Stochastic integration and financial mathematics

Philipp Harms*

June 6, 2019

Overview

These lecture notes form the core of an advanced master class held at the University of Freiburg in the summer term 2018. They cover stochastic integration and no-arbitrage theory in great generality. Nevertheless, they are reasonably short and self-contained thanks to some recent and elementary proofs of e.g. the Bichteler–Dellacherie theorem and the fundamental theorem of asset pricing. Some supplementary results are left as exercises and marked by [E], and some auxiliary results are collected in the appendix.

The notes are preliminary and incomplete: they don't cover the full extent of the lecture, and references to the literature are largely missing. Some parts are indebted to lecture notes of Josef Teichmann; this is gratefully acknowledged. Of course, the author takes full responsibility for any mistakes and would be glad to hear about any that you find.

Contents

*philipp.harms@stochastik.uni-freiburg.de

1 Stochastic integration

This section is devoted to the construction of the stochastic integral. We use an analytic approach based on the Emery topology, which is closely related to the notions of good integrators and L^0-valued vector measures. The advantage is that it is quick, fully general, and does not presuppose any semimartingale theory. The disadvantage is that it does not provide an explicit description of the integrands. An alternative approach is to define the stochastic integral separately for local martingales and finite variation processes [SC02]. A mixture of these approaches is presented in [Pro05], where integration with caglad integrands is treated using the notion of good integrators, whereas integration with predictable integrands is treated as in [SC02].

Setting. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, \mathbb{P})$ be filtered probability space satisfying the usual conditions, let $d \in \mathbb{N}$, and let $\cdot$ denote the Euclidean scalar product on $\mathbb{R}^d$. All processes are defined on $\Omega \times [0,1]$, and all stopping times take values in $[0,1] \cup \{\infty\}$, unless stated otherwise.

1.1 Good integrators

1.1.1 Prerequisites. Section 3.2 on gauges and Section 3.3 on convergence in probability are needed now.

1.1.2 Definition. Let $H, X\colon \Omega\times[0,1] \to \mathbb{R}^d$ with H caglad adapted and X cadlag adapted.

(i) H is *elementary predictable*, in symbols $H \in \mathcal{E}^d$, if it can be written as

$$H = h_0 \mathbb{1}_{[\![0]\!]} + \sum_{i=1}^{n} h_i \mathbb{1}_{(\!(T_i, T_{i+1}]\!]},$$

where $n \in \mathbb{N}$, $0 = T_0 = T_1 < T_2 \leq \cdots \leq T_{n+1} = 1$ are stopping times, and h_i are essentially bounded $\mathbb{R}^d$-valued $\mathcal{F}_{T_i}$-measurable random variables.

(ii) The *elementary indefinite integral* of $H \in \mathcal{E}^d$ against X is the real-valued cadlag adapted process

$$H \bullet X = h_0 \cdot X_0 + \sum_{i=1}^{n} h_i \cdot (X^{t_{i+1}} - X^{t_i}),$$

and the *elementary definite integral* is the real-valued random variable

$$(H \bullet X)_1 = h_0 \cdot X_0 + \sum_{i=1}^{n} h_i \cdot (X_{T_{i+1}} - X_{T_i}).$$

(iii) X is a *good integrator* if the elementary definite integral is continuous:

$$(\mathcal{E}^d, \|\cdot\|_\infty) \ni H \mapsto (H \bullet X)_1 \in (L^0, \llbracket\cdot\rrbracket_0),$$

where $\|H\|_\infty = \operatorname{ess\,sup}_{\omega\in\Omega} \sup_{t\in[0,1]} \sup_{j\in\{1,\dots,d\}} |H_t^j(\omega)|$.

1.1.3 Remark.

- The good integrator property is a very weak continuity requirement, as the $\|\cdot\|_\infty$ topology is strong and the $\llbracket\cdot\rrbracket_0$ topology weak.
- In the deterministic case the cadlag property of X corresponds to the cadlag property of cumulative distribution functions.
- The generalization to infinite time horizons works as follows: a process $X\colon \Omega \times [0,\infty) \to \mathbb{R}^d$ is a good integrator if its restriction to $\Omega \times [0,t]$ is a good integrator for each $t \in [0,\infty)$.
- The notion of good integrators remains the same if elementary integrands are replaced by simple integrands, i.e., elementary predictable processes H as in Definition 1.1.2 such that T_i and h_i take only finitely many values. E
- Good integrators are stable under stopping. Any local good integrator is a good integrator. Good integrators under $\mathbb{P}$ are also good integrators under $\mathbb{Q} \ll \mathbb{P}$. Good integrators with respect to $(\mathcal{F}_t)_{t\in[0,1]}$ are also good integrators with respect to any sub-filtration $(\mathcal{G}_t)_{t\in[0,1]}$ they are adapted to. E
- Finite variation processes are the only good integrators with respect to the larger set of elementary integrands consisting of all processes $H = \mathbb{1}_{\llbracket 0\rrbracket} + \sum_{i=1}^n h^i \mathbb{1}_{\rrbracket T_i, T_{i+1}\rrbracket}$ such that h^i is $\mathcal{F}_{T_{i+1}}$-adapted. E

1.2 Topologies on good integrators

1.2.1 Definition. Let $X\colon \Omega \times [0,1] \to \mathbb{R}^d$ be cadlag adapted, and let $H \in \mathcal{E}^d$.

(i) The good integrator and semimartingale (Emery) gauges of X and H are

$$\llbracket X\rrbracket_{\mathcal{I}^d} = \sup_{\substack{K\in\mathcal{E}^d\\ \|K\|_\infty\le 1}} \llbracket (K\bullet X)_1\rrbracket_0, \qquad \llbracket X\rrbracket_{\mathcal{S}^d} = \sup_{\substack{K\in\mathcal{E}^d\\ \|K\|_\infty\le 1}} \llbracket |K\bullet X|_1^*\rrbracket_0.$$

$$\llbracket H\rrbracket_X = \llbracket H\rrbracket_{X,\mathcal{I}} = \llbracket H\bullet X\rrbracket_{\mathcal{I}}, \quad \llbracket H\rrbracket_{X,\mathcal{S}} = \llbracket H\bullet X\rrbracket_{\mathcal{S}}.$$

(ii) $\mathcal{I}^d$ and $\mathcal{S}^d$ denote the sets of cadlag adapted processes X, which are finite for $\llbracket\cdot\rrbracket_{\mathcal{I}^d}$ and $\llbracket\cdot\rrbracket_{\mathcal{S}^d}$, respectively.

1.2.2 Lemma.

(i) $[\![\cdot]\!]_{\mathcal{I}^d} = [\![\cdot]\!]_{\mathcal{S}^d}$ *and* $\mathcal{I}^d = \mathcal{S}^d$ *is exactly the set of good integrators.*

(ii) $\mathcal{I}^d = \mathcal{S}^d$ *is a complete topological vector space under* $[\![\cdot]\!]_{\mathcal{I}^d} = [\![\cdot]\!]_{\mathcal{S}^d}$.

(iii) For any $X \in \mathcal{I}^d = \mathcal{S}^d$, $\mathcal{E}^d$ *is a topological vector space under* $[\![\cdot]\!]_{X,\mathcal{I}} = [\![\cdot]\!]_{X,\mathcal{S}}$.

Proof.

(i) We follow [Bic02, Lemma 2.3.2]. For any $\lambda > 0$ and $H \in \mathcal{E}^d$ define

$$T = \inf\{t \in [0,1]; |H \bullet X|_t > \lambda\}, \qquad K = \mathbb{1}_{[\![0,T]\!]}.$$

Then

$$\mathbb{P}[|H \bullet X|_1^* > \lambda] \leq \mathbb{P}[|H \bullet X|_T \geq \lambda] = \mathbb{P}[|HK \bullet X|_1 \geq \lambda].$$

This implies

$$\begin{aligned}\inf\{\lambda > 0; \mathbb{P}[|H \bullet X|_1^* > \lambda] \leq \lambda\} &\leq \inf\{\lambda > 0; \mathbb{P}[|HK \bullet X|_1^* \geq \lambda] \leq \lambda\} \\ &= \inf\{\lambda > 0; \mathbb{P}[|HK \bullet X|_1^* > \lambda] \leq \lambda\}.\end{aligned}$$

As $HK \in \mathcal{E}^d$ satisfies $\|HK\|_\infty \leq 1$, taking the supremum over all $H \in \mathcal{E}^d$ shows that $[\![X]\!]_{\mathcal{S}^d} \leq [\![X]\!]_{\mathcal{I}^d}$. The reverse inequality is trivial.

(ii) $\mathcal{I}^d = \mathcal{S}^d$ is a topological vector space by Lemma 3.2.4 because the gauge $[\![\cdot]\!]_{\mathcal{S}^d} = [\![\cdot]\!]_{\mathcal{I}^d}$ is subadditive, balanced, and finite. We show completeness following [Éme79, Théorème 1]. Let $(X^n)_{n\in\mathbb{N}}$ be Cauchy in $\mathcal{I}^d$. Then X^n converges uniformly in probability to some cadlag adapted X by (i) and by the completeness of the set of cadlag adapted processes with respect to uniform convergence in probability. To verify that $X \in \mathcal{I}^d$, note that $\{X^n; n \in \mathbb{N}\}$ is bounded in $\mathcal{I}^d$, i.e.,

$$\lim_{r\to 0} \sup_{n\in\mathbb{N}} [\![rX^n]\!]_{\mathcal{I}} = \lim_{r\to 0} \sup_{n\in\mathbb{N}} \sup_{\substack{H\in\mathcal{E}\\ |H|\leq 1}} [\![(rH \bullet X^n)_1]\!]_0 = 0.$$

As $[\![(rH \bullet X^n)_1]\!]_0 \to [\![(rH \bullet X)_1]\!]_0$, this implies

$$\lim_{r\to 0} \sup_{\substack{H\in\mathcal{E}\\ |H|\leq 1}} [\![(rH \bullet X)_1]\!]_0 = 0,$$

and we have shown that $X \in \mathcal{I}^d$.

(iii) The $[\![\cdot]\!]_{\mathcal{I}^d}$-finiteness of X implies the $[\![\cdot]\!]_{X,\mathcal{I}}$-finiteness of any $H \in \mathcal{E}^d$ because

$$\lim_{r\to 0} [\![rH]\!]_{X,\mathcal{I}^d} = \lim_{r\to 0} [\![rH \bullet X]\!]_{\mathcal{I}} = 0.$$

Thus, $[\![\cdot]\!]_{X,\mathcal{I}}$ is subadditive balanced and finite on $\mathcal{E}^d$, and Lemma 3.2.4 implies that $\mathcal{E}^d$ is a topological vector space.

□

1.3 Stochastic integrals

1.3.1 Definition. For any $X \in \mathcal{I}^d$, $L^1(X)$ denotes the closure of $\mathcal{E}^d$ with respect to $[\![\cdot]\!]_X$, and $L(X) = L^1_{\mathrm{loc}}(X)$.

1.3.2 Theorem. *Let* $X \in \mathcal{I}^d$.

(i) *The elementary definite stochastic integral extends uniquely to a continuous linear mapping*

$$L^1(X) \ni H \mapsto (H \bullet X)_1 \in L^0,$$

which is called the definite stochastic integral*.*

(ii) *The elementary indefinite stochastic integral extends uniquely to an isometry*

$$L^1(X) \ni H \mapsto H \bullet X \in \mathcal{I},$$

which is called the indefinite stochastic integral*.*

Proof. We follow [Bic02, Theorem 3.7.10].

(i) Continuity of the elementary definite integral follows from (ii) because evaluation at time one ($\mathcal{I} = \mathcal{S} \ni Y \mapsto Y_1 \in L^0$) is continuous. The extension exists because L^0 is complete.

(ii) The gauges on $\mathcal{E}^d$ and $\mathcal{I}^d$ are defined such that the elementary integral is isometric, and the extension exists because $\mathcal{I}^d$ is complete as shown in Lemma 1.2.2. □

1.3.3 Remark.

(i) What is lacking at this point is a characterization of $L^1(X)$ as a set of predictable processes determined by some integrability conditions. This requires some additional theory: either semimartingale theory [SC02] or Daniell's theory of integration[1] [Bic02]. Both approaches were read together in class.

(ii) The integral is as general as possible because the set of integrals is Emery-closed, i.e., for any $X \in \mathcal{I}^d$, the set $\{H \bullet X; H \in L^1(X)\}$ is closed in $\mathcal{I}$ thanks to the isometric property of the integral. E

(iii) The integral is more general than the component-wise integral: there are $X \in \mathcal{I}^2$ and $H = (H^1, H^2) \in L(X)$ with $H^1 \notin L(X^1)$ and $H^2 \notin L(X^2)$. E

(iv) The integral is associative, equivariant with respect to stopping, invariant under shrinkage of filtration, and invariant under absolutely continuous changes of measure. Moreover, the jumps of the integral are the integrand times the jumps of the integrator. These statements are obvious for elementary integrands and follow for general integrands by taking limits.

(v) The integral coincides with the Lebesgue–Stieltjes integral if X has finite variation. E

The following lemma provides a large class of integrands, which is sufficient for many purposes, including the definition of quadratic variation, Itô's formula, and stochastic differential equations.

1.3.4 Lemma. *For each $X \in \mathcal{I}^d$, the space of $\mathbb{R}^d$-valued caglad adapted processes with the topology of uniform convergence in probability is continuously embedded in $L^1(X)$.*

Proof. We claim that the topology of uniform convergence in probability on $\mathcal{E}^d$ is stronger than the $L^1(X)$ topology on $\mathcal{E}^d$. To prove the claim let $H^n \in \mathcal{E}^d$ satisfy $H^n \xrightarrow{up} 0$, and let $\epsilon > 0$. Choose $r \in (0, \infty)$ such that

[1] The integrals in Daniell's theory are not Emery-closed for $d \geq 2$, contrarily to what is stated between equations (3.10.2) and (3.10.3) in [Bic02].

