# Cheat Sheet: Mathematical Foundations for Finance

Steven Battilana, Autumn Semester 2017

## 1 Appendix: Some Basic Concepts and Results

### 1.1 Very Basic Things

**Def.**

(i) An empty product equals 1, i.e. $\prod_{j=1}^{0} s_j = 1$.

(ii) An empty sum equals 0, i.e. $\sum_{j=1}^{0} s_j = 0$.

**Def.** (geometric series)

(i) $s_n = a_0 \sum\limits_{k=0}^{n} q^k \overset{q\neq 1}{=} a_0 \frac{q^{n+1}-1}{q-1} = a_0 \frac{1-q^{n+1}}{1-q}$

(ii) $s = \sum\limits_{k=0}^{\infty} a_0 q^k \overset{|q|<1}{=} \frac{a_0}{q-1}$

**Def.** (conditional probabilities)

$$Q[C \cap D] = Q[C]Q[D|C]$$

**Def.** (compact)
Compactness is a property that generalises the notion of a subset of Euclidean space being closed (that is, containing all its limit points) and bounded (that is, having all its points lie within some fixed distance of each other). Examples include a closed interval, a rectangle, or a finite set of points (e.g. $[0,1]$).
**Def.** ($P$-trivial)
$\mathcal{F}_0$ is $P$-trivial iff $P[A] \in \{0,1\}$, $\forall A \in \mathcal{F}_0$.
**Useful rules**

- $E[e^Z] = e^{\mu + \frac{1}{2}\sigma^2}$, for $Z \sim \mathcal{N}(\mu, \sigma^2)$.
- Let $W$ be a BM, then $\langle W \rangle_t = t$, hence $d\langle W \rangle_s = ds$.

**Def. (Hilbert space)**
A Hilbert space is a vector space $H$ with an inner product $\langle f, g \rangle$such that the norm defined by

$$|f| = \sqrt{\langle f, g \rangle}$$

turns $H$ into a complete metric space. If the metric defined by the norm is not complete, then $H$ is instead known as an inner product space.
**Def. (cauchy-schwarz)**

- $|\langle u, w \rangle^2| \leq \langle u, u \rangle \cdot \langle w, w \rangle$
- $\sum_{j=1}^{N} u_j w_j \leq \left(\sum_{j=1}^{N} u_j^2\right)^{\frac{1}{2}} \left(\sum_{j=1}^{N} w_j^2\right)^{\frac{1}{2}}$
- $\int_a^b u(\tau) w(\tau) d\tau \leq \left(\int_a^b u^2(\tau) d\tau\right)^{\frac{1}{2}} \left(\int_a^b w^2(\tau) d\tau\right)^{\frac{1}{2}}$
- $|a(u,v)| \leq (a(u,v))^{\frac{1}{2}} (a(u,v))^{\frac{1}{2}}$

**Def. (Taylor)**

- $f(x \pm h) = \sum_{j=0}^{\infty} \frac{(\pm h)^j}{j!} \frac{d^j f}{dx^j}$
- $Tf(x;a) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n$

**Rem.** (uniform convergence on compacts in probability 1)
A mode of convergence on the space of processes which occurs often in the study of stochastic calculus, is that of *uniform convergence on compacts in probability* or *ucp convergence* for short.
First, a sequence of (non-random) functions $f_n : \mathbb{R}_+ \to \mathbb{R}$ converges uniformly on compacts to a limit $f$ if it converges uniformly on each bounded interval $[0,t]$. That is,

$$\sup_{s \leq t} |f_n(s) - f(s)| \to 0$$

as $n \to \infty$.
If stochastic processes are used rather than deterministic functions, then convergence in probability can be used to arrive at the following definition.
**Def. (uniform convergence on compacts in probability)**
A sequence of jointly measurable stochastic processes $X^n$ converges to the limit $X$ uniformly on compacts in probability if

$$P\left[\sup_{s \leq t} |X_s^n - X_s| > K\right] \to 0$$

as $n \to \infty, \forall t, K > 0$.
**Rem.** (uniform convergence on compacts in probability 2)
The notation $X^n \xrightarrow{ucp} X$ is sometimes used, and $X^n$ is said to converge ucp to $X$. Note that this definition does not make sense for arbitrary stochastic processes, as the supremum is over the uncountable index set $[0,t]$ and need not be measurable. However, for right or left continuous processes, the supremum can be restricted to the countable set of rational times, which will be measurable.
**Def.**

$$\mathcal{M}_d^2([0,a]) := \{M \in \mathcal{M}^2([0,a]) \mid t \mapsto M_t(\omega) \text{ RCLL for any } \omega \in \Omega\}$$

**Def.**

(i) We denote by $\mathcal{H}^2$ the vector space of all semimartingales vanishing at 0 of the form $X = M + A$ with $M \in \mathcal{M}_d^2(0,\infty)$ and $A \in FV$ (finite variation) predictable with total variation $V_\infty^{(1)}(A) = \int_0^\infty |dA_s| \in L^2(P)$.

(ii)

$$\|X\|_{\mathcal{H}^2}^2 = \|M\|_{M^2}^2 + \|V_\infty^{(1)}(A)\|_{L^2}^2 = E\left[[M]_\infty + \left(\int_0^\infty |dA|\right)^2\right]$$

**Thm. (dominated convergence theorem for stochastic integrals)**
Suppose that $X$ is a semimartingale with decomposition $X = M + A$ as above, and let $G^n, n \in \mathbb{N}$, and G be predictable processes. If

$$G_t^n(\omega) \to G_t(\omega) \quad \text{for any} t \geq 0, \text{ almost surely,}$$

and if there exists a process $H$ that is integrable w.r.t. $X$ such that $|G^n| \leq H$ for any $n \in \mathbb{N}$, then

$$G^n_{\cdot} X \to G.X \quad \text{u.c.p., as } n \to \infty.$$

If, in addition to the assumptions above, $X$ is in $\mathcal{H}^2$ and $\|H\|_X < \infty$ then even

$$\|G^n_{\cdot} X - G.X\|_{\mathcal{H}^2} \to 0 \quad \text{as } n \to \infty.$$

"**Def.**" (Wiki: **null set**)
In set theory, a null set $N \subset \mathbb{R}$ is a set that can be covered by a countable union of intervals of arbitrarily small total length. The notion of null set in set theory anticipates the development of Lebesgue measure since a null set necessarily has measure zero. More generally, on a given measure space $M = (X, \Sigma, \mu)$ a null set is a set $S \subset X$ s.t. $\mu(S) = 0$.
**Def. (power set)**
We denote by $2^\Omega$ the power set of $\Omega$; this is the family of all subsets of $\Omega$.
**Def.** (**$\sigma$-fiel** or **$\sigma$-algebra**)
A $\sigma$-field or $\sigma$-algebra on $\Omega$ is a family $\mathcal{F}$ of subsets of $\Omega$ which contains $\Omega$ and which is closed under taking complements and countable unions, i.e.

- $A \in \mathcal{F} \quad \Rightarrow \quad A^c \in \mathcal{F}$
- $A_i \in \mathcal{F},\ i \in \mathbb{N} \quad \Rightarrow \quad \bigcup_{i \in \mathbb{N}} A_i$ is in $\mathcal{F}$

**Rem.**
$\mathcal{F}$ is then also closed under countable intersections.
**Def. (finite $\sigma$-field)**
A $\sigma$-field is called finite if it contains only finitely many sets.
**Def. (measurable space)**
A pair $(\Omega, \mathcal{F})$ with $\Omega \neq \emptyset$ and $\mathcal{F}$ a $\sigma$-algebra on $\Omega$ is called a measurable space.
**Rem.**
$X$ is measurable (or more precisely Borel-measurable) if for every $B \in \mathcal{B}(\mathbb{R})$, we have $\{X \in B\} \in \mathcal{F}$.
**Def. (indicator function)**
For any subset $A$ of $\Omega$, the indicator function $I_A$ is the function defined by $I_A(\omega) := \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$
**Def.**
Let $\Omega \neq \emptyset$ and a function $X : \Omega \to \mathbb{R}$ (or more generally to $\Omega'$). Then $\sigma(X)$ is the smallest $\sigma$-field, say, on $\Omega$ s.t. $X$ is measurable with respect to $\mathcal{G}$ and $\mathcal{B}(\mathbb{R})$ (or $\mathcal{G}$ and $\mathcal{F}'$, respectively). We call $\sigma(X)$ the **$\sigma$-field generated by** $X$.

**Rem.**
We also consider a $\sigma$-field generated by a whole familiy of mappings; this is then the smallest $\sigma$-field that makes all the mappings in that family measurable.
**Def. (probability measure, probability space)**
If $(\Omega, \mathcal{F})$ is a masurable space, a probability measure on $\mathcal{F}$ is a mapping $P : \mathcal{F} \to [0,1]$ s.t. $P[\Omega] = 1$ and $P$ is $\sigma$-**additive**, i.e.

$$P\left[\bigcup_{i\in\mathbb{N}} A_i\right] = \sum_{i\in\mathbb{N}} P[A_i], \quad A_i \in \mathcal{F}, i \in \mathbb{N}, A_i \cap A_j = \emptyset, i \neq j.$$

The triple $(\Omega, \mathcal{F}, P)$ is then called a probability space.
**Def. ($P$-almost surely)**
A statement holds $P$-almost surely or $P$-a.s. if the set

$$A := \{\omega \mid \text{the statement does not hold}\}$$

is a $P$-nullset, i.e. has $P[A] = 0$. We sometimes use instead the formulation that a statement holds for $P$-almost all $\omega$.
**Notation**
E.g., $X \geq Y$ $P$-a.s. means that $P[X < Y] = 0$, or, equivalently $P[X \geq Y] = 1$. Notation:

$$P[X \geq Y] := P[\{X \geq Y\}] := P[\{\omega \in \Omega \mid X(\omega) \geq Y(\omega)\}].$$

**Def. (random variable)**
Let $(\Omega, \mathcal{F}, P)$ be a probability space and $X : \Omega \to \mathbb{R}$ a measurable function. We also say that $X$ is a (real-valued) random variable. If $Y$ is another random variable, we call $X$ and $Y$ equivalent if $X = Y$ $P$-a.s.
**Def. ($p$-integrable)**
We denote by $L^0$ or $L^0(\mathcal{F})$ the family of all equivalence classes of random variables on $(\Omega, \mathcal{F}, P)$. For $0 < p < \infty$, we denote by $L^p(P)$ the family of all equivalence classes of random variables $X$ which are $p$-integrable in the sense that $E[|X|^p] < \infty$, and we write then $X \in L^p(P)$ or $X \in L^p$ for short. Finally, $L^\infty$ is the family of all equivalence classes of random variables that are bounded by a constant $c$ (where the constant can depend on the random variable).
**Def. (Atom)**
If $(\Omega, \mathcal{F}, P)$ is a probability space, then an atom of $\mathcal{F}$ is a set $A \in \mathcal{F}$ with the properties that $P[A] > 0$ and that if $B \subseteq A$ is also in $\mathcal{F}$, then either $P[B] = 0$ or $P[B] = P[A]$. Intuitively, atoms are the 'smallest indivisible sets' in a $\sigma$-field. Atoms are pairwise disjoint up to $P$-nullsets.
**Def. (atomless)**
The space $(\Omega, \mathcal{F}, P)$ is called atomless if $\mathcal{F}$ contains no atoms; this can only happen if $\mathcal{F}$ is infinite. Finite $\sigma$-fields can be very conveniently described via their atoms because every set in $\mathcal{F}$ is then an union of atoms.
**Fatou's Lemma**
If $\{f_n\}$ is a sequence of nonnegative measurable functions, then

$$\int \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int f_n d\mu.$$

An example of an sequence of functions for which the inequality becomes strict is given by

$$f_n(x) = \begin{cases} 0, & x \in [-n, n] \\ 1, & \text{otherwise} \end{cases}$$

## 1.2 Conditional expectations: A survival kit

**Rem.**
Let $(\Omega, \mathcal{F}, P)$ be a probability space and $U$ a real-valued random variable, i.e. an $\mathcal{F}$-measurable mapping $U : \Omega \to \mathbb{R}$. Let $\mathcal{G} \subseteq \mathcal{F}$ be a fixed sub-$\sigma$-field of $\mathcal{F}$; the intuitive interpretation is that $\mathcal{G}$ gives us some partial information. The goal is then to find a prediction for $U$ on the basis of the information conveyed by $\mathcal{G}$, or a best estimate for $U$ that uses only information from $\mathcal{G}$.
**Def. (conditional expectation of $U$ given $\mathcal{G}$)**
A conditional expectation of $U$ given $\mathcal{G}$ is a real-valued random variable $Y$ with the following two properties:

(i) $Y$ is $\mathcal{G}$-measurable
(ii) $E[UI_A] = E[YI_A], \quad \forall A \in \mathcal{G}$

$Y$ is then called a version of the conditional expectation and is denoted by $Y = E[U|\mathcal{G}]$.
**Thm. 2.1**
Let $U$ be an integrable random variable, i.e. $U \in L^1(P)$. Then:

(i) There exists a conditional expectation $E[U|\mathcal{G}]$, and $E[U|\mathcal{G}]$ is again integrable.
(ii) $E[U|\mathcal{G}]$ is unique up to $P$-nullsets: If $Y, Y'$ are random variables satisfying the right above definition, then $Y' = Y$ $P$-a.s.

**Lem. (properties and computation rules)**
Next, we list properties of and computation rules for conditional expectations. Let $U, U'$ be integrable random variables s.t. $E[U|\mathcal{G}]$ and $E[U'|\mathcal{G}]$ exist. We denote by $b\mathcal{G}$ the set of all bounded $\mathcal{G}$-measurable random variables. Then we have:

(i) $E[UZ] = E[E[U|\mathcal{G}]Z], \quad \forall Z \in b$
(ii) Linearity: $E[aU1bU'|\mathcal{G}] = aE[U|\mathcal{G}] + bE[U'|\mathcal{G}], P\text{-a.s. } \forall a, b, \in \mathbb{R}$
(iii) Monotonicity: If $U \geq U'$ $P$-a.s., then $E[U|\mathcal{G}] \geq E[U'|\mathcal{G}]$ $P$-a.s.
(iv) Projectivity: $E[U|\mathcal{H}] = E[E[U|\mathcal{G}]|\mathcal{H}], \quad P$-a.s. $\forall\sigma$-field $\mathcal{H} \subseteq \mathcal{G}$
(v) $E[U|\mathcal{G}] = U, \quad P$-a.s. if $U \in \mathcal{G}$-measurable
(vi) $E[E[U|\mathcal{G}]] = E[U]$
(vii) $E[ZU|\mathcal{G}] \overset{(v)}{=} E[ZE[U|\mathcal{G}]|\mathcal{G}] \overset{(viii)}{=} ZE[U|\mathcal{G}], \quad P$-a.s. $\forall Z \in b\mathcal{G}$
(viii) $E[U|\mathcal{G}] = E[U], \quad P$-a.s. for $U$ independent of $\mathcal{G}$

**Rem.**

(i) Instead of integrability of $U$, one could also assume that $U \geq 0$; then analogous statements are true.
(ii) More generally, $(i)$ and $(vii)$ hold as soon as $U$ and $ZU$ are both integrable or both nonnegative.
(iii) If $U$ is $\mathbb{R}^d$-valued, one simply does everything component by component to obtain analogous results.

**Lem. 2.2**
Let $U, V$ be random variables s.t. $U$ is $\mathcal{G}$-measurable and $V$ is independent of $\mathcal{G}$. For every measurable function $F \geq 0$ on $\mathbb{R}^2$, then

$$E[F(U,V)|\mathcal{G}] = E[F(u,V)]|_{u=U} =: f(U).$$

Intuitively, one can compute the conditional expectation $E[F(U,V)|\mathcal{G}]$ by 'fixing the known value $U$ and taking the expectation over the independent quantity $V$. **Thm. 2.3**
Suppose $(U_n)_{n\in\mathbb{N}}$ is a sequence of random variables.

(i) If $U_n \geq X$ $P$-a.s. for all $n$ and some integrable random variable $X$, then

$$E[\liminf_{n\to\infty} U_n|\mathcal{G}] \leq \liminf_{n\to\infty} E[U_n|\mathcal{G}], \quad P\text{-a.s.}$$

(ii) If $(U_n)$ converges to some random variable $U$ $P$-a.s. and if $|U_n| \leq X$ $P$-a.s. for all $n$ and some integrable random variable $X$, then

$$E[\lim_{n\to\infty} U_n|\mathcal{G}] = E[U|\mathcal{G}] = \lim_{n\to\infty} E[U_n|\mathcal{G}], \quad P\text{-a.s.}$$

## 1.3 Stochastic processes and functions

**Def. (stochastic process)**
A (real-valued) stochastic process with index set $\mathcal{T}$ is a family of random variables $X_t, t \in \mathcal{T}$, which are all defined on the same probability space $(\Omega, \mathcal{F}, P)$. We often write $X = (X_t)_{t\in\mathcal{T}}$.
**Def.** (increment of stochastic process)
For any stochastic process $X = (X_k)_{k=0,1,\dots,T}$, we denote the increment from $k-1$ to $k$ of $X$ by

$$\Delta X_k := X_k - X_{k-1}.$$

**Rem.**
A stochastic process can be viewed as a function depending on two parameters, namely $\omega \in \Omega$ and $t \in \mathcal{T}$.
**Def. (trajectory)**
If we fix $t \in \mathcal{T}$, then $\omega \mapsto X_t(\omega)$ is simply a random variable. If we fix instead $\omega \in \Omega$, then $t \mapsto X_\omega(t)$ can be viewed as a function $\mathcal{T} \to \mathbb{R}$, and we often call this the path or the trajectory of the process corresponding to $\omega$.
**Def. (continuous)**
A stochastic process is continuous if all or $P$-almost all its trajectories are continuous functions.
**Def. (RCLL)**
A stochastic process is RCLL if all of $P$-almost all its trajectories are right-continuous (RC) functions admitting left limits (LL).
**Def.** (Wiki: signed measure)
Given a measurable space $(X, \Sigma)$, i.e. a set $X$ with a $\sigma$-algebra on it, an signed measure is a function $\mu : \Sigma \to \mathbb{R}$, s.t. $\mu(\emptyset) = 0$ and $\mu$ is $\sigma$-additive, i.e. it satisfies

$$\mu\left(\bigcup_{n=1}^\infty A_n\right) = \sum_{n=1}^\infty \mu(A_n)$$

where the series on the right must converge absolutely, for any sequence $A_1, A_2, \dots, A_n, \dots$ of disjoint sets in $\Sigma$.
**Def.** (Wiki: total variation in measure theory)
Consider a signed measure $\mu$ on a measurable space $(X, \Sigma)$, then it is possible to define two set functions $\overline{W}(\mu, \cdot)$ and $\underline{W}(\mu, \cdot)$, respectively

called upper variation and lower variation, as follows

$$\overline{W}(\mu, E) = \sup\{\mu(A) \mid A \in \Sigma, A \subset E\}, \quad \forall E \in \Sigma$$
$$\underline{W}(\mu, E) = \inf\{\mu(A) \mid A \in \Sigma, A \subset E\}, \quad \forall E \in \Sigma$$

The following clearly holds:

$$\overline{W}(\mu, E) \geq 0 \geq \underline{W}(\mu, E), \quad \forall E \in \Sigma$$

**Def.** (Wiki: bounded variation (finite variation))

(i) The total variation of a real-valued (or more generally complex-valued) function $f$, defined on an interval $[a, b] \subset \mathbb{R}$ is the quantity

$$V_a^b(f) = \sup_{P \in \mathcal{P}} \sum_{i=0}^{n_P - 1} |f(x_{i+1}) - f(x_i)|.$$

If $f$ is differentiable and its derivative is Riemann-integrable, its total variation is the vertical component of the arc-length of its graph, that is to say,

$$V_a^b(f) = \int_a^b |f'(x)| dx.$$

(ii) A real-valued function $f$ on the real line is said to be of bounded variation (BV function) on a chosen interval $[a, b] \subset \mathbb{R}$ if its total variation is finite, i.e.

$$f \in BV([a, b]) \quad \Leftrightarrow \quad V_a^b(f) < \infty.$$

It can be proved that a real function $f$ is of bounded variation in $[a, b]$ if and only if it can be written as the difference $f = f_1 - f_2$ of two non-decreasing functions on $[a, b]$: this result is known as the Jordan decomposition of a function and it is related to the Jordan decomposition of a measure.

**Def.** (Wiki: variation)
The variation (also called absolute variation) of the signed measure $\mu$ is the set function

$$|\mu|(E) = \overline{W}(\mu, E) + |\underline{W}(\mu, E)|, \quad \forall E \in \Sigma$$

and its total variation is defined as the value of this measure on the whole space of definition, i.e. $\|\mu\| = |\mu|(X)$.
**Def.** (planetmath.org: finite variation process)
In the theory of stochastic processes, the term finite-variation process is used to refer to a process $X_t$ whose paths are right-continuous and have finite total variation over every compact time interval, with probability one.
**Def. (finite variation)**
A stochastic process is of finite variation if all or $P$-almost all its trajectories are functions of finite variation.
**Def. (locally)**
A stochastic process has a property locally if there exists a sequence of stopping times $(\tau_n)_{n \in \mathbb{N}}$ increasing to $\infty$ $P$-a.s. s.t. when restricted to the stochastic interval $[\![0, \tau_n]\!] = \{(\omega, t) \in \Omega \times \mathcal{T} \mid 0 \leq t \leq \tau_n(\omega)\}$, the process has the property under consideration.
**Lem. (Brownian Motion and quadratic variation)**
If a continuous martingale $X$ has quadratic variation $[X]_t = t$, then it is a Brownian Motion.
**Def. (quadratic variation)**

$$[B]_t := \lim_{|\Pi| \to \infty} \sum_{\Pi} (B_{s_i^\Pi} - B_{s_{i-1}^\Pi})^2.$$

Note that the quadratic variation is itself a process $[B_t]_{t \geq 0}$. $\Pi$ here refers to a a partition of the interval $[0, t]$, which is supposed to get finer and finer. This limit taking process is the same one found when one defines the Riemann integral.

## 2 Financial Markets in Discrete Time

### 2.1 Basic setting

**Basic setting of the market**

- **Probability space:** $(\Omega, \mathcal{F}, \mathbb{P})$
- Finite discrete time horizon: $k = 0, 1, \ldots, T$
- Flow of information over time: **filtration** $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,\ldots,T}$.
  This is a family of $\sigma$-fields $\mathcal{F}_k \subseteq \mathcal{F}$ which is increasing.
- An ($\mathbb{R}^d$-valued) **stochastic process** in discrete time: $X = (X_k)_{k=0,1,\ldots,T}$ of ($\mathbb{R}^d$-valued) RVs which are all defined on the same probability space.
  This describes the random evolution over time of $d$ quantities.
- A stochastic process is called **adapted** (w.r.t. $\mathbb{F}$) if each $X_k$ is $\mathcal{F}_k$-measurable, i.e. observable at time $k$.
- A stochastic process is called **predictable** (w.r.t. $\mathbb{F}$) if each $X_k$ is even $\mathcal{F}_{k-1}$-measurable.

