Therefore, we conclude almost surely that, for every $t \in [0, \infty)$,

$$\int_0^t |X_s - a|\ \mathrm{d}L_s^a = 0,$$

which completes the proof. □

We are now prepared to present the Meyer-Tanaka formula, which generalizes Itô's formula to functions $f$ that are the difference of two convex functions. We emphasize that the above reasoning can equally be applied to a concave function, which is nothing more than the negative of a convex function.

**Theorem 5.9.** *Let $(X_t)_{t\in[0,\infty)}$ be a continuous semimartingale, let $(L_t^a)_{a\in\mathbb{R},t\in[0,\infty)}$ be the local times of $(X_t)_{t\in[0,\infty)}$, let $f\colon \mathbb{R} \to \mathbb{R}$ be the difference of two convex functions, and let $\mu_f$ denote the Riemann-Stieljes integral of $f'_-$. Then, for every $t \in [0, \infty)$,*

$$f(X_t) = f(X_0) + \int_0^t f'_-(X_s)\,\mathrm{d}X_s + \frac{1}{2}\int_0^t L_t^a \mu_f(\,\mathrm{d}a).$$

*Proof.* By linearity, it suffices to prove the theorem in the case that $f$ is convex. By a stopping time argument, we can assume without loss of generality that, almost surely, for some $N \in \mathbb{N}$,

$$\sup_{t\in[0,\infty)} \left(|X_t| + \langle X\rangle_t\right) < N.$$

Define the convex function $f^N\colon \mathbb{R} \to \mathbb{R}$ by

$$f^N(x) = \begin{cases} f(N) + f'_-(-N)(x + N) & \text{if } x \in (-\infty, -N], \\ f(x) & \text{if } x \in [-N, N], \\ f(N) + f'_-(N)(x - N) & \text{if } x \in [N, \infty). \end{cases}$$

The convexity of $f$ proves that $f^N$ is convex, and by definition $\mu_{f^N}$ has compact support and satisfies

$$f^N = f,\ \ (f^N)'_- = f'_N,\ \text{ and }\ \mu_{f^N} = \mu_f\ \text{ on }\ [-N, N].$$

Since $\mu_{f^N}$ as compact support, define the function

$$g(x) = \int_{\mathbb{R}} |x - a|\,\mu_{f^N}(\,\mathrm{d}a),$$

and observe by a direct computation using the compact support of $\mu_{f^N}$ that

$$g'_-(x) = \int_{\mathbb{R}} \left(\mathbf{1}_{\{x>a\}} - \mathbf{1}_{\{x\leq a\}}\right) \mu_{f^N}(\,\mathrm{d}a) = \int_{\mathbb{R}} \operatorname{sgn}(x-a)\mu_{f^N}(\,\mathrm{d}a).$$

Then, by Fubini's theorem, for every $\phi \in \mathrm{C}^\infty_c(\mathbb{R})$,

$$\int_{\mathbb{R}} g'_-(x)\phi'(x)\,\mathrm{d}x = \int_{\mathbb{R}}\int_{\mathbb{R}} \operatorname{sgn}(x-a)\phi'(x)\mu_{f^N}(\,\mathrm{d}a)\,\mathrm{d}x = \int_{\mathbb{R}}\int_{\mathbb{R}} \operatorname{sgn}(x-a)\phi'(x)\,\mathrm{d}x\mu_{f^N}(\,\mathrm{d}a).$$

Since as distributions $\operatorname{sgn}'(x) = 2\delta_0(x)$, it follows that

$$\int_{\mathbb{R}} g'_-(x)\phi'(x)\,\mathrm{d}x = -2\int_{\mathbb{R}} \phi(a)\mu_{f^N}(\,\mathrm{d}a) = -2\int_{\mathbb{R}} \phi(x)\mu_{f^N}(\,\mathrm{d}x).$$

That is, as a distribution,

$$\left(f^N - \frac{1}{2}g\right)'' = 0 \ \text{ on } \ \mathbb{R}.$$

This implies that there exists $a, b \in \mathbb{R}$ such that

$$f^N(x) = a + bx + \frac{1}{2}g(x) = a + bx + \frac{1}{2}\int_{\mathbb{R}} |x-a|\,\mu_{f^N}(\,\mathrm{d}a).$$

Therefore, by Tanaka's formula,

$$\begin{aligned}
f^N(X_t) &= a + bX_t + \frac{1}{2}\int_{\mathbb{R}} |X_t - a|\,\mu_{f^N}(\,\mathrm{d}a) \\
&= a + bX_t + \frac{1}{2}\int_{\mathbb{R}} \left(|X_0 - a| + \int_0^t \operatorname{sgn}(X_s - a)\,\mathrm{d}X_s + L^a_t\right)\mu_{f^N}(\,\mathrm{d}a).
\end{aligned}$$

Since

$$(f^N)'_-(x) = b + g'(x) = b + \frac{1}{2}\int_{\mathbb{R}} \operatorname{sgn}(x-a)\mu_{f^N}(\,\mathrm{d}a),$$

and since

$$f^N(0) = a + bX_0 + \frac{1}{2}\int_{\mathbb{R}} |X_0 - a|\,\mu_{f^N}(\,\mathrm{d}a),$$

it follows by Fubini's theorem that

$$\begin{aligned}
f^N(X_t) &= f(X_0) + b(X_t - X_0) + \frac{1}{2}\int_{\mathbb{R}}\int_0^t \operatorname{sgn}(X_s - a)\,\mathrm{d}X_s\mu_{f^N}(\,\mathrm{d}a) + \frac{1}{2}\int_{\mathbb{R}} L^a_t\mu_{f^N}(\,\mathrm{d}a) \\
&= f^N(X_0) + b(X_t - X_0) + \int_0^t \frac{1}{2}\int_{\mathbb{R}} \operatorname{sgn}(X_s - a)\mu_{f^N}(\,\mathrm{d}a)\,\mathrm{d}X_s + \frac{1}{2}\int_{\mathbb{R}} L^a_t\mu_{f^N}(\,\mathrm{d}a) \\
&= f^N(X_0) + \int_0^t \left(b + \frac{1}{2}\int_{\mathbb{R}} \operatorname{sgn}(X_s - a)\mu_{f^N}(\,\mathrm{d}a)\right)\,\mathrm{d}X_s + \frac{1}{2}\int_{\mathbb{R}} L^a_t\mu_{f^N}(\,\mathrm{d}a) \\
&= f^N(X_0) + \int_0^t (f^N)'(X_s)X_s + \frac{1}{2}\int_{\mathbb{R}} L^a_t\mu_{f^N}(\,\mathrm{d}a).
\end{aligned}$$

Therefore, since almost surely we have

$$\sup_{t\in[0,\infty)} \left(|X_t| + \langle X\rangle_t\right) < N,$$

and since the fact that the support of $\mathrm{d}L^a_t$ is almost surely contained in the set $\{X_t = a\}$ implies that $L^a_t = 0$ for every $|a| \geq N$, it follows by the construction of $f^N$ that

$$f(X_t) = f(X_0) + \int_0^t f'_-(X_s)\,\mathrm{d}X_s + \frac{1}{2}\int_{\mathbb{R}} L^a_t\mu_f(\,\mathrm{d}a),$$

which completes the proof. $\square$

The equivalence of Itô's formula and the Meyer-Tanaka formula for twice differentiable convex functions now yields the occupation formula that motivated our study of local times.

**Proposition 5.10.** *Let $(X_t)_{t\in[0,\infty)}$ be a continuous semimartingale and let $(L_t^a)_{a\in\mathbb{R},t\in[0,\infty)}$ be the local times of $(X_t)_{t\in[0,\infty)}$. Then, for every nonnegative Borel measurable function $\Phi\colon \mathbb{R}\to\mathbb{R}$,*

$$\int_{\mathbb{R}} \Phi(X_s)\,\mathrm{d}\langle X\rangle_s = \int_{\mathbb{R}} \Phi(a) L_t^a\,\mathrm{d}a.$$

*Proof.* Let $f\colon \mathbb{R}\to\mathbb{R}$ be a twice continuously differentiable convex function. Then, it follows from Itô's formula and the Tanaka-Meyer formula that, on the one hand,

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\,\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(X_s)\,\mathrm{d}\langle X\rangle_s,$$

and on the other hand that

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\,\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(a) L_t^a\,\mathrm{d}a,$$

where we have used the fact that $f$ being twice continuously differentiable implies that $f' = f'_-$ and that $\mu_f(\,\mathrm{d}a) = f''(a)\,\mathrm{d}a$. This implies that

$$\int_0^t f''(X_s)\,\mathrm{d}\langle X\rangle_s = \int_0^t f''(a) L_t^a\,\mathrm{d}a.$$

The proof now follows by smooth approximation. □

5.3. **Regularity of local times.** In this section, we will prove that the local times $(L_t^a)_{a\in\mathbb{R},t\in[0,\infty)}$ admit a modification that is continuous in time and Càdlàg in $a$. A process defined on $\mathbb{R}$ is called Càdlàg if it is right continuous with left limits.

**Definition 5.11.** Let $(X_a)_{a\in\mathbb{R}}$ be a stochastic process. We say that $(X_a)_{a\in\mathbb{R}}$ is Càdlàg if almost surely the map $a \to X_a$ is right continuous and satisfies

$$X_{a^-} = \lim_{h\to 0^+} X_{a-h} \;\text{ exists.}$$

That is, the left limit of $(X_a)_{a\in\mathbb{R}}$ exists for every $a \in \mathbb{R}$.

**Proposition 5.12.** *Let $(X_t)_{t\in\mathbb{R}}$ be a continuous semimartingale. Let $(L_t^a)_{a\in\mathbb{R},t\in[0,\infty)}$ be the local times of $(X_t)_{t\in[0,\infty)}$ Then, the local times $(L_t^a)_{a\in\mathbb{R},t\in[0,\infty)}$ admit a modification $(\hat{L}_t^a)_{a\in\mathbb{R},t\in[0,\infty)}$ that is continuous in time and Càdlàg in $a$. Furthermore,*

$$\hat{L}_t^a - \hat{L}_t^{a^-} = 2\int_0^t \mathbf{1}_{\{X_s=a\}}\,\mathrm{d}V_s = 2\int_0^t \mathbf{1}_{\{X_s=a\}}\,\mathrm{d}X_s.$$

*Proof.* Let $T \in (0,\infty)$. By construction, for every $a \in \mathbb{R}$ the local time $(L_t^a)_{t\in[0,\infty)}$ is a continuous function of time. We will therefore apply the Kolmogorov continuity criterion to the map

$$a \in \mathbb{R} \mapsto (L_t^a)_{t\in[0,T]} \in \mathrm{C}([0,T];\mathbb{R}),$$

where $\mathrm{C}([0,T];\mathbb{R})$ is the Banach space of continuous functions from $[0,T]$ to $\mathbb{R}$ equipped with the norm

$$\|f-g\|_{\mathrm{C}([0,T];\mathbb{R})} = \sup_{s\in[0,T]} |f(s)-g(s)|\,.$$

Let $(X_t = M_t + A_t)_{t\in[0,\infty)}$ be the semimartingale decomposition of $(X_t)_{t\in[0,\infty)}$ for a local martingale $(M_t)_{t\in[0,\infty)}$ and a process of bounded variation $(A_t)_{t\in[0,\infty)}$. By definition, for every $a \in \mathbb{R}$,

$$L_t^a = 2\left((X_t-a)_+ - (X_0-a)_+ - \int_0^t \mathbf{1}_{\{X_s>a\}}\,\mathrm{d}M_s - \int_0^t \mathbf{1}_{\{X_s>a\}}\,\mathrm{d}A_s\right).$$

We first observe that

$$(X_t-a)_+ - (X_0-a)_+,$$

is a continuous function of $t \in [0,\infty)$ and $a \in \mathbb{R}$. For every $a \in \mathbb{R}$, let $(\tilde{M}_t^a)_{t\in[0,\infty)}$ denote the continuous local martingale

$$\tilde{M}_t^a = \int_0^t \mathbf{1}_{\{X_s>a\}}\,\mathrm{d}M_s.$$

The Burkholder-Davis-Gundy inequality and the occupation formula prove that, for every $k \in (0,\infty)$, for every $a \le b \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}\left|\tilde{M}_t^a - \tilde{M}_t^b\right|^{2k}\right] &\le C_k\mathbb{E}\left[\left(\int_0^T \mathbf{1}_{\{a<X_s\le b\}}\,\mathrm{d}\langle M\rangle_t\right)^k\right] \\
&= C_k\mathbb{E}\left[\left(\int_0^T \mathbf{1}_{\{a<X_s\le b\}}\,\mathrm{d}\langle X\rangle_t\right)^k\right] \\
&= C_k\mathbb{E}\left[\left(\int_a^b L_\infty^x\,\mathrm{d}x\right)^k\right] \\
&= C_k(b-a)^k\mathbb{E}\left[\left(\frac{1}{b-a}\int_a^b L_\infty^x\,\mathrm{d}x\right)^k\right].
\end{aligned}$$