$[\![rX]\!]_{\mathcal{I}} \leq \epsilon$. Choose $N \in \mathbb{N}$ such that for all $n \geq N$: $\mathbb{P}[|H^n|_1^* > r] \leq \epsilon$. Then one has for all $n \geq N$ and $K \in \mathcal{E}$ with $|K| \leq 1$ that

$$\begin{aligned}\mathbb{P}[|H^n K \bullet X|_1 > 2\epsilon] &\leq \mathbb{P}[|H^n K \bullet X|_1 > \epsilon] \\ &\leq \mathbb{P}[|H^n|_1^* > r] + \mathbb{P}[r|K \bullet X|_1 > \epsilon] \leq 2\epsilon.\end{aligned}$$

This implies for all $n \geq N$ that $[\![H^n]\!]_X \leq 2\epsilon$. Thus, $[\![H^n]\!]_X \to 0$. This proves the claim. Now the lemma follows from the fact that the set of caglad adapted processes is the closure of $\mathcal{E}^d$ with respect to uniform convergence in probability. □

1.4 Semimartingales are good integrators

1.4.1 Prerequisites. Section 3.1 on discrete-time martingales is needed now.

1.4.2 Definition. A cadlag adapted process $X\colon \Omega \times [0,1] \to \mathbb{R}^d$ is a *semimartingale* if $X = M + A$ for a local martingale $M \in \mathcal{M}^d_{\text{loc}}$ and a finite variation process $A \in \mathcal{V}^d$. It is called *special* if A can be chosen predictable.

1.4.3 Remark. The semimartingale and good integrator properties are stable under stopping and localization. E

1.4.4 Lemma. *Let $X\colon \Omega \times [0,1] \to \mathbb{R}^d$ be cadlag adapted.*

(i) If X has finite variation, then X is a good integrator.

(ii) If X is a square integrable martingale, then X is a good integrator.

(iii) If X is a martingale, then X is a good integrator.

In particular, all semimartingales are local good integrators.

Proof. Let $H \in \mathcal{E}^d$ be as in Definition 1.1.2.

(i) Variation estimate: if X has finite variation and $X_0 = 0$, then

$$\begin{aligned}\operatorname{Var}(H \bullet X) &= \operatorname{Var}\left(\sum_{i=1}^n h_i \cdot (X^{T_{i+1}} - X^{T_i})\right) \\ &\leq \sum_{i=1}^n \sum_{j=1}^d |h_i^j| \operatorname{Var}\left(X^{j,T_{i+1}} - X^{j,T_i}\right) \leq \|H\|_\infty \sum_{j=1}^d \operatorname{Var}(X^j).\end{aligned}$$

(ii) Ito's inequality: if X is a square-integrable martingale, then

$$\begin{aligned}\mathbb{E}[(H \bullet X)_1^2] &= \mathbb{E}\left[(h_0 \cdot X_0 + \sum_{i=1}^n h_i \cdot (X_{T_{i+1}} - X_{T_i}))^2\right] \\ &= \mathbb{E}\left[(h_0 \cdot X_0)^2 + \sum_{i=1}^n (h_i \cdot (X_{T_{i+1}} - X_{T_i}))^2\right] \\ &\leq \mathbb{E}\left[\|h_0\|^2_{\mathbb{R}^d}\|X_0\|^2_{\mathbb{R}^d} + \sum_{i=1}^n \|h_i\|^2_{\mathbb{R}^d}\|X_{T_{i+1}} - X_{T_i}\|^2_{\mathbb{R}^d}\right] \\ &\leq d\|H\|^2_\infty \mathbb{E}\left[\|X_0\|^2_{\mathbb{R}^d} + \sum_{i=1}^n \|X_{T_{i+1}} - X_{T_i}\|^2_{\mathbb{R}^d}\right] = d\|H\|^2_\infty \mathbb{E}\left[\|X_1\|^2_{\mathbb{R}^d}\right].\end{aligned}$$

(iii) Burkholder's inequality: if X is a real-valued martingale ($d = 1$), then

$$\forall \alpha > 0: \qquad \alpha \mathbb{P}[|H \bullet X|_1^* \geq \alpha] \leq 18 \mathbb{E}[|X_1|].$$

This follows from Theorem 3.1.1.(v) by approximating the supremum over all t by a supremum over finitely many stopping times. Thus, one obtains for general d that

$$\begin{aligned}\alpha \mathbb{P}[|H \bullet X|_1^* \geq \alpha] &= \alpha \mathbb{P}\left[\left|\frac{H}{\|H\|_\infty} \bullet X\right|_1^* \geq \frac{\alpha}{\|H\|_\infty}\right] \\ &\leq \alpha \sum_{j=1}^d \mathbb{P}\left[\left|\frac{H^j}{\|H\|_\infty} \bullet X^j\right|_1^* \geq \frac{\alpha}{\|H\|_\infty d}\right] \leq 18\|H\|_\infty d \sum_{j=1}^d \mathbb{E}[|X_1^j|]. \ \square\end{aligned}$$

1.4.5 Remark. The proof of Lemma 1.4.4 shows: E

(i) If X is a local martingale and H is locally bounded, then $H \bullet X$ is a local martingale.

(ii) If X is a locally square integrable local martingale and H is locally bounded, then $H \bullet X$ is a locally square integrable local martingale.

Note however that there are locally unbounded integrands H and martingales X such that $H \bullet X$ is not a local martingale; cf. Section 3.9.

1.5 Good integrators are semimartingales

1.5.1 Prerequisites. Section 3.4 on Hahn–Banach and Section 3.5 on the L^1 version of Komlos are needed now.

1.5.2 Definition. Let $X \colon \Omega \times [0,1] \to \mathbb{R}^d$ be cadlag adapted.

(i) X is a *quasimartingale* if it has bounded mean variation

$$MV(X) := \sup_\pi MV(X,\pi) := \sup_\pi \mathbb{E} \sum_{i=0}^n \left\|\mathbb{E}[X_{t_{i+1}} - X_{t_i}]\right\|_{\mathbb{R}^d} < \infty,$$

where π denotes a partition $0 = t_0 \leq t_1 \leq \dots \leq t_{n+1} = 1$.

(ii) X is of class (D) if the set of all S_T, where T is a finite stopping time, is uniformly integrable.

1.5.3 Theorem (Bichteler, Dellacherie)**.** *Any good integrator is a semimartingale.*

Proof. We outline the proof of Beiglböck and Siorpaes [BS14].

- The process $J_t = \sum_{s \leq t} \Delta S_s \mathbb{1}_{\{|\Delta S_s| \geq 1\}}$ of cumulated big jumps has finite variation and is therefore a semimartingale and a good integrator. It remains to show that the locally bounded good integrator $S - J$ is a semimartingale. Thus, by localization, we may assume wlog. that S is bounded.
- [BS14, Lemma 4.1]: Let $n \in \mathbb{N}$, let π_n be the dyadic partition of $[0,1]$ with grid size 2^{-n}, and let H be the elementary predictable process which equals the sign of $\mathbb{E}[S_{t_{i+1}} - S_{t_i}|\mathcal{F}_{t_i}]$ on each interval $(t_i, t_{i+1}]$ of π_n. Then $MV(X,\pi_n) = (H^n \bullet S)_1$; in financial terms the mean variation is replicated by trading in S. The elementary integral $(H^n \bullet S)_1$ is bounded in probability because S is a good integrator. Thus, for each $\epsilon > 0$ there is $C > 0$ such $|H^n \bullet S|_1$ is bounded by C with probability at least $1 - \epsilon$. Equivalently, $MV(S^{T_n}, \pi_n) \leq C$, where T^n is the first dyadic time such that $|H^n \bullet S| \geq C - \|\Delta S\|_\infty$.

- [BS14, Lemma 4.2]: Some forward convex combinations of $(T_n)_{n\in\mathbb{N}}$ converges in L^1 to a stopping time T by the Komlos lemma, and S^T has bounded mean variation. Moreover, T is localizing in the sense that $\mathbb{P}[T<\infty]$ becomes small for small ϵ.
- By Rao's theorem, the bounded mean variation process S^T is the difference of two submartingales. The proof is elementary and uses again the L^1 version of the Komlos lemma [Kal06, Theorem 23.20].
- Any supermartingale is locally of class (D) [BS14, Lemma 5.2].
- By Doob–Meyer any submartingale of class (D) is a semimartingale. The proof uses Doob's decomposition (Theorem 3.1.1.(i)) in discrete time and extracts a continuous-time limit using the L^1 version of Komlos [BSV12]. □

1.5.4 Remark. The proofs of Rao's theorem [Kal06, Theorem 23.20] and the continuous-time Doob–Meyer decomposition [BSV12] are interesting and were read in class.

1.6 Riemann sums

1.6.1 Definition.

(i) A *random partition* is a finite sequence of stopping times $0=T_0\leq\dots\leq T_{n+1}=1$.

(ii) A sequence of random partitions *tends to the identity* if

$$\lim_{n\to\infty}\sup_k |T^n_{k+1}-T^n_k|=0.$$

(iii) A cadlag adapted process X *sampled* at a random partition π is the cadlag adapted process

$$X^\pi=\sum_k X_{T_k}\mathbb{1}_{[\![T_k,T_{k+1}[\![}.$$

In the following we write $X^n\xrightarrow{up}X$ if $|X^n-X|^*_1\to 0$ in probability.

1.6.2 Lemma. *Let $X\in\mathcal{I}^d$, let Y be $\mathbb{R}^d$-valued cadlag adapted, and let $(\pi_n)_{n\in\mathbb{N}}$ be a sequence of random partitions tending to the identity.*

(i) $Y^{\pi_n}_-\bullet X\xrightarrow{up}Y_-\bullet X$.

(ii) If $X^{\pi_n}\xrightarrow{up}X$, *then* $Y^{\pi_n}_-\bullet X^{\pi_n}\xrightarrow{up}Y_-\bullet X$.

Proof.

(i) Let $Y^k\xrightarrow{up}X$ for elementary $Y^k_-\in\mathcal{E}^d$. Then

$$\begin{aligned}[\![|(Y_--Y^{\pi_n}_-)\bullet X|^*_1]\!]_0&\leq[\![|(Y_--Y^k_-)\bullet X|^*_1]\!]_0\\&\quad+[\![|(Y^k_--(Y^k)^{\pi_n}_-)\bullet X|^*_1]\!]_0+[\![|((Y^k)^{\pi_n}_--Y^{\pi_n}_-)\bullet X|^*_1]\!]_0.\end{aligned}$$

Choose k large to make the first and third summands small. Then choose n large to make the second summand small using the right-continuity of S.

(ii) This follows from

$$Y^{\pi_n}_-\bullet(X-X^{\pi_n})=Y^{\pi_n}(X-X^{\pi_n})\xrightarrow{up}0.$$ □

1.6.3 Remark. E

(i) For any cadlag adapted process X there exists a sequence $(\pi_n)_{n\in\mathbb{N}}$ of random partitions such that $X^{\pi_n} \xrightarrow{up} X$.

(ii) There is a cadlag adapted process X and a sequence of random partitions such that $X^{\pi_n} \xrightarrow{up} X$ does *not* hold.

1.7 Quadratic variation

Recall from Lemma 1.3.4 that caglad adapted processes are integrable with respect to any good integrator.

1.7.1 Definition. Let $X, Y \in \mathcal{I}$. We use the convention $X_{0-} = Y_{0-} = 0$.

(i) The *quadratic variation* of X is the cadlag adapted process

$$[X] = [X, X] = X^2 - 2X_- \bullet X.$$

(ii) The *quadratic covariation* of X and Y is the cadlag adapted process

$$[X, Y] = XY - X_- \bullet Y - Y_- \bullet X.$$

1.7.2 Lemma. *Let $X \in \mathcal{I}$.*

(i) $[X]$ is cadlag increasing, $[X]_0 = X_0^2$, and $\Delta[X] = (\Delta X)^2$.

(ii) If $(\pi_n)_{n\in\mathbb{N}}$ is a sequence of random partitions tending to the identity, then

$$[X^{\pi_n}] = X_0^2 + \sum_i (X^{T_{i+1}^n} - X^{T_i^n})^2 \xrightarrow{up} [X].$$

(iii) For any stopping time T, $[X^T] = [X]^T$.

(iv) For any $H \in L(X)$, $[H \bullet X] = H^2 \bullet [X]$.

Proof.

(i) follows from $\Delta(X_- \bullet X) = X_- \Delta X$.

(ii) Assume wlog. $X_0 = 0$ by replacing X by $X - X_0$. Then Lemma 1.6.2 implies that

$$\begin{aligned}
&\sum_i (X^{T_{i+1}^n} - X^{T_i^n})^2 \\
&\quad = \sum_i \left((X^2)^{T_{i+1}^n} - (X^2)^{T_i^n}\right) - 2\sum_i X^{T_i^n}\left(X^{T_{i+1}^n} - X^{T_i^n}\right) \\
&\quad = X^2 - 2\sum_i X_{T_i^n}\left(X^{T_{i+1}^n} - X^{T_i^n}\right) \\
&\quad = X^2 - 2X_-^{\pi_n} \bullet X \xrightarrow{up} X^2 - 2X_- \bullet X = [X].
\end{aligned}$$

(iii) $[X^T] = (X^T)^2 - 2X_-^T \bullet X^T = (X^2)^T - 2(X_- \bullet X)^T = [X]^T$.