**Frictionless financial market**

- no transaction cost
- short-selling allowed
- investors are small, i.e. their trading does not affect stock prices

**Def.**

(i) $r_k$ describes the (simple) interest rate for the period $(k-1, k]$.
(ii) $Y_k$ is the growth factor for the time period $(k-1, k]$.
(iii) $\widetilde{S}^0$ models a bank account.
(iv) $\widetilde{S}^i$ models a stock.

### 2.2 Basic processes

**Def.** (discounted reference asset price)
The discounted price of the reference asset is $S_k^0 := \frac{\widetilde{S}_k^0}{\widetilde{S}_k^0} = 1$ at all times.
**Def.** (discounted asset price)
The discounted asset price $S = (S_k)_{k=0,1,2,\ldots,T}$ are given by $S_k := \frac{\widetilde{S}_k}{\widetilde{S}_k^0}$

**Trading strategy** $\varphi$

- A trading strategy is an $\mathbb{R}^d$-valued stochastic process $\varphi = (\varphi^0, \vartheta)$.
- $\varphi^0 = (\varphi_k^0)_{k=0,1,\ldots,T}$ denotes the riskless asset.
  $\varphi^0$ is $\mathbb{R}$-valued and adapted.
- $\vartheta = (\vartheta_k)_{k=0,1,\ldots,T}$ denotes the $d$ risky assets.
  $\vartheta$ is $\mathbb{R}^d$-valued and predictable.
- Initial value: $\varphi = (\varphi^0, \vartheta_0 \equiv 0)$
  (since there is no trading before time 0, i.e. investors start out without any shares)
- A trading strategy describes a dynamically evolving portfolio in the $d+1$ basic assets available for trade.

**Value process** $V$

- $V(\varphi) = (V_k(\varphi))_{k=0,1,\ldots,T}$ denotes the discounted value process of the strategy $\varphi$ and is given by

$$V_k(\varphi) := \underbrace{\varphi_k^0 S_k^0}_{\text{bank account}} + \underbrace{\vartheta_k \cdot S_k}_{\text{portfolio}} = \varphi_k^0 + \sum_{i=1}^d \vartheta_k^i S_k^i, \quad \text{for } k = 0, 1, \ldots, T \tag{1}$$

  where $S_k^0 = 1$.
- $V$ is $\mathbb{R}$-valued and adapted.
- Initial value: $V_0(\varphi) = \varphi_0^0 = C_0(\varphi)$.

**Cost process** $C$

- $(C_k(\vartheta))_{k=0,1,\ldots,T}$ denotes the discounted cost process associated to $\varphi$:

$$C_k(\varphi) := V_k(\varphi) - G_k(\varphi)$$

- By construction, $C_k(\varphi)$ describes the cumulative (total) costs for the strategy $\varphi$ on the time interval $[0, k]$.
- The initial cost for $\varphi$ at time $0$ comes from putting $\varphi_0^0$ into the bank account

$$C_0(\varphi) = \varphi_0^0 = V_0(\varphi).$$

- Incremental cost:

$$\Delta C_{k+1}(\varphi) := \underbrace{(\varphi^0_{k+1} - \varphi^0_k)\underbrace{S^0_k}_{=1}}_{\text{bank account}} + \underbrace{\sum_{i=1}^{d}(\vartheta^i_{k+1} - \vartheta^i_k)\, S^i_k}_{\text{portfolio}}$$

rewrite the above by adding and subtracting $\varphi^{tr}_{k+1} S_{k+1}$:

$$\begin{aligned}\boldsymbol{\Delta C_{k+1}(\varphi)} &= \varphi^0_{k+1} - \varphi^0_k + (\vartheta_{k+1} - \vartheta_k)^{tr} S_k \\ &= \varphi^0_{k+1} - \varphi^0_k + (\vartheta_{k+1} - \vartheta_k)^{tr} S_k \pm \varphi^{tr}_{k+1} S_{k+1} \\ &= \varphi^0_{k+1} + \vartheta^{tr}_{k+1} S_{k+1} - \varphi^0_k - \vartheta^{tr}_k S_k - \vartheta^{tr}_{k+1}\Delta S_{k+1} \\ &\overset{(1)}{=} V_{k+1}(\varphi) - V_k(\varphi) - \vartheta^{tr}_{k+1}\Delta S_{k+1} \\ &= \boldsymbol{\Delta V_{k+1}(\varphi) - \vartheta^{tr}_{k+1}\Delta S_{k+1}}\end{aligned}$$

**Gains process $G$**

- $(G_k(\vartheta))_{k=0,1,\dots,T}$ denotes the discounted gains process associated to $\vartheta$:

$$\begin{aligned}G_k(\vartheta) &:= \sum_{j=1}^{k} \vartheta^{tr}_j \Delta S_j, \quad k = 0,1,\dots,T \\ &= \vartheta \cdot S\end{aligned}$$

- $G$ is $\mathbb{R}$-valued and adapted.
- If we think of a continuous-time model where successive trading dates are infinitely close together, then the increment $\Delta S$ becomes a differential $dS$ and the sum becomes a stochastic integral

$$G(\vartheta) = \int \sum_{i=1}^{k} \vartheta^i dS^i, \quad k = 0,1,\dots,T$$

**Self-financing strategy**

- A trading strategy $\varphi = (\varphi^0, \vartheta)$ is called self-financing if its cost process $C(\varphi)$ is constant over time.
- **Proposition 2.1.** A self-financing strategy $\varphi = (\varphi^0, \vartheta)$ is uniquely determined by its initial wealth $V_0$ and its risky asset component $\vartheta$.
  In particular, any pair $(V_0, \vartheta)$ specifies in a unique way a self-financing strategy.
  If $\varphi = (\varphi^0, \vartheta)$ is self-financing, then $(\varphi^0_k)_{k=1,\dots,T}$ is automatically predictable.
- It then holds for the corresponding (incremental) **cost process:**

$$\begin{aligned}\Delta C_{k+1}(\varphi) &= (\varphi^0_{k+1} - \varphi^0_k)\, S^0_k + (\vartheta_{k+1} - \vartheta_k)\cdot S_k = 0, \quad P\text{-a.s.} \\ C(\varphi) &= C_0(\varphi) = V_0(\varphi) = \varphi^0_0\end{aligned}$$

- It then holds for the corresponding **value process:**

$$\begin{aligned}V_k(\varphi) &= V_0(\varphi) + G_k(\vartheta) = \varphi^0_0 + G_k(\vartheta) \\ &= \varphi^0_0 + \sum_{j=1}^{k} \vartheta^{tr}_j \Delta S_j\end{aligned}$$

Remarks:

- The notion of a strategy being self-financing is a kind of economic budget constraint.
- The notion of self-financing is numeraire (discounting) irrelevant, i.e. it does not depend on the units in which the calculations are done.

**A first introduction to stopping time**
**Notation:**

$$a \wedge b := \min(a,b)$$

**Def.** (stopped stochastic process)
The stochastic process $S^\tau = (S^\tau_k)_{k=0,1,\dots,T}$ defined by

$$S^\tau(\omega) := S_{k\wedge\tau}(\omega) := S_{k\wedge\tau(\omega)}(\omega)$$

is called the process $S$ stopped at $\tau$.
**Rem.**
$S^\tau$ could fail to be a stochastic process because $S^\tau_k = S_{k\wedge\tau}$ could fail to be a random variable, i.e. could fail to be measurable. But (in discrete time) this is not a problem if we assume that $\tau$ is measurable, which is mild and reasonable enough.
**Def.** (stopping time)
Note that $\vartheta_k = I_{\{k\le\tau\}}$ is $\mathcal{F}_{k-1}$-measurable for each $k$ iff $\{\tau \ge k\} \in \mathcal{F}_{k-1}$ for all $k$ or, equivalently by passing to complements,

$$\{\tau \le j\} \in \mathcal{F}_j, \quad \forall j.$$

By definition, this means that $\tau$ is a stopping time (w.r.t. $\mathbb{F}$).
**Example** (a doubling strategy, p.16)
We denote by the following the (random) time of the first stock price rise:

$$\tau := \inf\{k \mid Y_k = 1+u\} \wedge T$$

which is a stopping time because

$$\begin{aligned}\underbrace{\{\tau \le k\}}_{\in\mathcal{F}_k,\forall k} &= \underbrace{\{Y_1 = 1+u\}}_{\in\mathcal{F}_1\subseteq\mathcal{F}_k} \cup \dots \cup \underbrace{\{Y_k = 1+u\}}_{\in\mathcal{F}_k} \\ &= \{\max(Y_1,\dots,Y_k) \ge 1+u\} \in \mathcal{F}_k, \quad \forall k.\end{aligned}$$

With

$$\vartheta_k := \tfrac{1}{S_{k-1}} 2^{k-1} I_{\{k\le\tau\}}.$$

it follows that $\vartheta$ is predictable because each $\vartheta_k$ is $\mathcal{F}_{k-1}$-measurable.
Note that this uses $\{k \le \tau\} = \{\tau < k\}^c = \{\tau \le k-1\}^c$.

**Admissibility**

- For $a \in \mathbb{R}, a \ge 0$, a trading strategy $\varphi$ is called $a$-**admissible** if its value process $V(\varphi)$ is uniformly bounded from below by $-a$, i.e.

$$V_k(\varphi) \ge -a, \quad \mathbb{P}\text{-a.s.}, \quad \forall k \ge 0$$

- A trading strategy is called **admissible** if it is $a$-admissible for some $a \ge 0$.

Remark:

- An admissible strategy can be interpreted as a strategy having some credit line which imposes a lower bound on the associated value process. So one may make debts, but only within clearly defined limits.

## 2.3 Some important martingale results

**Def. (martingale, martingale property)**
Let $(\Omega, \mathcal{F}, Q)$ be a probability space with a filtration $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,\dots,T}$. A (real-valued) stochastic process $X = (X_k)_{k=0,1,\dots,T}$ is called martingale (w.r.t. $Q$ and $\mathbb{F}$) if it is adapted to $\mathbb{F}$, is $Q$-integrable in the sense that $X_k \in \mathcal{L}^1(Q), \forall k$, and satisfies the martingale property

$$E_Q[X_\ell|\mathcal{F}_k] = X_k, \quad Q\text{-a.s. for } k \le \ell.$$

**Def.** (supermartingale, submartingale)

(i) If we have " $\le$ " in the above definition (a tendency to go down), $X$ is called a supermartingale.

(ii) If we have " $\ge$ " in the above definition (a tendency to go up), $X$ is called a submartingale.

**Def. (local martingale, localising sequence)**
An adapted process $X = (X_k)_{k=0,1,\dots,T}$ null at $0$ (i.e. $X_0 = 0$) is called a local martingale (w.r.t. $Q$ and $\mathbb{F}$) if there exists a sequence of stoppint times $(\tau_n)_{n\in\mathbb{N}}$ increasing to $T$ s.t. $\forall n \in \mathbb{N}$, the stopped process $X^{\tau_n} = (X_{k\wedge\tau_n})_{k=0,1,\dots,T}$ is a $(Q,\mathbb{F})$-martingale. We then call $(\tau_n)_{n\in\mathbb{N}}$ a localising sequence.
**Thm. 3.1**
Suppose $X = (X_k)_{k=0,1,\dots,T}$ is an $\mathbb{R}^d$-valued martingale or local martingale null at $0$. For any $\mathbb{R}^d$-valued predictable process $\vartheta$, the stochastic integral process $\vartheta \cdot X$ defined by

$$\vartheta \cdot X_k := \sum_{j=1}^{k} \vartheta^{tr}_j \Delta X_j, \quad k = 0,1,\dots,T$$

is then a (real-valued) local martingale null at $0$. If $X$ is a martingale and $\vartheta$ is bounded, then $\vartheta \cdot X$ is even a martingale.
**Rem.**
In continuous time, the above theorem no longer holds.
**Rem.**
If we think of $X = S$ as discounted asset prices, then $\vartheta \cdot S = G(\vartheta)$

is the discounted gains process.

**Cor. 3.2**

For any martingale $X$ and any stopping time $\tau$, the stopped process $X^\tau$ is again a martingale. In particular, $E_Q[X_{k\wedge\tau}] = E_Q[X_0], \forall k$.

**Interpretation**

A martingale describes a fair game in the sense that one cannot predict where it goes next.

(i) Cor. 3.2 says that one cannot change this fundamental character by cleverly stopping the game.

(ii) Thm. 3.1 says that as long as one can only use information from the past, not even complicated clever betting will help.

**Thm. 3.3**

Suppose that $X$ is an $\mathbb{R}^d$-valued local $Q$-martingale null at $0$ and $\vartheta$ is an $\mathbb{R}^d$-valued predictable process. If the stochastic integral process $\vartheta \cdot X$ is uniformly bounded below (i.e. $\vartheta \cdot X \geq -b$ $Q$-a.s., $\forall k, b \geq 0$), then $\vartheta \cdot X$ is a $Q$-martingale.

## 2.4 An example: The multinomial model

**Def.** (multiplicative model)

The multiplicative model with i.i.d. returns is given by

$$\frac{\widetilde{S}_k^0}{\widetilde{S}_{k-1}^0} = 1 + r > 0, \quad \forall k,$$
$$\frac{\widetilde{S}_k^1}{\widetilde{S}_{k-1}^1} = Y_k, \quad \forall k,$$

where $\widetilde{S}_0^0 = 1, \widetilde{S}_0^1 = S_0^1 > 0$ is a constant, and $Y_1, ..., Y_T$ are i.i.d. and take the finitely many values $1 + y_1, ..., 1 + y_m$ with respective probabilities $p_1, ..., p_m$. We assume that all the probabilities $p_j$ are $> 0$ and that $y_m > y_{m-1} > ... > y_1 > -1$. This also ensures that $\widetilde{S}^1$ remains strictly positive.

**Rem.**

Intuition suggests that for a reasonable model, the sure factor $1 + r$ should lie between the minimal and maximal values $1+y_1$ and $1+y_m$ of the (uncertain) random factor.

**Def.** (canonical model)

The simplest and in fact canonical model for this setup is a path space. Let

$$\begin{aligned}\Omega &= \{1, ..., m\}^T \\ &= \{\omega = (x_1, ..., x_T) \mid x_k \in \{1, ..., m\} \text{ for } k = 1, ..., T\}\end{aligned}$$

be the set of all sequences of length $T$ formed by element of $\{1, ..., m\}$. Take $\mathcal{F} = 2^\Omega$, the family of all subsets of $\Omega$, and define $P$ by setting

$$P[\{\omega\}] = p_{x_1} p_{x_2} \cdot ... \cdot p_{x_T} = \prod_{k=1}^{T} p_{x_k}.$$

Finally, define $Y_1, ..., Y_T$ by

$$Y_k(\omega) := 1 + y_{x_k} \tag{2}$$

so that $Y_k(\omega) = 1 + y_j$ iff $x_k = j$. We take as filtration the one generated by $\widetilde{S}^1$ (or, equivalently, by $Y$) s.t.

$$\mathcal{F}_k = \sigma(Y_1, ..., Y_k), \quad k = 0, 1, ..., T.$$

**Def. (atom)**

A set $A \subseteq \Omega$ is an atom of $\mathcal{F}_k$ iff there exists a sequence $(\overline{x}_1, ..., \overline{x}_k)$ of length $k$ with elements $\overline{x}_i \in \{1, ..., m\}$ s.t. $A$ consists of all those $\omega \in \Omega$ that start with the substring $(\overline{x}_1, ..., \overline{x}_k)$, i.e.

$$A = A_{\overline{x}_1, ..., \overline{x}_k} := \{\omega = (x_1, ..., x_T) \in \{1, ..., m\}^T \mid x_i = \overline{x}_i \text{ for } i = 1, ..., k\}.$$

**Consequences**

(i) Each $\mathcal{F}_k$ is parametrised by substrings of length $k$ and therefore contains precisely $m^k$ atoms.

(ii) When going from time $k$ to time $k + 1$, each atom $A = A_{\overline{x}_1, ..., \overline{x}_k}$ from $\mathcal{F}_k$ splits into precisely $m$ subsets $A_1 = A_{\overline{x}_1, ..., \overline{x}_k, 1}, ..., A_m = A_{\overline{x}_1, ..., \overline{x}_k, m}$ that are atoms of $\mathcal{F}_{k+1}$.

(iii) The atoms of $\mathcal{F}_k$ are pairwise disjoint and their union is $\Omega$. Finally, each set $B \in \mathcal{F}_k$ is a union of atoms of $\mathcal{F}_k$; so the family $\mathcal{F}_k$ of events observable up to time $k$ consists of $2^{m^k}$ sets.

**Def. (one-step transition probabilities)**

For any atom $A = A_{\overline{x}_1, ..., \overline{x}_k}$ of $\mathcal{F}_k$, we then look at its $m$ successor atoms $A_1 = A_{\overline{x}_1, ..., \overline{x}_k, 1}, ..., A_m = A_{\overline{x}_1, ..., \overline{x}_k, m}$ of $\mathcal{F}_{k+1}$, and we define the one-step transition probabilities for $Q$ at the node corresponding to $A$ by the conditional probabilities

$$Q[A_j | A] = \frac{Q[A_j]}{Q[A]}, \quad \text{for } j = 1, ..., m.$$

Because $A$ is the disjoint union of $A_1, ..., A_m$, we have $0 \leq Q[A_j | A] \leq 1$ for $j = 1, .., m$ and $\sum_{j=1}^{m} Q[A_j | A] = 1$.

**Rem.**

The decomposition of factorisation of $Q$ in such a way that for every trajectory $\omega \in \Omega$, its probability $Q[\{\omega\}]$ is the product of the successive one-step transition probabilities along $\omega$.

**Rem.**

We can describe $Q$ equivalently either via its global weights $Q[\{\omega\}]$ or via its local transition behaviour.

**Def.** (independent growth rates)

The (coordinate) variables $Y_1, ..., Y_T$ from (2) are independent under $Q$ iff for each $k$, the one-step transition probabilities are the same for each node at time $k$, but they can still differ across date $k$.

**Def.**

$Y_1, ..., Y_T$ are i.i.d. under $Q$ iff at each node throughout the tree, the one-step transition probabilities are the same.

**Rem.**

Probability measures with this particular structure can therefore be described by $m - 1$ parameters; recall that the $m$ one-step transition probabilities at any given node must sum to $1$, which eliminates one degree of freedom.

## 2.5 Properties of the market

**Characterisation of financial markets via EMMs (equivalent martingale measures)** The description of a financial market model via EMMs can be summarized as follows:

- **Existence** of an EMM $\iff$ the market is **arbitrage-free** i.e. $\mathbb{P}_e(S) \neq \emptyset$ by the 1st FTAP
- **Uniqueness** of the EMM $\iff$ the market is **complete** i.e. $\#(\mathbb{P}_e(S)) = 1$ by the 2nd FTAP

### 2.5.1 Arbitrage

**1st Fundamental Theorem of Asset Pricing (FTAP)**

**Thm. 2.1 (Dalang/Morton/Willinger)**

(i) Consider a financial market model in finite discrete time.

(ii) Then $S$ is arbitrage-free iff there exists an EMM for $S$, i.e.

$$\boxed{\text{(NA)} \iff \mathbb{P}_e(S) \neq \emptyset}$$

- In other words: If there exists a probability measure $Q \approx P$ on $\mathcal{F}_T$ s.t. $S$ is a $Q$-martingale, then $S$ is arbitrage-free.
- *Limitations*: The most important of these assumptions are frictionless markets and small investors—and if one tries to relax these to allow for more realism, the theory even in finite discrete time becomes considerably more complicated and partly does not even exist yet.

**Cor. 2.2**

The multinomial model with parameters $y_1 < ... < y_m$ and $r$ is arbitrage-free iff $y_1 < r < y_m$.

**Cor. 2.3**

The binomial model with parameters $u > d$ and $r$ is arbitrage-free iff $d < r < u$. In that case, the EMM $Q^*$ for $\widetilde{S}^1 / \widetilde{S}^0$ is unique (on $\mathcal{F}_T$) and is given as in Cor. 1.4.

**Arbitrage opportunity**

**Def. (arbitrage opportunity, arbitrage-free)**

(i) An arbitrage opportunity is an admissible self-financing strategy $\varphi \mathrel{\widehat{=}} (0, \vartheta)$ with zero initial wealth, with $V_t(\varphi) \geq 0$, $P$-a.s. and with $P[V_T(\varphi) > 0] > 0$.

(ii) The financial market $(\Omega, \mathcal{F}, \mathbb{F}, P, S^0, S)$ or shortly $S$ is called arbitrage-free if there exist no arbitrage opportunities.

(iii) Sometimes one also says that $S$ satisfies (NA).

**Def.**

(i) (NA$_+$) := forbids to produce something out of nothing with $0$-admissible self-financing strategies.

(ii) (NA′) := forbids the same for all (not necessarily admissible) self-financing strategies.

(iii) Then we clearly have the implications:

$$(\text{NA}') \Rightarrow (\text{NA}) \Rightarrow (\text{NA}_+)$$

Note, for finite discrete time, the three concepts are all equivalent.

**Prop. 1.1**
For a financial market in finite discrete time, the following statements are equivalent:

(i) $S$ is arbitrage-free.

(ii) There exists no self-financing strategy $\varphi \mathrel{\widehat{=}} (0, \vartheta)$ with zero initial wealth and satisfying $V_T(\varphi) \geq 0$, $P$-a.s. and $P[V_T(\varphi) > 0] > 0$. In other words, $S$ satisfies (NA′).

(iii) For every (not necessarily admissible) self-financing strategy $\varphi$ with $V_0(\varphi) = 0$, $P$-a.s. and $V_T(\varphi) \geq 0$ $P$-a.s., we have $V_T(\varphi) = 0$ $P$-a.s.

(iv) For the space

$$\mathcal{G}' := \{G_T(\vartheta) \mid \vartheta \text{ is } \mathbb{R}^d\text{-valued and predictable}\}$$

of all final wealths that one can generate from zero initial wealth through self-financing trading, we have

$$\mathcal{G}' \cap L^0_+(\mathcal{F}_T) = \{0\}$$

where $L^0_+(\mathcal{F}_T)$ denotes the space of all nonnegative $\mathcal{F}_T$-measurable RVs, i.e. $L^0_+(\mathcal{F}_T) = \mathbb{R}^n_+$, $(L^0(\mathcal{F}_T) = \mathbb{R}^n)$.

**Rem.**

(i) The two sets $L^0_+(\mathcal{F}_T)$ and $\mathcal{G}'$ can be *separated by a hyperplane*, and the normal vector defining that hyperplane then yields (after suitable normalisation) the (density of the) desired EMM.

(ii) The existence of an EMM follows from the existence of a separating hyperplane between two sets.

(iii) The set $\mathbb{P}_e(S)$ is convex, it is either empty, or contains exactly one element, or contains infinitely (uncountably) many elements.

*Interpretation:* Absence of arbitrage is a natural economic/financial requirement for a reasonable model of a financial market, since there cannot exist "money pumps" (at least not for long).

**Def. (equivalent)**
Two probability measures $Q$ and $P$ on $(\Omega, \mathcal{F})$ are equivalent (on $\mathcal{F}$), written as $Q \approx P$ on $\mathcal{F}$, it they ahve the same nullsets (in $\mathcal{F}$), i.e. if for each set $A$ (in $\mathcal{F}$) we have $P[A] = 0$ iff $Q[A] = 0$.