Jensen's inequality and Fubini's theorem then prove that

$$
\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}\left|\tilde{M}^a_t-\tilde{M}^b_t\right|^{2k}\right] &\leq C_k(b-a)^k\mathbb{E}\left[\left(\frac{1}{b-a}\int_a^b L^x_\infty\,\mathrm{d}x\right)^k\right]\\
&\leq \mathbb{E}\left[\sup_{t\in[0,T]}\left|\tilde{M}^a_t-\tilde{M}^b_t\right|^{2k}\right]\\
&\leq C_k(b-a)^k\mathbb{E}\left[\frac{1}{b-a}\int_a^b (L^x_\infty)^k\,\mathrm{d}x\right]\\
&\leq C_k(b-a)^k\sup_{x\in\mathbb{R}}\mathbb{E}\left[(L^\infty_x)^k\right].
\end{aligned}
\tag{5.1}
$$

If for some $k\in(1,\infty)$ we have that

$$\sup_{x\in\mathbb{R}}\mathbb{E}\left[(L^\infty_x)^k\right]<\infty,$$

it then follows from (5.1) and the Kolmogorov continuity criterion that the map

$$a\in\mathbb{R}\to(\tilde{M}^a_t)_{t\in[0,T]}\in\mathrm{C}([0,T];\mathbb{R}),$$

admits a continuous modification. We will now show that we can always reduce to this case using a stopping time argument. Precisely, since for every $x\in\mathbb{R}$ and $t\in[0,\infty)$,

$$L^x_t=2\left((X_t-a)_+-(X_0-a)_+-\int_0^t \mathbf{1}_{\{X_s>a\}}\,\mathrm{d}X_s\right),$$

and since

$$|(X_t-x)_+-(X_0-x)_+|\leq|X-X_0|,$$

we have for every $k\in(1,\infty)$ using the Burkholder-Davis-Gundy inequality that there exists constants $c_{1,k},c_{2,k}\in(0,\infty)$ such that

$$
\begin{aligned}
\mathbb{E}(L^x_\infty)^k &\leq c_{1,k}\mathbb{E}\left[\sup_{t\in[0,\infty)}|X_t-X_0|^k+\sup_{t\in[0,T]}\left(\int_0^\infty \mathbf{1}_{\{X_s>x\}}\,\mathrm{d}M_s\right)^k+\left(\int_0^\infty|\,\mathrm{d}A_s|\right)^k\right]\\
&\leq c_{2,k}\mathbb{E}\left[\sup_{t\in[0,\infty)}|X_t-X_0|^k+\langle M\rangle_\infty^{\frac{k}{2}}+\left(\int_0^\infty|\,\mathrm{d}A_s|\right)^k\right].
\end{aligned}
\tag{5.2}
$$

We observe in particular that the righthand side of (5.2) is independent of $x\in\mathbb{R}$. Therefore, if we introduce the stopping times $(T_n)_{n\in\mathbb{N}}$ defined by

$$T_n=\inf\{t\in[0,\infty)\colon\sup_{s\in[0,t]}|X_s-X_0|+\int_0^t|\,\mathrm{d}A_s|+\langle M\rangle_t\geq n\},$$

it follows from (5.1) and (5.2) that the stopped process $(\tilde{M}^{T_n}_t)_{t\in[0,\infty)}$ admits a continuous modification for every $n\in\mathbb{N}$. Since $T_n\to\infty$ as $n\to\infty$ we conclude that $(\tilde{M}_t)_{t\in[0,\infty)}$ admits a continuous modification.

Let $(\hat{M}^a_t)_{a\in\mathbb{R},t\in[0,\infty)}$ denote the continuous modification of $(\tilde{M}^a_t)_{a\in\mathbb{R},t\in[0,\infty)}$. Let $(\hat{L}^a_t)_{a\in\mathbb{R},t\in[0,\infty)}$ be defined by

$$\hat{L}^a_t=2\left((X_t-a)_+-(X_0-a)_+-\hat{M}^a_t-\int_0^t\mathbf{1}_{\{X_s>a\}}\,\mathrm{d}A_s\right).$$

We observe that $(\hat{L}^a_t)_{a\in\mathbb{R},t\in[0,\infty)}$ is a modification of $(L^a_t)_{a\in\mathbb{R},t\in[0,\infty)}$. It remains to prove that the integral defined by the process of bounded variation is a Càdlàg process. For every $a\in\mathbb{R}$ and

$t \in [0, \infty)$, the dominated convergence theorem proves that

$$\lim_{b \to a^-} \int_0^t \mathbf{1}_{\{X_s > b\}} \, \mathrm{d}A_s = \int_0^t \mathbf{1}_{\{X_s \geq a\}} \, \mathrm{d}A_s.$$

Similarly, the dominated convergence proves that, for every $a \in \mathbb{R}$ and $t \in [0, \infty)$,

$$\lim_{b \to a^+} \int_0^t \mathbf{1}_{\{X_s > b\}} \, \mathrm{d}A_s = \int_0^t \mathbf{1}_{\{X_s > a\}} \, \mathrm{d}A_s.$$

It then follows from the continuity of $(\hat{M}_t^a)_{a \in \mathbb{R}, t \in [0,\infty)}$ and the continuity of $(X_t - a)_+$ that

$$\hat{L}_t^{a^-} = \lim_{b \to a^-} \hat{L}_t^b = 2 \left( (X_t - a)_+ - (X_0 - a)_+ - \hat{M}_t^a - \int_0^t \mathbf{1}_{\{X_s \geq a\}} \, \mathrm{d}A_s \right) \text{ exists.} \tag{5.3}$$

Similarly, we have that

$$\hat{L}_t^{a^+} = \lim_{b \to a^+} \hat{L}_t^b = 2 \left( (X_t - a)_+ - (X_0 - a)_+ - \hat{M}_t^a - \int_0^t \mathbf{1}_{\{X_s > a\}} \, \mathrm{d}A_s \right) = L_t^a. \tag{5.4}$$

Equations (5.3) and (5.4) prove that $(\hat{L}_t^a)_{a \in \mathbb{R}, t \in [0,\infty)}$ is a Càdlàg process. Furthermore, we have that

$$\hat{L}_t^a = \hat{L}_t^{a^-} = 2 \int_0^t \mathbf{1}_{\{X_s = a\}} \, \mathrm{d}V_s = 2 \int_0^t \mathbf{1}_{\{X_s = a\}} \, \mathrm{d}X_s,$$

where the final inequality follows from the fact that the occupation times formula proves that

$$2 \int_0^t \mathbf{1}_{\{X_s = a\}} \, \mathrm{d}\langle M \rangle_s = 0,$$

and therefore that

$$2 \int_0^t \mathbf{1}_{\{X_s = a\}} \, \mathrm{d}M_s = 0.$$

This completes the proof. □

The following corollary is an immediate consequence of Proposition 5.12.

**Corollary 5.13.** *Let $(X_t)_{t \in \mathbb{R}}$ be a continuous local martingale. Let $(L_t^a)_{a \in \mathbb{R}, t \in [0,\infty)}$ be the local times of $(X_t)_{t \in [0,\infty)}$ Then, the local times $(L_t^a)_{a \in \mathbb{R}, t \in [0,\infty)}$ admit a continuous modification $(\hat{L}_t^a)_{a \in \mathbb{R}, t \in [0,\infty)}$.*

## 6. The Girsanov theorem

We have seen in the previous sections that the class of semimartingales is stable with respect to addition, multiplication, and composition with twice-differentiable and convex functions. In this section, given a local martingale $(M_t)_{t\in[0,\infty)}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to a filtration $(\mathcal{F}_t)_{t\in[0,\infty($, we will analyze the effect of making an absolutely continuous change of measure.

Suppose that $\mathbb{Q}$ is a mutually absolutely continuous measure with respect to $\mathbb{P}$. Then, for each $t \in [0,\infty)$, the restriction $\mathbb{Q}|_{\mathcal{F}_t}$ is mutually absolutely continuous with $\mathbb{P}|_{\mathcal{F}_t}$. We will write $(D_t)_{t\in[0,\infty)}$ for the corresponding Radon-Nikodym derivatives. That is, for every $t \in [0,\infty)$ and $A \in \mathcal{F}_t$,

$$\mathbb{Q}(A) = \int_A D_t \, \mathrm{d}\mathbb{P}.$$

In general, if $(M_t)_{t\in[0,\infty)}$ is a local $\mathcal{F}_t$-martingale with respect to $\mathbb{P}$ then there is no reason to expect that $(M_t)_{t\in[0,\infty)}$ will be a local $\mathcal{F}_t$-martingale with respect to $\mathbb{Q}$. However, we will see that $(M_t)_{t\in[0,\infty)}$ is a semimartingale with respect to $\mathbb{Q}$, and its corresponding semimartingale decomposition is described by the Girsanov theorem to follow.

**Definition 6.1.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega, \mathcal{F})$. A pair of probability measures $(\mathbb{Q}, \mathbb{P})$ on $\mathcal{F}_\infty$ is called a *Girsanov pair* if $\mathbb{Q}$ and $\mathbb{P}$ are mutually absolutely continuous and if the Radon-Nikodym derivative $(D_t)_{t\in[0,\infty)}$ is a continuous $\mathbb{P}$-martingale.

**Remark 6.2.** Observe that the Radon-Nikodym derivative $(D_t)_{t\in[0,\infty)}$ is a $\mathbb{P}$-martingale by definition. However, the martingale $(D_t)_{t\in[0,\infty)}$ is not in general continuous. Therefore, in Definition 6.1, the essential assumption is that the process $(D_t)_{t\in[0,\infty)}$ is continuous. This assumption is not essential for the results to follow, although it is for us because we have thus far focused entirely on continuous martingales.

The following theorem is the Girsanov theorem, which describes the effect of making an absolutely continuous change of measure on pathspace. In particular, it describes how local $\mathbb{P}$-martingales are transformed into local $\mathbb{Q}$-martingales.

**Theorem 6.3.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega, \mathcal{F})$, and let $(\mathbb{P}, \mathbb{Q})$ be a Girsanov pair. Then, every continuous $\mathbb{P}$-semimartingale is a continuous $\mathbb{Q}$-semimartingale. Precisely, if $(M_t)_{t\in[0,\infty)}$ is a continuous $\mathbb{P}$-semimartingale then $(\tilde{M})_{t\in[0,\infty)}$ defined by*

$$\tilde{M}_t = M_t - \int_0^t D_s^{-1} \, \mathrm{d}\langle M, D\rangle_s,$$

*is a continuous $\mathbb{Q}$-semimartingale.*

*Proof.* We will first prove that the $\mathbb{P}$-martingale $(D_t)_{t\in[0,\infty)}$ is strictly positive $\mathbb{Q}$-almost surely. Let $T$ denote the stopping time

$$T = \inf\{s \in [0,\infty) \colon D_s = 0\}.$$

For every $N \in \mathbb{N}$ the optional stopping theorem applied to the bounded stopping times $N$ and $T \wedge N$ proves that, for the expectation $\mathbb{E}$ with respect to $\mathbb{P}$,

$$\mathbb{E}[M_N] = \mathbb{E}[M_{T\wedge N}] = \mathbb{E}[M_N \colon T > N].$$

Therefore, using the linearity of the expectation, the definition of $\mathbb{Q}$, and $\{T \le N\} \in \mathcal{F}_N$,

$$\mathbb{E}[M_N \colon T \le N] = \mathbb{Q}[\{T \le N\}] = 0.$$

Since $\{T < \infty\} = \cup_{N\in\mathbb{N}}\{T \le N\}$ we conclude that $\mathbb{Q}[\{T < \infty\}] = 0$.

We first observe that $(\tilde{M}_t)_{t\in[0,\infty)}$ is a $\mathbb{Q}$-local martingale if and only if $(\tilde{M}_t D_t)_{t\in[0,\infty)}$ is a $\mathbb{P}$-local martingale. We will therefore prove that $(\tilde{M}_t D_t)_{t\in[0,\infty)}$ is a $\mathbb{P}$-local martingale. For the stopping times $\{T_n\}_{n\in\mathbb{N}}$ defined above, we observe that $\left((D_t^{-1}\langle M,D\rangle_t)^{T_n}\right)_{t\in[0,\infty)}$ is a semimartingale as the produce of semimartingales. The integration-by-parts formula then proves that

$$\begin{aligned}(\tilde{M}D)_t^{T_n} &= M_0D_0 + \int_0^{T_n\wedge t}\tilde{M}_s\,\mathrm{d}D_s + \int_0^{T_n\wedge t} D_s\,\mathrm{d}\tilde{M}_s + \langle D,M\rangle_t \\ &= M_0D_0 + \int_0^{T_n\wedge t}\tilde{M}_s\,\mathrm{d}D_s + \int_0^{T_n\wedge t} D_s\,\mathrm{d}M_s.\end{aligned}$$

This completes the proof that $(\tilde{M}D)_t^{T_n})_{t\in[0,\infty)}$ is a $\mathbb{P}$-martingale, and therefore the proof that $(\tilde{M}_tD_t)_{t\in[0,\infty)}$ is a $\mathbb{P}$-local martingale. Hence, by the above, $(\tilde{M}_t)_{t\in[0,\infty)}$ is a $\mathbb{Q}$-local martingale. Since process of bounded variation remain processes of bounded variation with respect to an absolutely continuous chang eof measure, this completes the proof. $\square$

In the following proposition we characterize the Girsanov theorem in terms of the stochastic exponential. Precisely, we will show that the Radon-Nikodym process can be uniquely expressed as the stochastic exponential of a continuous local martingale $(L_t)_{t\in[0,\infty)}$, and we then use the process $(L_t)_{t\in[0,\infty)}$ to simplify the correction appearing due to the change of measure.