(iv) holds for elementary H by (iii), for $H \in L^1(X)$ by taking limits, and for $H \in L(X)$ by localization using (iii). ☐

1.7.3 Remark. E

(i) The above statements about $[X]$ extend to corresponding statements about $[X, Y]$ by polarization. In particular, $[X, Y]$ is a finite variation process.

(ii) The set of semimartingales is an algebra.

(iii) The jumps of any semimartingale are square summable.

1.7.4 Lemma. *Let $X \in \mathcal{I}$.*

(i) If X has finite variation, then $[X] = \sum_{s\leq \cdot}(\Delta X_s)^2$.

(ii) If X is a continuous local martingale, then $X^2 - [X]$ is a local martingale. Moreover, $X = X_0$ if and only if $[X] = X_0^2$.

(iii) If X is a locally square integrable local martingale, then $[X]$ is the unique finite variation process A such that $X^2 - A$ is a local martingale, $\Delta A = (\Delta X)^2$, and $A_0 = X_0^2$.

Proof.

(i) The stochastic integral coincides with the Lebesgue-Stieltjes integral, which satisfies the change of variables formula $X^2 = 2X_- \bullet X$.

(ii) $X^2 - [X] = 2X_\bullet X$ is a local martingale; see Remark 1.4.5. If $X = X_0$, then $[X] = X_0^2$. Conversely, if $[X] = X_0^2$, then $X^2 - X_0^2$ is a non-negative continuous local martingale with initial value zero, which implies that $X^2 - X_0^2$ is identically zero.

(iii) Existence: the process $A = [X]$ satisfies the stated properties because $X_- \bullet X$ is a local martingale; see Remark 1.4.5. Uniqueness: if A and B satisfy the stated properties, then $A - B$ is a continuous local martingale with initial value zero and paths of finite variation. Then $[A - B]$ vanishes by (i), which implies that $A - B$ vanishes by (ii). □

1.7.5 Example. E

(i) If $X_t = t$, then $[X] = 0$.

(ii) If X is Brownian motion, then $[X]_t = t$.

(iii) If X is a Poisson process, then $[X] = X$.

1.8 Burkholder–Davis–Gundy inqualities

1.8.1 Theorem. *For any $p \in [1, \infty)$ there exist constants a_p and b_p such that every local martingale X satisfies*

$$\mathbb{E}\big[[X,X]_1^{p/2}\big] \leq a_p \mathbb{E}\big[|X^p|_1^*\big], \qquad \mathbb{E}\big[|X^p|_1^*\big] \leq b_p \mathbb{E}\big[[X,X]_1^{p/2}\big],$$

Proof. This holds for every piecewise constant local martingale by [BS15], and for general local martingales by approximation as in Lemma 1.7.2. □

1.8.2 Remark.

(i) Theorem 1.8.1 gives control over the integral $H \bullet M$ for unbounded integrands H and local martingales M. As the proof is elementary, this can be taken as a starting point for defining stochastic integration; see [SC02].

(ii) The proof of Theorem 1.8.1 in [BS15] is elementary and interesting and was read in class. Financially speaking the key observation is that $|X^p|^*$ can be super-replicated by a constant times $[X]^{p/2}$ plus a stochastic integral, and vice versa.

1.9 Itô's formula

1.9.1 Theorem (Itô's formula)**.** *Let $X \in \mathcal{I}^d$, and let $f \in C^2(\mathbb{R}^d)$, and let X^i denote the i-th component of X. Then $f(X) \in \mathcal{I}$ and*

$$\begin{aligned} f(X) = f(X_0) &+ \sum_{i=1}^{d} (\partial_i f)(X_-) \bullet X + \frac{1}{2} \sum_{i,j=1}^{d} (\partial_{i,j} f)(X_-) \bullet [X^i, X^j] \\ &+ \sum_{s \leq \cdot} \Bigg(f(X) - f(X_-) - \sum_{i=1}^{d} (\partial_i f)(X_-) \Delta X^i \\ &\qquad - \frac{1}{2} \sum_{i,j=1}^{d} (\partial_{i,j} f)(X_-) \Delta X^i \Delta X^j \Bigg). \end{aligned}$$

Proof. We follow [Tei17, Lemma 5.8 and Theorem 5.10]. Let $(\pi_n)_{n \in \mathbb{N}}$ be a sequence of partitions tending to the identity such that $X^{\pi_n} \xrightarrow{up} X$; cf. Remark 1.6.3. As X^{π_n} is piecewise constant, Itô's formula with X replaced by X^{π_n} holds for all $n \in \mathbb{N}$ by direct inspection. It remains to take a limit $n \to \infty$ uniformly in probability.

The summands on the first line of Itô's formula converge uniformly in probability as $n \to \infty$ by Lemma 1.6.2. For any fixed $r > 0$, the sum $\sum_{s \leq \cdot}$ spanning the second and third line is split into a finite sum $\sum_A$ over large jumps with $|\Delta X| > r$ and a possibly countable sum $\sum_B$ of small jumps with $|\Delta X| \leq r$. As $\sum_A$ is a finite sum, it converges uniformly in probability as $n \to \infty$. Moreover, $\sum_B$ can be made arbitrarily small by choosing r sufficiently small thanks to the estimates $\sum_{s \leq \cdot} (\Delta X_s^i)^2 \leq [X^i]$ and

$$\begin{aligned} f(y) - f(x) - \sum_{i=1}^{d} (\partial_i f)(x)(y-x)^i - \frac{1}{2} \sum_{i,j=1}^{d} (\partial_{i,j} f)(x)(y-x)^i (y-x)^j \\ = o\big(\|y-x\|\big) \|y-x\|^2. \quad \square \end{aligned}$$

1.9.2 Remark. E

(i) Itô's formula can be rewritten as

$$\begin{aligned} f(X) = f(X_0) &+ \sum_{i=1}^{d} (\partial_i f)(X_-) \bullet X + \frac{1}{2} \sum_{i,j=1}^{d} (\partial_i \partial_j f)(X_-) \bullet [X^i, X^j]^c \\ &+ \sum_{s \leq \cdot} \Bigg(f(X) - f(X_-) - \sum_{i=1}^{d} (\partial_i f)(X_-) \Delta X^i \Bigg), \end{aligned}$$

where $[X]^c$ denotes the continuous part of $[X]$, which is defined as $[X]^c = [X] - \sum_{s \leq \cdot} \Delta [X]_s = \sum_{s \leq \cdot} (\Delta X_s)^2$.

(ii) Itô's formula can be used to derive and prove explicit solution formulas for the following stochastic differential equations: E

- Ornstein–Uhlenbeck: $dX_t = \kappa(\theta - X_t)dt + \sigma dW_t$ with $\kappa > 0$, $\theta, \sigma \in \mathbb{R}$, and W Brownian motion.
- Stochastic exponential: $dX_t = X_{t-} dZ_t$ with $Z \in \mathcal{I}$.

2 Financial mathematics

The cornerstone of modern mathematical finance is the fundamental theorem of asset pricing. It establishes an equivalence between arbitrage-properties of market models and the existence of so-called risk-neutral probabilities. Intuitively, the risk-neutral probability measure accounts for the risk preferences of the market participants by assigning higher weight to adverse extreme events. The resulting framework for option pricing and hedging has become the industry standard.

Setting. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, \mathbb{P})$ be filtered probability space satisfying the usual conditions, and let $d \in \mathbb{N}$. All processes are defined on $[0,1]\times\Omega$, and all stopping times take values in $[0,1]\cup\{\infty\}$. Recall that $L(S)$ denotes the set of all processes ϕ such that $\phi \bullet S$ is well-defined. It is assumed throughout this section that $\phi_0 = 0$ and therefore $(\phi \bullet S)_0 = 0$.

2.1 Financial markets

2.1.1 Definition.

(i) A *financial market* is a semimartingale $S = (S^1, \dots, S^d)$.

(ii) A *trading strategy* is a process $\phi = (\phi^1, \dots, \phi^d) \in L(S)$.

(iii) The *wealth process* associated to ϕ with initial wealth $V_0 \in L^0(\mathcal{F}_0)$ is the cadlag adapted process

$$V = V_0 + \phi \bullet S.$$

(iv) A *portfolio* is a triple (S, ϕ, V) as above.

2.1.2 Remark.

- There can be arbitrary short and long positions. There are no transaction costs, market impacts of trades, bid-ask spreads, and interest rates.
- The wealth process is defined such that all gains and losses are due to movements of the stock market.
- Discrete-time models can be treated in the framework by setting ϕ and S piecewise constant. Models with infinite time horizon work in a very similar way [SC02].

2.1.3 Example. Let W, B be Brownian motions, let B^H be fractional Brownian motion of Hurst index $H \in (0,1)$, let $0 = t_0 \le \dots \le t_{n+1} = 1$, let $a, b, g, \mu, \sigma \in \mathbb{R}$, and let $\beta > 0$.

- Bachelier: $dS_t = \mu dt + \sigma dW_t$.
- Black–Scholes: $dS_t = S_t(\mu dt + \sigma dW_t)$.
- Heston: $dS_t = S_t(\mu dt + \sqrt{V_t} dW_t)$, $dV_t = (b - \beta V_t)dt + \sigma\sqrt{V_t} dB_t$.
- Rough Bergomi: $dS_t = S_t(\mu dt + \sqrt{V_t} dW_t)$, $V_t = V_0 \exp(\sigma B_t^H - at)$.
- Cox–Ross–Rubinstein: $S_t = S_0 \prod_{t_i \le t} R_i$, $R_i \in \{g, b\}$.

2.1.4 Remark. To get a feel for how these models are used, one can download some stock prices (e.g. of the SNP 500 index) and estimate the volatility parameter σ in the Bachelier and Black–Scholes models (e.g. using the discrete approximations of Lemma 1.7.2). E

2.2 Admissible portfolios

2.2.1 Definition. A portfolio (S, ϕ, V) is called *admissible* if it is *a-admissible* for some $a \geq 0$, i.e.,

$$\mathbb{P}[V \geq -a] = 1.$$

2.2.2 Remark.

- Economically, the admissibility condition corresponds to a finite credit line.
- Without the admissibility condition, even the nicest markets admit arbitrages e.g. in the form of doubling strategies. E
- Due to the admissibility condition there may be no simple admissible trading strategies at all: consider e.g. asset prices whose jumps are unbounded from below. Thus, meaningful fundamental theorems of asset pricing require a larger class of strategies than just simple ones. E
- In the fundamental theorem of asset pricing, the main role of the admissibility condition is to ensure that wealth processes are supermartingales under the risk-neutral measure.

2.3 Self-financing portfolios

2.3.1 Definition. A portfolio (S, ϕ, V) is called *self-financing* if

$$V = \sum_i \phi^i S^i.$$

2.3.2 Remark.

- The self-financing condition means that all of the gains and losses are re-invested in the assets; no money is put into or taken out from the portfolio.
- In discrete time this means that the wealth does not change when the portfolio is rebalanced: E

$$\sum_i \phi^i_n S^i_n \xrightarrow{\text{rebalancing of } \phi} \sum_i \phi^i_{n+1} S^i_n \xrightarrow{\text{evolution of } S} \sum_i \phi^i_{n+1} S^i_{n+1}.$$

The self-financing condition can always be ensured by adding to the portfolio a constant asset $S^0 = 1$, which can be thought of as a bank account holding the gains and losses from trading.

2.3.3 Lemma. *There is a one-to-one correspondence between:*

(i) Portfolios $((S^1, \dots, S^d), (\phi^1, \dots, \phi^d), V)$*; and*

(ii) Self-financing portfolios $((S^0, \dots, S^d), (\phi^0, \dots, \phi^d), V)$ *with* $S^0 = 1$.

Proof.

(i)⇝(ii): Augment the market to $d + 1$ dimensions by setting

$$S^0 = 1, \qquad \phi^0 = V_- - \sum_{i=1}^d \phi^i S^i_-.$$

Then the resulting $(d + 1)$-dimensional market is self-financing because

$$V = V_- + \Delta V = \sum_{i=0}^d \phi^i S^i_- + \Delta V = \sum_{i=0}^d \phi^i S^i_- + \sum_{i=1}^d \phi^i \Delta S^i = \sum_{i=0}^d \phi^i S^i.$$

(i)⇝(ii): Discard the zero-th components of S and ϕ. □

2.4 Discounting

Discounting is a change of reference unit, also called numeraire. As a notational distinction, undiscounted quantities will carry a tilde.

2.4.1 Definition.

(i) A *numeraire* is a real-valued semimartingale $\tilde{S}^0$ which satisfies $\tilde{S}^0 > 0$, $\tilde{S}^0_- > 0$, and $\tilde{S}^0_0 = 1$.

(ii) A portfolio $(\tilde{S}, \tilde{\phi}, \tilde{V})$ is called *admissible with respect to* $\tilde{S}^0$ if there is $a \geq 0$ such that

$$\mathbb{P}[\tilde{V} \geq -a\tilde{S}^0] = 1.$$

2.4.2 Example. Common choices of numeraires are:

- Domestic or foreign currencies;
- Bank account processes, i.e., deposits with continuously (or, in practice, daily) compounded interest;
- Highly diversified market indices, which approximate the growth-optimal portfolio in the benchmark approach.

2.4.3 Lemma. *Let $\tilde{S}^0$ be a bank account process. Then the discounting relations*

$$S = \tilde{S}/\tilde{S}^0, \qquad \phi = \tilde{\phi}, \qquad V = \tilde{V}/\tilde{S}^0$$

establish a one-to-one correspondence between:

(i) Self-financing portfolios $((\tilde{S}^0, \dots, \tilde{S}^d), (\tilde{\phi}^0, \dots, \tilde{\phi}^d), \tilde{V})$ which are admissible with respect to $\tilde{S}^0$.