**Lem. 1.2**
If there exists a probability measure $Q \approx P$ on $\mathcal{F}_T$ s.t. $S$ is a $Q$-martingale, then $S$ is arbitrage-free.

**Cor. 1.3**
In the multinomial model with parameters $y_1 < ... < y_m$ and $r$, there exists a probability measure $Q \approx P$ s.t. $\widetilde{S}^1_0/\widetilde{S}^0$ is a $Q$-martingale iff $y_i < r < y_m$.

**Cor. 1.4**
In the binomial model with parameters $u > d$ and $r$, there exists a probability measure $Q \approx P$ s.t. $\widetilde{S}^1_0/\widetilde{S}^0$ is a $Q$-martingale iff $u > r > d$. In that case, $Q$ is unique (on $\mathcal{F}$) and characterised by the property that $Y_1, ..., Y_T$ are i.i.d. under $Q$ with parameter

$$Q[Y_k = 1 + u] = q^* = \tfrac{r-d}{u-d} = 1 - Q[Y_k = 1 + d].$$

---

### 2.5.2 Completeness

**Def. (Completeness of the market)**

- A financial market model (in finite discrete time) is called complete if *every payoff* $H \in L^0_+(\mathcal{F}_T)$ *is attainable.*
- Otherwise it is called *incomplete.*

**Thm. 2.1 (Valuation and hedging in complete markets)**
Consider a financial market model in finite discrete time and suppose that $\mathcal{F}_0$ is trivial and $S$ is arbitrage-free and *complete*. Then for every payoff $H \in L^0_+(\mathcal{F})$, there is a unique price process $V^H = (V^H_k)_{k=0,1,...,T}$ which admits no arbitrage. It is given by

$$V^H_k = E_Q[H|\mathcal{F}_k] = V_k(V_0, \vartheta), \quad \text{for } k = 0, 1, ..., T$$

for any EMM $Q$ for $S$ and any replicating strategy $\varphi \mathrel{\widehat{=}} (V_0, \vartheta)$ for $H$.

**2nd Fundamental Theorem of Asset Pricing (FTAP)**
**Thm. 2.2**

- Consider a financial market model in finite discrete time and assume that $S$ is arbitrage-free, $\mathcal{F}_0$ is trivial and $\mathcal{F}_T = \mathcal{F}$.
- Then $S$ is complete iff there is a unique EMM for $S$

$$\boxed{(\text{NA}) + \text{completeness} \iff \#(\mathbb{P}_e(S)) = 1,}$$

i.e. $\mathbb{P}_e(S)$ is a singleton.

Remark:

- If a financial market in discrete time is complete, then $\mathcal{F}_T$ is finite (i.e. completeness is quite restrictive).
- Completeness is only an assertion about $\mathcal{F}_T$-measurable quantities.

**Example: The binomial model**
**Recall:**
We recall that this model is described by parameters $p \in (0, 1)$ and $u > r > d > -1$; then we have $\widetilde{S}^0_k = (1+r)^k$ and $\widetilde{S}^1_k = S^1_0 \prod_{j=1}^k Y_j$ with $S^1_0 > 0$ and $Y_1, ..., Y_T$ i.i.d under $P$ taking values $1+u$ or $1+d$ with probability $p$ or $1-p$, respectively. The filtration $\mathbb{F}$ is generated by $\widetilde{S} = (\widetilde{S}^0, \widetilde{S}^1)$ or equivalently by $\widetilde{S}^1$ or by $Y$. Note that $\mathcal{F}_0$ is then trivial because $\widetilde{S}^0_0 = 1$ and $\widetilde{S}^1_0 = S^1_0$ is a constant. We also take $\mathcal{F} = \mathcal{F}_T$.
**Rem.**
We already know from Cor. 2.3 that this model is arbitrage-free and has a unique EMM for $S^1 = \widetilde{S}^1/\widetilde{S}^0$. Hence $S^1$ is complete by Thm. 2.2, and so every $H \in L^0_+(\mathcal{F}_T)$ is attainable, with a price process given by

$$V^H_k = E_{Q^*}[H|\mathcal{F}_k] \quad \text{for } k = 0, 1, ..., T,$$

where $Q^*$ is the unique EMM for $S^1$. We also recall from Cor. 2.3 that the $Y_j$ are under $Q^*$ again i.i.d., but with

$$Q^*[Y_1 = 1 + u] = q^* := \tfrac{r-d}{u-d} \in (0, 1).$$

All the above quantities $S^1, H, V^H$ are discounted with $\widetilde{S}^0$, i.e. expressed in units of asset 0. The undiscounted quantities are the stock price $\widetilde{S}^1 = S^1\widetilde{S}^0$, the payoff $\widetilde{H} := H\widetilde{S}^0_T$ and its price process $\widetilde{V}^{\widetilde{H}}$ with $\widetilde{V}^{\widetilde{H}}_k := V^H_k \widetilde{S}^0_k$ for $k = 0, 1, ..., T$.
**Cor. 3.1**
In te binomial model, te undiscounted arbitrage-free price process of any undiscountede payoff $\widetilde{H} \in L^0_+(\mathcal{F}_T)$ is given by

$$\widetilde{V}^{\widetilde{H}}_k = \widetilde{S}^0_k E_{Q^*}\left[\frac{\widetilde{H}}{\widetilde{S}^0_T}\middle|\mathcal{F}_k\right] = E_{Q^*}\left[\widetilde{H}\frac{\widetilde{S}^0_k}{\widetilde{S}^0_T}\middle|\mathcal{F}_k\right] = \frac{\widetilde{S}^0_k}{\widetilde{S}^0_T} E_{Q^*}[\widetilde{H}|\mathcal{F}_k]$$

for $k = 0, 1, ..., T$.

---

## 2.6 Pricing of contingent claims $H$

**Def. (general European option)**
A general European option of payoff or contingent claim is a random variable $H \in L^0_+(\mathcal{F}_T)$.
**Interpretation**
The interpretation is that $H$ describes the *net payoff* (in units of asset 0) that the owner of this instrument obtains at time $T$; so having $H \geq 0$ is natural and also avoids integrability issues. Since $H$ is $\mathcal{F}_T$-measurable, the payoff can depend on the entire information up to time $T$; and "European" means that the time for the payoff is fixed at the end $T$.
**Def. (European call option, net payoff)**

(i) A European call option on asset $i$, with *maturity* $T$ and *strike* $K$ gives its owner the right, but not the obligation, to buy at time $T$ one unit of asset $i$ for the price $K$, irrespective of what the actual asset price $S^i_T$ then is.

(ii) In monetary terms, any rational person will make use of (*exercise*) that right iff $S^i_T(\omega) > K$, then the net payoff is $S^i_T(\omega) - K$, more formal

$$H(\omega) = \max(0, S^i_T(\omega) - K) = (S^i_T(\omega) - K)^+.$$

As a random variable, this is clearly nonnegative and $\mathcal{F}_T$-measurable since $S^i$ is adapted.

**Example** (payoffs)

(i) If we want to bet on a reasonably stable asset price evolution, we might be interested in a payoff of the form $H = I_B$ with

$$B = \left\{ a \le \min_{i=1,\dots,d} \min_{k=0,1,\dots,T} S_k^i < \max_{i=1,\dots,d} \max_{k=0,1,\dots,T} S_k^i \le b \right\}$$

This option pays at time $T$ on unit of money iff all stocks remain between the levels $a$ and $b$ up to time $T$.

(ii) A payoff of the form

$$H = I_A g \left( \frac{1}{T} \sum_{k=1}^{T} S_k^i \right), \quad A \in \mathcal{F}_T, g \ge 0$$

gives a payoff which depends on the average price (over time) of asset $i$, but which is only due in case that a certain event $A$ occurs.

**Def. (Attainability)**

- A payoff $H \in L_+^0(\mathcal{F}_T)$ is called attainable if there exists an admissible self-financing strategy $\varphi \widehat{=} (V_0, \vartheta)$ with $V_T(\varphi) = H$ $P$-a.s.
- The stratety $\varphi$ is then said to **replicate** $H$ and is called a **replicating strategy** for $H$.

**Thm. 1.1 (Arbitrage-free valuation of attainable payoffs)**

- Consider a financial market in finite discrete time and suppose that $S$ is arbitrage-free and complete and $\mathcal{F}_0$ is trivial.
- Then for every payoff $H \in L_+^0(\mathcal{F}_T)$ has a unique price process $V^H = (V_k^H)_{k=0,1,\dots,T}$ which admits no arbitrage.
- $V^H$ is given by:

$$\boxed{V_k^H = E_Q\left[H | \mathcal{F}_k\right] = V_k(V_0, \vartheta)}$$

for $k = 0, 1, \dots, T$, for any EMM $Q$ for $S$ and for any replicating strategy $\varphi \widehat{=} (V_0, \vartheta)$ for $H$.

- *Rem.:* Because it involves no preferences, but only the assumption of absence of arbitrage, the valuation from this Thm. is often also called *risk-neutral valuation*, and an EMM $Q$ for $S$ is called a *risk-neutral measure*

**Thm. 1.2 (Characterisation of attainable payoffs)**

- Consider a financial market in finite discrete time and suppose that $S$ is arbitrage-free and $\mathcal{F}_0$ is trivial. For any payoff $H \in L_+^0(\mathcal{F}_T)$, the following are equivalent:
  (i) $H$ is *attainable.*
  (ii) $\sup_{Q \in \mathbb{P}_e(S)} E_Q[H] < \infty$ is attained in some $Q^* \in \mathbb{P}_e(S)$, i.e. the supremum is finite and a maximum.
  In other words, we have $\sup_{Q \in \mathbb{P}_e(S)} E_Q[H] = E_{Q^*}[H] < \infty$ for some $Q^* \in \mathbb{P}_e(S)$.
  (iii) The mapping $\mathbb{P}_e(S) \to \mathbb{R}$, $Q \mapsto E_Q[H]$ is constant, i.e. $H$ has the same and finite expectation under all EMMs $Q$ for $S$.

- Remark: Note that not all of these relationships necessarily hold for financial markets in infinite discrete time or continuous time. "2) $\Rightarrow$ 3)" in general only holds if $H$ is bounded.

**Approach to valuing and hedging payoffs** For a given payoff $H$ in a financial market in finite discrete time (with $\mathcal{F}_0$ trivial):

(i) Check if $S$ is arbitrage-free by finding at least one EMM $Q$ for $S$.
(ii) Find all EMMs $Q$ for $S$.
(iii) Compute $E_Q[H]$ for all EMMs $Q$ for $S$ and determine the supremum of $E_Q[H]$ over $Q$.
(iv) If the supremum is finite and a maximum, i.e. attained in some $Q^* \in \mathbb{P}_e(S)$, then $H$ is attainble and its price process can be computed as $V_k^H = E_Q[H | \mathcal{F}_k]$, for any $Q \in \mathbb{P}_(S)$.
If the supremum is not attained (or, equivalently for finite discrete time, there is a pair of EMMs $Q_1, Q_2$ with $E_{Q_1}[H] \ne E_{Q_2}[H]$), then $H$ is not attainable.

**Invariance of the risk-neutral pricing method under a change of numéraire**

- The risk-neutral pricing method is invariant under a change of numéraire, i.e. all assets can be priced under a risk-neutral method independent of the chosen asset used for discounting.
- Denote with $Q^{**}$ the EMM for $\hat{S}^0 := \frac{\tilde{S}^0}{\tilde{S}^1}$.
  Denote with $Q^*$ the EMM for $S^1 = \frac{\tilde{S}^1}{\tilde{S}^0}$.
- Then it holds for a financial market $(\tilde{S}^0, \tilde{S}^1)$ and an undiscounted payoff $\tilde{H} \in L_+^0(\mathcal{F}_T)$ that:

$$\tilde{S}_k^0 E_{Q^*}\left[ \frac{\tilde{H}}{\tilde{S}_T^0} \middle| \mathcal{F}_k \right] = \tilde{S}_k^1 E_{Q^{**}}\left[ \frac{\tilde{H}}{\tilde{S}_T^1} \middle| \mathcal{F}_k \right]$$

**EMMs in submarkets**

- If a market $(S^0, S^1, \dots, S^k)$ is (NA), i.e. there exists an EMM $Q$, then this EMM $Q$ is also an EMM for all submarkets.
  (e.g. for $(S^k, S^i), k \ne i$, for $(S^k, S^i, S^j), k \ne i \ne j$ etc.)
- If there exists a EMM $Q^j$ for a submarket $(S^0, S^j)$ which is not also an EMM for another submarket $(S^0, S^k)$, $j \ne k$, then the whole market $(S^0, S^1, \dots, S^k)$ is not (NA), i.e. it admits arbitrage.

## 2.7 Multiplicative model

- Suppose that we start with the RVs $r_1, \dots, r_T$ and $Y_1, \dots, Y_T$.
- Define the **bank account/riskless asset** by:

$$\tilde{S}_k^0 := \prod_{j=1}^{k} (1 + r_j), \qquad \frac{\tilde{S}_k^0}{\tilde{S}_{k-1}^0} = 1 + r_k, \qquad \tilde{S}_0^0 = 1$$

Remarks:
  - $\tilde{S}_k^0$ is $\mathcal{F}_{k-1}$-measurable (i.e. predictable).
  - $r_k$ denotes the rate for $(k-1, k]$.

- Define the **stock/risky asset** by:

$$\tilde{S}_k^1 := S_0^1 \prod_{j=1}^{k} Y_j, \qquad \frac{\tilde{S}_k^1}{\tilde{S}_{k-1}^1} = Y_k, \qquad \tilde{S}_0^1 = \text{const.}, \quad \tilde{S}_0^1 \in \mathbb{R}$$

Remarks:
  - $\tilde{S}_k^1$ is $\mathcal{F}_k$-measurable (i.e. adapted).
  - $Y_k$ denotes the growth factor for $(k-1, k]$.
  - The rate of return $R_k$ is given by $Y_k = 1 + R_k$.

### 2.7.1 Cox-Ross-Rubinstein (CRR) binomial model

**Assumptions**

- *Bank account/riskless asset:*
  Suppose all the $r_k \in \mathbb{R}$ are constant with value $r > -1$.
  This means that we have the same nonrandom interest rate over each period.
  Then the bank account evolves as $\tilde{S}_k^0$ for $k = 0, 1, \dots, T$.
- *Stock/risky asset:*
  Suppose that $Y_1, \dots Y_T \in \mathbb{R}$ are independent and only take two values, $1 + u$ with probability $p$, and $1 + d$ with probability $1 - p$ (i.e. all $Y_k$ are i.i.d.).
  Then the stock prices at each step moves either up (by a factor $1 + u$) or down (by a factor $1 + d$), i.e.

$$\frac{\tilde{S}_k^1}{\tilde{S}_{k-1}^1} = Y_k = \begin{cases} 1 + u, & \text{with probability } p \\ 1 + d, & \text{with probability } 1 - p \end{cases}$$

**Martingale property** The discounted stock price $\frac{\tilde{S}^1}{\tilde{S}^0}$ is a $\mathbb{P}$-martingale iff $r = pu + (1 - p)d$.

**EMM**

- In the binomial model, there exists a probability measure $\mathbb{Q} \approx \mathbb{P}$ s.t. $\frac{\tilde{S}^1}{\tilde{S}^0}$ is a $\mathbb{Q}$-martingale iff $u > r > d$.

- In that case, $\mathbb{Q}$ is unique (on $\mathcal{F}_T$) and characterised by the property that $Y_1, \dots, Y_T$ are i.i.d. under $\mathbb{Q}$ with parameter

$$q^* = \mathbb{Q}[Y_k = 1+u] = \frac{r-d}{u-d} \qquad (\nearrow \text{ up})$$
$$1-q^* = 1-\mathbb{Q}[Y_k = 1+d] = \frac{u-r}{u-d} \qquad (\searrow \text{ down})$$

**Arbitrage and completeness** The following statements are equivalent:

(i) $u > r > d$

(ii) $\exists$ a unique EMM $\mathbb{Q}^*$ for $\frac{\tilde{S}^1}{\tilde{S}^0}$ (on $\mathcal{F}_T$)

(iii) The market $S$ is (NA) and complete.

**Put-Call parity**

- Assuming $T = 1$, it holds that:

$$V_0^{C(K)} - V_0^{P(K)} = S_0^1 - \frac{K}{1+r}$$

**Pricing binomial contingent claims $H$**

- Assume time horizon $T = 1$, strike $K > 0$, and a payoff function $H(x, K) : \mathbb{R}_+ \to \mathbb{R}_+$ of a European style contingent claim with strike $K$.
- $H$ may be a European call function $C(x, K) = (x-K)^+$ or a European put function $P(x, K) = (K-x)^+$.
- Then $H$ can be replicated using a self-financing strategy $\varphi^{H(K)} = (V_0^{H(K)}, \vartheta^{H(K)})$ s.t.

$$V_1(\varphi^{H(K)} = \frac{H(\tilde{H}_1^1, K)}{1+r}, \qquad \mathbb{P} - a.s.$$

and $\varphi^{H(K)}$ is given by

$$V_0^{H(K)} = \frac{r-d}{u-d}\frac{H(S_0^1(1+u), K)}{1+r} + \frac{u-r}{u-d}\frac{H(S_0^1(1+d), K)}{1+r}$$
$$\vartheta_1^{H(K)} = \frac{H(1+u, K/S_0^1) - H(1+d, K/S_0^1)}{u-d}$$

- Note that this can also be expressed via the martingale pricing approach:

$$V_0^{H(K)} = \mathbb{E}_{\mathbb{Q}}\left[\frac{H(\tilde{S}_1^1, K)}{1+r}\right]$$

where

$$\mathbb{Q}\left[\tilde{S}_1^1 = S_1^0(1+u)\right] = q = \frac{r-d}{u-d}$$
$$\mathbb{Q}\left[\tilde{S}_1^1 = S_1^0(1+d)\right] = 1-q = \frac{u-r}{u-d}$$

**Binomial call pricing formula**

$$\tilde{V}_k^{\tilde{H}} = \tilde{S}_k^1 \mathbb{Q}^{**}[W_{k,T} > x] - \tilde{K}\frac{\tilde{S}_k^0}{\tilde{S}_T^0}\mathbb{Q}^*[W_{k,T} > x]$$
$$x = \frac{\log\frac{\tilde{K}}{\tilde{S}_k^1} - (T-k)\log(1+d)}{\log\frac{1+u}{1+d}}$$

Remark:

- This is the discrete analogue of the Black-Scholes formula.

---

### 2.7.2 Multinomial model

**EMM**

- IOT construct an EMM for $S^1$, it needs to hold that:

$$\mathbb{E}_{\mathbb{Q}}[S_1^1] = S_0^1$$
$$\iff \mathbb{E}_{\mathbb{Q}}[Y_1] = 1+r \qquad \iff \qquad \sum_{k=1}^m q_k(1+y_k) = 1+r$$

with the further conditions:

$$\sum_{k=1}^m q_i = 1, \qquad q_1, \dots, q_m \in (0,1)$$

**Arbitrage condition** The following statements are equivalent:

(i) $y_1 < r < y_m$

(ii) $\exists$ an EMM $\mathbb{Q} \approx \mathbb{P}$ s.t. $\frac{\tilde{S}^1}{\tilde{S}^0}$ is a $\mathbb{Q}$-martingale.

(iii) The market $S$ is (NA).

**Completeness** The multinomial model is

- **complete** whenever $m \leq 2$
(i.e. there are *no* nodes that allow for more than two possible stock price evolutions)
- **incomplete** whenever $m > 2$
(i.e. there is *at least one* node that allows for more than two possible stock price evolutions)

**Inequality of the payoffs of Asian and European call options**

- Consider a European call option $C_k^E = (\tilde{S}_k^1)^+$ and an Asian call option with

$$C_k^A = \left(\frac{1}{k}\sum_{j=1}^k \tilde{S}_j^1 - K\right)^+$$

- Then it holds for the $\mathbb{Q}$-expectation (risk-neutral) of the two payoffs that:

$$\mathbb{E}_{\mathbb{Q}}\left[\frac{C_k^A}{\tilde{S}_k^0}\right] \leq \mathbb{E}_{\mathbb{Q}}\left[\frac{C_k^E}{\tilde{S}_k^0}\right]$$

- *Interpretation:* Since the volatility of an Asian style contingent claim is lower than the one of a European style contingent claim, the Asian option bears lower risk and thus yields also lower profit.

**American options**

- Consider an American option with maturity $T$ and nonnegative adapted payoff process $U = (U_k)_{k=0,\dots,T}$.
- Then the arbitrage-free price process $\bar{V} = (\bar{V}_k)_{k=0,\dots,T}$ w.r.t. $\mathbb{Q}$ can be expressed as a backward recursive scheme such as:

$$\bar{V}_T = U_T$$
$$\bar{V}_k = \max(U_k, \mathbb{E}_{\mathbb{Q}}[\bar{V}_{k+1}|\mathcal{F}_k]), \qquad \text{for } k = 0, \dots, T-1$$

## 3 Martingales

### 3.1 Martingales

#### Martingales

- Let $(\Omega, \mathcal{F}, P)$ be a probability space with a filtration $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,\dots,T}$.
- A (real-valued) stochastic process $X = (X_k)_{k=0,1,\dots,T}$ is called a **martingale** (w.r.t. $P$ and $\mathbb{F}$) if:
  (i) $X$ is adapted to $\mathbb{F}$.
  (ii) $X$ is $P$-integrable in the sense that $X_k \in \mathcal{L}^1(P)$ for each $k$, i.e.
  $$E_P[|X|] < \infty$$
  (iii) $X$ satisfies the martingale property:
  $$E_P[X_l|\mathcal{F}_k] = X_k \qquad P\text{-a.s. for } k \leq l$$
- *Interpretation:* This means that the best prediction for the later value $X_l$ given the information $\mathcal{F}_k$ is just the current value $X_k$. Hence the changes in a martingale cannot be predicted.
  In other words, a martingale describes a fair game in the sense that one cannot predict where it goes next.
- A **supermartingale** is defined the same but with the property
  $$E_P[X_l|\mathcal{F}_k] \leq X_k$$
- A **submartingale** is defined the same but with the property
  $$E_P[X_l|\mathcal{F}_k] \geq X_k$$

#### Equivalent Martingale Measure (EMM)

**Def. (equivalent martingale measure (EMM))**

(i) An equivalent martingale measure (EMM) for $S$ is a probability measure $Q$ equivalent to $P$ on $\mathcal{F}_T$ s.t. $S$ is a $Q$-martingale.

(ii) We denote by $\mathbb{P}_e(S)$ or $\mathbb{P}_e$ the set of all EMMs for $S$.