**Proposition 6.4.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,F)$, and let $(D_t)_{t\in[0,\infty)}$ be a strictly positive $\mathcal{F}_t$-martingale. Then, there exists a unique continuous local martingale $(L_t)_{t\in[0,\infty)}$ such that*

$$D_t = \mathcal{E}(L)_t.$$

*Proof.* Since the martingale $(D_t)_{t\in[0,\infty)}$ is almost surely strictly positive, Itô's formula proves that

$$\log(D_t) = \log(D_0) + \int_0^t D_s^{-1}\,\mathrm{d}D_s - \frac{1}{2}\int_0^t D_s^{-2}\,\mathrm{d}\langle D\rangle_s.$$

Let $(L_t)_{t\in[0,\infty)}$ be defined by

$$L_t = \log(D_0) + \int_0^t D_s^{-1}\,\mathrm{d}D_s,$$

for which it follows that

$$\langle L\rangle_t = \int_0^t D_s^{-2}\,\mathrm{d}\langle D\rangle_s.$$

We therefore conclude that, after exponentiating the first equation,

$$D_t = \mathcal{E}(L)_t,$$

which completes the proof. $\square$

**Proposition 6.5.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(L_t)_{t\in[0,\infty)}$ be a continuous local $\mathbb{P}$-martingale. If for every $t\in[0,\infty)$ we have $\mathrm{d}\mathbb{Q}=\mathcal{E}(L)_t\,\mathrm{d}\mathbb{P}$ on $\mathcal{F}_t$, then for every continuous $\mathbb{P}$-martingale $(M_t)_{t\in[0,\infty)}$ the process $(\tilde{M}_t)_{t\in[0,\infty)}$ defined by*

$$\tilde{M}_t = M_t - \langle M,L\rangle_t,$$

*is a continuous local $\mathbb{Q}$-martingale. Furthermore, $\mathrm{d}\mathbb{P}=\mathcal{E}(-\tilde{L})_t\,\mathrm{d}\mathbb{Q}$ on $\mathcal{F}_t$.*

*Proof.* Let $(D_t)_{t\in[0,\infty)}$ be defined by $D_t=\mathcal{E}(L)_t$. In view of the Girsanov theorem, for the first statement it remains only to prove that

$$\int_0^t D_s^{-1}\,\mathrm{d}\langle M,D\rangle_s = \langle M,L\rangle_t.$$

Since by definition, for every $t\in[0,\infty)$,

$$L_t = \log(D_0) + \int_0^t D_s^{-1}\,\mathrm{d}D_s,$$

we have

$$\langle M, L\rangle_t = \int_0^t D_s^{-1}\,\mathrm{d}\langle M, D\rangle_s.$$

For the second statement, since by definition and the first statement

$$\tilde{L} = L - \langle \tilde{L}, L\rangle_t,$$

is a $\mathbb{Q}$-local martingale with $\langle \tilde{L}\rangle_t = \langle L\rangle_t$, we have that

$$\mathcal{E}(-\tilde{L}) = \exp(-L - \frac{1}{2}\langle L\rangle_t) = \mathcal{E}(L)_t^{-1}.$$

Therefore, $\mathrm{d}\mathbb{P} = \mathcal{E}(-\tilde{L})\,\mathrm{d}\mathbb{Q}$ on $\mathcal{F}_t$ for every $t \in [0,\infty)$. This completes the proof. $\square$

The transformation from $\mathbb{P}$-local martingales from $\mathbb{Q}$-local martingales will play an important role in our construction of weak solutions to stochastic differential equations. It is particularly important that for a Girsanov pair $(\mathbb{P},\mathbb{Q})$ the Girsanov transformation maps the continuous local $\mathbb{P}$-martingales bijectively onto the continuous local $\mathbb{Q}$-martingales.

**Definition 6.6.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(\mathbb{P},\mathbb{Q})$ be a Girsanov pair. The map from continuous $\mathbb{P}$-local martingales to continuous $\mathbb{Q}$-local martingales defined by

$$(M_t)_{t\in[0,\infty)} \mapsto (\tilde{M}_t)_{t\in[0,\infty)},$$

is called the *Girsanov transformation* from $\mathbb{P}$ to $\mathbb{Q}$.

**Proposition 6.7.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(\mathbb{P},\mathbb{Q})$ be a Girsanov pair. Then, the Girsanov transformation maps the space of continuous local $\mathbb{P}$-martingales bijectively onto the space of continuous local $\mathbb{Q}$-martingales.*

*Proof.* Let $(L_t)_{t\in[0,\infty)}$ be the unique continuous local $\mathbb{P}$-martingale satisfying $\mathrm{d}\mathbb{Q} = \mathcal{E}(L)_t\,\mathrm{d}\mathbb{P}$ on $\mathcal{F}_t$ for every $t \in [0,\infty)$. Then, the Girsanov transformation from $\mathbb{P}$ to $\mathbb{Q}$ is defined by

$$\tilde{M}_t = M_t - \langle M, L\rangle_t.$$

However, since $\mathrm{d}\mathbb{P} = \mathcal{E}(-\tilde{L})_t\,\mathrm{d}\mathbb{Q}$ on $\mathcal{F}_t$ for every $t \in [0,\infty)$ the inverse transformation is defined by

$$M'_t = M_t + \langle M, \tilde{L}\rangle_t = M_t + \langle M, L\rangle_t.$$

It is then clear that every continuous local $\mathbb{P}$-martingale satisfies

$$(\tilde{M})'_t = M_t.$$

This completes the proof. $\square$

In the next proposition, we prove that the Girsanov transformation commutes with stochastic integration. That is, the Girsanov transform of the stochastic integral is the same as the stochastic integral with respect to the Girsanov transform.

**Proposition 6.8.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(\mathbb{P},\mathbb{Q})$ be a Girsanov pair. Let $(M_t)_{t\in[0,\infty)}$ be a continuous local $\mathbb{P}$-martingale and let $H \in L^2_{loc}(M)$. Then,*

$$\widetilde{\left(\int_0^t H_s\,\mathrm{d}M_s\right)}_{t\in[0,\infty)} = \left(\int_0^t H_s\,\mathrm{d}\tilde{M}_s\right)_{t\in[0,\infty)}.$$

*Proof.* By density and stopping, it suffices to prove the statement for a bounded $(H_t)_{t\in[0,\infty)}$ and for an $L^2$-bounded $(M_t)_{t\in[0,\infty)}$. In this case, if $\mathrm{d}\mathbb{Q} = \mathcal{E}(L)_t\,\mathrm{d}\mathbb{P}$ on $\mathcal{F}_t$ for a continuous local martingale $(L_t)_{t\in[0,\infty)}$, for each $t\in[0,\infty)$,

$$\widetilde{\int_0^t H_s\,\mathrm{d}M_s} = \int_0^t H_s\,\mathrm{d}M_s - \int_0^t H_s\,\mathrm{d}\langle M, L\rangle_s = \int_0^t H_s\,\mathrm{d}\tilde{M}_s.$$

This completes the proof. ☐

The following theorem describes how Brownian motion is transformed by an absolutely continuous change of measure.

**Theorem 6.9.** *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration on $(\Omega,\mathcal{F})$, and let $(\mathbb{P},\mathbb{Q})$ be a Girsanov pair. Let $(B_t)_{t\in[0,\infty)}$ be a standard $\mathbb{P}$-Brownian motion. Then the process $(\tilde{B}_t)_{t\in[0,\infty)}$ defined by*

$$\tilde{B}_t = B_t - \langle L, B\rangle_t,$$

*is a standard $\mathbb{Q}$-Brownian motion.*

*Proof.* By the Girsanov theorem $(\tilde{B}_t)_{t\in[0,\infty)}$ is a continuous $\mathbb{Q}$-local martingale. The statement is now an immediate consequence of Levy's characterization of Brownian motion, using the fact that an absolutely continuous change of measure does not change the quadratic variation. ☐

Finally, in order to construct weak solutions to stochastic differential equations, we will oftentimes seek to change the measure ourselves. For this, it is useful to have a criterion that guarantees the stochastic exponential of a continuous local martingale is again a continuous local martingale. In general, the stochastic exponential is only a supermartingale. The following two conditions are known as Kazamaki's criterion and Novikov's criterion respectively.

**Proposition 6.10.** *Let $(L_t)_{t\in[0,\infty)}$ be a continuous local martingale. If $(\exp(\frac{1}{2}L_t))_{t\in[0,\infty)}$ is a uniformly integrable submartingale then the stochastic exponential $(\mathcal{E}(L)_t)_{t\in[0,\infty)}$ is a uniformly integrable martingale.*

*Proof.* The proof will be added after the fourth problem session. ☐

**Proposition 6.11.** *If $(L_t)_{t\in[0,\infty)}$ is a continuous local martingale which satisfies*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\langle L\rangle_\infty\right)\right] < \infty,$$

*then $(\mathcal{E}(L))_{t\in[0,\infty)}$ is a uniformly integrable martingale.*

*Proof.* The proof will be added after the fourth problem session. ☐

**Corollary 6.12.** *If $(L_t)_{t\in[0,\infty)}$ is a continuous local martingale which satisfies, for every $t\in[0,\infty)$,*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\langle L\rangle_t\right)\right] < \infty,$$

*then $(\mathcal{E}(L))_{t\in[0,\infty)}$ is a martingale.*

*Proof.* We apply Novikov's criterion applied to the stopped local martingale $(L_{s\wedge t})_{s\in[0,\infty)}$, for every $t\in[0,\infty)$. This completes the proof. ☐

## 7. Stochastic Differential Equations

Stochastic differential equations model quantities subject to random noise. The noise can come from collisions on the microscopic level, such as those experienced by a particle suspended in fluid, the molecules of a perfume being diffused by air particles, or the value of a portfolio subject to fluctuations in the stock market. These fluctuations are modeled, for instance, by a $d$-dimensional Brownian motion $(B_t)_{t\in[0,\infty)}$ and a diffusion coefficient $\sigma$ taking values in the space of $(d\times d)$-matrices. You can think of the coefficient $\sigma$ as shaping then noise, by increasing or decreasing its variance in certain directions. The situation just described leads to a stochastic differential equation of the type

$$\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t,$$

where in general the noise $(B_t)_{t\in[0,\infty)}$ and the solution $(X_t)_{t\in[0,\infty)}$ can have different dimensions. In this case, $\sigma$ will not be a square matrix.

However, such quantities are also subject to deterministic effects, which describe the local mean motion. That is, on average, particles suspended in a fluid move in the direction of the current. On average, the perfume will move in the direction of the wind. And, perhaps, the average appreciation of a stock will be governed by a deterministic interest rate. For a $d$-dimensional process, such motion is described by a drift $b$ taking values in $\mathbb{R}^d$, which leads to the more general equation

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t. \tag{7.1}$$

This is the prototypical model of the kind of stochastic differential equations that we will study in this course. Most generally, we will study equations of the type, for a $d$-dimensional process $(X_t)_{t\in[0,\infty)}$, a $m$-dimensional semimartingale $(Z_t)_{t\in[0,\infty)}$, and for $f$ taking values in the space of $(d\times m)$-matrices,

$$\mathrm{d}X_t = f(X_t)\,\mathrm{d}Z_t. \tag{7.2}$$

By choosing

$$Z_t = \begin{pmatrix} t \\ B_t \end{pmatrix},$$

and by choosing

$$f(x) = (b(x), \sigma(x)),$$

it follows that (7.1) can be written in the form (7.2).

Let $(B_t)_{t\in[0,\infty)}$ be a standard $d$-dimensional Brownian motion with respect to the filtration $(\mathcal{F}^B_t)_{t\in[0,\infty)}$ defined by

$$\mathcal{F}^B_t = \sigma(B_s\colon s\in[0,T]).$$

That is, the filtration $(\mathcal{F}^B_t)_{t\in[0,\infty)}$ is the filtration generated by the Brownian motion, and carries no additional information. We will first define in which senses we judge processes $(X_t)_{t\in[0,\infty)}$ and $(Y_t)_{t\in[0,\infty)}$ to be the same. Immediately following, we define in which senses a process $(X_t)_{t\in[0,\infty)}$ can be a solution of (7.1).

**Definition 7.1.** Let $(X_t)_{t\in[0,\infty)}$ and $(Y_t)_{t\in[0,\infty)}$ be $d$-dimensional continuous stochastic processes.