(ii) Admissible self-financing portfolios $((S^0, \dots, S^d), (\phi^0, \dots, \phi^d), V)$ with $S^0 = 1$.

Proof.

(i)⇝(ii): Given $(\tilde{S}, \tilde{\phi}, \tilde{V})$, define (S, ϕ, V) by the discounting relations, and let $D = 1/\tilde{S}^0$. Recall from the proof of Lemma 2.3.3 that the self-financing condition implies

$$\tilde{V}_- = \sum_{i=0}^{d} \tilde{\phi}^i \tilde{S}^i_-.$$

This implies that V is the discounted wealth process associated to ϕ:

$$\begin{aligned}
V = \tilde{V}D &= \tilde{V}_0 D + \tilde{V}_- \bullet D + D_- \bullet \tilde{V} + [\tilde{V}, D] \\
&= V_0 + \tilde{V}_- \bullet D + D_- \bullet (\tilde{\phi} \bullet \tilde{S}) + [\tilde{\phi} \bullet \tilde{S}, D] \\
&= V_0 + \tilde{V}_- \bullet D + \tilde{\phi} \bullet (D_- \bullet \tilde{S} + [\tilde{S}, D]) \\
&= V_0 + \tilde{V}_- \bullet D + \tilde{\phi} \bullet (D\tilde{S} - \tilde{S}_- \bullet D) \\
&= V_0 + \Big(\tilde{V}_- - \sum_{i=0}^{d} \tilde{\phi}^i \tilde{S}^i_-\Big) \bullet D + \tilde{\phi} \bullet (D\tilde{S}) \\
&= V_0 + \tilde{\phi} \bullet (D\tilde{S}) = V_0 + \phi \bullet S.
\end{aligned}$$

The self-financing and admissibility conditions of (S, ϕ, V) are trivial to check.

(ii)⇝(i): Similar to the above. □

2.4.4 Remark.

(i) Lemma 2.4.3 uses the general vector integral and breaks down for the component-wise integral. Indeed, there is a discounted self-financing portfolio (S, ϕ, V) and a bank account process $\tilde{S}^0$ such that the component-wise integral $\tilde{\phi} \bullet \tilde{S}$ does not exist. E

(ii) No-arbitrage theory is typically applied to discounted assets.

2.5 Fundamental theorem of asset pricing

2.5.1 Prerequisites. Section 3.9 on sigma-martingales is needed now.

2.5.2 Definition. Let S be a financial market.

(i) Claims which are *replicable* and *super-replicable* at zero initial wealth:

$$\begin{aligned} K &= \{f \in L^0; f = (\phi \bullet S)_1 \text{ for some admissible } \phi\} \cap L^\infty, \\ C &= \{f \in L^0; f \leq (\phi \bullet S)_1 \text{ for some admissible } \phi\} \cap L^\infty. \end{aligned}$$

(ii) S satisfies

$$\begin{aligned} \text{no arbitrage (NA)} &\Leftrightarrow C \cap L^\infty_+ = \{0\}, \\ \text{no free lunch with vanishing risk (NFLVR)} &\Leftrightarrow \overline{C} \cap L^\infty_+ = \{0\}, \\ \text{no free lunch (NFL)} &\Leftrightarrow \overline{C}^* \cap L^\infty_+ = \{0\}, \\ \text{completeness} &\Leftrightarrow (K \cap -K) + \mathbb{R} = L^\infty. \end{aligned}$$

Here $\overline{C}$ and $\overline{C}^*$ denote the closure and weak* closure in L^∞, respectively.

(iii) Equivalent *separating and sigma-martingale measures*:

$$\begin{aligned} \mathcal{M}^e_{\text{sep}}(S) &= \{\mathbb{Q} \sim \mathbb{P}; \mathbb{E}_\mathbb{Q}[f] \leq 0 \text{ for all } f \in C\}, \\ \mathcal{M}^e_\sigma(S) &= \{\mathbb{Q} \sim \mathbb{P}; S \text{ is } \mathbb{Q}\text{-sigma martingale}\}. \end{aligned}$$

Similarly, $\mathcal{M}^a_{\text{sep}}$ and $\mathcal{M}^a_\sigma$ for absolutely continuous measures.

2.5.3 Theorem (Fundamental theorem of asset pricing)**.** *Let S be a financial market.*

(i) S satisfies (NFLVR) $\Leftrightarrow \mathcal{M}^e_\sigma(S) \neq \emptyset$.

(ii) S satisfies (NFLVR) and is complete $\Leftrightarrow \mathcal{M}^e_\sigma(S) = \{\mathbb{Q}\}$ for some $\mathbb{Q}$.

Proof. The line of argument is as follows:

- Section 2.6: Identify the polars of C and K.
- Section 2.7: The bipolar theorem gives a preliminary version of the FTAP with $\mathcal{M}^e_{\text{sep}}(S)$ in place of $\mathcal{M}^e_\sigma(S)$ and under the assumption that $C = \overline{C}^*$.
- Section 2.8: Identify $\mathcal{M}^e_{\text{sep}}(S)$ and $\mathcal{M}^e_\sigma(S)$.
- Sections 2.9–2.12: Show that (NFLVR) implies $C = \overline{C}^*$. This is the difficult part. □

2.5.4 Remark.

(i) It is illuminating to formulate the contrapositions of (NA), (NFLVR), and (NFL): what is an arbitrage, what is a free lunch with vanishing risk, and what is a free lunch? E

(ii) The (NFLVR) condition strikes a balance between economic plausibility and mathematical usefulness:

- (NA) is economically convincing but mathematically insufficient because it does not imply $\mathcal{M}^e_\sigma(S) \neq \emptyset$ [DS94, Example 9.7.7].
- (NFL) mathematically sufficient but economically questionable because arbitrages in the weak* closure might be too complex to be siphoned off by traders in equilibrium.
- (NFLVR) is both mathematically and economically satisfactory thanks to the surprising theorem of Delbaen and Schachermayer that (NFLVR) implies (NFL).

(iii) Trivially, $C \subseteq \overline{C} \subseteq \overline{C}^*$ and (NFL) $\Rightarrow$ (NFLVR) $\Rightarrow$ (NA).

(iv) Sigma-martingales are needed because the arbitrage conditions do not imply the existence of an equivalent local martingale measure unless S is locally bounded [SC02, Example 5.3].

(v) A model can be complete, but not free of arbitrage.

(vi) Definition 2.5.2 depends on $\mathbb{P}$ only through its null sets.

(vii) If Ω is a finite set, then all strategies are admissible, and one has $C = \overline{C} = \overline{C}^*$. Indeed, $C = \langle c_1, \dots, c_n \rangle_{\text{cone}}$ for some $n \in \mathbb{N}$ and $c_1, \dots, c_n \in L^\infty = L^2$, and it can be shown by induction on n that C is closed. E

2.5.5 Example.

(i) Some common claims are:

- European call (put) options: the holder has the right to buy (sell) one unit of the asset S at time T at the preset strike K; the time-T value of the option is $(S_T - K)_+$ for calls and $(K - S_T)_+$ for puts.
- American call (put) options: a European call (put) which may be exercised at any time before T; the time-T value of the option is $\operatorname{ess\,sup}(S_\tau - K)_+$ for calls and $\operatorname{ess\,sup}(K - S_\tau)_+$ for puts, where the supremum is over $[0, T]$-valued stopping times τ.

(ii) No-arbitrage and completeness properties of some common models: E

- All models mentioned in Exercise 2.1.3 are free of arbitrage, provided that 1 is contained in the convex hull of $\{g, b\}$.
- The Bachelier, Black–Scholes and Cox–Ross–Rubinstein models are complete, whereas the Heston and Rough Bergomi models are incomplete.

These statements follow easily from the fundamental theorem of asset pricing if one skips the verification that the candidate measure changes from Girsanov's theorem are true martingales.

(iii) Option pricing in some common models:

- The Bachelier and Black–Scholes models admit explicit formulas for European call and put prices, which can be derived by integration over the density of the normal distribution. E
- The Heston model admits explicit formulas for Fourier payoffs, and general payoffs can be treated by Fourier inversion.
- In the Rough Bergomi model option prices can be calculated using Monte Carlo simulations.
- In the Cox–Ross–Rubinstein model option prices can be calculated by backwards induction (dynamic programming) on the binary tree of possible outcomes. E

2.6 Polars of replicable and super-replicable claims

2.6.1 Prerequisites. Section 3.7 on the bipolar theorem is needed now.

2.6.2 Lemma. *Let $\mathbb{Q}_0 \in \mathcal{M}^e_{\mathrm{sep}}(S)$, consider C and K as subsets of L^∞ with the weak* topology, and write $\overline{C}^*$ for the weak star closure of C.*

(i) The polar of C is

$$C^0 = \left\langle \frac{d\mathbb{Q}}{d\mathbb{P}}; \mathbb{Q} \in \mathcal{M}^a_{\mathrm{sep}}(S) \right\rangle_{\mathrm{cone}}.$$

(ii) If $C = \overline{C}^*$, then the polar of $K \cap -K$ is

$$(K \cap -K)^0 = \left\langle \frac{d\mathbb{Q}}{d\mathbb{P}}; \mathbb{Q} \in \mathcal{M}^a_{\mathrm{sep}}(S) \right\rangle_{\mathrm{vector}}.$$

Proof.

(i) $\subseteq$ Let $0 \neq z \in C^0 \subseteq L^1$. Then $\mathbb{P}[z \geq 0] = 1$ because $L^\infty_- \subseteq C$. Then $\mathbb{Q} := \frac{z}{\mathbb{E}_\mathbb{P}[z]}\mathbb{P} \in \mathcal{M}^a_{\mathrm{sep}}$.

$\supseteq$ C^0 is a cone, and $\{\frac{d\mathbb{Q}}{d\mathbb{P}}; \mathbb{Q} \in \mathcal{M}^a_{\mathrm{sep}}\} \subseteq C^0$.

(ii) $\subseteq$ Let

$$f \in \left\langle \frac{d\mathbb{Q}}{d\mathbb{P}}; \mathbb{Q} \in \mathcal{M}^a(S) \right\rangle^0_{\mathrm{vector}} \subseteq \left\langle \frac{d\mathbb{Q}}{d\mathbb{P}}; \mathbb{Q} \in \mathcal{M}^a(S) \right\rangle^0_{\mathrm{cone}} = C^{00} = \overline{C}^* = C.$$

Then $f \leq g$ for some $g \in K$. This implies $f = g$ $\mathbb{Q}_0$-a.s. because $0 = \mathbb{E}_{\mathbb{Q}_0}[f] \leq \mathbb{E}_{\mathbb{Q}_0}[g] \leq 0$ by separation. As $\mathbb{Q}_0 \sim \mathbb{P}$ one obtains $f \in K$. The same argument with f replaced by $-f$ yields $f \in K \cap -K$. As f was general, we have shown that

$$\left\langle \frac{d\mathbb{Q}}{d\mathbb{P}}; \mathbb{Q} \in \mathcal{M}^a(S) \right\rangle^0_{\mathrm{vector}} \subseteq K \cap -K.$$

By the bipolar theorem,

$$(K \cap -K)^0 \subseteq \left\langle \frac{d\mathbb{Q}}{d\mathbb{P}}; \mathbb{Q} \in \mathcal{M}^a(S) \right\rangle_{\mathrm{vector}},$$

where we have used that the right-hand side is $\sigma(L^1, L^\infty)$-closed.

$\supseteq$ $(K \cap -K)^0$ is a vector space, and $\{\frac{d\mathbb{Q}}{d\mathbb{P}}; \mathbb{Q} \in \mathcal{M}^a_{\mathrm{sep}}\} \subseteq (K \cap -K)^0$. □

2.7 Preliminary version of the FTAP

Hahn–Banach and the bipolar theorem yield the following preliminary version of the FTAP.

2.7.1 Theorem (Kreps, Yan)**.** *Let S be a financial market.*

(i) S satisfies (NFL) $\Leftrightarrow \mathcal{M}^e_{\mathrm{sep}}(S) \neq \emptyset$.

(ii) Assume that S satisfies (NFL) and $C = \overline{C}^*$. Then S is complete $\Leftrightarrow \mathcal{M}^e_{\mathrm{sep}}(S) = \{\mathbb{Q}\}$ for some $\mathbb{Q}$.

Proof.

(i) $\Leftarrow$ Let $\mathbb{Q} \in \mathcal{M}^e_{\text{sep}}(S)$. Then $\mathbb{E}_\mathbb{Q}[f] \leq 0$ for all $f \in C$ because $\mathbb{Q}$ is separating. This extends to all $f \in \overline{C}^*$ because the map $f \mapsto \mathbb{E}_\mathbb{Q}[f] = \mathbb{E}[\frac{d\mathbb{Q}}{d\mathbb{P}} f]$ is weak* continuous. This implies $\overline{C}^* \cap L^\infty_+ = \{0\}$.

$\Rightarrow$ By Hahn–Banach on L^∞ with the weak* topology, any nonzero $x \in L^\infty_+$ can be separated from $\overline{C}^*$: there is $z_x \in L^1$ such that $\mathbb{E}[z_x x] > 0$ and $\mathbb{E}[z_x f] \leq 0]$ for all $f \in C$. Then $z_x \geq 0$ because $L^\infty_- \subseteq C$, and we may normalize to $\mathbb{E}[z_x] = 1$. There are $x_1, x_2, \dots$ with $\mathbb{P} \bigcup_i [x_i > 0] = 1$: otherwise, the supremal = maximal such probability could be enlarged by adding $x = \mathbb{1}_{\bigcap_i [x_i = 0]}$ to the sequence. Set $z = \sum_i 2^{-i} z_{x_i}$. Then $\mathbb{P}[z > 0] = 1$, $\mathbb{E}_\mathbb{P}[z] = 1$, and $\mathbb{Q} := z\mathbb{P} \in \mathcal{M}^e_{\text{sep}}(S)$.