**Thm. (Radon-Nikodým)**
Let $(\Omega, \mathcal{F})$ and filtration $\mathbb{F} = (\mathcal{F}_{k=0,1,\dots,T}$ in finite discrete time. One $(\Omega, \mathcal{F})$, we have two probability measures $Q$ and $P$, and we assume that $Q \approx P$. Then the Radon-Nikodým theorem tells us that there exists a *density* $\mathcal{D} := \frac{dQ}{dP}$; this is a random variable $\mathcal{D} > 0$ $P$-a.s. (because $Q \approx P$) s.t. $Q[A] = E_P[\mathcal{D}I_A]$ for all $A \in \mathcal{F}$, or more generally

$$E_Q[Y] = E_P[Y\mathcal{D}], \quad \forall \text{RV } Y \geq 0.$$

One sometimes writes this as

$$\int_\Omega Y dQ = \int_\Omega Y \mathcal{D} dP.$$

**Rem.**

- The point of these fomulae is that the tell us how to compute $Q$-expectations in terms of $P$-expectations and vice versa.
- $\mathcal{D}$ is a random variable (density)
- $Z$ is a process (density process)

**Def.**

(i) $\mathcal{D} := \frac{dQ}{dP}$

(ii) $\mathcal{D} := \frac{dQ}{dP} := Z_0 \prod_{j=1}^T D_j$

**Def.**
We introduce the $P$-martingale $Z$ (sometimes denoted by $Z^Q$ or $Z^{Q;P}$) by

$$Z_k := E_P[\mathcal{D}|\mathcal{F}_k] = E_P\left[\frac{dQ}{dP}|\mathcal{F}_k\right], \quad \text{for } k = 0, 1, \dots, T.$$

Because $\mathcal{D} > 0$ $P$-a.s., the process $Z = (Z_k)_{k=0,1,\dots,T}$ is strictly positive in the sense that $Z_k > 0$ $P$-a.s. for each $k$, or also $P[Z_k > 0, \forall k] = 1$. $Z$ is called *density process* (of $Q$, w.r.t. $P$).

- $Q$ is an EMM for the stochastic process $(S_k)_{k\geq 0}$ iff $S$ is a martingale under $Q$ and if $Q$ is equivalent to $P$.
- The probability measure $Q$ is equivalent to $P$ ($Q \approx P$) iff:
  (i) $Q[A] > 0 \iff P[A] > 0$
  (ii) $Q[\Omega] = 1$
- A stochastic process $(S_k)_{k\geq 0}$ is a $Q$-martingale iff:
  (i) $S_k$ is adapted to the considered filtration.
  (ii) $S_k$ is integrable: $E_Q[|S_k|] < \infty$.
  (iii) Martingale property: $E_Q[S_{k+1}|\mathcal{F}_k] = S_k$.

#### Local martingale

- An adapted process $X = (X_k)_{k=0,1,\dots,T}$ null at 0 (i.e. with $X_0 = 0$) is called a local martingale (w.r.t. $P$ and $\mathbb{F}$) if there exists a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ increasing to $T$ s.t. for each $n \in \mathbb{N}$, the stopped process $X^{\tau_n} = (X_{k\wedge\tau_n})_{k=0,1,\dots,T}$ is a $(P, \mathbb{F})$-martingale.
- We then call $(\tau_n)_{n\in\mathbb{N}}$ a **localising sequence**.
- For any martingale $X$ and any stopping time $\tau$, the stopped process $X^\tau$ is again a martingale.
  In particular, $E_P[X_{k\wedge\tau}] = E_P[X_0]$, $\forall k$.

#### Martingale and BMs

**Def. (martingale, supermartingale, submartingale)**
As in descrete time, a martingale w.r.t. $P$ and $\mathbb{F}$ is a (real-valued) stochastic process $M = (M_t)$ s.t. $M$ is adapted to $\mathbb{F}$, $M$ is $P$-integrable in the sense that each $M_t$ is in $L^1(P)$, and the martingale property holds: for $s \leq t$, we have

$$E[M_t|\mathcal{F}_s] = M_s, \quad P\text{-a.s.}$$

If we have the inequality "$\leq$" instead of "$=$", then $M$ is a supermartingale; if we have "$\geq$", then $M$ is a submartingale. Of course, $\mathbb{F} = (\mathcal{F}_t)$ and $M = (M_t)$ should have the same time index set.
**Rem.**
Because our filtration satisfies the usual conditions, a general result from the theory of stochastic processes says that any martingale has a version with nice (RCLL, i.e. right-continuous with left limits) *trajectories*.
**Def. (local martingale, localising sequence)**
An adapted process $X = (X_t)_{t\geq 0}$ null at 0 (i.e. with $X_0 = 0$) is called a local martingale (w.r.t. $P$ and $\mathbb{F}$) if there exists a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ increasing to $\infty$ s.t. for each $n \in \mathbb{N}$, the stopped process $X^{\tau_n} = (X_{t\wedge\tau_n})_{t\geq 0}$ is a $(P, \mathbb{F})$-martingale. We then call $(\tau_n)_{n\in\mathbb{N}}$ a localising sequence.

### 3.2 Stopping times

**Stopping time** The stochastic process $S^\tau = (S_k^\tau)_{k=0,1,\dots,T}$ defined by

$$S_k^\tau(\omega) := S_{k\wedge\tau}(\omega) := S_{k\wedge\tau(\omega)}(\omega)$$

is called the process $S$ stopped at $\tau$. It clearly behaves like $S$ up to time $\tau$ and remains constant after time $\tau$.

**Thm 2.1 (Stopping theorem)**

- Suppose that $M = (M_t)_{t\geq 0}$ is a $(P, \mathbb{F})$-martingale with RC trajectories, and $\sigma, \tau$ are $\mathbb{F}$-stopping times with $\sigma \leq \tau$.
- If either $\tau$ is bounded by some $T \in (0, \infty)$ or $M$ is uniformly integrable, then $M_\tau$, $M_\sigma$ are both in $L^1(P)$ and
  $$E[M_\tau|\mathcal{F}_\sigma] = M_\sigma \qquad P\text{-a.s.}$$

**Rem.**

(i) For any RC martingale $M$ and any stopping time $\tau$, we have $E[M_{\tau\wedge t}|\mathcal{F}_s] = M_{\tau\wedge s}$ for $s \leq t$, i.e. the stopped process $M^\tau = (M_t^\tau)_{t\geq 0} = (M_{t\wedge t})_{t\geq 0}$ is again a martingale (because we have $E[M_t^\tau|\mathcal{F}_s] = M_s^\tau$).

(ii) If $M$ is an RC martingale and $\tau$ is any stopping time, then we always have for any $t \geq 0$ that $E[M_{\tau\wedge t}] = E[M_0]$. If either $\tau$ is bounded or $M$ is uniformly integrable, then we also obtain $E[M_\tau] = E[M_0]$.

**Cases of stopping times**

- Define the stopping time $\tau_a$ for $a \in \mathbb{R}, a > 0$ as:

  $\tau_a := \inf\{t \geq 0 | W_t > a\}$

  Then it holds that:

  - $\tau_{a_1} \leq \tau_{a_2}$, $P$-a.s. for $a_1 < a_2$.
  - $P[\tau_a < \infty] = 1$.
  - $W_{\tau_a} = a$, $P$-a.s.
  - $E\left[W_{\tau_{a_2}} \middle| \mathcal{F}_{\tau_{a_1}}\right] \neq W_{\tau_{a_1}}, \mathbb{P}$-a.s.
    i.e. the stopping theorem fails for $\tau = \tau_{a_2}$ and $\sigma = \tau_{a_1}$.
  - $\lim_{n\to\infty} \tau_n(\omega) = \infty, \mathbb{P}$-a.s.

- Define the stopping time $\rho_a$ for $a \in \mathbb{R}, a > 0$ as:

  $\rho_a := \sup\{t \geq 0 | W_t > a\}$

  Then it follows that $\rho_a = \infty$ with probability 1 under $P$.

**Def. (stopping time)**
Again exactly like in discrete time, a stopping time w.r.t. $\mathbb{F}$ is a mapping $\tau : \Omega \to [0, \infty]$ s.t. $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$.
**Def.** (events observable up to time $\sigma$)
We define for a stopping time $\sigma$, the $\sigma$-field of events observable up to time $\sigma$ as

$$\mathcal{F}_\sigma := \{A \in \mathcal{F} \mid A \cap \{\sigma \leq t\} \in \mathcal{F}_t, \text{ for all } t \geq 0\}.$$

(One must and can check that $\mathcal{F}_\sigma$ is a $\sigma$-field, and one has $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$ for $\sigma \leq \tau$.)
**Def.**
We also need to define $M_\tau$, the value $M$ at the stopping time $\tau$, by

$$(M_\tau)(\omega) := M_{\tau(\omega)}(\omega).$$

Note that this implicitly assumes that we have a random variable $M_\infty$, because $\tau$ can take the value $+\infty$.
**Def. (hitting times)**
One useful application of Prop. 2.2 is the computation of the Laplace transforms of certain *hitting times*. More precisely, let $W = (W_t)_{t\geq 0}$ be a BM and define for $a > 0, b > 0$ the stopping times

$$\tau_a := \inf\{t \geq 0 \mid W_t > a\},$$
$$\sigma_{a,b} := \inf\{t \geq 0 \mid W_t > a + bt\}.$$

## 3.3 Density processes/Girsanov's theorem

**Density in discrete time**

- Assume $(\Omega, \mathcal{F})$ and a filtration $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,\dots,T}$ in finite discrete time.
  On $(\Omega, \mathcal{F})$, we have two probability measures $Q$ and $P$, and we assume $Q \approx P$.

- **Radon-Nykodin theorem:** There exists a density

  $$\frac{dQ}{dP} := \mathcal{D}$$

  This is a RV $\mathcal{D} > 0$, $P$-a.s. s.t. for all $A \in \mathcal{F}_k$ and for all RVs $Y \geq 0$ it holds that:

  $$Q[A] = E_P[\mathcal{D} I_A] \iff E_Q[Y] = E_P[Y\mathcal{D}].$$

- This can also be written as

  $$\int_\Omega Y dQ = \int_\Omega Y \mathcal{D} dP$$

  This formula tells us how to compute $Q$-expectations in terms of $P$-expectations and vice-versa.

**Density process in discrete time**

- Assume the same setting as before.
- **Radon-Nykodin theorem:** The density process $Z$ of $Q$ w.r.t. $P$, or also called the $P$-martingale $Z$, is defined as

  $$Z_k := E_P[\mathcal{D}|\mathcal{F}_k] = E_P\left[\frac{dQ}{dP}\middle|\mathcal{F}_k\right] \quad \text{for } k = 0, 1, \dots, T$$

- Then for every $\mathcal{F}_k$-measurable RV $Y \geq 0$ or $Y \in \mathcal{L}^1(Q)$, it holds that

  $$E_Q[Y|\mathcal{F}_k] = E_P[Y Z_k|\mathcal{F}_k]$$

  and for every $k \in \{0, 1, \dots, T\}$ and any $A \in \mathcal{F}_k$, it holds that

  $$Q[A] = E_P[Z_k I_A]$$

- Properties:
  - $Z_k$ is a RV and $Z_k > 0$, $P$-a.s.
  - A process $N = (N_k)_{k=0,1,\dots,T}$ which is adapted in $\mathbb{F}$ is a $Q$-martingale iff the product $ZN$ is a $P$-martingale.
    (This tells us how martingale properties under $P$ and $Q$ are related to each other.)

- **Bayes formula:**
  If $j \leq k$ and $U_k$ is $\mathcal{F}_k$-measurable and either $\geq 0$ or in $\mathcal{L}^1(Q)$, then

  $$E_Q[U_k|\mathcal{F}_j] = \frac{1}{Z_j} E_P[Z_k U_k|\mathcal{F}_j] \qquad Q\text{-a.s.}$$

  This tells us how conditional expectations under $Q$ and $P$ are related to each other.

**Lem. 3.1**

(i) For every $k \in \{0, 1, ..., T\}$ and any $A \in \mathcal{F}_k$ or any $\mathcal{F}_k$-measurable random variable $Y \geq 0$ or $Y \in L^1(Q)$, we have

$$Q[A] = E_P[Z_k I_A] \quad \Leftrightarrow \quad E_Q[Y] = E_P[Z_k Y],$$

respectively. *(This means that $Z_k$ is the density of $Q$ w.r.t. $P$ on $\mathcal{F}_k$.)*

(ii) If $j \leq k$ and $U_k$ is $\mathcal{F}_k$-measurable and either $\geq 0$ or in $L^1(Q)$, then we have the *Bayes formula*

$$E_Q[U_k|\mathcal{F}_j] = \tfrac{1}{Z_j} E_P[Z_k U_k|\mathcal{F}_j] \quad Q\text{-a.s.}$$

*(This tells us how conditional expectations under $Q$ and $P$ are related to each other.)*
Written in terms of $D_k$, the Bayes formula for $j = k-1$ becomes

$$E_Q[U_k|\mathcal{F}_{k-1}] = E_P[D_k U_K|\mathcal{F}_{k-1}].$$

This shows that the rations $D_k$ play the role of "one-step conditional densities" of $Q$ with respect to $P$.

(iii) A process $N = (N_k)_{k=0,1,...,T}$ which is adapted to $\mathbb{F}$ is a $Q$-martingale iff the product $ZN$ is a $P$-martingale. *(This tells us how martingale properties under $P$ and $Q$ are related to each other.)*

**Proof Lem. 3.1**

(i)

$$E_Q[Y] = E_P[Y\mathcal{D}] \overset{(vi)}{=} E_P[E_P[Y\mathcal{D}|\mathcal{F}_k]]$$
$$\overset{Y\ \mathcal{F}_k\text{-meas.}}{=} E_P[Y E_P[\mathcal{D}|\mathcal{F}_k]] = E_P[Y Z_k]$$

□

(ii) a) LHS:

$$E_Q[N_k|\mathcal{F}_j] \overset{\text{if } Q\text{-martingale}}{=} N_j \quad \Rightarrow \quad N_j Z_j = E_P[N_k Z_k|\mathcal{F}_j]$$
$$\Rightarrow \quad NZ \ \ P\text{-martingale}$$

b) RHS:

$$\frac{1}{Z_j} E_P[N_k Z_k|\mathcal{F}_j] \overset{\text{if } P\text{-martingale}}{=} \frac{1}{Z_j} N_j Z_j$$
$$= N_j, \text{i.e. } E_Q[N_k|\mathcal{F}_j] = N_j$$
$$\Rightarrow \quad N \text{ is } Q\text{-martingale}$$

This concludes the proof for $(ii)$ □

**Def.**

$$D_k := \tfrac{Z_k}{Z_{k-1}}, \quad \text{for } k = 1, ..., T.$$

The process $D$ is adapted, strictly positive and satisfies by its definition

$$E_P[D_k|\mathcal{F}_{k-1}] = 1,$$

because $Z$ is a $P$-martingale.
**Rem.**
Again because $Z$ is a martingale and by Lem. 3.1,

$$E_P[Z_0] = E_P[Z_T] = E_P[Z_T I_\Omega] \overset{\text{Lem. 3.1 1)}}{=} Q[\Omega] = 1,$$

and we can recover $Z$ from $Z_0$ and $D$ via

$$Z_k = Z_0 \prod_{j=1}^k D_j, \quad \text{for } k = 0, 1, ..., T.$$

**Rem.**
To *construct an equivalent martingale measure* for a given process $S$, all we need are an $\mathcal{F}_0$-measurable random variable $Z_0 > 0$ $P$-a.s. with $E_P[Z_0] = 1$ and an adapted strictly positive process $D = (D_k)_{k=1,\dots,T}$ satisfying $E_P[D_k|\mathcal{F}_{k-1}] = 1$ for all $k$, and in addition $E_P[D_k(S_k - S_{k-1})|\mathcal{F}_{k-1}] = 0$ for all $k$.
**Def.** (i.i.d. returns)

$$\widetilde{S}_k^1 = S_0^1 \prod_{j=1}^{k} Y_j, \quad \widetilde{S}_k^0 = (1+r)^k,$$

where $Y_1, \dots, Y_T$ are $> 0$ and i.i.d. under $P$. **Rem.** (*construction of an EMM $Q$*)
Choose $D_k$ like $Y_k$ independend of $\mathcal{F}_{k-1}$. Then we must have $D_k = g_k(Y_k)$ for some measurable function $g_k$, and we have to choose $g_k$ in such a way that we get

$$1 = E_P[D_k|\mathcal{F}_{k-1}] = E_P[g_k(Y_k)]$$

and

$$1 + r = E_P[D_k Y_k|\mathcal{F}_{k-1}] = E_P[Y_k g_k(Y_k)].$$

(Note that these calculations both exploit the $P$ independence of $Y_k$ from $\mathcal{F}_{k-1}$.) If this choice is possible, we can then choose all the $g_k \equiv g_1$, because the $Y_k$ are (assumed) i.i.d. under $P$. To ensure that $D_k > 0$, we can impose $g_k > 0$.

### Density process in continuous time

- Suppose we have $P$ and a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$.
  Fix $T \in (0,\infty)$ and assume only that $Q \approx P$ on $\mathcal{F}_T$.
- If we have this for every $T < \infty$, we call $Q$ and $P$ *locally equivalent* and write $Q \overset{\text{loc}}{\approx} P$.
  For an infinite horizon, this is usually strictly weaker tahtn $Q \approx P$.

**Def. (density process)**
Let us for simplicity fix $T \in (0,\infty)$ and suppose that $Q \approx P$ on $\mathcal{F}_T$. Denote by

$$Z_t := E_P\left[\frac{dQ|_{\mathcal{F}_T}}{dP|_{\mathcal{F}_T}}\middle|\mathcal{F}_t\right] \qquad \text{for } 0 \leq t \leq T$$

the *density process* of $Q$ w.r.t. $P$ on $[0,T]$, choosing an RCLL version of this $P$-martingale on $[0,T]$.
**Rem.**
Since $Q \approx P$ on $\mathcal{F}_T$, we have $Z > 0$ on $[0,T]$, and because $Z$ is a $P$-(super)martingale, this implies that also $Z_- > 0$ on $[0,T]$ by the so-called *minimum principle for supermartingales*.
**Lem. 2.1**
Suppose that $Q \approx P$ on $\mathcal{F}_T$. Then

(i) For $s \leq t \leq T$ and every $U_t$ which is $\mathcal{F}_t$-measurable and either $\geq 0$ or in $L^1(Q)$, we have the *Bayes formula*

$$E_Q[U_t|\mathcal{F}_s] = \frac{1}{Z_s} E_P[Z_t U_t|\mathcal{F}_s] \qquad Q\text{-a.s.}$$

(ii) An adapted process $Y = (Y_t)_{0\leq t\leq T}$ is a (local) $Q$-martingale iff the product $ZY$ is a (local) $P$-martingale.

**Rem.**

- If $Q \overset{\text{loc}}{\approx} P$, we can use Lem. 2.1 for any $T < \infty$ and hence obtain a statement for processes $Y = (Y)_{t\geq 0}$ on $[0,\infty)$.
- One consequence of part 2) of Lem. 2.1 is also that $\frac{1}{Z}$ is a $Q$-martingale, on $[0,\infty]$ if $Q \approx P$ on $\mathcal{F}_T$, or even on $[0,\infty)$ if $Q \overset{\text{loc}}{\approx} P$.

**Rem.**
Suppose that $Q \overset{\text{loc}}{\approx} P$ with density process $Z$, then $Q \overset{\text{loc}}{\approx} P$ implies that $Z$ is a local martingale.

**Thm. 2.2 (Girsanov)**

(i) Suppose that $\mathbb{Q} \overset{\text{loc}}{\approx} \mathbb{P}$ with density process $Z$. If $M$ is a local $\mathbb{P}$-martingale null at 0, then

$$\widetilde{M} := M - \int \frac{1}{Z} d[Z,M]$$

is a local $\mathbb{Q}$-martingale null at 0.

(ii) In particular, every $\mathbb{P}$-semimartingale is also a $\mathbb{Q}$-semimartingale (and vice-versa, by symmetry).

**Def. (product rule)**

- $ZM = \int Z_- dM + \int M_- dZ + [Z,M]$
- $d(ZM) = Z_- dM + M_- dZ + d[Z,M]$

**Lem.**
Let $Z, M$ be of finite variation, then it holds

$$\langle M\rangle, \langle Z\rangle \quad \Rightarrow \quad \langle M, Z\rangle$$

is also of finite variation.
**Rem.**
When $[Z,M]$ is of finite variation, the following holds

$$\left[\frac{1}{Z}, [Z,M]\right] = \sum \Delta\left(\frac{1}{2}\right)\Delta[Z,M] = \int \Delta\left(\frac{1}{Z}\right) d[Z,M].$$

**Def.** (alternative stochastic Itô integral definition)
$\forall V$ "that are nice enough"

$$\langle \textstyle\int v dW, V\rangle = \textstyle\int v d\langle W, V\rangle$$

**Thm. 2.3 (Girsanov (continuous version))**

- Suppose that $\mathbb{Q} \overset{\text{loc}}{\approx} \mathbb{P}$ with *continuous* density process $Z$. Write $Z = Z_0 \mathcal{E}(L)$. If $M$ is a local $\mathbb{P}$-martingale null at 0, then

  $$\widetilde{M} := M - [L,M] = M - \langle L, M\rangle$$

  is a local $\mathbb{Q}$-martingale null at 0.
- Moreover, if $W$ is a $\mathbb{P}$-BM, then $\widetilde{W}$ is a $\mathbb{Q}$-BM.
- In particular, if $L = \int \nu dW$ for some $\nu \in L^2_{\text{loc}}(W)$, then $\widetilde{W} = W - \langle \int \nu dW, W\rangle = W - \int \nu_s ds$ so that the $\mathbb{P}$-BM $W = \widetilde{W} + \int \nu_s ds$ becomes under $Q$ a BM with (instantaneous) drift $\nu$.
- If we have a closer look at $W^*$ defined in the Black-Scholes chapter, we see that $W^*$ is a Brownian motion under the probability measure $Q^*$ given by

  $$\frac{dQ^*}{dP} := \mathcal{E}\left(-\int \lambda dW\right)_T = \exp\left(-\lambda W_T - \frac{1}{2}\lambda^2 T\right),$$

  whose density process w.r.t. $P$ is

  $$Z_t^* = \mathcal{E}\left(-\int \lambda dW\right)_t = \exp\left(-\lambda W_t - \frac{1}{2}\lambda^2 T\right) \quad \text{for } 0 \leq t \leq T.$$

# 4 Stochastic Integration and Calculus

## 4.1 Brownian motion

### Brownian motion

**Preliminery**
Throughout this chapter, we work on a probability space $(\Omega, \mathcal{F}, P)$. In particular, $\Omega$ cannot be finite or countable. A filtration $\mathbb{F} = (\mathcal{F}_t)$ in continuous time; this is like in discrete time a family of $\sigma$-fields $\mathcal{F}_t \subseteq \mathcal{F}$ with $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s \leq t$. The time parameter runs either through $t \in [0, T]$ with fixed time horizon $T \in (0, \infty)$ or through $t \in [0, \infty)$. In the later case, we define

$$\mathcal{F}_\infty := \bigvee_{t \geq 0} \mathcal{F}_t := \sigma\left(\bigcup_{t \geq 0} \mathcal{F}_t\right).$$

For technical reasons, we should also assume that $\mathbb{F}$ satisfies the so-called usual conditions of being right-continuous (RC) and $P$-complete.

**Def. (Brownian motion)**
A Brownian motion w.r.t. $P$ and a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ is a (real-valued) stochastic process $W = (W_t)_{t \geq 0}$ which [(BM0)] is adapted to $\mathbb{F}$, starts at $0$ (i.e. $W_0 = 0$ $P$-a.s.) and satisfies the following properties:

(BM0) *null at zero*
$W$ is adapted to $\mathbb{F}$ and null at 0 (i.e. $W_0 \equiv 0$, $P$-a.s.).