(i) We say that $(X_t)_{t\in[0,\infty)}$ and $(Y_t)_{t\in[0,\infty)}$ are *indistinguishable* if

$$\mathbb{P}\left(X_t = Y_t \ \text{ for every } \ t\in[0,\infty)\right) = 1.$$

(ii) We say that $(X_t)_{t\in[0,\infty)}$ is a *modification* of $(Y_t)_{t\in[0,\infty)}$ if

$$\mathbb{P}\left(X_t = Y_t\right) = 1 \ \text{ for every } \ t\in[0,\infty)$$

(iii) We say that $(X_t)_{t\in[0,\infty)}$ and $(Y_t)_{t\in[0,\infty)}$ have the same finite-dimensional distributions if, for every $0\le t_0 < t_1 < \ldots < t_n < \infty)$ and for every measurable $A\subseteq(\mathbb{R}^d)^n$,

$$\mathbb{P}((X_{t_1},\ldots,X_{t_n})\in\mathcal{A}) = \mathbb{P}((Y_{t_1},\ldots,Y_{t_n})\in A).$$

**Definition 7.2.** Let $d, m \in \mathbb{N}$, let $\mathrm{C}([0,\infty);\mathbb{R}^d)$ denote the space of continuous functions from $[0,\infty)$ to $\mathbb{R}^d$, and let

$$b\colon \Big([0,\infty)\times \mathrm{C}([0,\infty),\mathbb{R}^d)\Big)\rightarrow \mathbb{R}^d \;\;\text{and}\;\; \sigma\colon \Big([0,\infty)\times \mathrm{C}([0,\infty);\mathbb{R}^d)\Big)\rightarrow \mathbb{R}^{d\times m},$$

be measurable functions. A solution to the stochastic differential equation, for a probability measure $\mu$ on $\mathbb{R}^d$,

$$\begin{cases} \mathrm{d}X_t = b(t,X_\cdot)\,\mathrm{d}t + \sigma(t,X_\cdot)\,\mathrm{d}B_t & \text{in } (0,\infty), \\ X_0 = \mu, \end{cases}$$

is a $m$-dimensional Brownian motion $(B_t)_{t\in[0,\infty)}$ and a $d$-dimensional continuous semimartingale $(X_t)_{t\in[0,\infty)}$ which satisfy that, for every measurable $A\subseteq\mathbb{R}^d$,

$$\mathbb{P}(X_0\in A)=\mu(A),$$

and that, almost surely,

$$X_t = X_0 + \int_0^t b(s,X_\cdot)\,\mathrm{d}s + \int_0^t \sigma(s,X_\cdot)\,\mathrm{d}B_s.$$

We say that the solutions $(X_t,B_t)_{t\in[0,\infty)}$ is a *strong solution* if $(X_t)_{t\in[0,\infty)}$ is $\mathcal{F}^B_t$-adapted. If $(X_t)_{t\in[0,\infty)}$ is not $\mathcal{F}^B_t$-adapted, then we say the the solution $(X_t,B_t)_{t\in[0,\infty)}$ is a *weak solution.*

**Remark 7.3.** Observe in particular that the above definition applies to measurable maps

$$\tilde{b}\colon \Big([0,\infty)\times\mathbb{R}^d\Big)\rightarrow\mathbb{R}^d \;\;\text{and}\;\; \tilde{\sigma}\colon \Big([0,\infty)\times\mathbb{R}^d\Big)\rightarrow\mathbb{R}^{d\times m},$$

by defining $b(s,X_\cdot)=\tilde{b}(s,X_s)$ and $\sigma(s,X_\cdot)=\tilde{\sigma}(s,X_s)$. This is certainly the most important case that will be considered in this course, and yields the stochastic differential equation

$$\mathrm{d}X_t = \tilde{b}(t,X_t)\,\mathrm{d}t + \tilde{\sigma}(t,X_t)\,\mathrm{d}B_t,$$

which is a time-dependent version of the standard type (7.1) discussed above.

The following definition extends Definition 7.2 to the case of a general semimartingale $(Z_t)_{t\in[0,\infty)}$. The definition is identical, except that in this case a strong solution is adapted to the filtration $(\mathcal{F}^Z_t)_{t\in[0,\infty)}$ generated by $(Z_t)_{t\in[0,\infty)}$.

**Definition 7.4.** Let $d, m \in \mathbb{N}$, let $\mathrm{C}([0,\infty);\mathbb{R}^d)$ denote the space of continuous functions from $[0,\infty)$ to $\mathbb{R}^d$, and let

$$f\colon \Big([0,\infty)\times \mathrm{C}([0,\infty);\mathbb{R}^d)\Big)\rightarrow\mathbb{R}^{d\times m},$$

be a measurable function. Let $(Z_t)_{t\in[0,\infty)}$ be a continuous semimartingale. A solution to the stochastic differential equation, for a probability measure $\mu$ on $\mathbb{R}^d$,

$$\begin{cases} \mathrm{d}X_t = f(t,X_\cdot)\,\mathrm{d}Z_t & \text{in } (0,\infty), \\ X_0 = \mu, \end{cases}$$

is a $d$-dimensional continuous semimartingale $(X_t)_{t\in[0,\infty)}$ which satisfies that, for every measurable $A\subseteq\mathbb{R}^d$,

$$\mathbb{P}(X_0\in A)=\mu(A),$$

and that, almost surely,

$$X_t = X_0 + \int_0^t f(s,X_\cdot)\,\mathrm{d}Zs.$$

We say that the solutions $(X_t,Z_t)_{t\in[0,\infty)}$ is a *strong solution* if $(X_t)_{t\in[0,\infty)}$ is $\mathcal{F}^Z_t$-adapted. If $(X_t)_{t\in[0,\infty)}$ is not $\mathcal{F}^Z_t$-adapted, then we say the the solution $(X_t,Z_t)_{t\in[0,\infty)}$ is a *weak solution.*

We will now define two notions of uniqueness for solutions in the sense of Definition 7.2.

**Definition 7.5.** Let $d, m \in \mathbb{N}$, let $\mathrm{C}([0,\infty);\mathbb{R}^d)$ denote the space of continuous functions from $[0,\infty)$ to $\mathbb{R}^d$, and let

$$b\colon \Big([0,\infty)\times \mathrm{C}([0,\infty),\mathbb{R}^d)\Big) \to \mathbb{R}^d \ \text{ and } \ \sigma\colon \Big([0,\infty)\times \mathrm{C}([0,\infty);\mathbb{R}^d)\Big) \to \mathbb{R}^{d\times m},$$

be measurable functions. For a probability measure $\mu$ on $\mathbb{R}^d$, consider the stochastic differential equation

$$\begin{cases} \mathrm{d}X_t = b(t, X.)\,\mathrm{d}t + \sigma(t, X.)\,\mathrm{d}B_t & \text{in } (0,\infty), \\ X_0 = \mu. \end{cases} \tag{7.3}$$

(i) We say that there is *pathwise uniqueness* for (7.3) if whenever $(X_t, B_t)_{t\in[0,\infty)}$ and $(X'_t, B'_t)_{t\in[0,\infty)}$ are two solutions defined on the same probability space which satisfy that $X_0 = X'_0$ almost surely and that $(B_t)_{t\in[0,\infty)}$ and $(B'_t)_{t\in[0,\infty)}$ are indistinguishable $m$-dimensional Brownian motions, we have that $(X_t)_{t\in[0,\infty)}$ and $(X'_t)_{t\in[0,\infty)}$ are indistinguishable.

(ii) We say that there is *uniqueness in law* for (7.3) if whenever $(X_t, B_t)_{t\in[0,\infty)}$ and $(X'_t, B'_t)_{t\in[0,\infty)}$ are two solutions (defined for possibly different $m$-dimensional Brownian motions on possibly different probability spaces) for which $X_0 = X'_0$ in distribution, we have that $(X_t)_{t\in[0,\infty)}$ and $(X'_t)_{t\in[0,\infty)}$ have the same law.

7.1. **The space of continuous paths.** Let $\mathrm{C}([0,\infty);\mathbb{R}^d)$ denote the space of continuous functions $X\colon [0,\infty)\to\mathbb{R}^d$. This space comes equipped with the metric

$$d(X,Y)=\sum_{k=1}^{\infty}2^{-k}\left(\max_{s\in[0,k]}|X_s-Y_s|\wedge 1\right), \tag{7.4}$$

which induces the topology of locally uniform convergence. This is to say that a sequence of continuous paths $\{X_n\}_{n\in\mathbb{N}}$ converges to a path $X$ if and only if for every $T\in[0,\infty)$,

$$\lim_{n\to\infty}\left(\max_{s\in[0,T]}|X^n_s-X_s|\right)=0.$$

The space $\mathrm{C}([0,\infty);\mathbb{R}^d)$ comes equipped with the metric topology induced by (7.4) and the Borel $\sigma$-algebra $\mathcal{B}(\mathrm{C}([0,\infty);\mathbb{R}^d))$ generated by the metric topology. The Borel $\sigma$-algebra comes equipped with a natural filtration $\{\mathcal{F}_t\}_{t\in[0,\infty)}$. For every $t\in[0,\infty)$ the $\sigma$-algebra $\mathcal{F}_t$ is the sigma algebra generated by continuous paths up to time $t$, which can be defined using the projection map

$$\pi_t\colon\ \mathrm{C}([0,\infty);\mathbb{R}^d)\to\mathrm{C}([0,t];\mathbb{R}^d),$$

and letting $\mathcal{F}_t$ be defined by

$$\mathcal{F}_t=\pi_t^{-1}\left(\mathcal{B}(\mathrm{C}([0,t];\mathbb{R}^d))\right). \tag{7.5}$$

More explicitly, this definition if equivalent to defining

$$\mathcal{F}_t=\{A\in\mathcal{B}(\mathrm{C}([0,\infty);\mathbb{R}^d))\colon X_\cdot\in A\ \text{ if and only if }\ X_{\cdot\wedge t}\in A\},$$

where $X_{\cdot\wedge t}$ denotes the path $X$ stopped at time $t$ (which is again an element of $C([0,\infty);\mathbb{R}^d))$).

We can therefore view $\mathrm{C}([0,\infty);\mathbb{R}^d)$ equipped with the $\sigma$-algebra $\mathcal{B}(\mathrm{C}([0,\infty);\mathbb{R}^d))$ and filtration $\{F_t\}_{t\in[0,\infty)}$ as a filtered measurable space. Now, suppose that $(\Omega,\mathcal{G},\mathbb{P})$ is a probability space with a filtration $\{\mathcal{G}_t\}_{t\in[0,\infty)}$. Suppose that $(B_t)_{t\in[0,\infty)}$ be a $\mathcal{G}_t$-Brownian motion and suppose that $(X_t)_{t\in[0,\infty)}$ is a weak solution of the stochastic differential equation, defined for bounded measurable functions $\sigma$ and $b$,

$$\mathrm{d}X_t=\sigma(X_t)\,\mathrm{d}B_t+b(X_t)\,\mathrm{d}t\ \text{ in }\ (0,\infty)\ \text{ with }\ X_0=0. \tag{7.6}$$

The solution map is a measurable function $S\colon\Omega\to\mathrm{C}([0,\infty);\mathbb{R}^d)$ that is almost surely defined by $S(\omega)=X_\cdot(\omega)$. Observe as well that, using the definition, we have $S^{-1}(\mathcal{F}_t)\subseteq\mathcal{G}_t$ for every $t\in[0,\infty)$. We can then define the pushforward measure $P=S_*\mathbb{P}$ on $\mathrm{C}([0,\infty);\mathbb{R}^d)$ by

$$P(A)=\mathbb{P}[X_t(\omega)\in A],$$

and define the probability space $(\mathrm{C}([0,\infty)),\mathcal{B}(\mathrm{C}([0,\infty);\mathbb{R}^d)),P)$ with filtration $\{\mathcal{F}_t\}_{t\in[0,\infty)}$. And we can define a $\mathcal{F}_t$-Brownian motion $(\tilde{B}_t)_{t\in[0,\infty)}$ on $(\mathrm{C}([0,\infty)),\mathcal{B}(\mathrm{C}([0,\infty);\mathbb{R}^d)),P)$ by defining

$$\tilde{B}_t(X_\cdot)=\begin{cases}0 & \text{if }\ S_B^{-1}(X_\cdot)=\emptyset,\\ B_t(\omega) & \text{if }\ \omega\in S_B^{-1}(X_\cdot),\end{cases}$$

where $S_B\colon\Omega\to\mathrm{C}([0,\infty);\mathbb{R}^d)$ denotes the measurable map $S_B(\omega)=B_\cdot(\omega)$. Then, with respect to $P$, paths $X_\cdot\in\mathrm{C}([0,\infty);\mathbb{R}^d)$ almost surely satisfy, for every $t\in[0,\infty)$,

$$X_t=\int_0^t b(X_s)\,\mathrm{d}s+\int_0^t\sigma(X_s)\,\mathrm{d}\tilde{B}_s(X_\cdot). \tag{7.7}$$

We can therefore view a weak solution of (7.6) as a measure $P$ on $(\mathrm{C}([0,\infty)),\mathcal{B}(\mathrm{C}([0,\infty);\mathbb{R}^d))$ and an $\mathcal{F}_t$-Brownian motion $(\tilde{B}_t)_{t\in[0,\infty)}$ on $(\mathrm{C}([0,\infty)),\mathcal{B}(\mathrm{C}([0,\infty);\mathbb{R}^d),P)$ such that $P$-almost every path satisfies (7.7).

