UNIVERSITY OF ZURICH

# Mathematical Finance

**Theorem** (Itô's formula): Let $(X_t)_{t\geq 0} = (X_t^1, \ldots, X_t^d)_{t\geq 0}$ be an $\mathbb{R}^d$-valued semimartingale and let $F \in C^2(\mathbb{R}^d, \mathbb{R})$. Denote $X_{s-} := \lim_{u \nearrow s} X_u$, $\Delta X_s^j := X_s^j - X_{s-}^j$ the associated jump-process and $[X^j, X^k]^c$ denotes the quadratic variation of the continuous components of $X^j$ and $X^k$.
Then $(F(X_t))_{t\geq 0}$ is again a semimartingale and we have

$$F(X_t) - F(X_0) = \sum_{j=1}^d \int_0^t \frac{\partial F}{\partial x_j}(X_{s-})dX_s^j + \frac{1}{2}\sum_{j,k=1}^d \int_0^t \frac{\partial^2 F}{\partial x_j \partial x_k}(X_{s-})d[X^j, X^k]_s^c + \sum_{0<s\leq t}\left(F(X_s) - F(X_{s-}) - \sum_{j=1}^d \frac{\partial F}{\partial x_j}(X_{s-})\Delta X_s^j\right).$$

*Transcribed by*
*Marco Bertenghi*

Based on the lecture notes of
Prof. Martin SCHWEIZER (ETHZ)

July 20, 2019

# Contents

# 1 Introduction and Preliminaries

We start by recalling some important definitions and discuss the general setup we are working with in a heuristic manner.

**Definition 1.1.** *Given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{N}}, \mathbb{P})$ and a discrete-time stochastic process $(X_n)_{n\in\mathbb{N}}$. We say $(X_n)_{n\in\mathbb{N}}$ is predictable if $X_{n+1}$ is measurable with respect to the $\sigma$-algebra $\mathcal{F}_n$ for each $n \in \mathbb{N}$.*

**Definition 1.2.** *Given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{N}}, \mathbb{P})$, the predictable $\sigma$-field $\mathcal{P}$ on $\Omega \times (0, \infty)$ is generated by all adapted left-continuous processes. We call $H = (H_t)_{t\geq 0}$ predictable, if $H : \Omega \times (0, \infty) \to \mathbb{R}$ is $\mathcal{P}$-measurable.*

**Example 1.1.** *The following are easy examples of predictable processes.*

- *Every deterministic process is a predictable process.*
- *Every continuous-time adapted process that is left-continuous is predictable.*

We will now start with basic ideas, not all will be defined in detail here, consult later sections of this script.

**Setup**: We are given a **time horizon** $T \in (0, \infty)$, **trading dates** $t \in [0, T]$, a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ with the filtration satisfying the usual conditions (i.e. it is complete and right-continuous). Intuitively, $\mathcal{F}_t$ describes the information (events observable) up to time $t$.

Typically we are given one reference asset with positive price, $\widetilde{S_t^0} > 0$, for all $t$ which is adapted. Often called **bank account, bond, savings account** etc. Moreover, we are given $d$ **"risky" assets** with price process $\widetilde{S^i} = (\widetilde{S_t^i})_{0\leq t\leq T}$, so $\widetilde{S}$ is a $\mathbb{R}^d$-valued stochastic process that is also adapted. Intuitively, $\widetilde{S_t^i}$ is the time-$t$ price (in some units) of one unit (share) of asset $i$.

We agree on the following **simplifications**: express all prices in units of reference asset $\widetilde{S^0}$, so asset 0 has (new) unit price of 1 and $S^i := \widetilde{S^i}/\widetilde{S^0}$ are prices (in units of asset 0) of $d$ risky assets. Thus we usually start with model $(1, S)$ with $S$ being an adapted $\mathbb{R}^d$-valued stochastic process (almost always assume that $S$ is RCLL (**R**ight **C**ontinuous with **L**eft **L**imits) or in French càdlàg (**c**ontinue **à** **d**roite, **l**imite **à** **g**auche)), with $d + 1$ assets where $d \geq 1$.

Common terminology for the above simplifications is to say that $S$ is **asset-0-discounted**.

Here are two examples.

**Example 1.2** (Cox-Ross-Rubinstein/binomial model). *This is a discrete-time model, described by* $\widetilde{S_k^0} = (1+r)^k$ *and* $(d = 1)$ $\widetilde{S_k^1}/\widetilde{S_{k-1}^1}$ *i.i.d. with values* $1+u, 1+d$ *with probability* $p$ *respectively* $1-p$. *In other words, a multiplicative random walk with drift.*

**Remark 1.1.** We can naturally "transform" discrete into continuous-time models by taking the filtration $\mathbb{F}$ and processes to be piecewise constant.

**Example 1.3** (Black-Scholes model or Geometric Brownian Motion). *Bank account with (continuous compounding) interest rate* $r$, *so* $\widetilde{S_t^0} = e^{rt}$ *and* $(d = 1)$

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

*where* $W$ *denotes a Brownian motion. Thus*

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

*which satisfies by Itô's formula the SDE*

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

*More generally, we can discuss the Itô process model which has*

$$dS_t^i = S_t^i \left(b_t^i dt + \sum_{j=1}^n \sigma_t^{ij} dW_t^j\right)_{i=1,\dots,d}$$

*with* $\mathbb{R}^n$*-valued Brownian motion* $W$ *and predictable* $b$ *(*$\mathbb{R}^d$*-valued) and* $\sigma$ *(*$\mathbb{R}^{d\times n}$*-valued).*

We will now introduce further terminology/concepts in the same heuristic manner as before.

A **trading strategy** (or **dynamic portfolio**) is a stochastic process $\varphi = (\varphi_t)_{0\le t\le T}$, $\varphi = (\varphi^0, \vartheta)$ with $\varphi^0$ being $\mathbb{R}$-valued and $\vartheta$ being $\mathbb{R}^d$-valued. We interpret this as follows: at time $t$, we hold $\varphi_t^0$ units of asset 0 and $\vartheta_t^i$ units of asset $i$ where $i = 1, \ldots, d$.

A **value process** $V(\varphi) = (V_t(\varphi))_{0\le t\le T}$, in units of reference asset 0 is given by

$$V_t(\varphi) = \varphi_t^0 1 + \sum_{i=1}^d \vartheta_t^i S_t^i = \varphi_t^{\mathrm{tr}}(1 S_t)$$

which is the time-$t$ value of time-$t$ portfolio.