(BM1) *independent and stationary increments*
For $s \leq t$, the increment $W_t - W_s$ is independent (under $P$) of $\mathcal{F}_s$ and satisfies under $P$: $W_t - W_s \sim \mathcal{N}(0, t-s)$).

(BM2) *continuous sample paths*
$W$ has continuous trajectories, i.e. for $P$-a.a. $\omega \in \Omega$, the function $t \to W_t(\omega)$ on $[0, \infty)$ is continuous.

Remarks:

- Brownian motion in $\mathbb{R}^m$ is simply an adapted $\mathbb{R}^m$-valued stochastic process null at 0 and with the increment $W_t - W_s$ having the normal distribution $\mathcal{N}(0, (t-s)I_{m \times m})$, where $I_{m \times m}$ denotes the identity matrix.
- *Skript version:* Brownian motion in $\mathbb{R}^m$ is simply an adapted $\mathbb{R}^m$-valued stochastic process null at $0$ with (BM2) and s.t. (BM1) holds with $\mathcal{N}(0, t-s)$ replaced by $\mathcal{N}(0, (t-s)I_{m \times m})$, where $I_{m \times m}$ denotes the $m \times m$ identity matrix.

**Def.** (alternative definition of BM without any filtration)
There is also a definition of BM without any filtration $\mathbb{F}$. This is a (real-valued) stochastic process $W = (W_t)_{t \geq 0}$ which starts at $0$, satisfies (BM2) and instead of (BM1) the following property:

**(BM1′)** For any $n \in \mathbb{N}$ and any times $0 = t_0 < t_1 < \ldots < t_n < \infty$, the increments $W_{t_i} - W_{t_{i-1}}, i = 1, \ldots, n$, are independent (under $P$) and we have (under $P$) that $W_{t_i} - W_{t_{i-1}} \sim \mathcal{N}(0, t_i - t_{i-1})$, or $\mathcal{N}(0, (t_i - t_{i-1})I_{m \times m})$ if $W$ is $\mathbb{R}^m$-valued.

Instead of (BM1′), one also says (in words) that $W$ has *independent stationary increments* with a (specific) normal distribution.

### Transformations of BM

**Prop. 1.1**
Suppose $W = (W_t)_{t \geq 0}$ is a BM. Then:

(i) $W^1 := -W$ is a BM.

(ii) Restarting at a fixed time $T$:
$$W_t^2 := W_{T+t} - W_T$$
for $t \geq 0$ is a BM for any $T \in (0, \infty)$.

(iii) Rescaling in space and time:
$$W_t^3 := cW_{\frac{t}{c^2}}$$
for $t \geq 0$ is a BM for any $c \in \mathbb{R}$, $c \neq 0$.

(iv) Time-reversal:
$$W_t^4 := W_{T-t} - W_T$$
for $0 \leq t \leq T$ is a BM on $[0, T]$ for any $T \in (0, \infty)$.

(v) Inversion of small and large times:
$$W_t^5 := \begin{cases} tW_{\frac{1}{t}} & \text{for } t > 0 \\ 0 & \text{for } t = 0 \end{cases}$$
for $t \geq 0$ is a BM.

(vi) $W_t^6 := (W_t)^2 - t = 2\int_0^t W_s dW_s, \qquad t \geq 0$

(vii) $W_t^7 := \exp\left(\alpha W_t - \frac{1}{2}\alpha^2 t\right)$ for $t \geq 0$ and for any $\alpha \in \mathbb{R}$.

Note that we always use here the definition of BM without an exogenous filtration.

### Laws on BM

The next result gives some information about how *trajectories of BM* behave.

**Prop. 1.2**
Suppose $W = (W_t)_{t \geq 0}$ is a BM. Then:

(i) Law of large numbers:
$$\lim_{t \to \infty} \frac{W_t}{t} = 0, \qquad P\text{-a.s.}$$
i.e. BM grows asymptotically less than linearly (as $t \to \infty$).

(ii) (Global) law of the iterated logarithm (LIL):
With $\psi_{\text{glob}}(t) := \sqrt{2t \log(\log t)}$, it holds for $t \geq 0$ that:
$$\lim_{t \to \infty} \sup \frac{W_t}{\psi_{\text{glob}}(t)} = +1, \quad P\text{-a.s.}$$
$$\lim_{t \to \infty} \inf \frac{W_t}{\psi_{\text{glob}}(t)} = -1, \quad P\text{-a.s.}$$
i.e. for $P$-a.a. $\omega$, the function $t \mapsto W_t(\omega)$ for $t \to \infty$ oscillates precisely between $t \mapsto \pm\psi_{\text{glob}}(t)$.

(iii) (Local) law of the iterated logarithm (LIL):
With $\psi_{\text{loc}}(h) := \sqrt{2h \log(\log \frac{1}{h})}$, it holds for $t \geq 0$ that:
$$\lim_{h \searrow 0} \sup \frac{W_{t+h} - W_t}{\psi_{\text{loc}}(h)} = +1, \quad P\text{-a.s.}$$
$$\lim_{h \searrow 0} \inf \frac{W_{t+h} - W_t}{\psi_{\text{loc}}(h)} = -1, \quad P\text{-a.s.}$$
i.e. for $P$-a.a. $\omega$, to the right of $t$, the trajectory $u \mapsto W_u(\omega)$ around the level $W_t(\omega)$ oscillates precisely between $h \mapsto \pm\psi_{\text{loc}}(h)$.

**Prop. 1.3**
Suppose $W = (W_t)_{t \geq 0}$ is a BM. Then for $P$-a.a. $\omega \in \Omega$, the function $t \mapsto W_t(\omega)$ from $[0, \infty)$ to $\mathbb{R}$ is continuous, but *nowhere differentiable.*

**Def. (partition)**
Call a partition of $[0, \infty)$ any set $\Pi \subseteq [0, \infty)$ of time points with $0 \in \Pi$ and $\Pi \cap [0, T]$ finite for all $T \in [0, \infty)$. This implies that $\Pi$ is at most countable and can be ordered increasingly as $\Pi = \{0 = t_0 < t_1 < \ldots < t_m < \ldots < \infty\}$.
**Def. (mesh size, sum)**
The mesh size of $\Pi$ is then defined as $|\Pi| := \sup\{t_i - t_{i-1} \mid t_{i-1}, t_i \in \Pi\}$, i.e. the size of the biggest time-step in $\Pi$. For any partition $\Pi$ of $[0, \infty)$,
$$Q_T^\Pi := \sum_{t_i \in \Pi} (W_{t_i \wedge T} - W_{t_{i-1} \wedge T})^2$$
is then the *sum* up to time $T$ of the *squared increments* of BM along $\Pi$.
We expect, at least for $|\Pi|$ very small so that time points are close together, that $(W_{t_i \wedge T} - W_{t_{i-1} \wedge T})^2 \approx t_i \wedge T - t_{i-1} \wedge T$ and hence
$$Q_T^\Pi \approx \sum_{t_i \in \Pi} t_i \wedge T - t_{i-1} \wedge T2 \quad \text{for } |\Pi| \text{ small.}$$

**Thm. 1.4** (quadratic variation)
Suppose $W = (W_t)_{t \geq 0}$ is a BM. For any sequence $(\Pi_n)_{n \in \mathbb{N}}$ of partitions of $[0, \infty)$ which is refining (i.e. $\Pi_n \subseteq \Pi_{n+1} \forall n$) and satisfies $\lim_{n \to \infty} |\Pi_n| = 0$, we then have
$$P\left[\lim_{n \to \infty} Q_t^{\Pi_n} = t, \quad \text{for every } t \geq 0\right] = 1.$$

We express this by saying that along $(\Pi_n)_{n\in\mathbb{N}}$, the BM $W$ has (with probability 1) *quadratic variation* $t$ on $[0,t]$ for every $t \geq 0$, and we write $\langle W\rangle_t = t$.
(We sometimes also say that $P$-a.a. trajectories $W_\bullet(\omega) : [0,\infty) \to \mathbb{R}$ of BM have quadratic variation $t$ on $[0,t]$, for each $t \geq 0$.)

**Def. (finite variation)**
A function $g : [0,\infty) \to \mathbb{R}$ is of *finite variation* or has *finite 1-variation* if for every $T \in (0,\infty)$,

$$\sup_\Pi \sum_{t_i\in\Pi} |g(t_i \wedge T) - g(t_{i-1} \wedge T)| < \infty,$$

where the supremum is taken over all partitions $\Pi$ of $[0,\infty)$.

**Interpretation**
The interpretation is that the graph of $g$ has finite length on any time interval. More precisely, if we define the *arc length* of (the graph of) $g$ on the interval $[0,T]$ as

$$\sup_\Pi \sum_{t_i\in\Pi} \sqrt{(t_i \wedge T - t_{i-1} \wedge T)^2 + (g(t_i \wedge T) - g(t_{i-1} \wedge T))^2},$$

with the supremum again taken over all partitions $\Pi$ of $[0,\infty)$, then $g$ has finite variation on [0,T] iff it has finite arc length on $[0,T]$.

**Rem.**
Any monotonic (increasing or decreasing) function is clearly of finite variation, because the absolute values above disappear and we get a telescoping sum. Moreover, one can show that any function of finite variation can be written as the difference of two increasing functions (and vice versa).

**Rem.**
*Every continuous function $f$ which has nonzero quadratic variation along a sequence $(\Pi_n)$ as above must have infinite variation*, i.e. unbounded oscillations. We also recall that a classical result due to Lebesgue says that any function of finite variation is almost everywhere differentiable. So Prop. 1.3 implies that Brownian trajectories must have infinite variation, and Thm 1.4 makes this even quantitative.

**Prop. 2.2**
Suppose $W = (W_t)_{t\geq 0}$ is a $(P,\mathbb{F})$-Brownian motion. Then the following process are all $(P,\mathbb{F})$-martingales:

(i) $W$ itself.

(ii) $W_t^2 - t, t \geq 0$.

(iii) $e^{\alpha W_t - \frac{1}{2}\alpha^2 t}, t \geq 0$, for any $\alpha \in \mathbb{R}$.

**Prop. 2.3**
Let $W$ be a BM an $a > 0, b > 0$. Then for any $\lambda > 0$, we have

(i) $E[e^{-\lambda\tau_a}] = e^{-a\sqrt{2\lambda}}$

(ii) $E[e^{-\lambda\sigma_{a,b}}] = E[e^{-\lambda\sigma_{a,b}} I_{\sigma_{a,b}<\infty}] = e^{-a(b+\sqrt{b^2+2\lambda})}$.

**Rem.**

- In the proof of the above Prop. 2.3 the following was being used:

$$M_t := e^{\alpha W_t - \frac{1}{2}\alpha^2 t}, t \geq 0.$$

- For a general random variable $U \geq 0$, the function $\lambda \mapsto E[e^{-\lambda U}]$ for $\lambda > 0$ is called the *Laplace transform* of $U$.
- In mathematical finance, both $\tau_a$ and $\sigma_{a,b}$ come up in connection with a number of so-called *exotic options*. In particular, they are important for *barrier options* wohse payoff depends on whether or not a (upper or lower) level has been reached by a given time. When computing prices of such options in the Black-Scholes model, one almost immediately encounters the Laplace transforms in Prop. 2.3.

## 4.2 Markovian properties

**Rem.**
We have already seen in part 2) of Prop. 1.1 that for any fixed time $T \in (0,\infty)$, the process

$$W_{t+T} - W_T, t \geq 0, \quad \text{is again a BM}$$

if $(W_t)_{t\geq 0}$ is a BM. Moreover, one can show that the independence of increments of BM implies that

$$W_{t+T} - W_T, t \geq 0, \quad \text{is independent of } \mathcal{F}_T^0,$$

where $\mathcal{F}_T^0 = \sigma(W_s; s \leq T)$ is the $\sigma$-field generated by BM up to time $T$.

**Intuition**
Intuitively, this means that BM at any fixed time $T$ simply forgets its past up to time $T$ (with the only possible exception that it remembers its current position $W_T$ at time $T$), and starts afresh.

**Def. (Markov property)**
One consequence of the above remark is the following. Suppose that at some fixed time $T$, we are interested in the behaviour of $W$ after time $T$ and try to predict this on the basis of the past of $W$ up to time $T$, where "prediction" is done in the sense of a conditional expectation. Then we may as well forget about hte past and look ony at the current value $W_T$ at time $T$. A bit more precisely, we can express this, for functions $g \geq 0$ applied to the part of BM after time $T$, as

$$E[g(W_u; u \geq T) \mid \sigma(W_s; s \leq T)] = E[g(W_u; u \geq T) \mid \sigma(W_T)].$$

This is called the *Markov property* of BM.

**Rem.**

- BM has even the *strong Markov property*.
- **Generalised:** If we denote almost as above by $\mathbb{F}^W$ the filtration generated by $W$ (and made right-continuous, to be accurate), and if $\tau$ is a stopping time w.r.t. $\mathbb{F}^W$ and s.t. $\tau < \infty$ $P$-a.s., then

$$W_{t+\tau} - W_\tau, t \geq 0, \quad \text{is again a BM and independent of } \mathcal{F}_\tau^W.$$

- This includes the first remark in this subsection as special case and one can easily believe that it is even more useful than the above definition of Markov property.

## 4.3 Poisson processes

**Poisson processes**

- A Poisson process $N = (N_t)_{t\geq 0}$ with parameter $\lambda \in \mathbb{R}, \lambda > 0$ and w.r.t. $(\mathbb{P},\mathbb{F})$ is a real-valued stochastic process satisfying the following properties:

  (PP0) *null at zero*
  $N$ is adapted to $\mathbb{F}$ and null at 0 (i.e. $N_0 \equiv 0, \mathbb{P}$-a.s.).

  (PP1) *independent and stationary increments*
  For $0 \leq s < t$, the increment $N_t - N_s$ is independent (under $\mathbb{P}$) of $\mathcal{F}_s$ and follows (under $\mathbb{P}$) the Poisson distribution with parameter $\lambda(t-s)$, i.e. $N_t - N_s \sim \mathrm{Poi}(\lambda(t-s))$, i.e

  $$\mathbb{P}[N_t = k] = \frac{\lambda^k (t-s)^k}{k!} e^{-\lambda(t-s)}$$

  (PP2) *counting process*
  $N$ is a counting process with jumps of size 1, i.e. for $\mathbb{P}$-a.a. $\omega$, the function $t \mapsto N_t(\omega)$ is RCLL, piecewise constant and $\mathbb{N}_0$-valued, and increases by jumps of size 1.

- Important properties of Poisson processes: if $X \sim \mathrm{Poi}(\lambda)$, then:

  $$\mathbb{E}[X] = \lambda, \qquad \mathrm{Var}[X] = \lambda$$

  The quadratic variation of a Poisson process equals itself, i.e.

  $$[N]_t = N_t$$

- Examples of Poisson processes: the following Poisson processes are $(\mathbb{P},\mathbb{F})$-martingales:
  - **Compensated Poisson process:**

    $$\tilde{N}_t = N_t - \lambda t, \qquad t \geq 0$$

  - **Geometric Poisson process:**

    $$S_t = \exp\left(N_t \log(1+\sigma) - \lambda\sigma t\right), \qquad t \geq 0$$

    where $\sigma \in \mathbb{R}, \sigma > -1$.
  - Two cases of squared compensated Poisson processes:

    $$\left(\tilde{N}_t\right)^2 - N_t, \qquad \left(\tilde{N}_t\right)^2 - \lambda t, \qquad t \geq 0$$

    It follows that $[\tilde{N}]_t = N_t$.

## 4.4 Stochastic integration

**Def.** (recall discrete *stochastic integral*)
The trading gains or losses from a self-financing strategy $\varphi \widehat{=} (V_0, \vartheta)$ are described by the *stochastic integral*

$$G(\vartheta) = \vartheta \cdot S = \int \vartheta dS = \sum_j \vartheta_j^{\text{tr}} \Delta S_j = \sum_j \vartheta_j^{\text{tr}} (S_j - S_{j-1}).$$

**Rem.**

- Our goal in this section is to construct a stochastic integral process $H \cdot M = \int HdM$ when $M$ is a (real-valued) local martingale null at 0 and $H$ is a (real-valued) predictable process with a suitable integrability property (relative to $M$).
- To relate notation-wise to previous chapters we will the following notions:
  (i) $H \widehat{=} \vartheta$
  (ii) $M \widehat{=} S$

**Rem.**
*Throughout this chapter*, we work on a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathbb{F}(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions of right-continuity and $P$-completeness. If needed, we define $\mathcal{F}_\infty := \bigvee_{t\geq 0} \mathcal{F}_t$.
We also fix a (real-valued) *local martingale* $M = (M_t)_{t\geq 0}$ null at 0 and having RCLL trajectories. Since we want to define stochastic integrals $\int HdM$ and these are always over half-open intervals of the form $(a, b]$ with $0 \leq a < b \leq \infty$, the value of $M$ at 0 is irrelevant and it is enough to look at processes $H = (H_t)$ defined for $t > 0$.
**Def. (jump)**

(i) For any process $Y = (Y_t)_{t\geq 0}$ with RCLL trajectories, we denote by

$$\Delta Y_t := Y_t - Y_{t-} := Y_t - \lim_{s \nearrow t} Y_s$$

the jump of $Y$ at time $t > 0$.

(ii) $\Delta(\int H_r^2 d[M]_r)_t := \int_0^t H_r^2 d[M]_r - \lim_{s\nearrow t} \int_0^s H_r^2 d[M]_r$

**Optional quadratic variation/square bracket process**

- For any local martingale $M = (M_t)_{t\geq 0}$ null at 0, there exists a unique adapted increasing RCLL process $[M] = ([M]_t)_{t\geq 0}$ and having the property that $M^2 - [M]$ is also a local martingale.
- This process can be obtained as the quadratic variation of $M$ in the following sense.
  There exists a sequence $(\Pi_n)_{n\in\mathbb{N}}$ of partitions $[0, \infty)$ with $|\pi| \to 0$ as $n \to \infty$ s.t.

  $$P\left[ [M]_t(\omega) = \lim_{n\to\infty} \sum_{t_i \in \Pi_n} \left(M_{t_i \wedge t}(\omega) - M_{t_{i-1}\wedge t}(\omega)\right)^2 \forall t \geq 0 \right] = 1$$

  We call $[M]$ the optional quadratic variation or square bracket process of $M$.
- If $M$ satisfies $\sup_{0\leq s\leq t} |M_s| \in L^2$ for each $t \geq 0$ (and hence is in particular a martingale), then $[M]$ is integrable (i.e. $[M]_t \in L^1$ for every $t \geq 0$) and $M^2 - [M]$ is a martingale.

**Thm. 1.1**

(i) For any local martingale $M = (M_t)_{t\geq 0}$ null at 0, there exists a unique adapted increasing RCLL process $[M] = ([M]_t)_{t\geq 0}$ null at 0 with $\Delta[M] = (\Delta M)^2$ [*this is an important property*] and having the property $M^2 - [M]$ is also a local martingale.

(ii) This process can be obtained as the quadratic variation of $M$ in the following sense: There exists a sequence $(\Pi_n)_{n\in\mathbb{N}}$ of partitions of $[0, \infty)$ with $|\Pi_n| \to 0$ as $n \to \infty$ s.t.
*script:*

$$P\left[ [M]_t(\omega) = \lim_{n\to\infty} \sum_{t_i\in\Pi_n} (M_{t_i\wedge t}(\omega) - M_{t_{i-1}\wedge t}(\omega))^2, \ \forall t \geq 0 \right] = 1.$$

*lecture:*

$$P\left[ \{\omega : \exists [M]_t(\omega) = \lim_{n\to\infty} \sum_{t_i\in\Pi_n} (M_{t_i\wedge t}(\omega) - M_{t_{i-1}\wedge t}(\omega))^2, \ \forall t \geq 0\} \right] = 1.$$

We call $[M]$ the *optional quadratic variation* or *square bracket process* of $M$.

(iii) If $M$ satisfies $\sum_{0\leq s\leq T} |M_s| \in L^2$ for some $T > 0$ (and hence is in particular a martingale on $[0, T]$), then $[M]$ is integrable on $[0, T]$ (i.e. $[M]_t \in L^1$) and $M^2 - [M]$ is a martingale on $[0, T]$.

**(Optional) covariation process**

- For two local martingales $M, N$ null at 0, we define the (optional) covariation process $[M, N]$ by polarisation, i.e.

  $$[M, N] := \frac{1}{4} ([M + N] - [M - N])$$

- An alternative definition of the covariation process $[M, N]$

  $$[M, N]_t := \lim_{n\to\infty} \sum_{t_i\in\Pi_n} \left(M_{t_i\wedge t}(\omega) - M_{t_{i-1}\wedge t}(\omega)\right) \left(N_{t_i\wedge t}(\omega) - N_{t_{i-1}\wedge t}(\omega)\right), \quad \forall t \geq 0$$

- The operation $[\cdot, \cdot]$ is bilinear.
- From the characterisation of $[M]$ in Thm 1.1, it follows that the operation $[\cdot, \cdot]$ is bilinear, and also that $B = [M, N]$ is the unique adapted RCLL process $B$ null at 0, of finite variation with $\Delta B = \Delta M \Delta N$ and s.t. $MN - B$ is again a local martingale.

**Def. (predictable compensator)**
There exists a unique increasing *predictable* process $\langle M\rangle = (\langle M\rangle_t)_{t\geq 0}$ null at 0 s.t. $[M] - \langle M\rangle$, and therefore also $M^2 - \langle M\rangle$ $[= M^2 - [M] + [M] - \langle M\rangle]$, is a local martingale. The process $\langle M\rangle$ is called the *shapr bracket* (or sometimes the predictable variance) process of $M$.
Note, this might be useful since $[M]$ is not necessarily predictable.
**Rem. and Cor.**

(i) Any adapted process which is continuous is automatically locally bounded and therefore also locally square-integrable.

(ii) If $M$ is *continuous*, then so is $[M]$, because $\Delta[M] = (\Delta M)^2 = 0$. This implies then also that $[M] = \langle M\rangle$. In particular, for a Brownian motion $W$, we have $[W]_t = \langle W\rangle_t = t$ for all $t \geq 0$.

(iii) If bot $M$ and $N$ are locally square-integrable (e.g. if they are continuous), we also get $\langle M, N\rangle$ via polarisation, i.e.

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

(iv) If $M$ is $\mathbb{R}^d$-valued, then $[M]$ becomes a $d \times d$-matrix-valued process with entries $[M]^{ik} = [M^i, M^k]$. To work with that, one needs to establish more properties. The same applies to $\langle M\rangle$, if it exists.
In general the following holds: $[M, M] = [M]$.

**Def.** (**Set of all bounded elementary processes**)

- We denote by $b\mathcal{E}$ the set of all bounded elementary process of the form

  $$H = \sum_{i=0}^{n-1} h_i I_{(t_i, t_{i+1}]}$$

  with $n \in \mathbb{N}$, $0 \leq t_0 < t_1 < \ldots < t_n < \infty$ and each $h_i$ a bounded (real-valued) $\mathcal{F}_{t_i}$-measurable RV.