This formulation leads to the so-called martingale formulation of the equation (7.6). Restricting to the case of one-dimension, if $f$ is twice-continuously differentiable with bounded derivatives up to order two it follows from Itô's formula that

$$f(X_t) - f(X_0) - \int_0^t \left( b(X_s)f'(X_s) + \frac{1}{2}\sigma^2(X_s)f''(X_s) \right) \mathrm{d}s = \int_0^t \sigma(X_s)f'(X_s)\,\mathrm{d}\tilde{B}_s.$$

That is, the process

$$\left( f(X_t) - f(X_0) - \int_0^t \left( b(X_s)f'(X_s) + \frac{1}{2}\sigma^2(X_s)f''(X_s) \right) \mathrm{d}s \right)_{t\in[0,\infty)} \tag{7.8}$$

is a $P$-martingale. This is the so-called martingale problem associated to the coefficients $(b, \sigma^2)$. A solution to the martingale problem with initial data defined by a probability measure $\mu$ on $\mathbb{R}$ is a measure $P$ on $(\mathrm{C}([0,\infty)), \mathcal{B}(\mathrm{C}([0,\infty);\mathbb{R}^d))$ such that (i) for every measurable $A \subseteq \mathbb{R}$ we have $P[X_0 \in A] = \mu$ and (ii) the process (7.8) is a $P$-martingale for every twice-differentiable function $f$ with bounded derivatives up to order two. While we will not develop this approach in these notes, do keep it in mind for the future. The martingale problem provides a very general and powerful means for proving the existence of weak solutions to (7.6) and for proving that the solutions are unique in law. A full account of this approach can be found in the book *Multidimensional Diffusion Processes* by Stroock and Varadhan.

In the following section, we will use the following important proposition concerning measures on the space $(\mathrm{C}([0,\infty)), \mathcal{B}(\mathrm{C}([0,\infty);\mathbb{R}^d))$ equipped with the filtration $(\mathcal{F}_t)_{t\in[0,\infty)}$.

**Proposition 7.6.** *Let* $(\mathrm{C}([0,\infty)), \mathcal{B}(\mathrm{C}([0,\infty);\mathbb{R}^d))$ *be the space of continuous paths equipped with its Borel sigma algebra and let* $(\mathcal{F}_t)_{t\in[0,\infty)}$ *be the filtration* (7.5)*. Assume that for ever* $t \in [0,\infty)$ *there exists a probability measure* $Q_t$ *on the measurable space* $(\mathrm{C}([0,\infty);\mathbb{R}^d), \mathcal{F}_t)$*, and assume that the probability measures satisfy the compatibility condition, for every* $s \leq t \in [0,\infty)$*,*

$$Q_t|_{\mathcal{F}_s} = Q_s \;\; (\textit{that is, } \; Q_t(A) = Q_s(A) \; \textit{ for every } \; A \in \mathcal{F}_s).$$

*Then there exists a probability measure* $Q$ *on* $(\mathrm{C}([0,\infty)), \mathcal{B}(\mathrm{C}([0,\infty);\mathbb{R}^d))$ *such that, for every* $t \in [0,\infty)$*,*

$$Q|_{\mathcal{F}_t} = Q_t.$$

*Proof.* The statement can be found in Appendix 6 of Revuz and Yor, and the proof appears as special case of Theorem 1.3.5 in Stroock and Varadhan. $\square$

7.2. **Existence of weak solutions by change of measure and time.** In this section, we will construct weak solutions to stochastic differential equations of the form

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t, \tag{7.9}$$

using the Girsanov theorem and change of measure and using random time change. Precisely, given a one-dimensional Brownian motion $(B_t)_{t\in[0,\infty)}$ and a positive bounded function $\sigma\colon \mathbb{R}\to\mathbb{R}$ we will show that there exists a time change $(\tau_t)_{t\in[0,\infty)}$ and a new Brownian motion $(\tilde{B}_t)_{t\in[0,\infty)}$ such that the process $(X_t)_{t\in[0,\infty)}$ defined by

$$X_t = B_{\tau_t},$$

is a solution to the equation

$$\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}\tilde{B}_t. \tag{7.10}$$

Then, given a bounded measurable function $b\colon \mathbb{R}\to\mathbb{R}$, we will prove using the Girsanov theorem that there exists a new Brownian motion $(B'_t)_{t\in[0,\infty)}$ defined with respect to a new measure such that, with respect to this measure, the process $(X_t)_{t\in[0,\infty)}$ is a solution to the equation

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B'_t.$$

That is, beginning from a Brownian motion $(B_t)_{t\in[0,\infty)}$, we can construct weak solutions to (7.9) for a general class of coefficients. In the sections below, we will establish conditions that imply pathwise uniqueness holds for (7.9) and therefore using the theorem below that the weak solutions are in fact strong solutions.

**Theorem 7.7.** *Let $(B_t)_{t\in[0,\infty)}$ be a standard $n$-dimensional Brownian motion. Let $\sigma\colon\ \mathrm{C}([0,\infty);\mathbb{R}^d)\times[0,\infty)\to\mathbb{R}^{d\times n}$ and $b\colon\ \mathrm{C}([0,\infty),\mathbb{R}^d)\times[0,\infty)\to\mathbb{R}^d$ be measurable predictable functions. If pathwise uniqueness holds for the equation*

$$\begin{cases} \mathrm{d}X_t = b(X_\cdot,t)\,\mathrm{d}t + \sigma(X_\cdot,t)\,\mathrm{d}B_t & \text{in } (0,\infty),\\ X_0 = x, \end{cases} \tag{7.11}$$

*then uniqueness in law holds for* (7.11) *and every weak solution is a strong solution.*

We will now prove that above outline. In the first statement, we essentially use a random time change to transform a Brownian motion into a solution to a more general diffusion equation of the type (7.10).

**Proposition 7.8.** *Let $(B_t)_{t\in[0,\infty)}$ be a standard Brownian motion. Let $\sigma\colon\mathbb{R}\to\mathbb{R}$ be a bounded measurable function which satisfies, for some $\varepsilon\in(0,\infty)$,*

$$0<\varepsilon\leq\sigma.$$

*Then, for each $x\in\mathbb{R}$, the equation*

$$\begin{cases} \mathrm{d}X_t = \sigma^{\frac{1}{2}}(X_t)\,\mathrm{d}B_t & \text{in } (0,\infty),\\ X_0 = x \end{cases}$$

*has a weak solution and satisfies uniqueness in law.*

*Proof.* We assume without loss of generality that $x=0$. Since $\sigma$ is bounded and strictly positive, define the family of time changes $(\tau_t)_{t\in[0,\infty)}$ by

$$\tau_t = \inf\left\{s\in[0,\infty)\colon\ \int_0^s \sigma^{-1}(B_r)\,\mathrm{d}r \geq t.\right\}.$$

The boundedness and positivity of $\sigma$ proves that the process

$$t\in[0,\infty)\mapsto\int_0^{\tau_t}\sigma^{-\frac{1}{2}}(B_s)\,\mathrm{d}B_s,$$

is a local martingale. Furthermore, it follows by definition of $(\tau_t)_{t\in[0,\infty)}$ that the quadratic variation process satisfies

$$\langle\int_0^{\tau_\cdot} \sigma^{-\frac{1}{2}}(B_s)\,\mathrm{d}B_s\rangle_t = \int_0^{\tau_t} \sigma^{-1}(B_s)\,\mathrm{d}s = t.$$

Therefore, by Levy's characterization of Brownian motion, there exists a Brownian motion $(\tilde{B}_t)_{t\in[0,\infty)}$ such that

$$\tilde{B}_t = \int_0^{\tau_t} \sigma^{-\frac{1}{2}}(B_s)\,\mathrm{d}B_s.$$

Or, equivalently, that

$$\mathrm{d}\tilde{B}_t = \sigma^{-\frac{1}{2}}(B_{\tau_t})\,\mathrm{d}B_{\tau_t},$$

which implies that

$$\mathrm{d}B_{\tau_t} = \sigma^{\frac{1}{2}}(B_{\tau_t})\,\mathrm{d}\tilde{B}_t.$$

Hence, after defining the process $(X_t)_{t\in[0,\infty)}$ by $X_t = B_{\tau_t}$ we conclude that $(X_t, \tilde{B}_t)_{t\in[0,\infty)}$ is a solution to the equation

$$\begin{cases} \mathrm{d}X_t = \sigma^{\frac{1}{2}}(X_t)\,\mathrm{d}\tilde{B}_t \quad \text{in } (0,\infty), \\ X_0 = 0. \end{cases} \tag{7.12}$$

For the converse, the positivity of $\sigma$ proves that, as $t \to \infty$,

$$\langle X\rangle_t = \int_0^t \sigma(X_s)\,\mathrm{d}s \to \infty.$$

Therefore, by the Dambis-Dubins-Schwarz theorem, every solution to (7.12) is a time-changed Brownian motion. This completes the proof. □

In the next two propositions, we will use the Girsanov theorem to solve a general diffusion equation with drift. This will be achieved by changing the underlying measure. We assume without loss of generality that $(\Omega, \mathcal{F}, \mathbb{P})$ is the probability space $(\mathrm{C}([0,\infty)), \mathcal{B}(\mathrm{C}([0,\infty)), P)$ constructed in Section 7.1.

**Proposition 7.9.** *Let $(B_t)_{t\in[0,\infty)}$ be a standard $\mathcal{F}_t$-Brownian motion with respect to some filtration $(\mathcal{F}_t)_{t\in[0,\infty)}$. Let $\sigma\colon \mathbb{R}\to\mathbb{R}$ and $b\colon \mathbb{R}\to\mathbb{R}$ be bounded measurable functions. Then for each $x\in\mathbb{R}$ there exist weak solutions to the equation*

$$\begin{cases} \mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t \quad \text{in } (0,\infty), \\ X_0 = x \end{cases} \tag{7.13}$$

*if and only if there exist weak solutions to the equation*

$$\begin{cases} \mathrm{d}X_t = \sigma(X_t)b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t \quad \text{in } (0,\infty), \\ X_0 = x \end{cases} \tag{7.14}$$

*Furthermore, equation (7.13) satisfies uniqueness in law if and only if (7.14) satisfies uniqueness in law.*

*Proof.* Without loss of generality we assume that $x = 0$. Suppose that $(X_t)_{t\in[0,\infty)}$ is a weak solution of (7.13). Since $b$ is bounded, define the martingale $(L_t)_{t\in[0,\infty)}$ by

$$L_t = \int_0^t b(X_s)\,\mathrm{d}B_s$$

and using Proposition 7.6 we define the measure $\mathbb{Q}$ by $\mathrm{d}\mathbb{Q}|_{\mathcal{F}_t} = \mathcal{E}(L)_t\,\mathrm{d}\mathbb{P}|_{\mathcal{F}_t}$. The Girsanov theorem proves that the process $(\tilde{B}_t)_{t\in[0,\infty)}$ defined by

$$\tilde{B}_t = B_t - \langle L, B\rangle_t = B_t - \int_0^t b(X_s)\,\mathrm{d}s,$$

is a $\mathbb{Q}$-Brownian motion. And furthermore, since $(X_t)_{t\in[0,\infty)}$ solves (7.13), we have that

$$\begin{aligned}\mathrm{d}X_t &= \sigma(X_t)\,\mathrm{d}B_t \\ &= \sigma(X_t)\,\mathrm{d}B_t - \sigma(X_t)b(X_t)\,\mathrm{d}t + \sigma(X_t)b(X_t)\,\mathrm{d}t \\ &= \sigma(X_t)\,\mathrm{d}\tilde{B}_t + \sigma(X_t)b(X_t)\,\mathrm{d}t.\end{aligned}$$

That is, with respect to the measure $\mathbb{Q}$ we have that $(X_t, \tilde{B}_t)_{t\in[0,\infty)}$ is a solution of (7.14). For the reverse direction, if a process $(X_t)_{t\in[0,\infty)}$ is a solution of (7.14) with respect to the Brownian motion $(B_t)_{t\in[0,\infty)}$ we invert the previous transformation by replacing the process $(L_t)_{t\in[0,\infty)}$ by the process $(-L_t)_{t\in[0,\infty)}$. This completes the proof. $\square$

**Remark 7.10.** We emphasize that the measure $\mathbb{Q}$ constructed in the proof is in general singular with respect to $\mathbb{P}$, despite the fact that $\mathbb{Q}|_{\mathcal{F}_t}$ is mutually absolutely continuous with respect to $\mathbb{P}|_{\mathcal{F}_t}$ for every $t \in [0,\infty)$. Indeed, the Radon-Nikodym derivative is $\mathcal{E}(L)_t$. To see that $\mathbb{Q}$ can be singular, consider the two equations

$$\mathrm{d}X_t^1 = \mathrm{d}B_t \;\text{ and }\; \mathrm{d}X_t^2 = \mathrm{d}B_t + \mathrm{d}t \;\text{ with }\; X_0^1 = X_0^2 = 0,$$

for which we have $X_t^1 = B_t$ and $X_t^2 = B_t + t$.

By the Girsanov theorem, the laws of the processes $X^1$ and $X^2$ on $\mathrm{C}([0,T];\mathbb{R})$ are mutually absolutely continuous for every $T \in [0,\infty)$. However, if $Q^1$ denotes the pushforward measure of $X^1$ on $\mathrm{C}([0,\infty);\mathbb{R})$ and $Q^2$ the pushforward measure of $X^2$, then

$$0 = Q^1[\limsup_{t\to\infty} {}^{X_t}\!/_t = 1] \neq Q^2[\limsup_{t\to\infty} {}^{X_t}\!/_t = 1] = 1.$$

That is, $Q^1$ and $Q^2$ are singular measures on $\mathrm{C}([0,\infty);\mathbb{R})$. We should expect this, since the solution $X_t^2$ is almost surely running to infinity as $t\to\infty$ but the solution $X_t^1$ is a Brownian motion, and hence returns to zero infinity often as $t\to\infty$. The interesting statement is that on every finite interval $[0,T]$, the laws of the two solutions are mutually absolutely continuous.