**Remark 1.2.**

1. For ease of notation, often think of $d = 1$, (i.e. one risky asset), this is usually harmless.

2. In original (tilde) units, value is given by

$$\widetilde{V}_t(\varphi) = \varphi_t^{\mathrm{tr}} \widetilde{S}_t = \varphi_t^0 \widetilde{S_t^0} + \sum_{i=1}^d \vartheta_t^i \widetilde{S}_t^i, \text{ where } \widetilde{S}_t^i = S_t^i \widetilde{S_t^0}.$$

   So clearly, $\widetilde{V}(\varphi) = \widetilde{S}_0 V(\varphi)$.

**The costs of the strategy**: if we keep $\varphi$ constant between $t$ and $t + \Delta t$ and only change it from $\varphi_t$ to $\varphi_{t+\Delta t}$ at $t + \Delta t$ this gives on $(t, t + \Delta t]$ expenses or **cost-increment**

$$\begin{aligned} C_{t+\Delta t} - C_t &= (\varphi_{t+\Delta t} - \varphi_t)^{\mathrm{tr}}(1 S_{t+\Delta t}) \\ &= \varphi_{t+\Delta t}^0 - \varphi_t^0 + (\vartheta_{t+\Delta t} - \vartheta_t)^{\mathrm{tr}} S_{t+\Delta t} \\ &= \underbrace{\varphi_{t+\Delta t}^0 + \vartheta_{t+\Delta t}^{\mathrm{tr}} S_{t+\Delta t}}_{=V_{t+\Delta t}} \underbrace{- \varphi_t^0 - \vartheta_t^{\mathrm{tr}} S_t}_{=-V_t(\varphi)} - \varphi_t^{\mathrm{tr}} S_{t+\Delta t} + \vartheta_t^{\mathrm{tr}} S_t \\ &= V_{t+\Delta t}(\varphi) - V_t(\varphi) - \vartheta^{\mathrm{tr}}(S_{t+\Delta t} - S_t). \end{aligned}$$

Adding up and letting $\Delta t \to 0$ suggests that the natural definition of total cost on $[0, t]$ of strategy $\varphi$ is given by

$$C_t(\varphi) = V_t(\varphi) - \int_0^t \vartheta_u dS_u, \ 0 \le t \le T.$$

(Note: integrand $\vartheta$ is naturally evaluated at left end $t$ of interval $(t, t + \Delta t]$.)

**Remark 1.3.** In original units, costs on $[0,t]$ is

$$\widetilde{V}_t(\varphi) = \widetilde{V}_t(\varphi) - \int_0^t \vartheta_u d\widetilde{S_u},$$

note that then $\widetilde{C}(\varphi) \neq S^0 C(\varphi)$.

In **discrete time**, $S$ is piecewise constant and the previous integral $\int \vartheta dS$ reduces to a sum. Only condition needed for $\vartheta$ , on economic grounds, is that stock holdings $\vartheta_{k+1}$ on $[k, k+1)$ must be determined at beginning $k$ of interval $[k, k+1)$, so $\vartheta_{k+1}$ must be $\mathcal{F}_k$-measurable for all $k$, i.e. $\vartheta$ must be predictable (needed to exclude insiders or prophets). No extra conditions on $\varphi^0$, it can be adapted.

In **continuous time**, by analogy, still impose that $\vartheta$ is predictable and $\varphi^0$ is adapted. But to have $\int \vartheta dS$ well-defined, need that $S$ is a **semimartingale** and $\vartheta$ is $S$-integrable, we will discuss this in more detail later; if $S$ happens to be a continuous semimartingale, then we know from Brownian Motion and Stochastic Calculus lectures what a sufficient condition for $S$-integrability is.

**Remark 1.4.** $V(\varphi), C(\varphi), \int \vartheta dS$ are always real valued (clear from the interpretation). If $\vartheta$ and $S$ are $\mathbb{R}^d$-valued, then $\int \vartheta dS$ denotes the vector stochastic integral

$$\text{"} \int \sum \vartheta^i dS^i \text{"};$$

this may be different from

$$\sum \int \vartheta^i dS^i.$$

The latter may fail to be well-defined. Here $d > 1$ needs technical care.

**Definition 1.3.** *A strategy $\varphi = (\varphi^0, \vartheta)$ is called **self-financing** if $C(\varphi) \equiv C_0(\varphi)$, i.e. $C_t(\varphi) = C_0(\varphi)$ $\mathbb{P}$-almost surely for all $t$.*

We interpret the above definition as follows: after initial outlay of $C_0(\varphi) = V_0(\varphi)$, trading according to $\varphi$ generates neither expenses nor surplus.

**Lemma 1.1.**

1. *A strategy $\varphi = (\varphi^0, \vartheta)$ is self-financing if and only if $V(\varphi) = V_0(\varphi) + \int \vartheta dS$.*

2. *(Reparametrisation): There exists a bijection between self-financing strategies $\varphi = (\varphi^0, \vartheta)$ and pairs $(v_0, \vartheta) \in L^0(\mathcal{F}_0) \times \{$predictable, $S$-integrable processes$\}$. Explicitly, $v_0 = V_0(\varphi)$ and conversely $\varphi_0 = v_0 + \int \vartheta dS - \vartheta^{tr} S$.*

3. *If $\varphi = (\varphi^0, \vartheta)$ is self-financing, then $\varphi^0$ is predictable.*

**Remark 1.5.** Part 2) of the above lemma is extremely useful: Can specify self-financing strategies by giving initial wealth/capital $v_0$ and prescribing arbitrary trading pattern (predictable, $S$-integrable) in the $d$ risky assets. Self-financing requirement then automatically determines what one must do in asset 0. This is very often used in literature without explicit mention. Crucially needs existence of reference asset $\widetilde{S^0} > 0$.

*Proof.*

1. $$V(\varphi) \overset{\text{def}}{=} C(\varphi) + \int \vartheta dS \overset{\text{self-fin.}}{=} C_0(\varphi) + \int \vartheta dS = V_0(\varphi) + \int \vartheta dS.$$

2. Mapping $(v_0, \vartheta) \mapsto (\varphi^0, \vartheta)$ is injective and resulting $\varphi$ is self-financing, this is easy to check using 1) as
$$V(\varphi) = \varphi^0 1 + \vartheta^{\text{tr}} S = v_0 + \int \vartheta dS = V_0(\varphi) + \int \vartheta dS.$$
Easy to check that mapping is surjective, using $v_0 = V_0(\varphi)$ to recover a given $\varphi^0$.