(ii) $\Rightarrow$ Let S be complete, and let $A \in \mathcal{F}$. Then $\mathbb{1}_A = x + g$ for some $g \in K \cap -K$. Thus, all $\mathbb{Q}, \mathbb{Q}' \in \mathcal{M}^e_{\text{sep}}(S)$ satisfy

$$\mathbb{Q}'[A] = \mathbb{E}_{\mathbb{Q}'}[x + g] = x = \mathbb{E}_\mathbb{Q}[x + g] = \mathbb{Q}[A].$$

Thus, $\mathbb{Q}' = \mathbb{Q}$.

$\Leftarrow$ Let $\mathcal{M}^e_{\text{sep}}(S) = \{\mathbb{Q}\}$, let $f \in L^\infty$, and let $x = \mathbb{E}_\mathbb{Q}[f]$. Then $\mathbb{E}_\mathbb{Q}[f] = 0$. This implies $\mathbb{E}_{\mathbb{Q}'}[f] = 0$ for all $\mathbb{Q}' \in \mathcal{M}^a_{\text{sep}}$ by approximation

$$\mathcal{M}^e_{\text{sep}} \ni (1-\epsilon)\mathbb{Q}' + \epsilon\mathbb{Q} \xrightarrow[\epsilon \to 0]{} \mathbb{Q}.$$

Thus, $f \in \langle \frac{d\mathbb{Q}'}{d\mathbb{P}}; \mathbb{Q}' \in \mathcal{M}^a_{\text{sep}}(S) \rangle_{\text{vector}})^0 = (K \cap -K)^{00} = K \cap -K$. ☐

2.8 Separating and sigma-martingale measures

2.8.1 Prerequisites. Section 3.10 on semimartingale characteristics, Section 3.11 on the characteristics of sigma martingales, Section 3.12 on characteristics of stochastic integrals, and Section 3.13 on Girsanov's theorem are needed now.

2.8.2 Lemma (Delbaen, Schachermayer)**.** *Let S be a financial market.*

(i) $\mathcal{M}^e_\sigma(S) \neq \emptyset \Leftrightarrow \mathcal{M}^e_{\text{sep}}(S) \neq \emptyset$.

(ii) $\mathcal{M}^e_\sigma(S) = \{\mathbb{Q}\} \Leftrightarrow \mathcal{M}^e_{\text{sep}}(S) = \{\mathbb{Q}\}$.

(iii) S locally bounded $\Rightarrow \mathcal{M}^e_{\text{sep}}(S) = \mathcal{M}^e_\sigma(S) = \mathcal{M}^e_{\text{loc}}(S)$.

Proof. (i) and (ii) follow from (i)' and (ii)' below.

(i)' Claim: $\mathcal{M}^e_\sigma(S) \subseteq \mathcal{M}^e_{\text{sep}}(S)$. To see this, let $\mathbb{Q} \in \mathcal{M}^e_\sigma$ and $f \in C$. Then $f \leq (\phi \bullet S)_1$ for some admissible ϕ. By Ansel–Stricker, $\phi \bullet S$ is a $\mathbb{Q}$-local martingale, and by Fatou a $\mathbb{Q}$-supermartingale. Thus, $\mathbb{E}_\mathbb{Q}[f] \leq \mathbb{E}_\mathbb{Q}[(\phi \bullet S)_1] \leq 0$, and $\mathbb{Q} \in \mathcal{M}^e_{\text{sep}}(S)$.

(ii)' Claim: $\mathcal{M}^e_{\text{sep}}(S) \subseteq \overline{\mathcal{M}^e_\sigma(S)}$, where the closure is in total variation $\|\cdot\|_{\text{TV}}$. We sketch the proof of [Kab97, Addendum], restricting to $d = 1$. Let $\mathbb{Q} \in \mathcal{M}^e_{\text{sep}}(S)$, and let (b, c, K) be the differential characteristics of S with respect to A and h_d as in Lemma 3.10.5. Based on Girsanov (Lemma 3.13.2.(ii)), the Ansatz is to search for a measure $\mathbb{Q}' \sim \mathbb{Q}$ such that the differential characteristics under $\mathbb{Q}'$ become

$$b' = b + c\beta + \int h_d(x)(Y(x) - 1)K(dx), \qquad c' = c, \qquad K' = YK.$$

If Y is close to 1, then such a measure exists and is $\|\cdot\|_{\mathrm{TV}}$-close to $\mathbb{Q}$. In this case, by Lemma 3.11.1, S is a sigma-martingale under $\mathbb{Q}'$ if and only if

$$b + \int (xY(x) - h(x))K(dx) = 0. \qquad (*)$$

We distinguish three cases:

- If $b + \int (x - h(x))K(dx) = 0$, then $Y \equiv 1$ satisfies $(*)$.
- If $b + \int (x - h(x))K(dx) > 0$, then the support of K must be unbounded below: otherwise, there would be an admissible long position $\phi \geq 0$ with positive expected return $\mathbb{E}_{\mathbb{Q}}[(\phi \bullet S)_1] > 0$, a contradiction to the separation property of $\mathbb{Q}$. Thus, $(*)$ can be achieved by increasing $Y(x)$ for strongly negative x to a value slightly above 1.
- Similarly, if $b + \int (x - h(x))K(dx) < 0$, then the support of K must be unbounded below, and $(*)$ can be satisfied by increasing $Y(x)$ for strongly positive x to a value slightly above 1.

Y can be chosen predictably in (t, ω) using measurable selection techniques.

(iii) Claim: S locally bounded $\Rightarrow \mathcal{M}^e_{\mathrm{sep}}(S) = \mathcal{M}^e_\sigma(S) = \mathcal{M}^e_{\mathrm{loc}}(S)$. To see this, note from (i) that it suffices to show that $\mathcal{M}^e_{\mathrm{sep}}(S) \subseteq \mathcal{M}^e_{\mathrm{loc}}(S)$. Let $\mathbb{Q} \in \mathcal{M}^e_{\mathrm{sep}}$. By localization we may assume that S is bounded. Then all strategies $\phi = \pm \mathbb{1}_A \mathbb{1}_{(s,t]}$ with $A \in \mathcal{F}_s$ and $s, t \in [0,1]$ are admissible. By separation, $\mathbb{E}_{\mathbb{Q}}[(\pm \mathbb{1}_A \mathbb{1}_{(s,t]} \bullet S)_1] \leq 0$, and S is a $\mathbb{Q}$-martingale. □

As an aside, completeness is closely related to martingale representation:

2.8.3 Lemma. *Let S be a financial market, let the filtration satisfy $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_1 = \mathcal{F}$, and let $\mathbb{Q} \in \mathcal{M}^e_\sigma(S)$. Then $\mathcal{M}^e_\sigma(S) = \{\mathbb{Q}\}$ if and only if every $\mathbb{Q}$-sigma-martingale X can be represented as $X = X_0 + H \bullet S$ for some $H \in L(S)$.*

Proof.

$\Leftarrow$ Let $\mathbb{Q}' \in \mathcal{M}^e_\sigma$, let $A \in \mathcal{F}$, let $M_t = \mathbb{E}_{\mathbb{Q}}[\mathbb{1}_A | \mathcal{F}_t]$, and let $H \in L(S)$ such that $M = M_0 + H \bullet S$ under $\mathbb{Q}$. As M is bounded, $\mathbb{Q}'[A] = \mathbb{E}_{\mathbb{Q}'}[M_0 + (H \bullet S)_1] = M_0 = \mathbb{E}_{\mathbb{Q}'}[M_0 + (H \bullet S)_1] = \mathbb{Q}[A]$. Thus, $\mathbb{Q}' = \mathbb{Q}$.

$\Rightarrow$ See [SC02, Theorem 1.17]. □

2.9 No unbounded profit with bounded risk

2.9.1 Definition.

(i) Admissible and 1-admissible *wealth processes*:

$$\mathcal{X} = \{\phi \bullet S; \phi \text{ is admissible}\},$$
$$\mathcal{X}_1 = \{\phi \bullet S; \phi \text{ is 1-admissible}\}.$$

(ii) S satisfies

no unbounded profit with bounded risk (NUPBR) $\Leftrightarrow \{X_1; X \in \mathcal{X}_1\}$ is L^0-bounded,

predictable uniform tightness (PUT) $\Leftrightarrow \mathcal{X}_1$ is Emery-bounded.

2.9.2 Remark.

- (NUPBR) can be tested using only elementary strategies. This is not possible for (NA) and its variants.
- If S is cadlag adapted and $\mathcal{X}$ denotes the set of simple wealth processes, then (NUPBR) implies that S is a semimartingale [Kar13, Theorem 1.3].
- (NUPBR) fails if there exists $0 \neq \xi \in L^0_+$ which can be superreplicated at initial wealth x for all $x \in (0, \infty)$ [KKS16, Lemma A.1].

2.9.3 Lemma. *(NFLVR) $\Leftrightarrow$ (NA) + (NUPBR)*

Proof. We follow [Kab97, Lemma 2.2].

$\Rightarrow$ (NA) follows trivially. If (NUPBR) was violated, there would be $\alpha > 0$ and $X^n \in \mathcal{X}_1$ such that $\mathbb{P}[X^n_1 \geq n] \geq \alpha$. Set $f_n = (X^n_1/n) \wedge 1 \in C$. By the L^0 version of Komlos there are $g_n \in \langle f_n, f_{n+1}, \dots \rangle_{\text{conv}}$ such that $g_n \to g$ a.s. for some $g \geq 0$ with $\mathbb{P}[g > 0] =: 2\beta > 0$. By Egorov $g_n \to g$ uniformly on some set B with $\mathbb{P}[B] > 1 - \beta$. Thus, (NFLVR) fails because

$$C \ni g_n - g_n^+ \mathbb{1}_{\Omega \setminus B} = g_n \mathbb{1}_B - g_n^- \mathbb{1}_{\Omega \setminus B} \xrightarrow[n \to \infty]{L^\infty} g \mathbb{1}_B$$

and $\mathbb{P}[g\mathbb{1}_B > 0] \geq \beta > 0$.

$\Leftarrow$ If (NFLVR) failed, there would be $f_n \in C$ and $f \geq 0$ with $\mathbb{P}[f > 0] > 0$ such that $\|f_n - f\|_\infty \leq \frac{1}{n}$. By definition, $f_n \leq X^n_1$ for some $X^n \in \mathcal{X}$. Obviously, $X^n_1 \geq -\frac{1}{n}$. If there existed some $s \in [0, 1)$ with $\mathbb{P}[X^n_s < \frac{1}{n}] > 0$, then $\mathbb{1}_{\{X^n_s < -\frac{1}{n}\}} \mathbb{1}_{(s,1]} \bullet X^n$ would violate (NA). Thus, $X^n \geq -\frac{1}{n}$. By the L^0 version of Komlos we may assume that $X^n_1 \to g$ a.s. Thus, we have constructed $nX^n \in \mathcal{X}_1$ with $\mathbb{P}[nX^n_1 \to \infty] \geq \mathbb{P}[g > 0] \geq \mathbb{P}[f > 0] > 0$, a violation of (NUPBR). ☐

2.10 Supermartingale deflators

2.10.1 Definition. A *supermartingale deflator* for $1 + \mathcal{X}_1$ is a strictly positive cadlag adapted process with $D_0 \leq 1$ such that $D(1 + X)$ is a supermartingale for all $X \in \mathcal{X}_1$.

2.10.2 Lemma. *(NUPBR) $\Leftrightarrow$ there exists a supermartingale deflator.*

Proof. We skip the easier direction $\Leftarrow$, which is not needed subsequently, and sketch the proof in [KK07] under the assumption that $S > 0$ and $S_- > 0$. Then $S = \mathscr{E}(X)$ for some X with $\Delta X > -1$. For any trading strategy $\phi = (\phi^1, \dots, \phi^d)$ with wealth process $V = 1 + \phi \bullet S$, define the proportions of wealth invested in the assets as

$$\pi = \left(\frac{\phi^1 S^1_-}{V_-}, \dots, \frac{\phi^1 S^1_-}{V_-} \right).$$

Then $V = \mathscr{E}(\pi \bullet X)$ because

$$\frac{dV}{V_-} = \frac{\phi}{V_-} dS = \frac{\pi}{S_-} dS = \pi dX.$$

The Ansatz is to set $D = 1/\mathscr{E}(\rho \bullet X)$ for ρ to be determined. Let (b, c, K) be the differential characteristics of X with respect to A and h_d as in

Lemma 3.10.5. Then VD is a supermartingale deflator if and only if for all π,

$$(\pi-\rho)^\top \left(b - c\rho + \int \left(\frac{x}{1+\rho^\top x} - h_d(x)\right) K(dx)\right) \le 0, \qquad \mathbb{P}\otimes dA\text{-a.s.} \tag{$*$}$$

The left-hand side equals the (formal) directional derivative $\nabla g(\rho)(\pi-\rho)$, where $g(\pi)$ is defined such that $g(\pi)=0$ if and only if $\log V$ is a sigma-martingale:

$$g(\pi) = \pi^\top b - \frac{1}{2}\pi^\top c\pi + \int \left(\log(1+\pi^\top x) - \pi^\top h_d(x)\right) K(dx).$$

Up to technicalities, (NUPBR) implies for $\mathbb{P}\otimes dA$-a.e. (t,ω) that the function $\pi \mapsto g(\pi)$ is concave and bounded from above. The (non-unique) maximizer ρ satisfies $(*)$. Choosing ρ predictably in (t,ω) involves measurable selection techniques. □

2.11 Predictable uniform tightness

2.11.1 Lemma. *(NUPBR) ⇔ (PUT).*

Proof. We follow [CT15, Proposition 4.12].