**Def.** (**Stochastic integral**)

- For any stochastic process $X = (X_t)_{t\geq 0}$, the stochastic integral $\int HdX$ of $H \in b\mathcal{E}$ is defined as

  $$\int_0^t H_s dX_s := H \cdot X_t := \sum_{i=0}^{n-1} h_i \left(X_{t_{i+1}\wedge t} - X_{t_i\wedge t}\right) \quad \text{for } t \geq 0.$$

- If $X$ and $H$ are both $\mathbb{R}^d$-valued, the integral is still real-valued, and we simply replace products by scalar products everywhere.

**Lem 1.2** (**Isometry property**)

- Suppose $M$ is a square-integrable martingale (i.e. $M_t \in L^2$ for all $t \geq 0$).

- For every $H \in b\mathcal{E}$, the stochastic integral process $H \cdot M = \int HdM$ is then also a square-integrable martingale,
  - $[H \cdot M_t] = \left[\int_0^t H_s dX_s\right] = \int_0^t H_s^2 d[X]_s$

  and we have the isometry property

  - $$E\left[(H \cdot M_\infty)^2\right] = E\left[\left(\int_0^\infty H_s dM_s\right)^2\right]$$
  $$= E\left[\left(\sum_{i=0}^{n-1} h_i(M_{t_{i+1}} - M_{t_i})\right)^2\right]$$
  $$\overset{(*)}{=} E\left[\sum_{i=0}^{n-1} h_i^2([M]_{t_{i+1}} - [M]_{t_i})\right]$$
  $$= E\left[\int_0^\infty H_s^2 d[M]_s\right]$$

  Note that the last $d[M]$-integral can be defined $\omega$ by $\omega$, since $t \mapsto [M]_t(\omega)$ is increasing and hence of finite variation. But of course it is here also just a finite sum, because $H$ has such a simple form.

- $(*)$: Below we show the reasoning why this holds.
  From Thm. 1.1 we know:
  $$(M_t^2 - [M]_t)_{t \geq 0}$$
  is a local martingale (here in Lem. 1.2 it is even a martingale)
  $$\Rightarrow \quad E[M_{t_{i+1}}^2 - [M]_{t_{i+1}} | \mathcal{F}_{t_i}] = E[M_{t_{i+1}}^2 | \mathcal{F}_{t_i}] - E[[M]_{t_{i+1}} | \mathcal{F}_{t_i}]$$
  $$\overset{\text{mart. prop.}}{=} M_{t_i}^2 - [M]_{t_i}$$
  $$= 0$$

  $$\overset{M_{t_i}^2, [M]_{t_i}\ \mathcal{F}_{t_i}\text{-meas.}}{\Longleftrightarrow}$$
  $$E[M_{t_{i+1}}^2 - M_{t_i}^2 | \mathcal{F}_{t_i}] - E[[M]_{t_{i+1}} - [M]_{t_i} | \mathcal{F}_{t_i}] = 0$$
  $$E[M_{t_{i+1}}^2 - M_{t_i}^2 | \mathcal{F}_{t_i}] = E[[M]_{t_{i+1}} - [M]_{t_i} | \mathcal{F}_{t_i}].$$

**Rem.**

(i) The argument in the proof of Lem. 1.2 actually shows that the process $(H \cdot M)^2 - \int H^2 d[M]$ is a martingale.

(ii) It is "not very difficult" to argue that
$$\Delta(\textstyle\int H^2 d[M] = (\Delta(H \cdot M))^2 \text{ for } H \in b\mathcal{E},$$
by exploiting that $H$ is piecewise constant and $\Delta[M] = (\Delta M)^2$.

(iii) In view of Thm. 1.1 and the uniqueness there, the combination of these two properties can also be formulated as saying that
$$[H \cdot M] = \left[\int HdM\right] = \int H^2 d[M], \quad \text{for } H \in b\mathcal{E}.$$

(iv) This is the proof for $(ii)$:
$$\Delta\left(\int H_r^2 d[M]_r\right)_t = \int_0^t H_r^2 d[M]_r - \lim_{s \nearrow t} \int_0^s H_r^2 d[M]_r$$
$$= \sum_{i=0}^{n-1} h_i^2([M]_{t_{i+1} \wedge t} - [M]_{t_i \wedge t})$$
$$- \lim_{s \nearrow t} \sum_{i=0}^{n-1} h_i^2([M]_{t_{i+1} \wedge s} - [M]_{t_i \wedge s})$$
$$\overset{t_i \leq s < t \leq t_{i+1}}{=} h_i^2([M]_t - \lim_{s \nearrow t}[M]_s)$$
$$= h_i^2 \Delta[M]_t$$
$$\overset{\Delta[M] = (\Delta M)^2}{=} h_i^2(\Delta M_t)^2$$
$$= (h_i \Delta M_t)^2$$
$$= (h_i(M_t - \lim_{s \nearrow t} M_s))^2$$
$$\overset{t_i \leq s < t \leq t_{i+1}}{=} \left(\sum_{i=0}^{n-1} h_i(M_{t_{i+1} \wedge t} - M_{t_i \wedge t}) - \lim_{s \nearrow t} \sum_{i=0}^{n-1} h_i(M_{t_{i+1} \wedge s} - M_{t_i \wedge s})\right)^2$$
$$= \left(H \cdot M_t - \lim_{s \nearrow t} H \cdot M_s\right)^2$$
$$= (\Delta(H \cdot M)_t)^2$$
□

**Def. (product space)**
$$\overline{\Omega}. = \Omega \times (0, \infty).$$

**Def. (predictable $\sigma$-field)**

(i) We define the *predictable $\sigma$-field* $\mathcal{P}$ on $\overline{\Omega}$ as the $\sigma$-field generated by all adapted left-continuous processes.

(ii) We call a stochastic process $H = (H_t)_{t>0}$ predictable if it is $\mathcal{P}$-measurable when viewed as a mapping $H : \overline{\Omega} \to \mathbb{R}$.
Note, as a consequence, every $H \in b\mathcal{E}$ is then predictable since it is adapted and left-continuous.

**Def.**
We define the (possibly infinite) measure $P_M := P \otimes [M]$ on $(\overline{\Omega}, \mathcal{P})$ by setting
$$E_M[Y] := E\left[\int_0^\infty Y_s(\omega) d[M]_s(\omega)\right], \quad \text{for } Y \geq 0 \text{ predictable;}$$
the inner integral is defined $\omega$-wise as a Lebesgue-Stieltjes integral because $t \mapsto [M]_t(\omega)$ is increasing, null at 0 and RCLL and so can be viewed as the distribution function of a (possibly infinite) $\omega$-dependent measure on $(0, \infty)$.

**Def.**
We introduce the space
$$L^2(M) := L^2(M, P) := L^2(\overline{\Omega}, \mathcal{P}, P_M)$$
$$= \Big\{\text{all (equivalence classes of) predictable } H = (H_t)_{t>0} \text{ s.t.}$$
$$\|H\|_{L^2(M)} := (E_M[H^2])^{\frac{1}{2}} = \left(E\left[\int_0^\infty H_s^2 d[M]_s\right]\right)^{\frac{1}{2}} < \infty\Big\}.$$
(As usual, by taking equivalence classes, we identify $H$ and $H'$ if they agree $P_M$-a.e. (almost everywhere $\widehat{=}$ " a.s.) on $\overline{\Omega}$.)

**Def.**
We define the space $\mathcal{M}_0^2$ as the space of all RCLL martingales $N = (N_t)_{t \geq 0}$ null at 0 which satisfy $\sup_{t \geq 0} E[N_t^2] < \infty$.

**Lem. 1.2 (restate Lem. 1.2. with above notations)**
For a fixed square-integrable martingale $M$, the mapping $H \mapsto H \cdot M$ is linear and gies from $b\mathcal{E}$ to the space $\mathcal{M}_0^2$ of all RCLL martingales $N = (N_t)_{t \geq 0}$ null at 0 which satisfy $\sup_{t \geq 0} E[N_t^2] < \infty$. **Lem. (Doob's inequality)**
The last assertion is true because each $H \cdot M$ remains constant after some $t_n$ given by $H \in b\mathcal{E}$, and because *Doob's inequality* gives for any martingale $N$ and any $t \geq 0$ that
$$E\left[\sup_{0 \leq s \leq t} |N_s|^2\right] \leq 4E[|N_t|^2].$$

**Rem.**

(i) Now the martingale convergence theorem implies that each $N \in \mathcal{M}_0^2$ admits a limit $N_\infty = \lim_{t \to \infty} N_t$ $P$-a.s., and we have $N_\infty \in L^2$ by Fatou's lemma, and the process $(N_t)_{0 \leq t \leq \infty}$ defined up to $\infty$, i.e. on the closed interval $[0, \infty]$, is still a martingale.

(ii) Doob's maximal inequality implies that two martingales $N$ and $N'$ which have the same final value, i.e. $N_\infty = N'_\infty$ $P$-a.s., mus coincide. Therefore we can identify $N \in \mathcal{M}_0^2$ with its limit $N_\infty \in L^2(\mathcal{F}_\infty, P)$, and so $\mathcal{M}_0^2$ becomes a *Hilbert space* with the norm
$$\|N\|_{\mathcal{M}_0^2} = \|N_\infty\|_{L^2} = (E[N_\infty^2])^{\frac{1}{2}}$$
and the scalar product
$$(N, N')_{\mathcal{M}_0^2} = (N_\infty, N'_\infty)_{L^2} = E[N_\infty, N'_\infty].$$

**Cor.**
Because of the above remark the mapping $H \mapsto H \cdot M$ from $b\mathcal{E}$ to $\mathcal{M}_0^2$ is *linear* and an *isometry* because of Lem. 1.2 says that for $H \in b\mathcal{E}$,
$$\|H \cdot M\|_{\mathcal{M}_0^2} = (E[(H \cdot M_\infty)^2])^{\frac{1}{2}}$$
$$= \left(E\left[\int_0^\infty H_s^2 d[M]_s\right]\right)^{\frac{1}{2}}$$
$$= \|H\|_{L^2(M)}.$$

**Rem.**
By general principles, this mapping can therefore be uniquely extended to the closure of $b\mathcal{E}$ in $L^2(M)$. In other words, we can define

a stochastic integral process $H \cdot M$ for every $H$ that can be approximated, w.r.t. the norm $\|\cdot\|_{L^2(M)}$, by processes from $b\mathcal{E}$, and the resulting $H \cdot M$ is again a martingale in $\mathcal{M}_0^2$ and still satisfies the isometry property described in the equation above.
Prop. 1.3
Suppose that $M$ is in $\mathcal{M}_0^2$. Then:

(i) $b\mathcal{E}$ is *dense* in $L^2(M)$, i.e. the closure of $b\mathcal{E}$ in $L^2(M)$ is $L^2(M)$.

(ii) For every $H \in L^2(M)$, the stochastic integral process $H \cdot M = \int HdM$ is well defined, in $\mathcal{M}_0^2$ and satisfies the above equation.

**Rem.**
Let $M \in \mathcal{M}_0^2$, we then have $E[|M_t|^2] < \infty$ for every $t \geq 0$ s.t. every $M \in \mathcal{M}_0^2$ is also a square-integrable martingale.
However, the converse is not true; Brownian motion $W$ for example is a martingale and has $E[W_t^2] = t$. So $\sup_{t\geq 0} E[W_t^2] = \infty$ which means that BM is not in $\mathcal{M}_0^2$.
**Def. (locally square-integrable, stochastic interval)**

(i) We call a local martingale $M$ null at $0$ *locally square-integrable* and write $M \in \mathcal{M}_{0,\text{loc}}^2$ if there is a sequence of stopping times $\tau_n \nearrow \infty$ $P$-a.s. s.t. $M^{\tau_n} \in \mathcal{M}_0^2$ for each $n$.

(ii) We say for a predictable process $H$ that $H \in L^2_{\text{loc}}(M)$ if there exists a sequence of stopping times $\tau_n \nearrow \infty$ $P$-a.s. s.t. $HI_{]\!]0,\tau_n]\!]} \in L^2(M)$ for each $n$. Here we use the *stochastic interval* notation $]\!]0,\tau_n]\!] := \{(\omega,t) \mid 0 < t \leq \tau_n(\omega)\}$.

**Def.** (another definition of **stochastic integral**)

- For $M \in \mathcal{M}_{0,\text{loc}}^2$ and $H \in L^2_{\text{loc}}(M)$, defining the stochastic integral is straightforward, we simply set
$$H \cdot M := (HI_{]\!]0,\tau_n]\!]}) \cdot M^{\tau_n}, \quad \text{on } ]\!]0,\tau_n]\!]$$
which gives a definition on all of $\overline{\Omega}$ since $\tau_n \nearrow \infty$, s.t. $]\!]0,\tau_n]\!]$ increases to $\overline{\Omega}$.
- The only piont we need to check is that this definition is *consistent*, i.e. tht the definition on $]\!]0,\tau_{n+1}]\!] \supseteq ]\!]0,\tau_n]\!]$ does not clash with the defintion on $]\!]0,\tau_n]\!]$. This can be done by using the properties of stachastic integrals.
- Of course, $H \cdot M$ is then in $\mathcal{M}_{0,\text{loc}}^2$.

**Rem.**
If $M$ is $\mathbb{R}^d$-valued with components $M^i$ that are all in $\mathcal{M}_{0,\text{loc}}^2$, one can also define the so-called *vector stochastic integral* $H \cdot M$ for $\mathbb{R}^d$-valued predictable processes in a suitable space $L^2_{\text{loc}}(M)$; the result is then a real-valued process. However, one **warning** is indicated: $L^2_{\text{loc}}(M)$ is not obtained by just asking that each component $H^i$ should be in $L^2_{\text{loc}}(M^i)$ and then setting $H \cdot M = \sum_i H^i \cdot M^i$. In fact, it can happen that $H \cdot M$ is well defined whereas the individual $H^i \cdot M^i$ are not. So the *intuition* for the multidimensional case is that
$$\int HdM = \int \sum_i H^i dM^i \neq \sum_i \int H^i dM^i.$$

**Def. (continuous** local martingale, **locally bounded)**

(i) $M$ is a *continuous* local martingale null at $0$, briefly written as $M \in \mathcal{M}_{0,\text{loc}}^c$.
This includes in particular the case of a Brownian motion $W$.

(ii) Then $M$ is in $\mathcal{M}_{0,\text{loc}}^2$ because it is even *locally bounded*:
For the stopping times
$$\tau_n := \inf\{t \geq 0 \mid |M_t| > n\} \nearrow is\infty \quad P\text{-a.s.},$$
We have by continuity that $|M^{\tau_n}| \leq n$ for each $n$, because
$$|M_t^{\tau_n}| = |M_{t\wedge\tau_n}| = \begin{cases} |M_t| \leq n, & t \leq \tau_n \\ |M_{\tau_n}| = n, & t > \tau_n. \end{cases}$$

(iii) The set $L^2_{\text{loc}}(M)$ of nice integrands for $M$ can here be explicitly described as
$$L^2_{\text{loc}} = \{\text{all predictable processes } H = (H_t)_{t>0} \text{ s.t.} \\ \int_0^t H_s^2 d\langle M\rangle_s < \infty,\ P\text{-a.s. } \forall t \geq 0\}.$$
Finally, the resulting stochastic integral $H \cdot M = \int HdM$ is then, also a continuous local martingale, and of course null at $0$.

**Properties**

- (Local) Martingale properties
  - If $M$ is a local martingale and $H \in L^2_{\text{loc}}(M)$, then $\int HdM$ is a local martingale in $\mathcal{M}_{0,\text{loc}}^2$. If $H \in L^2(M)$, then $\int HdM$ is even a martingale in $\mathcal{M}_0^2$.
  - If $M$ is a local martingale and $H$ is predictable and locally bounded $(*)$, then $\int HdM$ is a local martingale.
  $(*)$ : (which means that there are stopping times $\tau_n \nearrow \infty$ $P$ - a.s. s.t. $HI_{]\!]0,\tau_n]\!]}$ is bounded by a constant $c_n$, say, for each $n \in \mathbb{N}$)
  - If $M$ is a martingale in $\mathcal{M}_0^2$ and $H$ is predictable and bounded, then $\int HdM$ is again a martingale in $\mathcal{M}_0^2$.
  - *Warning:* If $M$ is a martingale and $H$ is predictable and bounded, then $\int HdM$ need not be a martingale; this is in striking contrast to the situation in discrete time.
- Linearity
If $M$ is a local martingale and $H, H'$ are in $L^2_{\text{loc}}(M)$ and $a, b \in \mathbb{R}$, then also $aH + bH'$ is in $L^2_{\text{loc}}(M)$ and
$$(aH+bH')\cdot M = (aH)\cdot M + (bH')\cdot M = a(H\cdot M) + b(H'\cdot M).$$
- Associativity If $M$ is a local martingale and $H \in L^2_{\text{loc}}(M)$, then we already know that $H \cdot M$ is again a local martingale. Then a predictable process $K$ is in $L^2_{\text{loc}}(H\cdot M)$ iff $KH$ is in $L^2_{\text{loc}}(M)$, and then
$$K \cdot (H \cdot M) = (KH) \cdot M,$$
i.e.
$$\int Kd(\textstyle\int HdM) = \int KHdM.$$
- Behaviour under stopping
  - Suppose that $M$ is a local martingale, $H \in L^2_{\text{loc}}(M)$ and $\tau$ is a stopping time. Then $M^\tau$ is a local martingale by the stopping theorem, $H$ is in $L^2_{\text{loc}}(M^\tau)$, $HI_{]\!]0,\tau]\!]}$ is in $L^2_{\text{loc}}(M)$, and we have
  $$(H \cdot M)^\tau = H \cdot (M^\tau) = (HI_{]\!]0,\tau]\!]}) \cdot M = (HI_{]\!]0,\tau]\!]}) \cdot (M^\tau).$$
  - In words: A stopped stochastic integral is computed by either first stopping the integrator and then integrating, or setting the integrand equal to $0$ after the stopping time and then integrating, or combining the two.
- Quadratic variation and covariation
  - Suppose that $M, N$ are local martingales, $H \in L^2_{\text{loc}}(M)$ and $K \in L^2_{\text{loc}}(N)$. Then
  $$\left[\int HdM, N\right] = \int Hd[M,N]$$
  and
  $$\left[\int HdM, \int KdN\right] = \int HKd[M,N].$$
  - The covariation process of two stochastic integrals is obtained by integrating the product of the integrands w.r.t. the covariation process of the integrators.
  - In particular, $[\int HdM] = \int H^2 d[M]$. (We have seen this already for $H \in b\mathcal{E}$ in the remark after Lem. 1.2)
- Jumps
Suppose $M$ is a local martingale and $H \in L^2_{\text{loc}}(M)$. Then we already know that $H \cdot M$ is in $\mathcal{M}_{0,\text{loc}}^2$ and therefore RCLL. Its jumps are given by
$$\Delta\left(\int HdM\right)_t = H_t \Delta M_t, \quad \text{for } t > 0,$$
where $\Delta Y_t := Y_t - Y_{t-}$ again denotes the jump at time $t$ of a process $Y$ with trajectories which are RCLL.

## 4.5 Extension to semimartingales

**Def. (semimartingale, special semimartingale)**

(i) A *semimartingale* is a stochastic process $X = (X_t)_{t\geq 0}$ that can be decomposed as $X = X_0 + M + A$, where $M$ is a local martingale null at $0$ and $A$ is an adapted process null at $0$ and having RCLL trajectories of finite variation.

(ii) A semimartingale $X$ is called *special* if there is such a decomposition where $A$ is in addition predictable.

**Rem. (canonical decomposition, continuous semimartingale, optional quadratic variation)**

(i) If $X$ is a special semimartingale, the decomposition with $A$ predictable is *unique* and called the *canonical decomposition*. The uniqueness result uses that any local martingale which is predictable and of finite variation must be constant.

(ii) If $X$ is a *continuous* semimartingale, both $M$ and $A$ can be chosen continuous as well. Therefore $X$ is special because $A$ is then predictable, since $A$ is adapted and continuous.

(iii) If $X$ is a semimartingale, then we define its *optional quadratic variation* or *square bracket* process $[X] = ([X]_t)_{t\geq 0}$ via

$$[X] := [M] + 2[M, A] + [A] := [M] + 2\sum \Delta M \Delta A + \sum(\Delta A)^2.$$

One can show that this is well defined and does not depend on the chosen decomposition of $X$. Moreover, $[X]$ can also be obtained as a quadratic variation similarly as in Thm. 1.1. However, $X^2 - [X]$ is no longer a local martingale, but only a semimartingale in general.

**Def. (stochastic integral for semimartingale)**
If $X$ is a semimartingale, we can define a stochastic integral $H \cdot X = \int HdX$ at least for any process $H$ which is predictable and locally bounded. We simply set

$$H \cdot X := H \cdot M + H \cdot A,$$

where $H \cdot M$ is as is the previous section and $H \cdot A$ is defined $\omega$-wise as a Lebesgue-Stieltjes integral.

**Properties Rem.**

- The resulting stochastic integral then has all the *properties* from the previous section except those that rest in an essential way on the (local) martingale property.
- The isometry property for example is of course lost.
- We still have, for $H$ predictable and locally bounded:
  - $H \cdot X$ is a semimartingale.
  - If $X$ is special with canonical decomposition $X = X_0 + M + A$, then $H \cdot X$ is also special, with canonical decomposition $H \cdot X = H \cdot M + H \cdot A$.
    (This uses the non-obvious fact that if $A$ is predictable and of finite variation and $H$ is predictable and locally bounded, the pathwise defined integral $H \cdot A$ can be chosen to be predictable again.)
  - *linearity*: same formula as before.
  - *associativity*: same formula as before.
  - *behaviour under stopping*: same formula as before.
  - *quadratic variation and covariation*: same formula as before.
  - *jumps*: same formula as before.
  - If $X$ is *continuous*, then so is $H \cdot X$; this is clear from $\Delta(H \cdot X) = H\Delta X = 0$.

**Thm.** (sort of **dominated convergence theorem, continuity property**)

(i) If $H^n, n \in \mathcal{N}$, are predictable processes with $H^n \to 0$ pointwise on $\overline{\Omega}$ and $|H^n| \leq |H|$ for some locally bounded $H$, then $H^n \cdot X \to 0$ uniformly on compacts in probability, which means that

$$\sup_{0\leq s\leq t} |H^n \cdot X_s| \to 0 \quad \text{in probability as } n \to \infty, \forall t \geq 0.$$

(ii) This can also be viewed as a *continuity property* of the stochastic integral operator $H \mapsto H \cdot X$, since (pointwise and locally bounded) convergence of $(H^n)$ implies convergence of $(H^n \cdot X)$, in the sense of above formula.

**Rem.** (further **properties**)

(i) If $X$ is a semimartingale and $f$ is a $C^2$-function, then $f(X)$ is again a semimartingale.

(ii) If $X$ is a semimartingale w.r.t. $P$ and $R$ is a probability measure equivalent to $P$, then $X$ is a semimartingale w.r.t. $R$.
This will follow from *Girsanov's theorem*, which even gives a decomposition of $X$ under $R$.

(iii) If $X$ is any adapted process with RC trajectories, we can always define the (elementary) stochastic integral $H \cdot X$ for processes $H \in b\mathcal{E}$. If $X$ is s.t. this mapping on $b\mathcal{E}$ also has the continuity property from the above Thm. for any sequence $(H^n)_{n\in\mathbb{N}} \in b\mathcal{E}$ converging pointwise to $0$ and with $|H^n| \leq 1$ for all $n$, then $X$ must in fact be a semimartingale.