3. Needs one result from stochastic calculus. For any RCLL process $Y = (Y_t)_{0 \le t \le T}$ denote by $\Delta Y_t := Y_t - Y_{t^-} = Y_t - \lim_{s \to t, s < t} Y_s$ the jump of $Y$ at $t$. From stochastic integration theory, we need
$$\Delta \left( \int \vartheta dS \right)_t = \vartheta_t^{\text{tr}} \Delta S_t = \vartheta_t^{\text{tr}} (S_t - S_{t^-}).$$
So we get from 2) that
$$\begin{aligned} \varphi_t^0 &= v_0 + \int_0^t \vartheta dS - \vartheta_t^{\text{tr}} S_t = v_0 + \int_0^{t^-} \vartheta_u dS_u + \Delta \left( \int \vartheta dS \right)_t - \vartheta_t^{\text{tr}} S_t \\ &= v_0 + \int_0^{t^-} \vartheta_u dS_u - \vartheta_t^{\text{tr}} S_{t^-} \end{aligned}$$

We have $v_0$ is $\mathcal{F}_0$-measurable and thus predictable, $\int_0^{t-} \vartheta_u dS_u$ is LC, adapted and thus predictable, $\vartheta_t^{\text{tr}}$ is predictable by assumption, $S_{t-}$ is LC and adopted, hence predictable, which concludes that $\varphi_t^0$ is predictable.

□

The previous lemma establishes that in the model $(1, S)$, we can identify self-financing $\varphi$ with $(v_0, \vartheta) \in L^0(\mathcal{F}_0) \times \{\text{predictable, } S\text{-integrable processes}\}$, via $V(\varphi) = v_0 + \int \vartheta dS$. We use the notation

$$\varphi \mathrel{\hat{=}} (v_0, \vartheta).$$

**Remark 1.6** (Important)**.** Take $v_0 = 0$ and look at $\varphi \mathrel{\hat{=}} (0, \vartheta)$. Then

$$G(\vartheta) := \int \vartheta dS = 0 + \int \vartheta dS = V((0, \vartheta)) = V(\varphi)$$

is value process of a self-financing strategy starting with zero initial wealth and trading according to $\vartheta$. In particular

$$G_T(\vartheta) = \int_0^T \vartheta_u dS_u$$

is final wealth one can generate via self-financing trading using $\vartheta$ from 0 initial capital. This will come up very often.

**Remark 1.7.**

- Note that $\varphi \mathrel{\hat{=}} (0, \vartheta)$ and $\vartheta$ are different objects and will be distinguished - $\varphi$ is self-financing strategy, whereas $\vartheta$ is integrand for $S$.
- Also note that the identification needs the existence of $S^0 \equiv 1$.

---

During our treatise so far we have made many **implicit assumptions**. Let us list them here for clarification:

- Can trade **continuously in time**.
- All agents have **same information**.
- Prices for buying and selling are both given by $S$: no **transaction costs/frictionless trading**.
- The integrand $\vartheta$ is $\mathbb{R}^d$-valued, $\varphi^0$ is $\mathbb{R}$-valued, so $\vartheta_t^i, \varphi_t^0$ can take arbitrary values in $\mathbb{R}$, no **trading constraints** (like e.g. minimal lot size, integer number of units, short sales of stocks ($\vartheta_t^i < 0$) and borrowing reference asset ($\varphi_t^0 < 0$) allowed.
- Asset prices are exogenously given by fixed model $(1, S)$, do not react to trading strategies: **small investors/price takers**. In consequence, "book value" $V(\varphi)$ is also market/liquidation value.
- Probability measure $\mathbb{P}$ and hence law of $S$ are known, **no uncertainty about underlying probability model** (of course, the price evaluation $t \mapsto S_t(\omega)$ is still unknown, because $\omega \in \Omega$ is unknown).

Self-financing strategies are a reasonable requirement. But one cannot allow **all** self-financing $\varphi$; some restrictions are needed as the next example demonstrates.

**Example 1.4.** *Take $d = 1$, and $S = W$ a standard Brownian motion. For ease of exposition, work on infinite time horizon $[0, \infty)$.*

**Exercise 1.1.** *Construct similar example on $[0, T]$ and with $S > 0$.*

*Define $\tau := \inf\{t \geq 0 : W_t = 1\}$. This $\tau$ is a stopping time and (LIL: Law of Iterated Logarithm) $\tau < \infty$ $\mathbb{P}$-a.s. Define*

$$\vartheta := 1_{(\!(0,\tau]\!]}$$

*(notation $(\!(0,\tau]\!] := \{(\omega, t) \in \Omega \times [0,\infty) : 0 < t \leq \tau(\omega)\}$). We notice that $\vartheta$ is adapted and left-continuous, hence predictable, moreover it is bounded. So $\vartheta$ is $S$-integrable, and $\varphi \mathrel{\hat{=}} (0, \vartheta)$ has wealth given by*

$$V(\varphi) = \int \vartheta dS = W^\tau - W_0 = W^\tau = W_{t \wedge \tau}$$

*and in particular final value given by*

$$V_\infty(\varphi) = W_\infty^\tau = W_{\tau \wedge \infty} = W_\tau \equiv 1.$$

*Thus $\varphi$ starts from $0$, is self-financing and ends up with wealth $1$, which is clearly a* ***money pump****!*

Let us discuss the **problem** that has become evident in the previous example: The value process $V(\varphi) = W^\tau$ is unbounded from below.

- Interpretation: must be able to **borrow unlimited** amounts!

Consequently we should probably impose lower bound on wealth for "good" self-financial strategies.

- Why is $W^\tau$ unbounded from below?

Suppose for contradiction that $W^\tau \geq -a$ for some constant $a$. Then $W^\tau$ is a martingale (because BM $W$ is a martingale), hence a supermartingale and uniformly bounded from below by $a$. But then $W^\tau$ is closable on $[0, \infty]$ as super-martingale and so we can apply stopping theorem on $[0, \infty]$ to obtain

$$1 = \mathbb{E}(W_\tau) = \mathbb{E}(W^\tau_\infty) \leq \mathbb{E}(W^\tau_0) = \mathbb{E}(W_0) = 0$$

which gives a contradiction, hence $W^\tau = W_{t\wedge\tau}$ is indeed unbounded from below.

We conclude this introduction section by giving an overview of the **main topics** of this course:

1. **Arbitrage theory**: give precise mathematical description of idea that money pumps should be impossible, and characterize those models which satisfy this. Other models are not reasonable from economic/financial perspective.

2. **Pricing and hedging**: Start with reasonable model $(1, S)$. Fix $H \in L^0(\mathcal{F}_T)$ and view this as a random payoff to make at $T$ (due to financial contract). Can we find a self-financing $\varphi \mathrel{\hat{=}} (v_0, \vartheta)$ with $V_T(\varphi) = H$ (or perhaps $\geq H$, or close to $H$)? If yes, what is (perhaps minimal) required initial wealth $v_0$?

3. **Optimal investment**: Suppose again that $(1, S)$ is reasonable. Given initial wealth $v_0$, what is best investment strategy, i.e. which self-financing $\varphi \mathrel{\hat{=}} (v_0, \vartheta)$ produces "best" final wealth $V_T(\varphi) = v_0 + \int_0^T \vartheta_u dS_u$? This is a control problem over variable $\vartheta$, and it requires (subject) criterion for comparing final wealths.