- ⇐ The evaluation map $(\mathcal{X}_1 \ni X \mapsto X_1 \in L^0)$ is continuous, where $\mathcal{X}_1$ carries the Emery topology; cf. Lemma 1.2.2.
- ⇒ Lemma 2.10.2 provides a supermartingale deflator D. Then $\mathcal{Z} := \{Z = D(1+X); X \in \mathcal{X}_1\}$ consists of nonnegative supermartingales. Recall Burkholder's inequality: for every nonnegative supermartingale Z and every $H \in \mathcal{E}$ with $\|H\|_\infty \le 1$ we have

  $$\forall \alpha \ge 0: \qquad \alpha\mathbb{P}[|H \bullet Z|_1^* \ge \alpha] \le 9\mathbb{E}[Z_0].$$

  This gives Emery-boundedness of $\mathcal{Z}$. We want to deduce Emery-boundedness of $\mathcal{X}_1$. Note that D is a semimartingale because $\mathcal{X}_1$ contains zero. Thus, one obtains using integration by parts that

  $$1 + X = 1 + \frac{1}{D_-} \bullet Z + Z_- \bullet \frac{1}{D} + \left[\frac{1}{D}, Z\right].$$

  Recall that Emery-boundedness of $\mathcal{Z}$ implies uniform boundedness in probability of $\mathcal{Z}$. Thus, Emery boundedness of the sets $\{\frac{1}{D_-} \bullet Z; Z \in \mathcal{Z}\}$ and $\{Z_- \bullet \frac{1}{D}; Z \in \mathcal{Z}\}$ follows from the good integrator property. Emery-boundedness of $\{\left[\frac{1}{D}, Z\right]; Z \in \mathcal{Z}\}$ follows similarly using the estimate

  $$\left[\frac{1}{D}, Z\right] = \frac{1}{2}\left(\left[\frac{1}{D}+Z, \frac{1}{D}+Z\right] - [Z,Z] - \left[\frac{1}{D}, \frac{1}{D}\right]\right). \qquad \square$$

2.12 Weak star closedness of superreplicable claims

This section is the core of the fundamental theorem of asset pricing and concludes its proof.

2.12.1 Definition.

(i) Unbounded claims which are (super-)replicable:

$$K_0 = \{f \in L^0; f = (\phi \bullet S)_1 \text{ for some admissible } \phi\},$$
$$C_0 = \{f \in L^0; f \le (\phi \bullet S)_1 \text{ for some admissible } \phi\}.$$

(ii) Unbounded claims which are (super-)replicable using 1-admissible strategies:

$$K_0^1 = \{f \in L^0; f = (\phi \bullet S)_1 \text{ for some 1-admissible } \phi\},$$
$$C_0^1 = \{f \in L^0; f \leq (\phi \bullet S)_1 \text{ for some 1-admissible } \phi\}.$$

2.12.2 Lemma. *(NFLVR) ⇒* $C = \overline{C}^*$.

2.12.3 Remark.

- This result hinges on the Emery-closedness of the set of stochastic integrals.
- It implies the equivalence of (NFLVR) and (NFL) and concludes the proof of the FTAP (Theorem 2.5.3).
- To see the structure of the proof, draw the sets K, C, K_0, C_0, K_0^1, C_0^1 in the special case where $\Omega = \{g, b\}$ and $S_t(\omega) = \mathbb{1}_{\{1\}}(t)\big(\mathbb{1}_{\{g\}}(\omega) - \mathbb{1}_{\{b\}}(\omega)\big)$. Moreover, for some fixed $f \in C_0^1$, draw the set $K_0^1 \cap L^\infty_{\geq f}$ and the maximal elements in this sets.

Proof. We follow [CT15, Section 3]. Let $\widehat{K_0^1}$ denote the L^0 closure of K_0^1.

- It suffices to show that C_0 is Fatou closed by Lemma 3.8.2. Take $-1 \leq f_n \in C_0$ converging a.s. to f. We want to show that $f \in C_0$. Thus, we need to construct $h \in K_0$ with $h \geq f$.
- Note: We certainly find $h_n \in K_0$ with $h_n \geq f_n$, but there is no reason why these h_n should converge. Indeed, there are typically many $h_n \in K_0$ which dominate f_n, corresponding to different ways of wasting money. The idea is to stop the spending spree and ask for near-optimal results.
- We claim that $\widehat{K_0^1} \cap L^0_{\geq f}$ has a maximal element h. By (NUPBR) this set is bounded and closed. Thus, it contains its essential supremum if it is non-empty. This is indeed the case, as we verify now. There are $g_n = X_1^n$ with $g_n \geq f_n \geq -1$ for some $X^n \in \mathcal{X}$. Then $X^n \geq -1$ because $\mathbb{P}(X_s^n < -1) > 0$ for some $s \in [0,1)$ implies $(\mathbb{1}_{\{X_s^n<-1\}}\mathbb{1}_{(s,1]} \bullet X^n)_1 \gneq 0$ in violation with (NA). Therefore, g_n belongs to $K_0^1 \cap L^0_{\geq f}$. By Komlos, some forward convex combinations of g_n converge a.s. to some limit in $\widehat{K_0^1} \cap L^0_{\geq f}$.
- Let h be a maximal element in $\widehat{K_0^1} \cap L^0_{\geq f}$, and let $X^n \in \mathcal{X}_1$ such that $X_1^n \to h$ a.s. We claim that X^n converges uniformly in probability to some X. Otherwise, there would be $\alpha > 0$ and sequences n_k, m_k of natural numbers such that $\mathbb{P}[|(X^{n_k} - X^{m_k})^+|_1^* > \alpha] \geq \alpha$. Set

  $$T_k = \inf\{t \in [0,1]; X^{n_k} - X^{m_k} > \alpha\}, \quad Y^k = \mathbb{1}_{[\![0,T_k]\!]} \bullet X^{m_k} + \mathbb{1}_{(\!(T_k,T]\!]} \bullet X^{n_k}.$$

  Then $\mathbb{P}[T_k < \infty] \geq \alpha$, $Y^k \in \mathcal{X}_1$, and

  $$Y_1^k = \underbrace{X_1^{m_k}\mathbb{1}_{\{T_k=\infty\}} + X_1^{n_k}\mathbb{1}_{\{T_k<\infty\}}}_{\to h} + \underbrace{(X_{T_k}^{m_k} - X_{T_k}^{n_k})\mathbb{1}_{\{T_k<\infty\}}}_{\mathbb{P}[\cdot>\alpha]\geq\alpha}.$$

  Some forward convex combinations of Y_1^k converge to an element of $\widehat{K_0^1}$ strictly greater than h by Komlos, in violation of the maximality of h.
- (PUT) implies that X^n Emery-converge to X [CT15, Theorem 5.1].
- Then $X \in \mathcal{X}_1$ by Emery-closedness of the set of stochastic integrals. Thus, $f \leq h = X_1 \in K_0$ and $f \in C_0$. □

2.12.4 Remark. The proof of [CT15, Theorem 5.1] was read in class.

3 Auxiliary results

3.1 Martingales in discrete time

3.1.1 Theorem. *Let $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{N}}, \mathbb{P})$ be a filtered probability space with a complete filtration, let $X\colon \Omega \times \mathbb{N} \to \mathbb{R}$ be an adapted process, let X^* be the running supremum of X given by $X_n^* = \sup_{k\leq n} X_k$, and let $S, T\colon \Omega \to \mathbb{N} \cup \{\infty\}$ be stopping times.*

(i) Doob's decomposition: *if X_n is integrable for each $n \in \mathbb{N}$, then $X = M - A$ for a martingale M and a predictable process A with $A_0 = 0$.*

(ii) Martingale transforms: *if X is a martingale (supermartingale) and H is predictable (predictable non-negative) such that the random variables $(H \bullet X)_n$, $n \in \mathbb{N}$, are integrable, then $H \bullet X$ is a martingale (supermartingale).*

(iii) Doob's optional sampling: *if X is a supermartingale (martingale, resp.) and $S \leq T$, then X_S and X_T are integrable and satisfy $X_S \geq \mathbb{E}[X_T|\mathcal{F}_S]$ ($X_S \geq \mathbb{E}[X_T|\mathcal{F}_S]$, resp.).*

(iv) Doob's maximal inequality, weak type: *if X be a supermartingale and $\alpha > 0$, then*

$$\mathbb{P}[|X|_n^* \geq \alpha] \leq \frac{c}{\alpha} \sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|],$$

where $c = 1$ if $X \geq 0$, $X \leq 0$, or X is a martingale, and $c = 3$ otherwise.

(v) Doob's maximal inequality, strong type: *if X is a non-negative supermartingale and $p, q \in (1, \infty)$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$, then*

$$\forall n \in \mathbb{N}: \qquad \||X|_n^*\|_{L^p} \leq q\mathbb{E}[|X_n|].$$

(vi) Burkholder's maximal inequality: *let X be a supermartingale, let H be predictable, and let $\alpha > 0$. If X is nonnegative, then*

$$\alpha\mathbb{P}[|H \bullet X|_1^* \geq \alpha] \leq 9\mathbb{E}[|X_0|],$$

and if X is a martingale, then

$$\alpha\mathbb{P}[|H \bullet X|_1^* \geq \alpha] \leq 18\mathbb{E}[|X_1|].$$

Proof.

(i) See [Mey72, Theorem I.16].

(ii) See [Mey72, Theorem II.1].

(iii) See [Mey72, Theorem II.3].

(iv) See [Mey72, Theorem II.5].

(v) See [Mey72, Theorem II.8].

(vi) See the proof of [Mey72, Theorem II.47]. □

3.2 Gauges

3.2.1 Definition. Let V be a vector space.

(i) A *gauge* is a function $\llbracket\cdot\rrbracket\colon V \to \overline{\mathbb{R}}_+$.

(ii) The gauge $\llbracket\cdot\rrbracket$ is *subadditive* if

$$\forall x, y \in V: \qquad \llbracket x+y\rrbracket \le \llbracket x\rrbracket + \llbracket y\rrbracket.$$

(iii) The gauge $\llbracket\cdot\rrbracket$ is *balanced* if

$$\forall x \in V, \forall |\lambda| \le 1: \qquad \llbracket \lambda x\rrbracket \le \llbracket x\rrbracket.$$

(iv) The gauge $\llbracket\cdot\rrbracket$ is *absolute homogeneous* if

$$\forall x \in V, \forall \lambda \in \mathbb{R}: \qquad \llbracket \lambda x\rrbracket = |\lambda|\llbracket x\rrbracket.$$

(v) The gauge $\llbracket\cdot\rrbracket$ *induces* the pseudo-metric $(x,y) \mapsto \llbracket x-y\rrbracket$ and corresponding topology.

(vi) A vector $x \in V$ is called $\llbracket\cdot\rrbracket$-*finite* if

$$\lim_{r\to 0}\llbracket rx\rrbracket = 0,$$

and $\llbracket\cdot\rrbracket$ is called finite if this holds for all $x \in V$.

(vii) A vector $x \in V$ is called $\llbracket\cdot\rrbracket$-*negligible* if $\llbracket x\rrbracket = 0$.

3.2.2 Remark. The word gauge is used with different meanings in different contexts. Our use of the word is consistent with [Bic02].

3.2.3 Lemma. *For any topological vector space V, the following are equivalent:*

(i) The topology of V is induced by a metric.

(ii) The topology of V is induced by a subadditive gauge.

Proof. (i) follows trivially from (ii), and (ii) follows from (i) and the Birkhoff–Kakutani theorem. □

3.2.4 Lemma. *Let $\llbracket\cdot\rrbracket$ be a gauge on a vector space V.*

(i) If $\llbracket\cdot\rrbracket$ is subadditive, then addition $V \times V \to V$ is $\llbracket\cdot\rrbracket$-continuous.

(ii) If $\llbracket\cdot\rrbracket$ is subadditive, balanced, and finite, then scalar multiplication $\mathbb{R} \times V \to V$ is $\llbracket\cdot\rrbracket$-continuous.

Thus, $\llbracket\cdot\rrbracket$ induces a vector space topology under the assumptions of (ii).

Proof. Let $x_n \to x$, $y_n \to y$, $\lambda_n \to \lambda$, and $N \in \mathbb{N}$ such that $\sup_n |\lambda_n| \le N$. Under the assumptions of (i),

$$\llbracket (x_n + y_n) - (x+y)\rrbracket \le \llbracket x_n - x\rrbracket + \llbracket y_n - y\rrbracket \to 0.$$

Moreover, under the additional assumptions of (ii),

$$\begin{aligned}\llbracket \lambda_n x_n - \lambda x\rrbracket &\le \llbracket \lambda_n N^{-1} N(x_n - x)\rrbracket + \llbracket (\lambda_n - \lambda)x\rrbracket \\ &\le N\llbracket x_n - x\rrbracket + \llbracket (\lambda_n - \lambda)x\rrbracket \to 0,\end{aligned}$$

where the second inequality is obtained as follows: first, the factor $\lambda_n N^{-1}$ is discarded using the balanced property of the gauge, and then the factor N is pulled out using an N-fold application of the triangle inequality. □

3.3 The gauge space L^0

3.3.1 Setting. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let V be the vector space of measurable functions $f : \Omega \to \mathbb{R}$.