**Rem.**
The above result implies that if we start with any model where $S$ is *not a semimartingale*, there will be *arbitrage* of some kind.
**Lem.**
The family of semimartingales is *invariant* under a transformation by a $C^2$-function, i.e. $f(X)$ is a semimartingale whenever $X$ is a semimartingale and $f \in C^2$.

## 4.6 Stochastic calculus

**Good to know**

**Def. (weak convergence of probability measure)**
Suppose $\mu_j$ is a sequence of measures on $\mathbb{R}$. By the definition of weak convergence of measures, $\mu_j$ weak converges to $\mu$ means that for any bounded continuous function $f$, there holds that

$$\int_{\mathbb{R}} f\mu_j \quad \xrightarrow{n\to\infty} \quad \int_{\mathbb{R}} f\mu.$$

**Thm.** (Wiki: **Lebesgue's Dominated Convergence Theorem**)
Let $\{f_n\}$ be a sequence of real-valued measurable functions on a measure space $(S, \Sigma, \mu)$. Suppose that the sequence converges pointwise to a function $f$ and is dominated by some integrable function $g$ in the sense that

$$|f_n(x)| \leq g(x)$$

$\forall n$ in the index set of the sequence and $x \in S$. Then $f$ is integrable and

$$\lim_{n\to\infty} \int_S |f_n - f| d\mu = 0$$

which also implies

$$\lim_{n\to\infty} \int_S f_n d\mu = \int_S f d\mu.$$

**Rem.**
The statement "$g$ is integrable" is meant in the sense of Lebesgue, i.e

$$\int_S |g| d\mu < \infty.$$

**Lem.**
One can use that any continuous local martingale of finite variation is constant.
**Throughout this chapter**
We work on a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)$ satisfying the usual conditions of right-continuity and $P$-completeness. For all local martingales, we then can and tacitly do choose a version with RCLL trajectories. For the time parameter $t$, we have either $t \in [0, T]$ with a fixed time horizon $T \in (0, \infty)$ or $t \geq 0$. In the latter case, we set

$$\mathcal{F}_\infty := \bigvee_{t\geq 0} \mathcal{F}_t := \sigma\left(\bigcup_{t\geq 0} \mathcal{F}_t\right).$$

### Itôs formula I

**Def. (classical chain rule from analysis)**
If $x \in C^1$ and $f : \mathbb{R} \to \mathbb{R}$ is in $C^1$, the composition $f \circ x : [0, \infty) \to \mathbb{R}, t \to f(x(t))$ is again in $C^1$ and

(i) its derivative is given by

$$\frac{d}{dt}(f \circ x)(t) = \frac{df}{dx}(x(t)) \frac{dx}{dt}(t),$$

(ii) in formal differential notation

$$d(f \circ x)(t) = f'(x(t))dx(t),$$

(iii) in integral form

$$f(x(t)) - f(x(0)) = \int_0^t f'(x(s))dx(s).$$

In this last form, the chain rule can be extended to the case where $f \in C^1$ and $x$ is *continuous* and of *finite variation*.

**Def. (quadratic variation)**
Suppose $A$ is an adapted process null at $0$ with RCLL trajectories of finite variation. For any such $A$, the *quadratic variation* (along any fixed, i.e. non-random, sequence $(\Pi_n)_{n\in\mathbb{N}}$ of partitions of $[0, \infty)$ with $\lim_{n\to\infty} |\Pi_n| = 0$) is given by the sum of the squared jumps of $A$, i.e.

$$[A]_t = \sum_{0<s\leq t} (\Delta A_s)^2 = \sum_{0<s\leq t} (A_s - A_{s-})^2 \quad \text{for } t \geq 0$$

By polarisation, then, we have for any semimartingale $Y$ that

$$[A, Y]_t = \sum_{0<s\leq t} \Delta A_s \Delta Y_s \quad \text{for } t \geq 0.$$

**Rem.**

- The quadratic variation of a general semimartingale $X = X_0 + M + A$ has the form

$$\begin{aligned}[X] &= [M + A] \\ &= [M] + [A] + 2[M, A] \\ &= [M] + \sum_{0<s\leq \cdot} (\Delta A_s)^2 + 2 \sum_{0<s\leq \cdot} \Delta M_s \Delta A_s.\end{aligned}$$

- If $A$ is continuous, we obtain that $[X] = [M]$, even if $X$ (hence $M$) is only RCLL.
- A continuous semimartingale $X$ with canonical decomposition $X = X_0 + M + A$ therefore has the quadratic variation $[X] = \langle X\rangle = [M] = \langle W\rangle$ which is again continuous.

**Lem. (a simple result from analysis)**
Any continuous function of finite variation has zero quadratic variation along any sequence $(\Pi_n)_{n\in\mathbb{N}}$ of partitions of $[0, \infty)$ whose mesh size $|\Pi_n|$ goes to $0$ as $n \to \infty$.

**Thm. 1.1 (Itô's formula I)**

- Suppose $X = (X_t)_{t\geq 0}$ is a continuous real-valued semimartingale and $f : \mathbb{R} \to \mathbb{R}$ is in $C^2$ (i.e. $f$ is twice continuously differentiable).
- Then $f(X) = (f(X_t))_{t\geq 0}$ is again a continuous (real-valued) semimartingale, and we explicitly have $P$-a.s.:

(i) Integral form:

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)d\langle X\rangle_s$$

for all $t \geq 0$.

(ii) Differential form:

$$\begin{aligned}df(X_t) &= f'(X_t)dX_t + \frac{1}{2}f''(X_t)d\langle X\rangle_t \\ &= f'(X_t)dX_t + \frac{1}{2}f''(X_t)d\langle M\rangle_t\end{aligned}$$

for all $t \geq 0$.
We used that $\langle X\rangle = \langle M\rangle$.

Remarks:

- The $dX$-integral is a stochastic integral. It is well-defined since $X$ is a semimartingale and $f'(X)$ is adapted and continuous, hence predictable and locally bounded.
The $d\langle X\rangle$-integral is a classical Lebesgue-Stieltjes integral since $\langle X\rangle$ has increasing trajectories. It is also well-defined since $f''(X)$ is also predictable and locally bounded.
- In comparison to the classical chain rule, the $d\langle X\rangle$-integral is an extra second-order term coming from the quadratic variation of $X$. Hence Itô's formula can be viewed as an extension of the chain rule.
- The important message of this formula is that when one is dealing with stochastic models, a simple linear approximation is not good enough, since one also has to account for the second-order behaviour of $X$.
- $\langle X\rangle_t = \langle M\rangle_t$
- To see the financial relevance of Itô's formula, think of $X$ as some underlying financial asset and of $Y = f(X)$ as a new product obtained from the underlying by a possibly nonlinear transformation $f$. Then the formula show us how the product reacts to changes in the underlying. The important message of Thm. 1.1 is then that when using stochastic models (for $X$), a simple linear approximation is not good enough; one must also account for the second-order behaviour of $X$.

### Itô's formula II

**Thm. 1.2 (Itô's fomula II)**
Suppose $X = (X_t)_{t\geq 0}$ is a general $\mathbb{R}^d$-valued semimartingale and $f : \mathbb{R}^d \to \mathbb{R}$ is in $C^2$. Then $f(X) = (f(X_t))_{t\geq 0}$ is again a (real-valued) semimartingale and we explicitly have $P$-a.s. for all $t \geq 0$.

(i) If $X$ has continuous trajectories:

$$\begin{aligned}f(X_t) =& f(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x^i}(X_s)dX_s^i \\ &+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s)d\langle X^i, X^j\rangle_s\end{aligned}$$

or in more compact notation, using subscripts to denote partial derivatives,

$$df(X_t) = \sum_{i=1}^d f_{x^i}(X_t)dX_t^i + \frac{1}{2}\sum_{i,j=1}^d f_{x^i x^j}(X_t)d\langle X^i, X^j\rangle_s$$

(ii) if $d = 1$ (so that $X$ is real-valued, but not necessarily continuous):

$$\begin{aligned}f(X_t) =& f(X_0) + \int_0^t f'(X_{s-})dX_s + \frac{1}{2}\int_0^t f''(X_{s-})d[X]_s \\ &+ \sum_{0<s\leq t} \Big(f(X_s) - f(X_{s-}) \\ &\qquad - f'(X_{s-})\Delta X_s - \frac{1}{2}f''(X_{s-})(\Delta X_s)^2\Big).\end{aligned}$$

- If a stochastic process $X_t = f(t, W_t)$ with $f : \mathbb{R}^2 \to \mathbb{R}$ in $C^{1,2}$ (i.e. once continuously differentiable in time $t$ and twice continuously differentiable in $W_t$), then:

$$\boxed{\begin{aligned}X_t = X_0 &+ \underbrace{\int_0^t \frac{\partial f}{\partial w}(W_s, s)dW_s}_{\text{local } (P,\mathbb{F}) \text{ martingale}} \\ &+ \int_0^t \left(\frac{\partial f}{\partial t}(W_s, s) + \frac{1}{2}\frac{\partial^2 f}{\partial w^2}(W_s, s)\right) ds\end{aligned}}$$

Note that $X$ is a (continuous) local $(P, \mathbb{F})$-martingale iff

$$\int_0^t \left(\frac{\partial f}{\partial t}(W_s, s) + \frac{1}{2}\frac{\partial^2 f}{\partial w^2}(W_s, s)\right) ds = 0, \qquad \forall t \geq 0$$

- **Itô's formula with jumps**
If $d = 1$, $X$ real-valued but not necessarily continuous, then it holds that:

$$\begin{aligned}f(X_t) = f(X_0) &+ \int_0^t f'(X_{s-})dX_s + \frac{1}{2}\int_0^t f''(X_{s-})d[X]_s \\ &+ \sum_{0<s\leq t} \Big(f(X_s) - f(X_{s-}) \\ &\qquad - f'(X_{s-})\Delta X_s - \frac{1}{2}f''(X_{s-})(\Delta X_s)^2\Big)\end{aligned}$$

- If $f : \mathbb{R} \to \mathbb{R}$ is in $C^2$, $\alpha, \beta \in \mathbb{R}$ and the semimartingale $X = (X_t)_{t\geq 0}$ is given by $X_t = \alpha t + \beta N_t$, then:

$$f(X_t) = f(X_0) + \alpha \int_0^t f'(X_{s-})ds + \sum_{0<s\leq t} (f(X_s) - f(X_{s-}))$$

- One frequently *simplification* of Thm. 1.2 (i) arises if one or several of the components of $X$ are of finite variation. If $X^k$, say, is of finite variation, then we know from Lem. (a simple result from analysis) that $\langle X^k \rangle \equiv 0$ and hence also $\langle X^i, X^k \rangle \equiv 0$ for all $i$ by Cauchy-Schwarz.

**Def. (stochastic differential equation (SDE), geometric Brownian motion (GBM))**

- Wiki: A *stochastic differential equation (SDE)* is a differential equation in which one or more of the terms is a stochastic process, resulting in a solution which is also a stochastic process. SDEs are used to model various phenomena such as unstable stock prices or physical systems subject to thermal fluctuations. Typically, SDEs contain a variable which represents random white noise calculated as the derivative of Brownian motion.
- A typical SDE is of the form

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t.$$

- The solution of this SDE is given by the *geometric Brownian motion (GBM)*

$$X_t = X_0 \exp\left(\sigma W_t + \left(\mu - \frac{1}{2}\sigma^2\right)t\right) \quad \text{for } t \geq 0.$$

**Def. (Stochastic exponential)**
For a general real-valued semimartingale $X$ null at 0, the stochastic exponential of $X$ is defined as the unique solution $Z$ of the SDE

$$dZ_t = Z_{t-}dX_t, \qquad Z_0 = 1$$

and it follows that the unique solution to this SDE is:

$$Z_t := \mathcal{E}(X) = 1 + \int_0^t Z_{s-}dX_s \qquad \forall t \geq 0$$

$$\boxed{\mathcal{E}(X)_t = \exp\left(X_t - \frac{1}{2}[X]_t\right)}$$

- **Yor's formula:**

$$\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X + Y + [X, Y])$$

**Itô process**

- An Itô process is a stochastic process of the form

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s, \qquad \forall t \geq 0$$

where $W$ is some Brownian motion and $\mu$ and $\sigma$ are predictable processes.
- More generally, $X, \mu, W$ could be vector-valued and $\sigma$ could be matrix-valued.
- For any $C^2$ function $f$, the process $f(X)$ is again an Itô process, and Itô's formula gives

$$f(X_t) = f(X_0) + \int_0^t \left(f'(X_s)\mu_s + \frac{1}{2}f''(X_s)\sigma_s^2\right) ds + \int_0^t f'(X_s)\sigma_s dW_s$$

**Itô's representation theorem**

**Def. ($P$-augmented filtration)**
We start with a BM $W = (W_t)_{t\geq 0}$ in $\mathbb{R}^m$ defined on a probability space $(\Omega, \mathcal{F}, P)$ without an a priori filtration. We define

$$\mathcal{F}_t^0 := \sigma(W_s, s \leq t) \quad \text{for } t \geq 0,$$
$$\mathcal{F}_\infty^0 := \sigma(W_s, s \geq 0),$$

and construct the filtration $\mathbb{F}^W = (\mathcal{F}_t^W)_{0\leq t\leq\infty}$ by adding to each $\mathcal{F}_t^0$ all subsets of $P$-nullsets in $\mathcal{F}_\infty^0$ to obtain $\mathcal{F}_t^W$. This so-called *$P$-augmented filtration* $\mathcal{F}^W$ is then $P$-complete (in $(\Omega, \mathcal{F}_\infty^0, P)$, to be accurate) by construction.
**Rem.**

- One can show, by using the string Markov property of BM, that $\mathbb{F}^W$ is also automatically right-continuous [RC] (so that it satisfies the usual conditions).
- We usually call $\mathbb{F}^W$, somewhat misleadingly, the filtration generated by $W$.
- One can show that $W$ is also a BM w.r.t. $\mathbb{F}^W$; the key point is to argue that $W_t - W_s$ is still independet of $\mathcal{F}_s^W \supseteq \mathcal{F}_s^0$, even though $\mathcal{F}_s^W$ contains some sets from $\mathcal{F}_\infty^0$.
- If one works on $[0, T]$, one replaces $\infty$ by $T$; then $\mathcal{F}_s^0$ is not needed separately since we use the $P$-nullsets from the "last" $\sigma$-field $\mathcal{F}_T^0$.

**Thm. 3.1 (Itô's representation theorem)**

- Suppose that $W = (W_t)_{t\geq 0}$ is a $\mathbb{R}^m$-valued BM.
- Then every RV $H \in L^1(\mathcal{F}_\infty^W, P)$ has a unique representation as

$$H = E[H] + \int_0^\infty \psi_s dW_s, \qquad P\text{-a.s.}$$

for an $\mathbb{R}^m$-valued integrand $\psi \in L^2_{\text{loc}}(W)$.
- $\psi$ has the additional property that $\int \psi dW$ is a $(P, \mathbb{F}^W)$-martingale on $[0, \infty]$ (and is thus uniformly integrable).

**Rem.**
The assumptions on $H$ say that $H$ is integrable and $\mathcal{F}_\infty^W$-measurable. The latter means *intuitively* that $H(\omega)$ can depend in a measurable way on the entire trajectory $W.(\omega)$ of BM, but not on any other source of randomness.
**Cor. 3.2**
Suppose the filtration $\mathbb{F} = \mathbb{F}^W$ is generated by a BM $W$ in $\mathbb{R}^m$. Then:

(i) Every (real-valued) local $(P, \mathbb{F}^W)$-martingale $L$ is of the form $L = L_0 + \int \gamma dW$ for some $\mathbb{R}^m$-valued process $\gamma \in L^2_{\text{loc}}(W)$.

(ii) Every local $(P, \mathbb{F}^W)$-martingale is continuous.

**Thm. 3.3 (Dudley)**
Suppose $W = (W_t)_{t\geq 0}$ is a BM w.r.t. $P$ and $\mathbb{F} = (\mathcal{F}_{t\geq 0}$. As usual, set

$$\mathcal{F}_\infty := \bigvee_{t\geq 0} \mathcal{F}_t = \sigma\left(\bigcup_{t\geq 0} \mathcal{F}_t\right).$$

Then every $\mathcal{F}_\infty$-measurable random variable $H$ with $|H| < \infty$ $P$-a.s. (e.g. every $H \in L^1(\mathcal{F}_\infty, P)$) can be written as

$$H = \int_0^\infty \psi_s dW_s \quad P\text{-a.s.}$$

for some integrand $\psi \in L^2_{\text{loc}}(W)$.
**Rem.**
It is almost immediately clear that the integrand $\psi$ in Thm. 3.3 cannot be nice. In fact:

(i) In Thm. 3.3, the stochastic integral process $\int \psi dW$ is of course a *local martingale, but* in general *not a martingale* on $[0, \infty]$; if it were, it would have constant expectation 0, which would imply that $E[H] = 0$.

(ii) In Thm. 3.3, the representation by $\psi$ is *not unique*.

**Itô product formula**

- Define the stochastic process $Z = XY$, where $X$ and $Y$ are two continuous real-valued semimartingales.
- Then $Z$ can be written as the sum of stochastic integrals:

$$Z_t - Z_0 = \int_0^t Y_s dX_s + \int_0^t X_s dY_s + \int_0^t d[X, Y]_s$$

**General properties/results**

- Any continuous, adapted process $H$ is also predictable and locally bounded.
  It furthermore holds for any predictable, locally bounded process $H$ that $H \in L^2_{\text{loc}}(W)$.
- Let $f : \mathbb{R} \to \mathbb{R}$ be an arbitrary continuous convex function. Then the process $(f(W_t))_{t\geq 0}$ is integrable and is a $(P, \mathbb{F})$-submartingale.
- Given a $(P, \mathbb{F})$-martingale $(M_t)_{t\geq 0}$ and a measurable function $g : \mathbb{R}_+ \to \mathbb{R}$, the process

  $$(M_t + g(t))_{t\geq 0}$$

  is a:
  - $(P, \mathbb{F})$-supermartingale iff $g$ is decreasing;
  - $(P, \mathbb{F})$-submartingale iff $g$ is increasing.
- A continuous local martingale of finite variation is identically constant (and hence vanishes if it is null at 0).
- For a function $f : \mathbb{R} \to \mathbb{R}$ in $C^1$, the stochastic integral $\int_0^\cdot f'(W_s)dW_s$ is a continuous local martingale.
  Furthermore, for $f \in C^2$ it holds that $f(W)$ is a continuous local martingale iff $\int_0^\cdot f''(W_s)ds = 0$.
- If a predictable process $H = (H_t)_{t\geq 0}$ satisfies

  $$E\left[H_s^2 ds\right] < \infty, \qquad \forall T \geq 0$$

  then $\int H dW_s$ is a square-integrable martingale.
- If $f : \mathbb{R} \to \mathbb{R}$ is bounded and continuous, then the stochastic integral $\int f(W)dW$ is a square-integrable martingale.
- If a process $H = (H_t)_{t\geq 0}$ is predictable and the map $s \mapsto E[H_s^2]$ is continuous, then the stochastic integral $\int H dW$ is a square-integrable martingale.
- If $f : \mathbb{R} \to \mathbb{R}$ is polynomial, then the stochastic integral $\int f(W)dW$ is a square-integrable martingale.

# 5 Black-Scholes Formula

## 5.1 Black-Scholes (BS) model

**Rem.**
The *Black-Scholes model* or *Samuelson model* is the continuous-time analogue os the Cox-Ross-Rubinstein binomial model we have seen at length in earlier chapters.
**Def.**
Throughout this we will use the following setting.
A fixed time horizon $T \in (0, \infty)$ and a probability space $(\Omega, \mathcal{F}, P)$ on which there is a Brownian motion $W = (W_t)_{0\leq t\leq T}$. We take as *filtration* $\mathbb{F} = (\mathcal{F}_{0\leq t\leq T}$ the one generated by $W$ and augmented by the $P$-nullsets form $\mathcal{F}_T^0 := \sigma(W_s; s \leq T)$ s.t. $\mathbb{F} = \mathbb{F}^W$ satisfies the usual conditions under $P$.
**Def.** (undiscounted **financial market model**)
The *financial market model* has two basic traded assets: a bank account with constant continuously compounded *interest rate* $r \in \mathbb{R}$, and a risky asset (usually called stock) habing two parameters $\mu \in \mathbb{R}$ and $\sigma > 0$. Undiscounted prices are given by

$$\widetilde{S}_t^0 = e^{rt}$$

$$\widetilde{S}_t^1 = S_0^1 \exp\left(\sigma W_t + \left(\mu - \frac{1}{2}\sigma^2\right) t\right)$$

with a constant $S_0^1 > 0$.
**Cor.**
Applying Itô's formula to the above equations yields

$$d\widetilde{S}_t^0 = \widetilde{S}_t^0 r dt,$$

$$d\widetilde{S}_t^1 = \widetilde{S}_t^1 \mu dt + \widetilde{S}_t^1 \sigma dW_t,$$

which can be rewritten as

$$\frac{d\widetilde{S}_t^0}{\widetilde{S}_t^0} = r dt,$$

$$\frac{d\widetilde{S}_t^1}{\widetilde{S}_t^1} = \mu dt + \sigma dW_t,$$

This means that the bank account has a *relative price change* $(\widetilde{S}_{t+dt}^0 - \widetilde{S}_t^0)/\widetilde{S}_t^0$ of $r dt$ over a short time period $(t, t+dt]$; so $r$ is the growth rate of the bank account. In the same way, the relative price change of the stock has a part $\mu dt$ giving a growth at rate $\mu$, and a second part $\sigma dW_t$ "with mean 0 and variance $\sigma^2 dt$" that causes random fluctuations. We call $\mu$ the *drift* (rate) and $\sigma$ the (instantaneous) *volatility* of $\widetilde{S}^1$.
**Def.** (discounted **financial market model**)
We pass to quantities *discounted* with $\widetilde{S}^0$; so we have $\S^0 = \widetilde{S}^0/\widetilde{S}^0 \equiv 1$, and $S^1 = \widetilde{S}^1/\widetilde{S}^0$ is by the undiscounted financial market model given by

$$S_t^1 = S_0^1 \exp\left(\sigma W_t + \left(\mu - r - \frac{1}{2}\sigma^2\right) t\right).$$

We obtain via Itô's formula that $S^1$ solves the SDE

$$dS_t^1 = S_t^1((\mu - r)dt + \sigma dW_t).$$

For later use, we observe that this gives

$$d\langle S^1\rangle_t = (S_t^1)^2 \sigma^2 dt$$

for the *quadratic variation* of $S^1$, since $\langle W\rangle_t = t$.
**Rem.**
As in discrete time, we should like to have an *equivalent martingale measure* for the discounted stock price process $S^1$. To get an idea how to find this, we rewrite

$$\begin{aligned} dS_t^1 &= S_t^1((\mu - r)dt + \sigma dW_t). \\ \Leftrightarrow \quad dS_t^1 &= S_t^1 \sigma \left(dW_t + \frac{\mu - r}{\sigma} dt\right) = S_t^1 \sigma dW_t^*, \end{aligned}$$

with $W^* = (W_t^*)_{0\leq t\leq T}$ defined by

$$W_t^* := W_t + \frac{\mu - r}{\sigma} t = W_t + \int_0^t \lambda ds \quad \text{for } 0 \leq t \leq T.$$

**Def.** (**market price of risk** or **Sharpe ratio**)
The quantity

$$\lambda := \tfrac{\mu - r}{\sigma}$$

is often called the instantaneous *market price of risk* or infinitesimal *Sharpe ratio* of $S^1$.
**Rem.**

$$\frac{\text{mean portfolio return} - \text{risk-free rate}}{\text{standard deviation of portfolio return}} = \text{Sharpe ratio.}$$

**Rem.**
By looking at Grisanov's theorem, we see that $W^*$ is a Brownian motion under the probability measure $Q^*$ given by

$$\frac{dQ^*}{dP} := \mathcal{E}\left(-\int \lambda dW\right)_T = \exp\left(-\lambda W_T - \frac{1}{2}\lambda^2 T\right),$$

whose density process w.r.t. $P$ is

$$Z_t^* = \mathcal{E}\left(-\int \lambda dW\right)_t = \exp\left(-\lambda W_t - \frac{1}{2}\lambda^2 T\right) \quad \text{for } 0 \leq t \leq T.$$

**Rem.**
By

$$dS_t^1 = S_t^1 \sigma \left(dW_t + \frac{\mu - r}{\sigma} dt\right) = S_t^1 \sigma dW_t^*,$$

the stochastic integral process

$$S_t^1 = S_0^1 + \int_0^t S_u^1 \sigma dW_u^*$$

is then a continuous local $Q^*$-martingale like $W^*$; it is even a $Q^*$-martingale since we get

$$\begin{aligned} & S_t^1 = S_0^1 \exp\left(\sigma W_t + \left(\mu - r - \frac{1}{2}\sigma^2\right) t\right) \\ \Leftrightarrow \quad & S_t^1 = S_0^1 \exp\left(\sigma W_t^* - \frac{1}{2}\sigma^2 t\right) \end{aligned}$$

by Itô's formula, and so we can use Proposition 2.2 under $Q^*$.