## 2 First ideas in arbitrage theory

**Basic idea**: in a reasonable model of a financial market, money pumps should not exist. What is the mathematical formulation for this idea? How do we characterize such models?

**Setup**: Given a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ over $[0, T]$ with $S^0 \equiv 1$ and $S$ being $\mathbb{R}^d$-valued RCLL.

By Lemma 1.1 any $\mathbb{R}^d$-valued, predictable $\vartheta$ with $\int \vartheta dS$ well defined (which may need extra condition on $S$ and $\vartheta$) induces self-financing $\varphi \mathrel{\hat{=}} (0, \vartheta)$ with wealth

$$V_t(\varphi) = \int_0^t \vartheta_u dS_u = G_t(\vartheta),\ 0 \leq t \leq T.$$

**Definition 2.1.** *Call strategy $\varphi$ a-**admissible** with $a \geq 0$ if $V(\varphi) \geq -a$, meaning $V_t(\varphi) \geq -a$ $\mathbb{P}$-a.s. for all $t$ (if we can choose RCLL version for $V(\varphi)$ this is equivalent to $\mathbb{P}(V_t(\varphi) \geq -a, \forall t) = 1$). Call strategy $\varphi$ **admissible** if $a$-admissible for some $a \geq 0$.*

- **Interpretation**: trader using $\varphi$ (admissible as above) has some bound on his wealth/debts.

**Remark 2.1.**

- For $\varphi \mathrel{\hat{=}} (0, \vartheta)$, strategy $\varphi$ is $a$-admissible iff integrand $\vartheta$ is $a$-admissible, meaning that $\int \vartheta dS \geq -a$. Then write

$$\vartheta \in \Theta^a_{\text{adm}} \text{ and } \Theta_{\text{adm}} = \bigcup_{a \geq 0} \Theta^a_{\text{adm}}.$$

- For general $\varphi \mathrel{\hat{=}} (v_0, \vartheta)$ only clear connection between admissibility for $\varphi$ and admissibility of $\vartheta$ is if $v_0$ is constant. In general, $v_0 \in L^0(\mathcal{F}_0)$ and so one cannot say much, unless $\mathcal{F}_0$ is trivial.

**Definition 2.2.** *We say $\vartheta$ is a **simple integrand**, denoted by $\vartheta \in \boldsymbol{b\mathcal{E}}$ if it is of the form*

$$\vartheta = \sum_{i=1}^n h_i 1_{(\!(\tau_{i-1}, \tau_i]\!]},\ n \in \mathbb{N}$$

*with stopping times $0 \leq \tau_0 \leq \tau_1 \leq \tau_n \leq T$ and $h_i \in L^\infty(\mathcal{F}_{\tau_{i-1}}, \mathbb{R}^d)$, i.e. $h$ is $\mathbb{R}^d$-valued, $\mathcal{F}_{\tau_{i-1}}$-measurable and bounded. If all $\tau_i$ (but not $h_i$) are deterministic $t_i$, call $\vartheta$ **very simple** and write $\vartheta \in \boldsymbol{b\mathcal{E}}_{det}$.*

**Remark 2.2.**

- If we take any $\vartheta \in \mathbf{b}\mathcal{E}$, then for any $\mathbb{R}^d$-valued process $S$, the integral $\int \vartheta dS$ is well-defined as

$$G.(\vartheta) = \int_0^{\cdot} \vartheta dS = \sum_{i=1}^{n} h_i^{\mathrm{tr}}(S^{\tau_i} - S^{\tau_{i-1}})$$

so that

$$G_T(\vartheta) = \sum_{i=1}^{n} h_i^{\mathrm{tr}}(S_{\tau_i} - S_{\tau_{i-1}}).$$

If $S$ is adapted RCLL, then so is $G(\vartheta)$.

- For model with finite discrete time $k = 0, 1, \dots, T \in \mathbb{N}$, we have

$$\mathbf{b}\mathcal{E} = \{\text{all bounded } \mathbb{R}^d\text{-valued predictable processes}\}.$$

Let us now come back to our discussion of arbitrage opportunities:

- For **general** $S$**, simple arbitrage opportunity** is $\vartheta \in \mathbf{b}\mathcal{E}$, admissible, with $G_T(\vartheta) \in L^0_+ \setminus \{0\}$ i.e. $G_T(\vartheta) \geq 0$ $\mathbb{P}$-a.s. with $\mathbb{P}(G_T(\vartheta) > 0) > 0$, start from initial capital $v_0 = 0$, manage to keep debts bounded, trade via $\varphi \stackrel{\wedge}{=} (0, \vartheta)$ in self-financing way, end up with final wealth $V_T(\varphi) = G_T(\vartheta) \geq 0$ and even positive ($> 0$) with positive probability.

- For **semimartingale** $S$**, (general) arbitrage opportunity** is integrand $\vartheta$ for $S$ ($\mathbb{R}^d$-valued, predictable, $S$-integrable, so that $\int \vartheta dS$ is well-defined) which is admissible, with $G_T(\vartheta) \in L^0_+ \setminus \{0\}$ i.e. $G_T(\vartheta) \geq 0$ $\mathbb{P}$-a.s. with $\mathbb{P}(G_T(\vartheta) > 0) > 0$.

We now give mathematical sound definitions for the absence of arbitrage conditions. Let us introduce the notation $G_T(\Theta) := \{G_T(\vartheta) : \vartheta \in \Theta\}$.

**Absence of arbitrage conditions:**

For general $S$ we impose the following conditions:

- $(\mathrm{NA}_{\mathrm{elem}})$: $G_T(\mathbf{b}\mathcal{E}) \cap L^0_+ = \{0\}$.
- $(\mathrm{NA}^{\mathrm{adm}}_{\mathrm{elem}})$: $G_T(\mathbf{b}\mathcal{E} \cap \Theta_{\mathrm{adm}}) \cap L^0_+ = \{0\}$.
- $(\mathrm{NA}_{\mathrm{det}})$: $G_T(\mathbf{b}\mathcal{E}_{\mathrm{det}}) \cap L^0_+ = \{0\}$.

For semimartingale $S$ we only impose:

- (NA) $G_T(\Theta_{\mathrm{adm}}) \cap L^0_+ = \{0\}$.

**Remark 2.3.** Notice that a better notation for (NA) would be $(\mathrm{NA}^{\mathrm{adm}})$ because we insist on the integrand $\vartheta$ being admissible (i.e. $a$-admissible for some $a \geq 0$). But in literature this is never done, hence we do so here as well.