3.3.2 Definition. $L^0(\Omega) = (V, \llbracket\cdot\rrbracket_0)$ with $\llbracket\cdot\rrbracket_0$ as in Lemma 3.3.3 below.

3.3.3 Lemma.

(i) The following gauges on V are subadditive, balanced, and finite:

$$\llbracket f\rrbracket_0 = \inf\{\lambda \in \mathbb{R}; \mathbb{P}[|f| > \lambda] \le \lambda\},$$
$$\llbracket f\rrbracket_\odot = \mathbb{E}[|f| \wedge 1].$$

(ii) For each $\alpha \in \mathbb{R}$, the following gauge on V is absolute homogeneous and finite:

$$\|f\|_{[\alpha]} = \inf\{\lambda \in \mathbb{R}; \mathbb{P}[|f| > \lambda] \le \alpha\}.$$

(iii) The gauges $\llbracket\cdot\rrbracket_0$ and $\llbracket\cdot\rrbracket_\odot$ and the family of gauges $\{\|\cdot\|_{[\alpha]}; \alpha \in \mathbb{R}\}$ generate the same vector space topology on V thanks to the following relations:

$$\llbracket f\rrbracket_\odot \le 2\llbracket f\rrbracket_0,$$
$$\llbracket f\rrbracket_\odot < 1 \Rightarrow \llbracket f\rrbracket_0^2 \le \llbracket f\rrbracket_\odot$$
$$\llbracket f\rrbracket_0 = \inf\{\alpha; \|f\|_{[\alpha]} \le \alpha\}.$$

Proof. See [Bic02, Appendix A.8] □

3.3.4 Remark. It is illuminating to spell out the meaning of the following concepts in terms of the gauges $\llbracket\cdot\rrbracket_0$, $\llbracket\cdot\rrbracket_\odot$, and $\|\cdot\|_{[\alpha]}$: E

(i) Continuity of sequences in L^0.

(ii) Boundedness of sets in L^0.

(iii) Continuity of linear maps from a topological vector space into L^0.

(iv) Boundedness of linear maps from a topological vector space into L^0.

3.4 Hahn–Banach

3.5 Komlos for L^1 spaces

3.5.1 Lemma (Komlos, L^1 version)**.** *Let $(f_n)_{n\in\mathbb{N}}$ be a uniformly integrable sequence of random variables. Then there exist forward convex combinations $\tilde{f}_n \in \langle f_n, f_{n+1}, \dots\rangle_{\text{conv}}$ such that $(\tilde{f}_n)_{n\in\mathbb{N}}$ converges in L^1.*

Proof. See [BSV12, Lemma 2.1]; the proof was read in class. □

3.6 Komlos for L^0 spaces

3.6.1 Lemma (Komlos, L^0 version)**.** *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of non-negative random variables.*

(i) There exist forward convex combinations $g_n \in \langle f_n, f_{n+1}, \dots\rangle_{\text{conv}}$ such that $(g_n)_{n\in\mathbb{N}}$ converges a.s. to a random variable g with values in $[0, \infty]$.

(ii) $\mathbb{P}[g < \infty] = 1$ if $\langle f_1, f_2, \dots\rangle_{\text{conv}}$ is bounded in L^0.

(iii) $\mathbb{P}[g > 0] > 0$ if there is $\alpha > 0$ such that $\mathbb{P}[f_n \ge \alpha] \ge \alpha > 0$.

Proof. We follow [Kab97, Lemma A].

(i) Let
$$J_n := \inf\left\{\mathbb{E}[e^{-g}]; g \in \langle f_n, f_{n+1}, \dots\rangle_{\text{conv}}\right\}.$$
Then J_n increases to some $J \le 1$. Take $g_n \in \langle f_n, f_{n+1}, \dots\rangle_{\text{conv}}$ with $\mathbb{E}[e^{-g}] \le J_n + \frac{1}{n}$. Let $\epsilon > 0$, and let
$$\begin{aligned} A &= \{(x,y) \in \mathbb{R}^2_+; |x-y| < \epsilon\}, \\ B &= \{(x,y) \in \mathbb{R}^2_+; x \wedge y \ge \tfrac{1}{\epsilon}\}, \\ C &= \mathbb{R}^2_+ \setminus (A_\epsilon \cup B_\epsilon). \end{aligned}$$
By convexity there is $\delta > 0$ such that
$$e^{-(x+y)/2} \le (e^{-x} + e^{-y})/2 - \delta \mathbb{1}_C(x,y).$$
Therefore,
$$\underset{\sim}{J_{n\wedge m}} \le \mathbb{E}[e^{-(g_n+g_m)/2}] \le (\underset{\sim}{\mathbb{E}[e^{-g_n}]} + \underset{\sim}{\mathbb{E}[e^{-g_m}]})/2 - \delta\mathbb{P}[(g_n, g_m) \in C],$$
where the highlighted expressions converge to J as $n, m \to \infty$. It follows that
$$\lim_{n,m\to\infty} \mathbb{P}[(g_n, g_m) \in C] = 0.$$
Then e^{-g_n} is Cauchy in L^1 by the following estimate:
$$\begin{aligned} \mathbb{E}[|e^{-g_n} - e^{-g_m}|] &\le \epsilon\mathbb{P}[(g_n,g_m) \in A] + 2e^{-1/\epsilon}\mathbb{P}[(g_n,g_m) \in B] + \mathbb{P}[(g_n,g_m) \in C] \\ &\le \epsilon + 2e^{-1/\epsilon} + \mathbb{P}[(g_n,g_m) \in C]. \end{aligned}$$
Thus, it has an a.s. convergent subsequence.

(ii) Clear by the completeness of L^0.

(iii) For any $g = \sum_n \lambda_n f_n \in \langle f_1, f_2, \cdots\rangle_{\text{conv}}$ one has
$$\begin{aligned} \mathbb{E}[e^{-g}] &\le \sum_n \lambda_n \mathbb{E}[e^{-f_n}] \\ &\le \sum_n \lambda_n \big(\mathbb{P}[f_n < \alpha] + e^{-\alpha}\mathbb{P}[f_n \ge \alpha]\big) \\ &\le (1-\alpha) + \alpha e^{-\alpha} < 1. \end{aligned}$$
□

3.7 Bipolar theorem

The following theorem is typically applied with $W = V^*$, where V is locally convex, and W carries the weak-* topology. The general form of the theorem will be useful later on.

3.7.1 Theorem (Bipolar theorem)**.** *Let V, W be locally convex topological vector spaces, and let $\langle\cdot,\cdot\rangle\colon V \times W \to \mathbb{R}$ be a continuous bilinear mapping which satisfies the non-degeneracy condition*
$$\begin{aligned} \{x \in V; \langle x, y\rangle = 0\ \forall y \in W\} = \{0\}, \\ \{y \in W; \langle x, y\rangle = 0\ \forall x \in V\} = \{0\}. \end{aligned}$$
The polars of $A \subseteq V$ and $B \subseteq W$ are defined as
$$\begin{aligned} A^0 &= \{b \in W; \langle a, b\rangle \le 1\ \forall a \in A\}, \\ B^0 &= \{a \in A; \langle a, b\rangle \le 1\ \forall b \in B\}. \end{aligned}$$

The weak dual topology wd *on* V *is coarsest topology such that* $\langle\cdot,\cdot\rangle\colon V\to W^*$ *is continuous, where* W^* *carries the weak-* topology; similarly with* V *and* W *interchanged. Then*

$$A^{00}=\overline{\langle A\cup\{0\}\rangle_{conv}}^{\mathrm{wd}}.$$

Proof.

- Auxiliary claim: The following map is a linear isomorphism,

  $$(W,\mathrm{wd})\ni w\mapsto\langle\cdot,w\rangle\in(V,\mathrm{wd})^*.$$

  Injectivity follows from the non-degeneracy of the pairing. To prove surjectivity, let $\ell\in(V,\mathrm{wd})^*$. By the continuity of ℓ some neighborhood basis of (V,wd) is mapped into the unit ball in $\mathbb{R}$, i.e., there exist $n\in\mathbb{N}$, $w_1,\dots,w_n\in W$, and $t_1,\dots,t_n\in\mathbb{R}$ such that

  $$\forall v\in V:\qquad|\ell(v)|\le t_1|\langle v,w_1\rangle|+\dots+t_n|\langle v,w_n\rangle|.$$

  This implies (verify!) that there exist $\alpha_1,\dots,\alpha_n\in\mathbb{R}$ such that

  $$\ell(v)=\alpha_1t_1\langle v,w_1\rangle+\dots+\alpha_nt_n\langle v,w_n\rangle,$$

  which proves surjectivity.

- $\supseteq$ A^{00} is weak dual closed, convex, and contains $A\cup\{0\}$.
- $\subseteq$ Let $x\in V\setminus\overline{\langle A\cup\{0\}\rangle_{\mathrm{conv}}}^{w}$. We claim that $x\notin A^{00}$. By Hahn–Banach, there is $\ell\in V^*$ and $s\in\mathbb{R}$ such that

  $$\forall a\in A\cup\{0\}:\qquad\ell(a)<s<\ell(x).$$

  Wlog. $s=1$ after rescaling of ℓ by $2/(\ell(x)+s)$. By non-degeneracy $\ell=\langle\cdot,w\rangle$ for some unique $w\in W$. Then $w\in A^0$ by the first and $x\notin A^{00}$ by the second inequality above. □

3.8 Fatou convergence

3.8.1 Definition. A subset C of L^0 is *Fatou closed* if for every sequence $(f_n)_{n\in\mathbb{N}}$ uniformly bounded from below and such that $f_n\to f$ almost surely, we have $f\in C$.

3.8.2 Lemma. *Let* C_0 *be a Fatou closed convex cone in* L^0*, and let* $C=C\cap L^\infty$*. Then* C *is* $\sigma(L^\infty,L^1)$ *closed.*

Proof. Let $\bigcirc L^\infty$ denote the unit ball in L^∞, and let C_0 be Fatou closed.

- $C\cap\bigcirc L^\infty$ is L^2 closed because L^2 convergence implies almost sure convergence of a subsequence.
- $C\cap\bigcirc L^\infty$ is Mackey closed, where the Mackey topology is defined as the topology of uniform convergence on $\sigma(L^1,L^\infty)$-compact absolutely convex subsets of L^1. Indeed, the Mackey topology is finer than the L^2 topology on L^∞. To see this, note that $\bigcirc L^2$ is relatively weakly compact in L^1 by de la Vallée–Poussin. Thus, letting A be the $\sigma(L^1,L^\infty)$ closure of $\bigcirc L^2$, the set $\{g\in L^\infty;\sup_{f\in A}\langle g,f\rangle_{L^\infty,L^1}\le 1\}$, which is contained in $\bigcirc L^2$, is Mackey open.
- $C\cap\bigcirc L^\infty$ is $\sigma(L^\infty,L^1)$ closed. Indeed, all compatible topologies (i.e., locally convex topologies with the same dual) have the same convex closed sets [Jar12, Proposition 8.2.5]. The weak* topology $\sigma(L^\infty,L^1)$ is the coarsest [Jar12, Theorem 8.1.2] and the Mackey topology $\mu(L^\infty,L^1)$ the finest compatible topology [Jar12, Proposition 8.5.5].
- C is $\sigma(L^\infty,L^1)$ closed by Krein–Smulian [DS58, Theorem V.5.7] and the convexity of C. □

3.9 Sigma-martingales

3.9.1 Definition. A *sigma-martingale* is a semimartingale of the form $H \bullet M$ with $M \in \mathcal{M}_{\text{loc}}$ and $H \in L(M)$.

3.9.2 Remark. E

(i) Every local martingale is a sigma-martingale, but not the other way round.

(ii) X is a sigma-martingale if and only if there is a sequence $(D_n)_{n\in\mathbb{N}}$ of predictable sets such that $\bigcup_n D_n = [0,1] \times \Omega$ and $\mathbb{1}_{D_n} \bullet X$ is a uniformly integrable martingale for each $n \in \mathbb{N}$.

3.9.3 Lemma (Ansel, Stricker)**.** *Every sigma-martingale which is bounded from below is a local martingale.*

Proof. We follow [DP07]. By localization it is sufficient to show for any $C > 0$, $M \in \mathcal{M}$, and $H \in L(M)$ with $H \bullet M \geq -C$ that $H \bullet M \in \mathcal{M}_{\text{loc}}$. Let $H^n = H\mathbb{1}_{\{|H|\leq n\}}$ for each $n \in \mathbb{N}$. After passing to a subsequence one has

$$|(H^n - H) \bullet M|_1^* \xrightarrow[n\to\infty]{a.s.} 0.$$

For all $m, n \in \mathbb{N}$ define

$$T_n := \inf\{t \in [0,1] : |H \bullet M|_t > n \text{ or } |(H - H^n) \bullet M|_t > 1\} \xrightarrow[n\to\infty]{a.s.} \infty,$$
$$S_m := \inf_{n\geq m} T_n \xrightarrow[m\to\infty]{a.s.} \infty.$$

Let T be a finite stopping time. For each $n \geq m$, $(H^n \bullet M)_{T\wedge S_m}$ is bounded from below: indeed, as the jumps of $H^n \bullet M$ are less extreme than those of $H \bullet M$,

$$\begin{aligned}(H^n \bullet M)_{T\wedge S_m} &= (H^n \bullet M)_{(T\wedge S_m)-} + \Delta(H^n \bullet M)_{T\wedge S_m}\\ &\geq (H^n \bullet M)_{(T\wedge S_m)-} + \Delta(H \bullet M)_{T\wedge S_m} \wedge 0\\ &\geq -(m+1) - (m+C).\end{aligned}$$

This implies via Fatou that

$$\mathbb{E}[(H \bullet M)_{T\wedge S_m}] \leq \liminf_{n\to\infty} \mathbb{E}[(H^n \bullet M)_{T\wedge S_m}] = 0.$$

Thus, $(H \bullet M)_{T\wedge S_m}$ is integrable. As $H \bullet M$ is bounded strictly before S_m, $\Delta(H \bullet M)_{T\wedge S_m}$ is also integrable. This leads to the following integrable upper bound:

$$\begin{aligned}(H^n \bullet M)_{T\wedge S_m} &= (H^n \bullet M)_{(T\wedge S_m)-} + \Delta(H^n \bullet M)_{T\wedge S_m}\\ &\leq (H^n \bullet M)_{(T\wedge S_m)-} + \Delta(H \bullet M)_{T\wedge S_m} \vee 0\\ &\leq (m+1) + |\Delta(H \bullet M)_{T\wedge S_m}|.\end{aligned}$$

By dominated convergence one obtains

$$\mathbb{E}[(H \bullet M)_{T\wedge S_m}] = \lim_{n\to\infty} \mathbb{E}[(H^n \bullet M)_{T\wedge S_m}] = 0.$$

As this holds for all finite stopping times, $(H \bullet M)^{S_m} \in \mathcal{M}$. $\square$

3.10 Semimartingale characteristics

3.10.1 Definition.

(i) $\mathcal{P}$ denotes the *predictable sigma-algebra* on $[0,1] \times \Omega$.