**Rem. (unique equivalent martingale measure)**
In the script on page 118 and 119 we have shown that in the Black-Scholes model, there is a *unique equivalent martingale measure*, which is given explicitly by $Q^*$. So we expect that the Black-Scholes model is not only "arbitrage-free", but also "complete" in a suitable sense.

**Def.**

(i) Take any $H \in L^0_+(\mathcal{F}_T)$ and view $H$ as a random *payoff* (in discounted units) due at time $T$. Recall that $\mathbb{F}$ is generated by $W$ and that $W_t^* = W_t + \lambda t, 0 \le t \le T$, is a $Q^*$-Brownian motion.

(ii) Because $\lambda$ is deterministic, $W$ and $W^*$ generate the same filtration, and so we can also apply Itô's representation theorem with $Q^*$ and $W^*$ instead of $P$ and $W$. So if $H$ is also in $L^1(Q^*)$, the $Q^*$-martingale $V_t^* := E_{Q^*}[H|\mathcal{F}_t], 0 \le t \le T$, can be represented as

$$V_t^* = E_{Q^*}[H] + \int_0^t \psi_s^H dW_s^* \quad \text{for } 0 \le t \le T,$$

with some unique $\psi^H \in L^2_{\text{loc}}(W^*)$ s.t. $\int \psi^H dW^*$ is a $Q^*$-martingale.

**Rem.** (trading strategy, self-financing)
If we define for $0 \le t \le T$

$$\vartheta_t^H := \frac{\psi_t^H}{S_t^1 \sigma},$$

$$\eta_t^H := V_t^* - \vartheta_t^H S_t^1$$

(which are both predictable because $\psi^H$ is), then we can interpret $\varphi^H = (\vartheta^H, \eta^H)$ as a *trading strategy* whose discounted value process is given by

$$V_t(\varphi^H) = \vartheta_t^H S_t^1 + \eta_t^H S_t^0 = V_t^* \quad \text{for } 0 \le t \le T,$$

and which is *self-financing* in the (usual) sense that

$$V_t(\psi^H) = V_t^* = V_0^* + \int_0^t \psi_u^H dW_u^* = V_0(\psi^H) + \int_0^t \vartheta_u^H dS_u^1, 0 \le t \le T.$$

Moreover,

$$V_T(\varphi^H) = V_T^* = H \text{ a.s.}$$

shows that the strategy $\varphi^H$ *replicates* $H$, and

$$\int \vartheta^H dS^1 = V(\varphi^H) - V_0(\varphi^H) = V^* - E_{Q^*}[H] \ge -E_{Q^*}[H]$$

(because $V^* \ge 0$, since $H \ge 0$) shows that $\vartheta^H$ is admissible (for $S^1$) int the usual sense.

**Rem.**

- In summary, every $H \in L^1_+(\mathcal{F}_T, Q^*)$ is attainable in the sense that it can be replicated by a dynamic strategy trading in the stock and the bank account in such a way that the strategy is self-financing and admissible, and its value process is a $Q^*$-martingale.
- In that sense, we can say that the Black-Scholes model is complete.
- By the same arguments as in discrete time, we then also obtain the arbitrage-free value at time $t$ of any payoff $H \in L^1_+(\mathcal{F}_T, Q^*)$ as its conditional expectation

$$V_t^H = V_t^* = E_{Q^*}[H|\mathcal{F}]$$

  under the unique equivalent martingale measure $Q^*$ for $S^1$.
- This is in perfect parallel to the results we have seen for the CRR binomial model.

**Rem.**

(i) All the above computations and results are in *discounted* units.

(ii) Itô's representation theorem gives the *existence* of a strategy, but does not tell us how it looks.

(iii) The SDE $dS_t^1 = S_t^1((\mu - r)dt + \sigma dW_t)$ for discounted prices is

$$\frac{dS_t^1}{S_t^1} = (\mu - r)dt + \sigma dW_t$$

and this is rather restrictive since $\mu, r, \sigma$ are all constant. An obvious *extension* is to allow the coefficients $\mu, r, \sigma$ to be (suitably integrable) predictable processes, or possibly funcitonals of $S$ or $\tilde{S}$. This brings up several issue which are enlisted in the script on page 121.

**BS model (undiscounted, historical measure $\mathbb{P}$)**

$$\tilde{S}_t^0 = e^{rt} \qquad \tilde{S}_t^1 = \tilde{S}_0^1 \exp\left(\sigma W_t + \left(\mu - \frac{1}{2}\sigma^2\right) t\right)$$

$$\frac{d\tilde{S}_t^0}{\tilde{S}_t^0} = rdt \qquad \frac{d\tilde{S}_t^1}{\tilde{S}_t^1} = \mu dt + \sigma dW_t$$

**BS model (discounted, historical measure $\mathbb{P}$)**

$$S_t^0 = 1 \qquad S_t^1 = S_0^1 \exp\left(\sigma W_t + \left(\mu - r - \frac{1}{2}\sigma^2\right) t\right)$$

$$\frac{dS_t^1}{S_t^1} = (\mu - r)dt + \sigma dW_t$$

**BS model (discounted, risk-neutral measure $\mathbb{Q}$)**

$$dS_t^1 = S_t^1 \sigma \left(dW_t + \frac{\mu - r}{\sigma} dt\right) = S_t^1 \sigma dW_t^*$$

$$S_t^1 = S_0^1 + \int_0^t S_u^1 \sigma dW_u^* = S_0^1 \exp\left(\sigma W_t^* - \frac{1}{2}\sigma^2 t\right)$$

where

$$W_t^* := W_t + \frac{\mu - r}{\sigma} t = W_t + \int_0^t \lambda ds$$

**Market price of risk**

- The market price of risk or infinitemsimal **Sharpe ratio** of $S^1$ is defined as

$$\lambda^* = \frac{\mu - r}{\sigma}$$

## 5.2 Markovian payoffs and PDEs

**Rem.** (martingale approach, PDE approach)

- The presentation in the previous subsection is often called the *martingale approach* to valuing options, for obvious reasons.
- If one has more structure for the payoff $H$, an alternative method involves the use of partial differential equations (PDEs) and is thus called the *PDE approach*.

**Def.** (used throughout this subsection)
Suppose that the (discounted) payoff is of the form $H = h(S_T^1)$ for some measurable function $h \ge 0$ on $\mathbb{R}_+$. We also suppose that $H$ is in $L^1(Q^*)$; here, $H = (\tilde{S}_T^1 - \tilde{K})^+ / \tilde{S}_T^0 = (S_T^1 - \tilde{K}e^{-rT})^+$.

**Rem. (value process)**
We start with the *value process*. Since we have $V_t^* = E_{Q^*}[H|\mathcal{F}_t] = E_{Q^*}[h(S_T^1)|\mathcal{F}_t]$, we look at

$$S_t^1 = S_0^1 \exp(\sigma W^* - \tfrac{1}{2}\sigma^2 t)$$

and write

$$S_T^1 = S_t^1 \frac{S_T^1}{S_t^1} = S_t^1 \exp(\sigma(W_T^* - W_t^*) - \tfrac{1}{2}\sigma^2(T-t)).$$

In the last term, the first factor $S_t^1$ is obviously $\mathcal{F}_t$-measurable. Moreover, $W^*$ is a $Q^*$-Brownian motion w.r.t. $\mathbb{F}$, and so in the second factor, $W_T^* - W_t^*$ is under $Q^*$ independent of $\mathcal{F}_t$ and has an $\mathcal{N}(0, T-t)$-distribution.

**Def.**
Therefore we get

$$V_t^* = E_{Q^*}[h(S_T^1)|\mathcal{F}_t] = v(t, S_t^1)$$

with the function $v(t,x)$ given by

$$
\begin{aligned}
v(t,x) &= E_{Q^*}\left[h\left(x\exp\left(\sigma(W_T^*-W_t^*)-\frac{1}{2}\sigma^2(T-t)\right)\right)\right]\\
&= E_{Q^*}\left[h\left(xe^{\sigma\sqrt{T-t}Y-\frac{1}{2}\sigma^2(T-t)}\right)\right]\\
&= \int_{-\infty}^{\infty} h\left(xe^{\sigma\sqrt{T-t}y-\frac{1}{2}\sigma^2(T-t)}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}dy,
\end{aligned}
$$

where $Y\sim(0,1)$ under $Q^*$.

**Rem.**

This already gives a fairly precise structural description of $V_t^*$ as a function of ($t$ and) $S_t^1$, instead of a general $\mathcal{F}_t$-measurable random variable.

**Rem. (strategy)**

As explained in the script on page 123,

$$dV_t^* = v_x(t,S_t^1)dS_t^1 + (v_t(t,S_t^1)+\tfrac{1}{2}v_{xx}(t,S_t^1)\sigma^2(S_t^1)^2)dt$$

and

$$V_t(\varphi^H) = V_t^* = V_0(\varphi^H)+\int_0^t \vartheta_u^H dS_u^1$$

yield

$$v_x(t,S_t^1)dS_t^1 = dV_t^* = \vartheta_t^H dS_t^1$$

s.t. we obtain the *strategy* explicitly as

$$\vartheta_t^H = \frac{\partial v}{\partial x}(t,S_t^1),$$

i.e. as the spatial derivative of $v$, evalueated along the trajectories of $S^1$.

**Def.** (discounted **PDE**)

In fact, the vanishing of the $dt$-term means that the function $v_t(t,x)+\frac{1}{2}v_{xx}(t,x)\sigma^2x^2$ must vanish along the trajectories of the space-time process $(t,S_t^1)_{0<t<T}$. But each $S_t^1$ is by

$$S_t^1 = S_0^1\exp(\sigma W_t^* - \tfrac{1}{2}\sigma^2 t)$$

lognormally distributed and hence has all of $(0,\infty)$ in its support. so the support of the space-time process contains $(0,T)\times(0,\infty)$, and so $v(t,x)$ must satisfy the (linear, second-order) *PDE*

$$0=\frac{\partial v}{\partial t}+\frac{1}{2}\sigma^2x^2\frac{\partial^2 v}{\partial x^2},\quad \text{on } (0,T)\times(0,\infty).$$

Moreover, the definition of $v$ via

$$V_t^* = E_{Q^*}[h(S_T^1)|\mathcal{F}_t] = v(t,S_t^1)$$

gives the *boundary condition*

$$v(T,\cdot)=h(\cdot)\quad \text{on } (0,\infty),$$

because $v(T,S_T^1)=V_T^*=H=h(S_T^1)$ and the support of the distribution of $S_T^1$ contains $(0,\infty)$.

**Rem.**

So if we cannot compute the integral in

$$\int_{-\infty}^{\infty} h\left(xe^{\sigma\sqrt{T-t}y-\frac{1}{2}\sigma^2(T-t)}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}dy$$

explicitly, we can at least obtain $v(t,x)$ *numerically* by solving the above PDE.

**Def.** (discounted **PDE**)

If the *undiscounted* payoff is $\widetilde{H}=\widetilde{h}(\widetilde{S}_T^1)$ and the undiscounted value at time $t$ is $\widetilde{v}(t,\widetilde{S}_t^1)$, we have the relations

$$\widetilde{h}(\widetilde{S}_T^1)=\widetilde{h}(e^{rT}\widetilde{S}_T^1)=\widetilde{H}=e^{rT}H=e^{rT}h(S_T^1)$$

and

$$\widetilde{v}(t,\widetilde{x})=e^{rt}v(t,\widetilde{x}e^{-rt}).$$

For the function $\widetilde{v}$, we then obtain from

$$0=\frac{\partial v}{\partial t}+\frac{1}{2}\sigma^2x^2\frac{\partial^2 v}{\partial x^2},\quad \text{on } (0,T)\times(0,\infty)$$

the PDE

$$0=\frac{\partial\widetilde{v}}{\partial t}+r\widetilde{x}\frac{\partial\widetilde{v}}{\partial\widetilde{x}}+\frac{1}{2}\sigma^2\widetilde{x}^2\frac{\partial^2\widetilde{v}}{\partial\widetilde{x}^2}-r\widetilde{v}$$

with the boundary condition

$$\widetilde{v}(T,\cdot)=\widetilde{h}(\cdot).$$

## 5.3 Black-Scholes PDE

**Black-Scholes PDE**

$$0=\frac{\partial\tilde{v}}{\partial t}+r\tilde{x}\frac{\partial\tilde{v}}{\partial\tilde{x}}+\frac{1}{2}\sigma^2\tilde{x}^2\frac{\partial^2\tilde{v}}{\partial\tilde{x}^2}-r\tilde{v},\qquad \tilde{v}(T,\cdot)=\tilde{h}(\cdot)$$

## 5.4 Black-Scholes formula for option pricing

**Martingale pricing approach**

- The discounted arbitrage-free value at time $t$ of any discounted payoff $H\in L_+^1(\mathcal{F}_T,\mathbb{Q}^*)$, $H_T=H(\tilde{S}_T^0,\tilde{S}_T^1)$, is given by

  $$\boxed{V_t^* = \mathbb{E}_{\mathbb{Q}}[H|\mathcal{F}_t] := \vartheta(t,S_t^1)}$$

- Then the discounted payoff $H$ can be hedged via the replicating strategy $(V_0,\vartheta)$ s.t.

  $$V_0+\int_0^T \vartheta_u dS_u^1 = H(\tilde{S}_T^0,\tilde{S}_T^1)$$

  Using Itô's representation theorem the replicating strategy can be expressed as

  $$\boxed{\begin{aligned}&V_t^*=\vartheta(t,S_t^1)=V_0+\int_0^t\vartheta_s dS_s^1+\underbrace{\text{cont. FV process}}_{\text{"usually" vanishes}}\\ &V_0=\vartheta(0,S_0^1),\qquad \vartheta_t=\frac{\partial\vartheta}{\partial x}(t,S_t^1)\end{aligned}}$$

**Rem. (European call option)**

In the special case of a *European call option*, the value process and the corresponding strategy can be computed explicitly, and this has found widespread use in industry.

**Def.**

Suppose the undiscounted strike price is $\widetilde{K}$ s.t. the undiscounted payoff is

$$\widetilde{H}=(\widetilde{S}_T^1-\widetilde{K})^+.$$

Then $H=\widetilde{H}/\widetilde{S}_T^0=(S_T^1-\widetilde{K}e^{-rT})^+=:(S_T^1-K)^+$. **Rem.**

We obtain form

$$
\begin{aligned}
v(t,x) &= E_{Q^*}\left[h\left(x\exp\left(\sigma(W_T^*-W_t^*)-\frac{1}{2}\sigma^2(T-t)\right)\right)\right]\\
&= E_{Q^*}\left[h\left(xe^{\sigma\sqrt{T-t}Y-\frac{1}{2}\sigma^2(T-t)}\right)\right]\\
&= \int_{-\infty}^{\infty} h\left(xe^{\sigma\sqrt{T-t}y-\frac{1}{2}\sigma^2(T-t)}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}dy,
\end{aligned}
$$

that the discounted value of $H$ at time $t$ is

$$V_t^* = E_{Q^*}\left[\left(xe^{\sigma\sqrt{T-t}Y-\frac{1}{2}\sigma^2(T-t)}-K\right)^+\right]\Bigg|_{x=S_t^1}.$$

Because we have $Y\sim\mathcal{N}(0,1)$ under $Q^*$, an elementary computation yields for $x>0, a>0$ and $b\geq 0$ that

$$E_{Q^*}\left[\left(xe^{aY-\frac{1}{2}a^2}-b\right)^+\right]=x\Phi\left(\frac{\log\left(\frac{x}{b}\right)+\frac{1}{2}a^2}{a}\right)-b\Phi\left(\frac{\log\left(\frac{x}{b}\right)-\frac{1}{2}a^2}{a}\right.$$

where

$$\Phi(y)=Q^*[Y\leq y]=\int_{-\infty}^{y}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz$$

is the *cumulative distribution function* (CDF) of the standard normal distribution $\mathcal{N}(0,1)$.

**Def. (Black-Scholes formula)**

Plugging in the above formula $x=S_t^1, a=\sigma\sqrt{T-t}, b=K$ and then passing to undiscounted quantities therefore yields the famous *Black-Scholes formula* in the form

$$\widetilde{V}_t^{\widetilde{H}}=\widetilde{v}(t,\widetilde{S}_t^1)=\widetilde{S}_t^1\Phi(d_1)-\widetilde{K}e^{-r(T-t)}\Phi(d_2)$$

with

$$d_{1,2} = \frac{\log\left(\frac{\tilde{S}_t^1}{\tilde{K}}\right) + \left(r \pm \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}.$$

**Rem.**
Note that the drift $\mu$ of the stock does not appear here; this is analogous to the result that the probability $p$ of an up move in the CRR binomial model does not appear in the binomial option pricing formula.
**Rem. (replicating strategy)**
To compute the *replicating strategy*, we recall that the stock price holdings at time $t$ are given by

$$\vartheta_t^H = \frac{\partial v}{\partial x}(t, S_t^1).$$

Moreover, $v(t,x) = e^{-rt}\tilde{v}(t, xe^{rt})$ s.t.

$$\frac{\partial v}{\partial x}(t,x) = e^{-rt}\frac{\partial \tilde{v}}{\partial \tilde{x}}(t, xe^{rt})e^{rt} = \frac{\partial \tilde{v}}{\partial \tilde{x}}(t, xe^{rt}).$$

Computing the above derivative explicitly gives

$$\vartheta_t^H = \frac{\partial \tilde{v}}{\partial \tilde{x}}(t, \tilde{S}_t^1) = \Phi(d_1) = \Phi\left(\frac{\log\left(\frac{\tilde{S}_t^1}{\tilde{K}}\right) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right),$$

which always lies between $0$ and $1$.

**Black-Scholes formula for a European call option**

$$\tilde{V}_t^{\tilde{H}} = \tilde{v}(t, \tilde{S}_t^1) = \tilde{S}_t^1\Phi(d_1) - \tilde{K}e^{-r(T-t)}\Phi(d_2)$$

$$d_{1,2} = \frac{\log\left(\frac{\tilde{S}_t^1}{\tilde{K}}\right) + \left(r \pm \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$$

$$\Phi(y) = Q^*[Y \le y] = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

Remarks:

- $\Phi$ denotes the CDF of the standard normal distribution $\mathcal{N}(0,1)$.
- Note that the drift $\mu$ of the stock does not appear here. This is analogous to the result that the probability $p$ of an up move in the CRR binomial model does not appear in the binomial option pricing formula.

**Replicating strategy for a European call option**

$$\nu_t^H = \frac{\partial \tilde{v}}{\partial \tilde{x}}(t, \tilde{S}_t^1) = \Phi(d_1)$$

**Greeks** The derivatives of the option price w.r.t. the various parameters, i.e. the sensitivities of the option price w.r.t. to the parameters, are called Greeks.

## 6 Appendix

**Markov's inequality** For $X$ any nonnegative integrable RV and $a \in \mathbb{R}, a > 0$:

$$\mathbb{P}[X \ge a] \le \frac{\mathbb{E}[X]}{a}$$

**Chebyshev's inequality** For $X$ an integrable RV with finite expected value $\mu \in \mathbb{R}$ and finite non-zero variance $\sigma^2, \sigma \in \mathbb{R}$ and for any real number $k > 0$:

$$\mathbb{P}[|X - \mu| \ge k\sigma] \le \frac{1}{k^2}$$

**Jensen's inequality** For $X$ a RV and $f$ a convex function, it holds that

$$f(\mathbb{E}[X]) \le \mathbb{E}[f(X)]$$

**Sets/families**

- $L_+^0$: familiy of all nonnegative RVs

**Correlation and Independence**

- Let $X$ and $Y$ be two RVs.
- Then $X, Y$ are **uncorrelated** iff

  $$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

- Then $X, Y$ are **independent** iff

  $$\mathbb{P}[X = x, Y = y] = \mathbb{P}[X = x] \cdot \mathbb{P}[Y = y]$$

  Note that independence of $X, Y$ implies that $X, Y$ are uncorrelated (but not vice-versa!).

**Independence of equations**

- If there is e.g. a system of equations such as

  $$\begin{cases} a_1x + b_1y = c_1 \\ a_2x + b_2y = c_2 \\ a_3x + b_3y = c_3 \end{cases}$$

  then this system admits a solution iff

  $$\det\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = 0$$

**Fubini's lemma**

$$\mathbb{E}\left[\int_0^T H_s^2 ds\right] = \int_0^T \mathbb{E}\left[H_s^2\right] ds$$

*intentionally left blank*

## Abbreviations

| | |
|---|---|
| **a.a.** | almost all |
| **a.s.** | almost surely |
| **BM** | Brownian motion |
| **CDF** | cumulative distribution function |
| **iff** | if and only if |
| **IOT** | in order to |
| **PDE** | partial differential equation |
| **PDF** | probability density function |
| **RCLL** | right-continuous with left limits |
| **RV** | random variable |
| **SDE** | stochastic differential equation |
| **s.t.** | such that |
| **w.r.t.** | with respect to |

## Disclaimer

- This summary is work in progress, i.e. neither completeness nor correctness of the content are guaranteed by the author.
- This summary may be extended or modified at the discretion of the readers.
- Source: Lecture Mathematical Foundations for Finance, autumn semester 2015/16, ETHZ (lecture notes, script and exercises). Copyright of the content is with the lecturers.
- The layout of this summary is mainly based on the summaries of several courses of the BSc ETH ME from Jonas LIECHTI.