Giving a sufficient condition is easy, as described by the Lemma below:

**Lemma 2.1.** *Suppose $S$ is adapted $\mathbb{R}^d$-valued. If there exists a probability measure $Q \approx \mathbb{P}$ such that $S$ is a local $Q$-martingale, then both $(NA^{adm}_{elem})$ and (NA) hold. However, $(NA_{det})$ (and hence also $(NA_{elem})$) can fail.*

**Exercise 2.1.** *Counterexample for $(NA_{det})$.*

*Proof.* Notice that $(\mathrm{NA}^{\mathrm{adm}}_{\mathrm{elem}})$ follows from (NA), because $\mathbf{b}\mathcal{E} \cap \Theta_{\mathrm{adm}} \subset \Theta_{\mathrm{adm}}$.

We first show that $S$ is actually a semimartingale under the assumption of the Lemma. Indeed, $S$ is by assumption a $Q$-local martingale, and $Q \approx \mathbb{P}$, so the claim follows from general **Girsanov Theorem**. Thus we can talk about (NA) (which is only defined for semimartingales).

From stochastic calculus, if $S$ is a semimartingale and $\vartheta$ is predictable, $S$-integrable, then $G(\vartheta) = \int \vartheta dS$ is well-defined and again a semimartingale. But if $S$ is actually a local martingale, then $\int \vartheta dS$ must **not** be a local martingale (counterexample by Emery, uses $S$ with jumps)

However, if $S$ is a local martingale, $\vartheta$ is predictable, $S$-integrable and if $G(\vartheta) = \int \vartheta dS$ is uniformly bounded from below (i.e. if $\vartheta \in \Theta_{\text{adm}}$), then $\int \vartheta dS$ is again a local martingale (Ansel/Stricker). Then, of course, by Fatou $\int \vartheta dS$ is also a supermartingale.

We now argue (NA): $S \in \mathcal{M}_{\text{loc}}(Q)$ and $\vartheta \in \Theta_{\text{adm}}$ (this is same under $Q$ since $\mathbb{P} \approx Q$), thus $\int \vartheta dS \geq -a$ almost surely by admissibility. By our excursion above this shows that $\int \vartheta dS$ is a $Q$-supermartingale. So

$$\mathbb{E}_Q(G_T(\vartheta)) \leq \mathbb{E}_Q(G_0(\vartheta)) = 0.$$

If now also $G_T(\vartheta) \geq 0$ $\mathbb{P}$-a.s., then also $Q$-a.s. (since $Q \approx \mathbb{P}$), so we must have $G_T(\vartheta) = 0$ $Q$-a.s. hence $\mathbb{P}$-a.s. and so (NA) holds. □

**Remark 2.4.** Above supermartingale argument will come up again several times.

**Definition 2.3.** *An equivalent (local) martingale measure for $S$ is a probability measure $Q \approx \mathbb{P}$ such that $S$ is a $Q$-(local) martingale. $Q$ is then called an E(L)MM. (Sometimes only ask for $Q \approx \mathbb{P}$ on $\mathcal{F}_T$).*

With this definition in mind, we can rephrase Lemma 2.1:

**Lemma 2.2** (Rephrasing of Lemma 2.1)**.** *Suppose there exists an ELMM for $S$, then $S$ satisfies (NA).*

A natural question is then to ask if this sufficient condition is also necessary?

- **Answer**: For **finite discrete time** yes, see later. In general, the answer however is **no**! To get necessary and sufficient condition need something stronger than (NA).

We will now illustrate problems by giving two examples:

## 2.1 A counterexample in infinite discrete time

Take $(Y_n)_{n\in\mathbb{N}}$ under $\mathbb{P}$-independent with values $\pm 1$ and $\mathbb{P}(Y_n = \pm 1) = 1/2(1+\alpha_n)$. Define $S_0 = 1, \Delta S_n = S_n - S_{n-1} = \beta_n Y_n$. So $S$ is a binary random walk with drift. Take $\mathbb{F} = \mathbb{F}^Y = \mathbb{F}^S$ and $\mathcal{F} = \mathcal{F}_\infty$ if needed.

Since $\Delta S_n$ only takes values $\pm\beta_n$, the only way to get $S$ a $Q$-martingale is to set $Q(Y_{n+1} = +1 \mid \mathcal{F}_n) = 1/2$ for all $n \in \mathbb{N}$. So the sequence $(Y_n)_{n\in\mathbb{N}}$ must be i.i.d. under $Q$ with $Q(Y_n = +1) = 1/2$. This is the only candidate for an EMM.

**Theorem 2.1** (Kakutani's dichotomy theorem)**.** *In the above setting, $Q \approx \mathbb{P}$ if and only if $\sum_{n=1}^{\infty} \alpha_n^2 < \infty$, otherwise, $Q \perp \mathbb{P}$.*

(Note: $Q \overset{\text{loc}}{\approx} \mathbb{P}$ in the sense that $Q \approx \mathbb{P}$ on $\mathcal{F}_n$ for all $n \in \mathbb{N}$, thus we need an infinite horizon to create problems.)

Thus if we choose $(\alpha_n)_{n\in\mathbb{N}}$ such that $\sum_{n=1}^{\infty}\alpha_n^2 = +\infty$, then there is **no** EMM for $S$. The role of $\beta_n$ so far is not important, if we choose $\sum_{n=1}^{\infty}\beta_n < \infty$, then $S$ is bounded, and then $S \in \mathcal{M}_{\text{loc}}$ is the same as $S \in \mathcal{M}$. Hence EMM is the same as ELMM, and so there can also **exist no** ELMM for $S$.

**Claim**: Although $S$ does not admit any ELMM, it still satisfies (NA$_{\text{elem}}$).

- Why not look at (NA)? Infinite horizon means that we should look at $G_\infty(\vartheta)$, this is complicated for it is an infinite sum! For $\vartheta \in \mathbf{b}\mathcal{E}$, $G(\vartheta)$ is a finite sum and $G_\infty(\vartheta)$ is no problem.