**Proposition 7.11.** *Let $(B_t)_{t\in[0,\infty)}$ be a standard Brownian motion. Let $\sigma\colon \mathbb{R}\to\mathbb{R}$ and $b\colon\mathbb{R}\to\mathbb{R}$ be bounded measurable functions and assume that $\sigma$ satisfies, for some $\varepsilon\in(0,\infty)$,*

$$0 < \varepsilon \leq \sigma.$$

*Then for each $x\in\mathbb{R}$ there exist weak solutions to the equation*

$$\begin{cases} \mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t & \text{in } (0,\infty), \\ X_0 = x \end{cases} \tag{7.15}$$

*if and only if there exist weak solutions to the equation*

$$\begin{cases} \mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t & \text{in } (0,\infty), \\ X_0 = x \end{cases} \tag{7.16}$$

*Furthermore, equation* (7.15) *satisfies uniqueness in law if and only if* (7.16) *satisfies uniqueness in law.*

*Proof.* The proof is an immediate consequence of Proposition 7.9 applied to the drift

$$\tilde{b} = \sigma^{-1} b,$$

where the positivity of $\sigma$ is used to ensure that this function is bounded. This completes the proof. $\square$

We are now prepared to present the main theorem of this section, which proves that for strictly positive diffusion coefficients we can always find weak solutions to equations of the form (7.9). And furthermore, such equations satisfy uniqueness in law.

**Theorem 7.12.** *Let $(B_t)_{t\in[0,\infty)}$ be a standard Brownian motion. Let $\sigma\colon\mathbb{R}\to\mathbb{R}$ and $b\colon\mathbb{R}\to\mathbb{R}$ be bounded measurable functions and assume that $\sigma$ satisfies, for some $\varepsilon\in(0,\infty)$,*

$$0<\varepsilon\leq\sigma.$$

*Then for each $x\in\mathbb{R}$ there exist a weak solution to the equation*

$$\begin{cases} \mathrm{d}X_t = b(X_t)\,\mathrm{d}t+\sigma(X_t)\,\mathrm{d}B_t \quad \textit{in } (0,\infty), \\ X_0 = x \end{cases} \tag{7.17}$$

*Furthermore, equation* (7.17) *satisfies uniqueness in law.*

*Proof.* The proof is an immediate consequence of Proposition 7.8 and Proposition 7.11. That is, assuming without loss of generality that $x=0$, given a Brownian motion $(B_t)_{t\in[0,\infty)}$ we first define the time change $(\tau_t)_{t\in[0,\infty)}$ by

$$\tau_t=\inf\left\{s\in[0,\infty)\colon\ \int_0^t\sigma^{-2}(B_s)\,\mathrm{d}s=1\right\}.$$

Then by Proposition 7.8 there exists a Brownian motion $(\tilde{B}_t)_{t\in[0,\infty)}$ such that the process $(X_t)_{t\in[0,\infty)}$ defined by $X_t=B_{\tau_t}$ satisfies $X_0=0$ and is a weak solution of

$$\mathrm{d}X_t=\sigma(X_t)\,\mathrm{d}\tilde{B}_t.$$

Then we define the martingale $(L_t)_{t\in[0,\infty)}$ by

$$L_t=\int_0^t\sigma^{-1}(X_s)b(X_s)\,\mathrm{d}\tilde{B}_s,$$

and the corresponding measure $\mathbb{Q}$ by $\mathrm{d}\mathbb{Q}_{\mathcal{F}_t}=\mathcal{E}(L)_t\,\mathrm{d}\mathbb{P}_{\mathcal{F}_t}$. It then follows from the Girsanov theorem and Proposition 7.11 that the process $(B'_t)_{t\in[0,\infty)}$ defined by

$$B'_t=\tilde{B}_t-\int_0^t\sigma^{-1}(X_s)b(X_s)\,\mathrm{d}s,$$

is a $\mathbb{Q}$-Brownian motion and that $(X_t,B'_t)_{t\in[0,\infty)}$ satisfies $X_0=0$ and is a weak solution to

$$\mathrm{d}X_t=\sigma(X_t)\,\mathrm{d}B'_t+b(X_t)\,\mathrm{d}t.$$

Since Brownian motion satisfies uniqueness in law, we conclude that (7.17) satisfies uniqueness in law. □

**Remark 7.13.** We emphasize that the solutions constructed in Theorem 7.12 are in general weak solutions, but not strong solutions, despite the fact that we started with a strong solution. That is, despite the fact that we started with a Brownian motion. The reason for this is that the time-change introduces a new filtration, and the application of the Girsanov theorem changes the underlying measure and thereby introduces a new Brownian motion.

**Remark 7.14.** The conclusion of Theorem 7.12 remains true in higher dimensions, however in this case the assumption that $\sigma$ is positive is replaced by the assumption that the matrix $\sigma\sigma^t$ is uniformly elliptic. This is to say that, for some constants $\lambda\leq\Lambda\in(0,\infty)$, as symmetric matrices,

$$\lambda I\leq\sigma\sigma^t\leq\Lambda I.$$

Or, in terms of eigenvalues, this condition states that the eigenvalues of $\sigma\sigma^t$ are bounded below by $\lambda$ and above by $\Lambda$. The proof in the higher dimensional case is significantly more difficult.

7.3. **Existence and uniqueness: Lipschitz continuous coefficients.** In this section, we will prove pathwise uniqueness for solutions of the stochastic differential equation

$$\mathrm{d}X_t = f(t, X.)\,\mathrm{d}Z_t,$$

for a continuous semimartingale $(Z_t)_{t\in[0,\infty)}$, and a for a Lipschitz continuous function $f$ on the space of continuous paths. We recall that Lipschitz continuity is also the condition required to prove existence and uniqueness for ordinary differential equations of the type

$$\mathrm{d}X_t = f(X_t)\,\mathrm{d}t.$$

Indeed, the equation

$$\mathrm{d}X_t = \sqrt{X_t}\,\mathrm{d}t \;\text{ with }\; X_0 = 0,$$

has infinitely many solutions. In the sections to follow, we will prove that, at least for one-dimensional equations, the Lipschitz continuity can be relaxed for the case of Brownian motion due to the regularizing effect of its quadratic variation. We will present the proof of existence and uniqueness immediately following Grönwall's inequality.

**Proposition 7.15.** *Let $\phi\colon [0,\infty) \to [0,\infty)$ be a nonnegative, locally bounded function which satisfies, for some $a, b \in [0,\infty)$, for every $t \in [0,\infty)$,*

$$\phi(t) \leq a + b\int_0^t \phi(s)\,\mathrm{d}s.$$

*Then, for every $t \in [0,\infty)$,*

$$\phi(t) \leq a\exp(bt).$$

*Proof.* Let $t \in [0,\infty)$. By assumption,

$$\phi(t) \leq a + b\int_0^t \phi(s)\,\mathrm{d}s \leq a + b\int_0^t \left(a + \int_0^s \phi(r)\,\mathrm{d}r\right).$$

Therefore,

$$\phi(t) \leq a + abt + \int_0^t\int_0^s \phi(r)\,\mathrm{d}r\,\mathrm{d}s.$$

Proceeding inductively, it follows for every $n \in \mathbb{N}$ that

$$\phi(t) \leq \sum_{k=0}^n \frac{a(bt)^k}{k!} + b^{n+1}\int_0^t\int_0^{t_1}\dots\int_0^{t_n} \phi(t_{n+1})\,\mathrm{d}t_{n+1}\,\mathrm{d}t_n\dots\,\mathrm{d}t_1,$$

where the local boundedness of $\phi$ proves that, for some $c \in (0,\infty)$ depending on $t \in [0,\infty)$,

$$b^{n+1}\int_0^t\int_0^{t_1}\dots\int_0^{t_n} \phi(t_{n+1})\,\mathrm{d}t_{n+1}\,\mathrm{d}t_n\dots\,\mathrm{d}t_1 \leq \frac{cb^{n+1}t^{n+1}}{(n+1)!}.$$

We therefore conclude that, after passing to the limit $n \to \infty$,

$$\phi(t) \leq \sum_{k=0}^\infty \frac{a(bt)^k}{k!} = a\exp(bt),$$

which completes the proof. □

**Theorem 7.16.** *Let $m, d \in \mathbb{N}$. Let $(Z_t)_{t\in[0,\infty)}$ be a $m$-dimensional continuous $\mathcal{F}_t$-semimartingale. Let*

$$f\colon \Big([0,\infty) \times \mathrm{C}([0,\infty), \mathbb{R}^d)\Big) \to \mathbb{R}^{d\times m},$$

*be a locally bounded measurable function that is Lipschitz continuous in the sense that, for some $K \in (0,\infty)$,*

$$|f(t, X.) - f(t, Y.)| \leq K\sup_{s\in[0,t]}|X_s - Y_s|,$$

*for every $t \in [0,\infty)$ and $X, Y \in \mathrm{C}([0,\infty);\mathbb{R}^d)$. For every $x \in \mathbb{R}^d$, here exists a unique up to being indistinguishable $\mathcal{F}^Z_t$-adapted process $(X_t)_{t\in[0,\infty)}$ such that*

$$X_t = x + \int_0^t f(s, X_\cdot)\,\mathrm{d}Z_s.$$

*Proof.* We will consider the case $m = d = 1$. The additional difficulties in the general setting are only notational. The proof follows by Picard iteration. Let $(X^0_t)_{t\in[0,\infty)}$ denote the process $X^0_t = x$ for every $t \in [0,\infty)$. Then, for every $n \in \mathbb{N}$, define inductively the process $(X^n_t)_{t\in[0,\infty)}$ by

$$X^n_t = x + \int_0^t f(s, X^{n-1}_\cdot)\,\mathrm{d}Z_s.$$

Let $(Z_t = M_t + A_t)_{t\in[0,\infty)}$ denote the semimartingale decomposition of $(Z_t)_{t\in[0,\infty)}$, where $(A_t)_{t\in[0,\infty)}$ is a process of bounded variation and $(M_t)_{t\in[0,\infty)}$ is a continuous local martingale. We will first consider the case that the measures $\mathrm{d}\langle M\rangle_t$ and $|\,\mathrm{d}A_t|$ are dominated by the Lebesgue measure. The general case will then follow using a time-change argument.

Fix $t \in [0,\infty)$. We first observe that, if $n = 1$, for each $s \in [0,t]$,

$$\left|X^1_s - X^0_s\right| \leq 2\left(\left|\int_0^s f(r,x)\,\mathrm{d}M_r\right|^2 + \left|\int_0^s f(r,x)\,\mathrm{d}A_r\right|^2\right),$$

and, if $n \in \{2,3,\ldots\}$, for each $s \in [0,t]$,

$$\begin{aligned}\left|X^n_s - X^{n-1}_s\right|^2 \leq{}& 2\left|\int_0^t \left(f(r,X^{n-1}_\cdot) - f(r,X^{n-2}_\cdot)\right)\,\mathrm{d}M_r\right|^2 \\ &+ 2\left|\int_0^t \left(f(r,X^{n-1}_\cdot) - f(r,X^{n-2}_\cdot)\right)\,\mathrm{d}A_r\right|^2.\end{aligned}$$

It therefore follows that, for each $n \in \{2,3,\ldots\}$,

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[0,t]}\left|X^n_t - X^{n-1}_t\right|^2\right] \leq{}& 2\mathbb{E}\left[\sup_{s\in[0,t]}\left|\int_0^r \left(f(r,X^{n-1}_\cdot) - f(r,X^{n-2}_\cdot)\right)\,\mathrm{d}M_r\right|^2\right] \\ &+ 2\mathbb{E}\left[\sup_{s\in[0,t]}\left|\int_0^s \left(f(r,X^{n-1}_\cdot) - f(r,X^{n-2}_\cdot)\right)\,\mathrm{d}A_r\right|^2\right].\end{aligned}$$

The Burkholder-David-Gundy inquality with $p = 2$ (or, really, Doob's inequality in this case) and Hölder's inequality prove that

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[0,t]}\left|X^n_s - X^{n-1}_s\right|^2\right] \leq{}& 8\mathbb{E}\left[\int_0^t \left(f(s,X^{n-1}_\cdot) - f(s,X^{n-2}_\cdot)\right)^2\,\mathrm{d}\langle M\rangle_s\right] \\ &+ 2t\mathbb{E}\left[\int_0^t \left(f(s,X^{n-1}_\cdot) - f(s,X^{n-2}_\cdot)\right)^2\,|\,\mathrm{d}A_s|\right].\end{aligned}$$