(ii) A *truncation function* on a topological vector space V is a bounded function $h\colon V \to V$ which is equal to the identity on a neighborhood of zero.

(iii) For any semimartingale $X\colon [0,1] \times \Omega \to \mathbb{R}^d$, truncation function $h_d\colon \mathbb{R}^d \to \mathbb{R}^d$, and $t \in [0,1]$, let

$$\hat{X}_t = X_t - \sum_{s \leq t} (\Delta X_s - h_d(\Delta X_s)), \qquad \check{X}_t = \sum_{s \leq t} (\Delta X_s - h_d(\Delta X_s)).$$

3.10.2 Remark.

- $\hat{X}$ has bounded jumps, and $\check{X}$ has finite variation.
- A typical choice of truncation function is $h_d(x) = x \mathbb{1}_{\{\|x\| \leq 1\}}$.
- If X is special, one can (at least formally) set $h_d(x) = x$.

3.10.3 Definition (Characteristics, integrated form)**.**

(i) The *drift* of X is the compensator of the special semimartingale $\hat{X}$, i.e., the unique predictable finite-variation process $B\colon [0,1] \times \Omega \to \mathbb{R}^d$ such that $\hat{X} - X_0 - B \in \mathcal{M}^d_{\text{loc}}$.

(ii) The *volatility* is the continuous quadratic covariation of X, i.e., the continuous finite-variation process $C\colon [0,1] \times \Omega \to \mathbb{R}^{d \times d}$ given by

$$C^{i,j}_t = [X^i, X^j]_t - \sum_{s \leq t} \Delta X^i_s \Delta X^j_s = \langle X^{i,c}, X^{j,c} \rangle_t = [X^{i,c}, X^{j,c}]_t = [X^i, X^j]^c_t.$$

(iii) The *jump measure* of X is the compensator of the jumps of X, i.e., the unique transition kernel ν from $(\Omega, \mathcal{F})$ to $([0,1] \times \mathbb{R}^d, \mathcal{B}([0,1] \times \mathbb{R}^d))$ such that for all $W \in \mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$ the process

$$(W * \nu)_t(\omega) := \int_{[0,t] \times \mathbb{R}^d} W(\omega, t, x) \nu(\omega, dt, dx)$$

is predictable non-decreasing and compensates the process

$$(W * \mu^X)_t(\omega) := \sum_{s \leq t} W(\omega, s, \Delta X_s(\omega)) \mathbb{1}_{\{\Delta X_s(\omega) \neq 0\}}.$$

We call (B, C, ν) the characteristics of X with respect to h_d.

3.10.4 Remark.

- The characteristics give a predictable forecast of the behavior of X.
- Under some Lipschitz conditions the characteristics determine X uniquely, but not in general.
- Only the drift depends on the choice of truncation function; the volatility and jump measure do not.

3.10.5 Lemma (Characteristics, differential form)**.** *Let X be a semimartingale with characteristics (B, C, ν). Then*

$$B = b \bullet A, \qquad C = C \bullet A, \qquad \nu = K \; dA,$$

where:

(i) *A is predictable, non-decreasing, and locally integrable;*

(ii) *b is predictable with values in* $\mathbb{R}^d$*;*

(iii) *c is predictable with values in the positive semi-definite* $(d \times d)$*-matrices;*

(iv) *K is a transition kernel from* $([0,1] \times \Omega, \mathcal{P})$ *into* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ *which satisfies*

$$K_{\omega,t}(\{0\}) = 0, \qquad \int (\|x\|^2 \wedge 1) K_{\omega,t}(dx) \leq 1,$$
$$\Delta A_t(\omega) > 0 \Rightarrow b_t(\omega) = \int h(x) K_{\omega,t}(dx),$$
$$\Delta A_t(\omega) K_{\omega,t}(\mathbb{R}^d) \leq 1.$$

We call (b, c, K) *the* differentiable characteristics *of* X *with respect to* A *and* h_d.

Proof. See [JS03, Proposition II.2.9]. □

3.10.6 Remark. Lévy processes are semimartingales with constant and deterministic differential characteristics with respect to $A_t = t$. For example, in $d = 1$ with truncation function h_1: E

- The process $X_t = t$ has differential characteristics $(1, 0, 0)$.
- Brownian motion X_t has differential characteristics $(0, 1, 0)$.
- Poisson processes X_t have differential characteristics $(\int h_1(x)K(dx), 0, K)$. Standard Poisson process have $K = \delta_1$.

3.11 Characteristics of sigma martingales

The following lemma provides a predictable characterization of semimartingale properties in terms of the characteristics.

3.11.1 Lemma.

(i) *X is a special semimartingale if and only if*

$$\int (\|x\|^2 \wedge \|x\|) K(dx) \in L(A).$$

(ii) $X \in \mathcal{M}^d_{\mathrm{loc}}$ *if and only if*

$$\int (\|x\|^2 \wedge \|x\|) K(dx) \in L(A)$$

and

$$b + \int (x - h(x)) K(dx) = 0 \qquad \mathbb{P} \otimes dA\text{-}a.s.$$

(iii) $X \in \mathcal{M}^d_\sigma$ *if and only if*

$$\int (\|x\|^2 \wedge \|x\|) K(dx) < \infty \qquad \mathbb{P} \otimes dA\text{-}a.s.$$

and

$$b + \int (x - h(x)) K(dx) = 0 \qquad \mathbb{P} \otimes dA\text{-}a.s.$$

Proof. See [JS03, Proposition II.2.29]. □

3.12 Characteristics of stochastic integrals

The following lemma describes how the characteristics transform under stochastic integration: the drift is multiplied by the integrand, the volatility is multiplied from the left and right by the integrand, and the jump measure is pushed forward along the multiplication map corresponding to the integrand.

3.12.1 Lemma. *Let X be a semimartingale with differential characteristics (b, c, K) and A as in Lemma 3.10.5, let $h_1 \colon \mathbb{R} \to \mathbb{R}$ be a truncation function, and let $H \in L(X)$. Then $H \bullet X$ has differential characteristics (b', c', K') with respect to A and h_1, where*

$$\begin{aligned} b' &= Hb + \int (h_1(Hx) - Hh_d(x))K(dx), \\ c' &= \sum_{i,j=1}^{d} H^i c^{i,j} H^j, \\ K'(B) &= \int \mathbb{1}_B(Hx)K(dx), \qquad \forall B \in \mathcal{B}(\mathbb{R}). \end{aligned}$$

Proof. This follows from the more general result [JS03, Proposition IX.5.3]. □

3.13 Girsanov's theorem

Girsanov's theorem describes how the characteristics are affected by changes of measures: the drift is adjusted by some multiple β of the volatility, and the jump measure is multiplied by some density function Y.

3.13.1 Definition. Let X be a semimartingale with jump measure μ^X as in Lemma 3.10.5.

(i) $\mathbb{P} \otimes \mu^X$ denotes the measure on $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$ given by

$$\forall W \in \mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d): \qquad (\mathbb{P} \otimes \mu^X)(W) = \mathbb{E}[(W * \mu^X)_1].$$

(ii) For any nonnegative $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}^d)$-measurable function W, the conditional expectation $\mathbb{E}_{\mathbb{P}\otimes\mu^X}[W|\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)]$ is the $\mathbb{P} \otimes \mu^X$-a.e. unique $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable function W' satisfying

$$\forall U \in \mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d): \qquad \mathbb{E}_{\mathbb{P}\otimes\mu^X}[WU] = \mathbb{E}_{\mathbb{P}\otimes\mu^X}[W'U].$$

3.13.2 Lemma. *Let X be a semimartingale with differential characteristics (b, c, K) and A as in Lemma 3.10.5, let $h_1 \colon \mathbb{R} \to \mathbb{R}$ be a truncation function, let $\mathbb{Q} \ll \mathbb{P}$, and let Z be the density process given by $Z_t = \frac{d\mathbb{Q}|\mathcal{F}_t}{d\mathbb{P}|\mathcal{F}_t}$.*

(i) There exists a $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable function Y and a predictable $\mathbb{R}^d$-valued process β such that X has differential characteristics (b', c', K') with respect to A and h_d under $\mathbb{Q}$, where

$$b' = b + c\beta + \int h_d(x)(Y(x) - 1)K(dx), \qquad c' = c, \qquad K' = YK.$$

(ii) The coefficients Y and β are related to the density process Z as follows: for all $i \in \{1, \ldots, d\}$,

$$YZ_- = \mathbb{E}_{\mathbb{P}\otimes\mu^X}[Z|\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)], \qquad [Z, X^{i,c}] = \Big(\sum_{j=1}^{d} c^{i,j} \beta^j Z_-\Big) \bullet A,$$

where $X^{i,c}$ denotes the continuous local martingale part of the i-th component of X.

Proof. See [JS03, Theorem III.2.24]. □

References

[Bic02] K. Bichteler. *Stochastic integration with jumps.* Vol. 89. Cambridge University Press, 2002.

[BS14] M. Beiglböck and P. Siorpaes. "Riemann-integration and a new proof of the Bichteler–Dellacherie theorem". In: *Stochastic Processes and their Applications* 124.3 (2014), pp. 1226–1235.

[BS15] M. Beiglböck and P. Siorpaes. "Pathwise versions of the Burkholder-Davis-Gundy inequality". In: *Bernoulli* 21.1 (2015), pp. 360–373.

[BSV12] M. Beiglboeck, W. Schachermayer, and B. Veliyev. "A short proof of the Doob–Meyer theorem". In: *Stochastic Processes and their Applications* 122.4 (2012), pp. 1204–1209.

[CT15] C. Cuchiero and J. Teichmann. "A convergence result for the Emery topology and a variant of the proof of the fundamental theorem of asset pricing". In: *Finance and Stochastics* 19.4 (2015), pp. 743–761.

[DP07] M. De Donno and M. Pratelli. "On a lemma by Ansel and Stricker". In: *Séminaire de Probabilités XL*. Vol. 1899. Lecture Notes in Math. Springer, Berlin, 2007, pp. 411–414.

[DS58] N. Dunford and J. T. Schwartz. *Linear operators, part 1.* Interscience Publishers New York, London, 1958.

[DS94] F. Delbaen and W. Schachermayer. "A general version of the fundamental theorem of asset pricing". In: *Mathematische Annalen* 300 (1994), pp. 463–520.

[Éme79] M. Émery. "Une topologie sur l'espace des semimartingales". In: *Séminaire de probabilités (Strasbourg)* 13 (1979), pp. 260–280.

[Jar12] H. Jarchow. *Locally convex spaces.* Springer Science & Business Media, 2012.

[JS03] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes.* Second. Vol. 288. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Berlin: Springer-Verlag, 2003.

[Kab97] Y. M. Kabanov. "On the FTAP of Kreps–Delbaen–Schachermayer". In: *Statistics and control of stochastic processes (Moscow, 1995/1996).* World Sci. Publ., River Edge, NJ, 1997, pp. 191–203.

[Kal06] O. Kallenberg. *Foundations of modern probability.* Springer Science & Business Media, 2006.

[Kar13] C. Kardaras. "On the closure in the Emery topology of semimartingale wealth-process sets". In: *The annals of applied probability* (2013), pp. 1355–1376.

[KK07] I. Karatzas and C. Kardaras. "The numéraire portfolio in semimartingale financial models". In: *Finance and Stochastics* 11.4 (2007), pp. 447–493.

[KKS16] Y. Kabanov, C. Kardaras, and S. Song. "No arbitrage of the first kind and local martingale numéraires". In: *Finance and Stochastics* 20.4 (2016), pp. 1097–1108.

[Mey72] P.-A. Meyer. *Martingales and Stochastic Integrals.* Vol. I. Springer, 1972.

[Pro05] P. E. Protter. *Stochastic integration and differential equations.* Vol. 21. Stochastic Modelling and Applied Probability. Berlin: Springer-Verlag, 2005.

[SC02] A. N. Shiryaev and A. S. Chernyı. "A vector stochastic integral and the fundamental theorem of asset pricing". In: *Tr. Mat. Inst. Steklova* 237.Stokhast. Finans. Mat. (2002), pp. 12–56.

[Tei17] J. Teichmann. *Foundations of martingale theory and stochastic calculus from a finance perspective.* Lecture Notes (ETH Zurich). 2017.