Before we prove the claim, we need:

**Exercise 2.2.** *There exists an arbitrage opportunity in* $\boldsymbol{b}\mathcal{E}$ *if and only if there exists an arbitrage opportunity with* $\vartheta$ *of the form* $\vartheta = h1_{(\!(\sigma,\tau]\!]}$ *with* $\sigma \le \tau$ *stopping times and* $h \in L^\infty(\mathcal{F}_0)$ *(no admissibility here).*

Now choose $\beta_n = 3^{-n}$ so that for every $m$ we have $\sum_{k=m+1}^{\infty}\beta_k < \beta_m$. This ensures that for $n > m$ the both conditions below hold:

$$\begin{cases}\operatorname{sign}(S_n - S_m) & = \operatorname{sign}(Y_{m+1})\\ \operatorname{sign}(g(S_n - S_m)) & = \operatorname{sign}(gY_{m+1}), \text{ for any RV } g.\end{cases}$$

(Write $S_n - S_m = \sum_{k=m+1}^{\infty}\beta_k Y_k$, since the terms of the sequence of the sum are decreasing, only the first term is responsible for the sign)

Take $\vartheta = h1_{(\!(\sigma,\tau]\!]}$ and compute:

$$G_\infty(\vartheta) = \int_0^\infty \vartheta_u dS_u = h(S_\tau - S_\sigma).$$

It is **our goal** to show that if $G_\infty(\vartheta) \ge 0$ $\mathbb{P}$-a.s., then in fact $G_\infty(\vartheta) = 0$ $\mathbb{P}$-a.s., this then implies that we have (NA$_{\text{elem}}$).

Define for any $m \in \mathbb{N}$ the set $A_m := \{\sigma = m, \tau > m\} \in \mathcal{F}_m$. We then have:

$$\operatorname{sign}(G_\infty(\vartheta)) = \operatorname{sign}(h(S_\tau - S_\sigma)) = \operatorname{sign}(hY_{m+1}) \text{ on } A_m.$$

So if $G_\infty(\vartheta) \ge 0$ $\mathbb{P}$-a.s., we have $1_{A_m}\operatorname{sign}(hY_{m+1}) \ge 0$ for all $m \in \mathbb{N}$. We **claim** that this implies $h = 0$ $\mathbb{P}$-a.s. and then of course $G_\infty(\vartheta) = 0$ $\mathbb{P}$-a.s.

Indeed, $h$ is $\mathcal{F}_0$-measurable and $\sigma = m$ on $A_m$ (it can be shown that) this implies $h1_{A_m}$ is $\mathcal{F}_m$-measurable. Now, $0 \le 1_{A_m}\operatorname{sign}(h1_{A_m}Y_{m+1})$ and because $h1_{A_m}$ is $\mathcal{F}_m$-measurable and $A_m \in \mathcal{F}_m$, $h1_{A_m}Y_{m+1}$ must have the same sign on both $A_m \cap \{Y_{m+1} = +1\}$ and $A_m \cap \{Y_{m+1} = -1\}$. But product $h1_{A_m}Y_{m+1}$ has on $A_m$ a unique sign ($\ge 0$) and so we must have $h1_{A_m} = 0$ $\mathbb{P}$-a.s., this holds for all $m$ so that $h = 0$ $\mathbb{P}$-a.s. $\square$

## 2.2 A counterexample in continuous time

Start with a Brownian motion $W = (W_t)_{0\le t\le T}$ and take as $\mathbb{G}$ natural (augmented) filtration of $W$. For $k(t) := 1/\sqrt{T-t}$, define

$$Z_t := \begin{cases} \mathcal{E}\left(-\int_0^t k dW\right), & \text{on } [0,T) \\ Z_T := 0, & \text{for } t = T \end{cases}$$

so that

$$Z_t = \exp\left(-\int_0^t k(s)dW_s - \frac{1}{2}\int_0^t k^2(s)ds\right),\ 0 \le t < T.$$

With $\tau := \inf\{t \in [0,T] : Z_t \ge 2\} \wedge T$, $Z^\tau$ is a bounded martingale starting at 1, and

$$\begin{cases} Z_\tau = 2, & \text{if } \tau < T \\ Z_\tau = Z_T = 0, & \text{if } \tau = T \end{cases}$$

so that by the stopping theorem we get $\mathbb{E}(Z_\tau) = \mathbb{E}(Z_0) = 1$ which readily implies $\mathbb{P}(\tau < T) = 1/2$, because $\mathbb{E}(Z_\tau) = E(2 \cdot 1_{\tau<T} + 0 \cdot 1_{\tau=T}) = 1$. Now define $S = (S_t)_{0\le t\le T}$ as

$$S_t = \begin{cases} W_t + \int_0^t k(s)ds, & \text{for } t \le \tau \\ S_\tau, & \text{for } t \ge \tau \end{cases}$$

and take (augmented) filtration $\mathcal{F}_t := \mathcal{G}_{t\wedge T}$, $0 \le t \le T$, so that (up to nullsets) $\mathbb{F} = \mathbb{F}^S = \mathbb{F}^{W^\tau}$.

Now Brownian motion has a martingale representation in its own filtration. Using this for $W^\tau$, we can conclude that all $(\mathbb{F},\mathbb{P})$-local martingales must be stochastic integrals of $W^\tau$. By **Girsanov's theorem**, the only $Q \ll \mathbb{P}$ which makes $S = S^\tau$ into a local $Q$-martingale must remove drift $k$ and hence must have density process $Z^\tau$ and $\frac{dQ}{d\mathbb{P}} = Z_\tau$.

But $Z_\tau = 0$ on $\{\tau = T\}$ which has probability 1/2 so $\mathbb{P}(Z_\tau = 0) > 0$ and thus $Q \not\approx \mathbb{P}$ (only $Q \ll \mathbb{P}$). Therefore there cannot exist any ELMM for $S$.

**Goal:** Want to show that $S$ satisfies (NA) under $\mathbb{P}$.

**Remark 2.5.**

- Note that $S$ is a local $Q$-martingale, so by supermartingale argument used in Lemma 2.1, $S$ satisfies (NA) under $Q$.
- Note also that admissibility and $L^0$, hence also (NA), depend on underlying measure via its nullsets, so if $Q \not\approx \mathbb{P}$, there are differences.

Back to our goal: First note that if $\vartheta \in \Theta_{\mathrm{adm}}(\mathbb{P})$ with $G_T(\vartheta) \geq 0$ $\mathbb{P}$-a.s., then also $G_T(\vartheta) \geq 0$ $Q$-a.s. (by $Q \ll \mathbb{P}$) and $\vartheta \in \Theta_{\mathrm{adm}}(Q)$ (same reason). Due to NA(Q), this implies $G_T(\vartheta) = 0$ $Q$-a.s.

- But how about NA($\mathbb{P}$)?

Notice that

$$\frac{dQ}{d\mathbb{P}}\bigg|_{\mathcal{F}_t} = Z_t^\tau = Z_{t\wedge\tau} > 0, \text{ for any } t < T$$

by definition of $Z$. So this means that $Q \approx \mathbb{P}$ on $\mathcal{F}_t$ for any $t < T$.