It then follows from the Lipschitz condition and the assumption that $\mathrm{d}\langle M\rangle_t \leq \mathrm{d}t$ and $|\,\mathrm{d}A_t| \leq dt$ that

$$\mathbb{E}\left[\sup_{s\in[0,t]}\left|X^n_s - X^{n-1}_s\right|^2\right] \leq 2K^2(4+t)\int_0^t \mathbb{E}\left[\sup_{r\in[0,s]}\left|X^{n-1}_s - X^{n-2}_s\right|^2\right]\,\mathrm{d}s.$$

Applying this inequality inductively, using the fact that $t \in [0,\infty)$ was arbitrary, we conclude using the definition of $(X_t^0)_{t\in[0,\infty)}$ that, for every $n \in \mathbb{N}$,

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[0,t]}\left|X_s^n - X_s^{n-1}\right|^2\right] &\leq \left(2K^2(4+t)\right)^n \int_0^t \int_0^{t_1} \cdots \int_0^{t_n} \mathbb{E}\left[\sup_{r\in[0,s]}\left|X_r^1 - X_r^0\right|^2\right] \mathrm{d}s \\ &\leq \frac{\left(2K^2(4+t)\right)^n t^n}{n!}\mathbb{E}\left[\sup_{s\in[0,t]}\left|X_s^1 - x\right|^2\right].\end{aligned}$$

Since for every $t \in [0,\infty)$ the local boundedness of $f$, the definition of $(X_t^1)_{t\in[0,\infty)}$, and the fact that $\mathrm{d}\langle M\rangle_t < \infty$ imply that

$$\mathbb{E}\left[\sup_{s\in[0,t]}\left|X_s^1 - x\right|^2\right] < \infty,$$

we conclude that

$$\sum_{n=1}^{\infty}\mathbb{E}\left[\sup_{s\in[0,t]}\left|X_s^n - X_s^{n-1}\right|^2\right] < \infty.$$

Therefore, almost surely, we have that

$$\sum_{n=1}^{\infty}\sup_{s\in[0,t]}\left|X_s^n - X_t^n\right| < \infty,$$

which implies that there exists a $\mathcal{F}_t^Z$-measurable continuous process $(X_t)_{t\in[0,\infty)}$, which is continuous since it is the uniform limit of continuous processes, and which is $\mathcal{F}_t^Z$-measurable since each of the $(X_t^n)_{t\in[0,\infty)}$ are $\mathcal{F}_t^Z$-measurable, such that, almost surely, for every $t \in [0,\infty)$, as $n \to \infty$,

$$X_\cdot^n \to X_\cdot \ \text{ uniformly on } \ [0,t]. \tag{7.18}$$

Since a repetition of the above estimates prove that

$$\lim_{n\to\infty}\int_0^t f(s, X_\cdot^n)\,\mathrm{d}Z_s = \int_0^t f(s, X_\cdot)\,\mathrm{d}Z_s, \tag{7.19}$$

it follows from (7.18) and (7.19) that $(X_t)_{t\in[0,\infty)}$ is a solution of

$$\begin{cases}\mathrm{d}X_t = f(t, X_\cdot)\,\mathrm{d}Z_t & \text{in } \ (0,\infty), \\ X_0 = x\end{cases} \tag{7.20}$$

To prove uniqueness, suppose that $(X_t)_{t\in[0,\infty)}$ and $(Y_t)_{t\in[0,\infty)}$ are two solutions of (7.20). For every $n \in \mathbb{N}$ let $T_n$ denote the $\mathcal{F}_t^Z$-stopping time

$$T_n = \inf\{t \in [0,\infty)\colon \ |X_t| \geq n \ \text{ or } \ |Y_t| \geq n\}.$$

It then follows that the stopped processes $(X_t^{T_n})_{t\in[0,\infty)}$ and $(Y_t^{T_n})_{t\in[0,\infty)}$ are solutions of (7.20) for the semimartingale $(Z_t^{T_n})_{t\in[0,\infty)}$. It then follows by a repetition of the above estimates that

$$\mathbb{E}\left[\sup_{s\in[0,t]}\left|X_s^{T_n} - Y_s^{T_n}\right|^2\right] \leq 2K^2(4+t)\int_0^t \mathbb{E}\left[\sup_{s\in[0,r]}\left|X_r^{T_n} - Y_r^{T_n}\right|^2\right] \mathrm{d}r.$$

Grönwall's inequality therefore proves that, for every $t \in [0,\infty)$ and $n \in \mathbb{N}$,

$$\mathbb{E}\left[\sup_{s\in[0,t]}\left|X_s^{T_n} - Y_s^{T_n}\right|^2\right] = 0.$$

By letting $t \to \infty$ and $n \to \infty$, the monotone convergence theorem proves that

$$\mathbb{E}\left[\sup_{s\in[0,\infty)} |X_s - Y_s|^2\right] = 0,$$

which proves that $(X_t)_{t\in[0,\infty)}$ and $(Y_t)_{t\in[0,\infty)}$ are indistinguishable.

It remains to treat the general case. For this, define the time change $\{\mathrm{C}_t\}_{t\in[0,\infty)}$ by

$$C_t = \inf\left\{s \in [0,\infty) \colon t + \int_0^s |\,\mathrm{d}A_r| + \langle M\rangle_t \geq t.\right\}$$

Consider the time changed process $(\tilde{Z}_t)_{t\in[0,\infty)}$ by

$$\tilde{Z}_t = Z_{C_t},$$

and define the function

$$\tilde{f}\colon\ ([0,\infty) \times \mathrm{C}([0,\infty);\mathbb{R})) \to \mathbb{R}$$

by $\tilde{f}(s, X.) = f(C_s, X.)$. Since it follows by definition that $C_t \leq t$ for every $t \in [0,\infty)$, we have by definition that

$$\left|\tilde{f}(t,X.) - \tilde{f}(t,Y.)\right| = |f(C_t,X.) - f(C_t,Y.)| \leq K \sup_{s\in[0,C_t]} |X_s - Y_s| \leq K \sup_{s\in[0,t]} |X_s - Y_s|\,,$$

and that $\tilde{f}$ is locally bounded. Therefore, since $(\tilde{Z}_t = \tilde{A}_t + \tilde{M}_t)_{t\in[0,\infty)}$ for $(\tilde{A}_t = A_{C_t})_{t\in[0,\infty)}$ and $(\tilde{M}_t = M_{C_t})_{t\in[0,\infty)}$ is a semimartingale satisfying by definition of the time change $(C_t)_{t\in[0,\infty)}$ that

$$\mathrm{d}\langle\tilde{M}\rangle_t \leq\ \mathrm{d}t\ \text{ and }\ \left|\mathrm{d}\tilde{A}_t\right| \leq dt,$$

there exists a unique solution $(\tilde{X}_t)_{t\in[0,\infty)}$ to the stochastic differential equation

$$\begin{cases} \mathrm{d}\tilde{X}_t = \tilde{f}(t,\tilde{X}.)\,\mathrm{d}\tilde{Z}_t & \text{in } (0,\infty), \\ \tilde{X}_0 = x. \end{cases}$$

Define the inverse process $(A_t)_{t\in[0,\infty)}$ by

$$A_t = \inf\{s \in [0,\infty) \colon C_s = t\}.$$

It then follows by the change of variables formula that the process $(X_t)_{t\in[0,\infty)}$ defined by

$$X_t = \tilde{X}_{A_t},$$

is a solution of the stochastic differential equation

$$\begin{cases} \mathrm{d}X_t = f(t,X.)\,\mathrm{d}Z_t & \text{in } (0,\infty), \\ X_0 = x, \end{cases}$$

This completes the proof. □

The following theorem proves that the solution can be constructed to depend continuously on the initial condition as well as time under the additional assumption that the coefficient $f$ is globally bounded.

**Theorem 7.17.** *Let $m, d \in \mathbb{N}$. Let $(Z_t)_{t\in[0,\infty)}$ be a $m$-dimensional continuous $\mathcal{F}_t$-semimartingale. Let*

$$f\colon\ \left([0,\infty) \times \mathrm{C}([0,\infty),\mathbb{R}^d)\right) \to \mathbb{R}^{d\times m},$$

*be a bounded measurable function that is Lipschitz continuous in the sense that, for some $K \in (0,\infty)$,*

$$|f(t,X.) - f(t,Y.)| \leq K \sup_{s\in[0,t]} |X_s - Y_s|\,,$$

*for every* $t \in [0,\infty)$ *and* $X, Y \in \mathrm{C}([0,\infty);\mathbb{R}^d)$. *Then there exists a unique* $\mathcal{F}_t^Z$*-adapted process* $\{X_t^x\}_{t\in[0,\infty),x\in\mathbb{R}^d}$ *that is continuous is both variables* $(x,t) \in \mathbb{R}^d \times [0,\infty)$ *such that almost surely, for every* $x \in \mathbb{R}^d$ *and* $t \in [0,\infty)$,

$$X_t^x = x + \int_0^t f(s, X_\cdot^x)\,\mathrm{d}Z_s. \tag{7.21}$$

*Proof.* The issue is proving continuity in space. For every $x \in \mathbb{R}$ there exists a continuous in time solution $(X_t^x)_{t\in[0,\infty)}$ of (7.21). Let $t \in (0,\infty)$. We will consider the solution as a map from $\mathbb{R}$ to the Banach space fo continuous functions $\mathrm{C}([0,T];\mathbb{R})$ equipped with the norm

$$\|f\|_{\mathrm{C}([0,t];\mathbb{R})} = \sup_{s\in[0,t]} |f(s)|\,.$$

We will then apply the Kolmogorov continuity criterion to the map

$$x \in \mathbb{R} \mapsto (X_s^x)_{s\in[0,t]} \in \mathrm{C}([0,t];\mathbb{R}).$$

Let $p \in [2,\infty)$. The inequality $|a+b+c|^p \leq 3^{p-1}\left(|a|^p + |b|^p + |c|^p\right)$ and equation (7.21) prove that, for each $x, y \in \mathbb{R}^d$,

$$\begin{aligned}
&\sup_{s\in[0,t]} |X_s^x - X_s^y|^p \\
&\leq 3^{p-1}\,|x-y|^p \\
&\quad + 3^{p-1}\left(\sup_{s\in[0,t]}\left|\int_0^s \mu(r,X_r^x) - \mu(r,X_r^y)\,\mathrm{d}r\right|^p + \sup_{s\in[0,t]}\left|\int_0^s \sigma(r,X_r^x) - \sigma(r,X_r^y)\,\mathrm{d}B_r\right|^p\right).
\end{aligned}$$

The Burkholder-Davis-Gundy inequality, Hölder's inequality, and $p \in [2,\infty)$ prove that there exists $C_p \in (0,\infty)$ such that

$$\begin{aligned}
\mathbb{E}\left[\sup_{s\in[0,t]}\left|\int_0^s \sigma(r,X_r^x) - \sigma(r,X_r^y)\,\mathrm{d}B_r\right|^p\right] &\leq C_p\mathbb{E}\left[\left(\int_0^t \left((\sigma(s,X_s^x) - \sigma(s,X_s^y)\right)^2\,\mathrm{d}s\right)^{\frac{p}{2}}\right] \\
&\leq C_pK_p^2\mathbb{E}\left[\left(\int_0^t |X_s^x - X_s^y|^2\,\mathrm{d}s\right)^{\frac{p}{2}}\right] \\
&\leq C_pK_2^p t^{\frac{p-2}{p}}\int_0^t \mathbb{E}\left[\sup_{s\in[0,r]}|X_s^x - X_s^y|^p\right]\mathrm{d}r.
\end{aligned}$$

Similarly, it follows by Jensen's inequality and $p \in [2,\infty)$ that

$$\begin{aligned}
\sup_{s\in[0,t]}\left|\int_0^s \mu(r,X_r^x) - \mu(r,X_r^y)\,\mathrm{d}r\right|^p &\leq t^{p-1}\int_0^t |\mu(r,X_r^x) - \mu(r,X_r^y)|^p\,\mathrm{d}r \\
&\leq t^{p-1}K_2^p\int_0^t \sup_{s\in[0,r]}|X_s^x - X_s^y|^p\,\mathrm{d}r.
\end{aligned}$$

The previous three inequalities prove that, for every $p \in [2,\infty)$ and $t \in [0,\infty)$,

$$\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[0,t]}|X_s^x - X_s^y|^p\right] \\
&\leq 3^{p-1}\left(|x-y|^p + \left(C_pK_2^p t^{\frac{p-2}{p}} + K_2^p t^{p-1}\right)\int_0^t \mathbb{E}\left[\sup_{s\in[0,r]}|X_s^x - X_s^y|^p\right]\mathrm{d}r\right).
\end{aligned}$$

Grönwall's inequality proves that, for every $p \in [2,\infty)$ and $t \in [0,\infty)$ there exists a constant $c(t,p) \in (0,\infty)$ that is locally bounded in $t$ and $p$ such that

$$\mathbb{E}\left[\sup_{s\in[0,t]} |X^x_s - X^y_s|^p\right] \leq c(t,p)\,|x-y|^p\,.$$

The Kolmoogorov continuity criterion that there exists a continuous modification $(\hat{X}^x_s)_{s\in[0,t]}$ of the process $x \in \mathbb{R} \mapsto (X^x_s)_{s\in[0,t]} \in \mathrm{C}([0,T];\mathbb{R})$. It follows that the modification solves (7.21), because the zero set is independent of time. That is, for every $x \in \mathbb{R}$ we have

$$\mathbb{P}[\hat{X}^x_t = X^x_t \ \text{ for every } \ t \in [0,\infty)] = 1.$$

This completes the proof. $\square$

7.4. **Uniqueness by local times: Hölder continuous coefficients in one-dimension.** In the previous section, for instance, we proved the pathwise well-posedness of stochastic differential equations of the form

$$\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}B_t,$$

for coefficients $b$ and $\sigma$ that were locally bounded and globally Lipschitz continuous. In general, these assumptions cannot be relaxed for the drift $b$, since the equations

$$\dot{x}(t) = \sqrt{x(t)} \text{ and } \dot{x}(t) = x(t)^2,$$

are either have non-unique solutions or blow-up in finite time. However, in this section, we will prove using local times that the these assumptions can be relaxed for the diffusion coefficient $\sigma$. In particular, we will prove that the equation

$$\begin{cases} \mathrm{d}X_t = \sqrt{X_t}\,\mathrm{d}B_t & \text{in } (0,\infty), \\ X_0 = 0 \end{cases} \tag{7.22}$$

is well-posed and that zero is its unique solution.