Fix $\epsilon > 0$, set $\sigma := \inf\{t \in [0,T] : G_t(\vartheta) \geq \epsilon\} \wedge T$ and $\vartheta' := 1_{[\![0,\sigma]\!]}\vartheta$. Then $G(\vartheta') = (G(\vartheta))^\sigma$ so that $\vartheta' \in \Theta_{\mathrm{adm}}(\mathbb{P}) \cap \Theta_{\mathrm{adm}}(Q)$ just like $\vartheta$.
Moreover $G_T(\vartheta') = G_\sigma(\vartheta') = \epsilon$ on $\{\sigma < T\}$, hence, due to NA($Q$), we must have $Q(\sigma < T) = 0$ which means that $Q$-a.s., we have $G_.(\vartheta) \leq \epsilon$. In particular, $G_t(\vartheta) \leq \epsilon$ $Q$-a.s. for all $t < T$, hence also $G_t(\vartheta) \leq \epsilon$ $\mathbb{P}$-a.s. for all $t < T$, and because

$$G(\vartheta) = \int \vartheta dS$$

is ($\mathbb{P}$-a.s.) continuous like $S = W$, we get $G_T(\vartheta) \leq \epsilon$ $\mathbb{P}$-a.s.

Also we have by assumption $G_T(\vartheta) \geq 0$ $\mathbb{P}$-a.s. and since $\epsilon > 0$ was arbitrary we conclude that $G_T(\vartheta) = 0$ $\mathbb{P}$-a.s. In other words $G_T(\Theta_{\mathrm{adm}}(\mathbb{P})) \cap L^0_+(\mathbb{P}) = \{0\}$ so we have NA($\mathbb{P}$). □

**Remark 2.6.** Details for counterexamples can be found in Dalbaen/Schachenmayer Proposition 1.7. and Example 9.7.7.

## 2.3 The case of finite discrete time

**Goal**: Recapitulate version of FTAP (Fundamental Theorem of Asset Pricing) for finite discrete time, known from Introduction to Mathematical Finance (proofs been presented there), moreover we want to give an overview of the structure.

**Setup**: $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,\dots,T}$ and $S^0 = 1,\ S = (S_k)_{k=0,1,\dots,T}$ being $\mathbb{R}^d$-valued and $\mathbb{F}$-adapted.

**Notation**: We introduce the following notation

- $\Theta := m\mathcal{P} := \{\vartheta = (\vartheta_k)_{k=1,\dots,T},\ \vartheta \text{ is } \mathbb{R}^d\text{-valued predictable}\}$.
- $$G_\cdot(\vartheta) := \int_0^\cdot \vartheta dS = \sum_{k=1}^{\cdot} \vartheta_k^{\mathrm{tr}} \Delta S_k = \sum_{k=1}^{\cdot} \vartheta_k^{\mathrm{tr}} (S_k - S_{k-1}).$$
- $\mathcal{G} := G_T(\Theta) = \{G_T(\vartheta) : \vartheta \in \Theta\}$.
- $\Theta_{\mathrm{adm}} = \{\vartheta \in \Theta : G_\cdot(\vartheta) \geq -a \text{ for some } a \geq 0\}$.
- $\mathcal{G}_{\mathrm{adm}} = G_T(\Theta_{\mathrm{adm}})$.
- $\mathcal{C}^0 := \mathcal{C} := \mathcal{G} - L^0_+ = \{H = G_T(\vartheta) - Y : \vartheta \in \Theta, Y \geq 0\}$, all the payoffs one can dominate/superreplicate from zero initial wealth via self-financing trading.
- $\mathcal{C}^0_{\mathrm{adm}} = \mathcal{G}_{\mathrm{adm}} - L^0_+$ (analogous, with strategy even admissible).

**Recall**:

- Basic no-arbitrage condition (NA) is given by $\mathcal{G}_{\mathrm{adm}} \cap L^0_+ = \{0\}$.
- E(L)MM for $S$ is a probability measure $Q$ with $Q \approx \mathbb{P}$ such that $S$ is a (local) $Q$-martingale.

Let us denote by $\overline{\cdot}^{L^0}$ the closure in $L^0$, i.e. closure under convergence in probability.

**Theorem 2.2.** *For financial markets in finite discrete time the following are equivalent:*

1. *(NA).*
2. $\mathcal{C}^0_{adm} \cap L^0_+ = \{0\}$.
3. $\mathcal{G} \cap L^0_+ = \{0\}$.
4. $\mathcal{C}^0 \cap L^0_+ = \{0\}$.
5. $\mathcal{C}^0 \cap L^0_+ = \{0\}$ *and* $\mathcal{C}^0 = \overline{\mathcal{C}^0}^{L^0}$, *i.e.* $\mathcal{C}^0$ *is closed in* $L^0$.
6. $\overline{\mathcal{C}^0}^{L^0} \cap L^0_+ = \{0\}$.
7. *There exists an EMM $Q$ for $S$ with* $\frac{dQ}{d\mathbb{P}} \in L^\infty$.
8. *There exists an EMM $Q$ for $S$.*
9. *There exists an ELMM $Q$ for $S$.*

**Corollary 2.1** (Dalang-Morton-Willinger)**.** *In finite discrete time, a financial market is arbitrage free (in the sense that NA holds) if and only if it admits an equivalent martingale measure.*

**Comments on the proof**: Obviously, Corollary 2.1 is just 1) $\iff$ 8) from Theorem 2.2.

In the proof of Theorem 2.2, many easy small steps, one argument and two main ideas plus steps;

- 1) $\iff$ 2), 3) $\iff$ 4) are elementary.
- 1) $\implies$ 3) needs argument (see IMF); **only works in finite discrete time**!
- 4) $\implies$ 5) one major argument, argues that the arbitrage (here (NA)) implies that $\mathcal{C}^0$ is closed (here in $L^0$). In that form, **only works in finite discrete time**.
- 5 $\implies$ 6, 7) $\implies$ 8), 8) $\implies$ 9) are all clear.
- 9) $\implies$ 1) standard supermartingale argument from 2.1.

- 6) $\implies$ 7) second major argument, proves existence of EMM via **separation argument**. First change from $\mathbb{P}$ to $R \approx \mathbb{P}$ such that all $S_k$ are in $L^1(R)$. Call $R$ again $\mathbb{P}$, note that (NA) for $R$ equals (NA) for $\mathbb{P}$, that is we can without loss of generality assume that we work in $L^1$.

  To verify (NA) for $\mathbb{P}$, define $\mathcal{C}^1 := \mathcal{C}^0 \cap L^1$ to get $\mathcal{C}^1 \subset L^1$. $\mathcal{C}^1$ is a convex cone, $\mathcal{C}^1 \geq -L^1_+$ and $\mathcal{C}^1$ is closed in $L^1$ (for norm topology) hence by convexity also weakly closed in $L^1$.