Uniqueness for equations like (7.22) will be obtained the local time of the solution at zero. The idea is that on the one hand a solution $(X_t)_{t\in[0,\infty)}$ of (7.22) begins at zero and therefore if $(X_t)_{t\in[0,\infty)}$ is non-constant we expect that $L^0_t(X) > 0$ for every $t \in (0,\infty)$. However, on the other hand, the quadratic variation process

$$\mathrm{d}\langle X\rangle_t = X_t\,\mathrm{d}t,$$

vanishes whenever $X_t = 0$. Therefore, time isn't running when $X_t = 0$, and we expect that $L^0_t(X) = 0$ for all $t \in [0,\infty)$. That is, from the point of the view of the local time, the process $(X_t)_{t\in[0,\infty)}$ spends no time at zero. The only way to reconcile these two perspectives is for $(X_t)_{t\in[0,\infty)}$ to be constantly equal to zero.

**Proposition 7.18.** *Let $(B_t)_{t\in[0,\infty)}$ be a one-dimensional Brownian motion and let $b, \sigma\colon [0,\infty)\times\mathbb{R} \to \mathbb{R}$ be predictable processes. Let $(X^1_t)_{t\in[0,\infty)}$ and $(X^2_t)_{t\in[0,\infty)}$ satisfy that $X^1_0 = X^2_0$ almost everywhere and that, for each $i \in \{1,2\}$,*

$$\mathrm{d}X^i_t = b(X^i_t)\,\mathrm{d}t + \sigma(X^i_t)\,\mathrm{d}B_t. \tag{7.23}$$

*Then $X^1 \vee X^2$ is a solution of (7.23) if and only if $L^0_t(X^2 - X^1)$ vanishes identically.*

*Proof.* We observe using Tanaka's formula applied to the semimartingale $(X^1_t - X^2_t)_{t\in[0,\infty)}$ with $a = 0$ that

$$X^1_t \vee X^2_t = X^1_t + (X^2_t - X^1_t)_+ = X^1_t + \int_0^t \mathbf{1}_{\{X^2_t > X^1_t\}}\,\mathrm{d}(X^2_t - X^1_t) + L^0_t(X^2 - X^1).$$

Since, for each $i \in \{1,2\}$ we have that

$$\mathrm{d}X^i_t = b(X^i_t, t)\,\mathrm{d}t + \sigma(X^i_t, t)\,\mathrm{d}B_t,$$

it follows that, since $X^1_0 = X^2_0$ almost everywhere,

$$\begin{aligned} X^1_t \vee X^2_t &= X^1_0 + \int_0^t \mathbf{1}_{\{X^1_s \geq X^2_s\}} b(X^1_s, s) + \mathbf{1}_{\{X^2_s > X^1_s\}} b(X^2_s, s)\,\mathrm{d}s \\ &\quad + \int_0^t \mathbf{1}_{\{X^1_s \geq X^2_s\}} \sigma(X^1_s, s) + \mathbf{1}_{\{X^2_s > X^1_s\}} \sigma(X^2_s, s)\,\mathrm{d}B_s + L^0_t(X^2 - X^1) \\ &= (X^1_0 \vee X^2_0) + \int_0^t b(X^1_s \vee X^2_s, s)\,\mathrm{d}s + \int_0^t \sigma(X^1_x \vee X^2_s, s)\,\mathrm{d}B_s + L^0_t(X^1 - X^2). \end{aligned}$$

We therefore conclude that $(X^1_t \vee X^2_t)_{t\in[0,\infty)}$ is a solution of (7.23) if and only if $L^0_t(X^2 - X^1)$ vanishes identically. This completes the proof. $\square$

We will now prove that uniqueness in law and the vanishing of the local times implies pathwise uniqueness.

**Proposition 7.19.** *Let $(B_t)_{t\in[0,\infty)}$ be a one-dimensional Brownian motion and let $b,\sigma\colon [0,\infty)\times\mathbb{R}\to\mathbb{R}$ be predictable processes. Suppose that we have uniqueness in law for the equation*

$$\mathrm{d}X_t^i = b(X_t^i)\,\mathrm{d}t + \sigma(X_t^i)\,\mathrm{d}B_t, \tag{7.24}$$

*and that any two solutions $(X^1)_{t\in[0,\infty)}$ and $(X_t^2)_{t\in[0,\infty)}$ that satisfy $X_0^1 = X_0^2$ almost everywhere satisfy $L_t^0(X^2 - X^1) = 0$. Then, pathwise uniqueness holds for* (7.24)*.*

*Proof.* Suppose that $(X^1)_{t\in[0,\infty)}$ and $(X_t^2)_{t\in[0,\infty)}$ are two solutions of (7.24) defined with respect to the same Brownian motion that satisfy $X_0^1 = X_0^2$ almost everywhere. Then, since $L_t^0(X^2 - X^1)$ vanishes identically, we have by Proposition 7.18 that $X^1 \vee X^2$ is also a solution. Hence, by uniqueness in law, it follows that the law of $(X_t^1)_{t\in[0,\infty)}$ and the law of $(X_t^1 \vee X_t^2)_{t\in[0,\infty)}$ are the same, which occurs if and only if $(X_t^1)_{t\in[0,\infty)}$ and $(X_t^2)_{t\in[0,\infty)}$ are indistinguishable. □

In the following to statements we are essentially imposing a Hölder continuity condition on $\sigma$. That is, for $\rho(x) = |x|^\alpha$ we are requiring that

$$|\sigma(x) - \sigma(y)| \le |x-y|^\alpha\,,$$

or, equivalently, that

$$|\sigma(x) - \sigma(y)|^2 \le |x-y|^{2\alpha}\,.$$

The integrability condition requires that $2\alpha \ge 1$ and therefore that $\alpha \ge 1/2$. So, in particular, the theorem applies to $\sigma(x) = \sqrt{x}$.

**Proposition 7.20.** *Let $\rho\colon [0,\infty)\to[0,\infty)$ be a nondecreasing function that satisfies, for every $\varepsilon\in(0,1)$,*

$$\int_0^\varepsilon \frac{1}{\rho(s)}\,\mathrm{d}s = \infty. \tag{7.25}$$

*Let $(X_t)_{t\in[0,\infty)}$ be a continuous semimartingale such that for some $\varepsilon\in(0,1)$, for every $\theta\in(0,\infty)$,*

$$\int_0^t \mathbf{1}_{\{0<X_s\le\varepsilon\}}\frac{1}{\rho(X_s)}\,\mathrm{d}\langle X\rangle_s < \infty. \tag{7.26}$$

*Then $L_t^0(X) = 0$ almost surely.*

*Proof.* By the occupation formula,

$$\int_0^t \mathbf{1}_{\{0<X_s\le\varepsilon\}}\frac{1}{\rho(X_s)}\,\mathrm{d}\langle X\rangle_s = \int_0^\varepsilon \frac{1}{\rho(s)} L_t^a\,\mathrm{d}a.$$

By the right continuity of $L_t^a$ in $a$, it follows from assumption (7.25) that the righthand side of this equality is infinite on the set $\{L_t^0 > 0\}$. We therefore conclude by (7.26) that $L_t^0 = 0$ almost surely, for every $t\in(0,\infty)$. Since $L_t^0$ is continuous and increasing in $t$, this implies that $L_t^0$ almost surely vanishes identically. □

**Proposition 7.21.** *Let $\sigma, b^1, b^2\colon [0,\infty)\times\mathbb{R}\to\mathbb{R}$ be a predictable processes. Assume that there exists a nondecreasing $\rho\colon[0,\infty)\to[0,\infty)$ which satisfies* (7.25) *such that, for all $(x,t),(y,t)\in\mathbb{R}\times[0,\infty)$,*

$$|\sigma(x,t) - \sigma(y,t)|^2 \le \rho(|x-y|).$$

*Then, for a one-dimensional Brownian motion $(B_t)_{t\in[0,\infty)}$, if for each $i\in\{1,2\}$ the process $(X_t^i)_{t\in[0,\infty)}$ is a solution of*

$$\mathrm{d}X_t^i = b(X_t^i,t)\,\mathrm{d}t + \sigma(X_t^i,t)\,\mathrm{d}B_t,$$

*we have that $L_t^0(X^1 - X^2) = 0$ for every $t\in[0,\infty)$.*

*Proof.* Since for each $i \in \{1,2\}$ we have that

$$\mathrm{d}X^i_t = b(X^i_t, t)\,\mathrm{d}t + \sigma(X^i_t, t)\,\mathrm{d}B_t,$$

it follows that

$$\mathrm{d}\langle X^1 - X^2\rangle_t = \left|\sigma(X^1_t, t) - \sigma(X^2_t, t)\right|^2\,\mathrm{d}t \le \rho(\left|X^1_t - X^2_t\right|)\,\mathrm{d}t.$$

Therefore, for every $t \in (0,\infty)$,

$$\int_0^t \mathbf{1}_{\{X^1 > X^2\}} \frac{1}{\rho(X^1_s - X^2_x)}\,\mathrm{d}\langle X^1 - X^2\rangle_s \le \int_0^t \mathbf{1}_{\{X^1 > X^2\}}\,\mathrm{d}t \le t.$$

□

We may now state the main theorem of this section.

**Theorem 7.22.** *Let* $\rho\colon [0,\infty) \to [0,\infty)$ *be a nondecreasing function which satisfies* (7.25)*, let* $(B_t)_{t\in[0,\infty)}$ *be a one-dimensional Brownian motion, and let* $\sigma, b\colon [0,\infty) \times \mathbb{R} \to \mathbb{R}$ *be predictable functions. Then, pathwise uniqueness for the equation*

$$\mathrm{d}X_t = b(X_t, t)\,\mathrm{d}t + \sigma(X_t, t)\,\mathrm{d}B_t,$$

*under the following two conditions.*

*(i) Both* $\sigma(t,x) = \sigma(x)$ *and* $b(t,x) = b(x)$ *are time-independent,* $|\sigma(x) - \sigma(y)|^2 \le \rho(|x-y|)$*there exists* $\varepsilon > 0$ *such that* $0 < \varepsilon \le \sigma$*, and* $\sigma$ *and* $b$ *are bounded.*

*(ii)* $|\sigma(x,t) - \sigma(y,t)|^2 \le \rho(|x-y|)$ *and* $b$ *is locally Lipschitz continuous.*

*Proof.* In case (i), these conditions prove that the solution is unique in law and that the local time at zero of the difference of any two solutions must vanish. Therefore, the solution is pathwise unique. In case (ii), we may assume without loss of generality by a stopping time argument that $\sigma$ is bounded and that $b$ is globally Lipschitz continuous. It then follows by Proposition 7.21 that, for any two solutions $(X^1_t)_{t\in[0,\infty)}$ and $(X^2_t)_{t\in[0,\infty)}$ defined with respect to the same Brownian motion that satisfy $X^1_0 = X^2_0$ almost surely, we have almost surely that $L^0_t(X^1 - X^2) = 0$ for every $t \in [0,\infty)$. Therefore, by the Tanaka formula applied to $X^1_t - X^2_t$ at $a = 0$,

$$\left|X^1_t - X^2_t\right| = \int_0^t \mathrm{sgn}(X^1_t - X^2_t)\,\mathrm{d}(X^1_t - X^2_t).$$

The boundedness of $\sigma$ and the equation then prove that

$$\left|X^1_t - X^2_t\right| - \int_0^t (b(X^1_s, t) - b(X^2_s, s))\,\mathrm{d}s \ \text{ is a martingale.}$$

Therefore, for every $t \in [0,\infty)$, there exists $c \in (0,\infty)$ depending on the Lipschitz constant of $b$ such that

$$\mathbb{E}[\left|X^1_t - X^2_t\right|] \le c\int_0^t \mathbb{E}[\left|X^1_s - X^2_s\right|]\,\mathrm{d}s.$$

It then follows by Grönwall's inequality that $\mathbb{E}[\left|X^1_t - X^2_t\right| = 0$ for every $t \in [0,\infty)$. Therefore, by continuity, we conclude that the processes $(X^1_t)_{t\in[0,\infty)}$ and $(X^2_t)_{t\in[0,\infty)}$ are indistinguishable. This completes the proof. □