  **Kreps-Yan theorem** gives the existence of a $Q \approx \mathbb{P}$ with $\frac{dQ}{d\mathbb{P}} \in L^\infty$ such that $\mathbb{E}_Q(Y) \leq 0$, for all $Y \in \mathcal{C}^1$ (see Appendix A on the lecture website). Take $\vartheta^{(\pm)} := \pm 1_P e^i$ with $P := A_k \times \{k\} \in \mathcal{P}$ (predictable set) for $A_k \in \mathcal{F}_{k-1}$, so $\vartheta^{(\pm)}$ is in $\Theta$. Then

$$\mathcal{C}^1 \ni G_T(\vartheta^{(\pm)}) = \pm 1_A(S^i_k - S^i_{k-1}).$$

  So $\mathbb{E}_Q(1_A(S^i_k - S^i_{k-1})) = 0$ for all $A \in \mathcal{F}_{k-1}$, for all $k$ and $i$ and so $S$ is a $Q$-martingale. We conclude that $Q$ is EMM with $\frac{dQ}{d\mathbb{P}} \in L^\infty$.

For a general model in continuous time, main ideas, results and arguments are similar, but more advanced.

- Do not work in $L^1$ (that depends on $\mathbb{P}$), but in $L^\infty$, so topology and closure change.

- We need a stronger condition than (NA) or $\mathcal{C}^0_{\text{adm}} \cap L^0_+ = \{0\}$, namely

$$\overline{\mathcal{C}^0_{\text{adm}} \cap L^\infty}^{L^\infty} \cap L^\infty_+ = \{0\},$$

  thus we must exclude not only "direct money pumps", but also their limits.

- Again show that this no-arbitrage condition implies closedness of $\mathcal{C}^0_{\text{adm}} \cap L^\infty$, but for weak-$*$ topology on $L^\infty$. Because $\mathcal{C}^0_{\text{adm}}$, via $\mathcal{G}_{\text{adm}}$, involves stochastic integrals of $S$, things become more technical.

- Again use Kreps-Yan theorem, now for $L^\infty$, to get $Q \approx \mathbb{P}$, now just with $\frac{dQ}{d\mathbb{P}} \in L^1(\mathbb{P})$ with $\mathbb{E}_Q(Y) \leq 0$, for all $Y \in \mathcal{C}^0_{\text{adm}} \cap L^\infty$.

- Need extra work to show that (even close to $Q$) there exists $Q' \approx Q \approx \mathbb{P}$ such that $S$ is under $Q'$ a $\sigma$-martingale (i.e. $S - S_0 = \int \Psi dM$ for integrand $\Psi > 0$ and some $M \in \mathcal{M}_{\text{loc}}(Q, \mathbb{R}^d)$. Also technically difficult (involves semimartingale characterisation).

  - So we need semimartingales and stochastic integrals **first**!

# 3 Stochastic integration and semimartingales

Start with $(\Omega, \mathcal{F}, \mathbb{P})$, $T \in (0, \infty)$, $\mathbb{F} = (\mathcal{F})_{0 \le t \le T}$ with usual conditions. For sub-/super-/ martingales, always choose RCLL version.

**Generic notations**: $S$ integrator or semimartingale, $H$ integrand, $X$ is a generic process.

**Definition 3.1.** *A process $X = (X_t)_{0 \le t \le T}$* ***satisfies property $\mathcal{E}$ locally*** *if there exists a sequence of $[0, T]$-valued stopping times $(\tau_n)_{n \in \mathbb{N}}$ with $\tau_n \nearrow T$ stationarily, and such that each $X^{\tau_n} = (X_{t \wedge \tau_n})_{0 \le t \le T}$ satisfies property $\mathcal{E}$. We say $\tau_n$ is a localising sequence.*

**Remark 3.1.** $\tau_n \nearrow T$ stationarily means $\tau_n \nearrow T$ $\mathbb{P}$-a.s. and $n \mapsto \tau_n(\omega)$ becomes constant ($\equiv T$) for $n \ge n_0(\omega)$. Equivalently, $\tau_n \nearrow T$ $\mathbb{P}$-a.s. and $\mathbb{P}(\tau_n = T) \to 1$ as $n \to \infty$. (For example, $\tau_n := T - \frac{1}{n}$ does not satisfy $\tau_n \nearrow T$ stationarily.)

**Definition 3.2.** *$S = (S_t)_{0 \le t \le T}$ is called a* ***semimartingale*** *if $S$ is adapted, RCLL and $S - S_0 = M + A$ with $M$ a RCLL local martingale null at $0$ and $A$ adapted RCLL with FV (finite variation) trajectories, null at 0.*

**Recall**: $H \in \mathbf{b}\mathcal{E}$ means that $H$ is of the form

$$H = \sum_{i=0}^{n} h_i 1_{(\!(\tau_i, \tau_{i+1}]\!]},$$

with $n \in \mathbb{N}$, stopping times $0 \le \tau_0 \le \tau_1 \le \cdots \le \tau_{n+1} \le T$ and $h_i \in L^\infty(\mathcal{F}_{\tau_i}, \mathbb{R}^d)$

**Definition 3.3.** *For any $\mathbb{R}^d$-valued $S = (S_t)_{0 \le t \le T}$, define (stochastic integrals) as map $I_S : \boldsymbol{b\mathcal{E}} \to L^0$ by*

$$I_S(H) := \sum_{i=0}^{n} h_i^{tr} (S_{\tau_{i+1}} - S_{\tau_i}) =: \int_0^T H_u dS_u =: H \cdot S_T$$

**Definition 3.4.** *We say $S = (S_t)_{0 \le t \le T}$ is a* ***good integrator*** *if $S$ is adapted RCLL, $\mathbb{R}^d$-valued such that $I_S : (\boldsymbol{b\mathcal{E}}, \| \cdot \|_\infty) \to L^0$ is continuous. In other words, if $H^n \to H$ uniformly (in $(\omega, t)$) and $H^n, H$ are in $\boldsymbol{b\mathcal{E}}$, then $I_S(H^n) \to I_S(H)$ in probability as $n \to \infty$.*

**Interpretation**: think of $S$ as discounted asset price and of $(v_0, H)$ as a (simple) self-financing strategy with initial wealth $v_0$. Then final wealth is given by $v_0 + \int_0^T H_u dS_u = v_0 + I_S(H)$, and then continuity means that small changes in strategies only creates small changes in final wealth.

**First main goal**: prove that semimartingales and good integrators are the same thing (Bichteler-Dellacherie theorem), in full generality.

**Intuition**: martingales are constant on average and thus stochastic analogues of constant functions. Semimarintgales are stochastic analogues of FV functions, and thus it may be useful to look at processes whose variation on average is finite. These are so-called quasimartingales.

**Definition 3.5.** *Let $X$ be integrable, i.e. $X_t \in L^1$ for all $0 \leq t \leq T$. For non-random partition $\pi = \{0 = t_0 \leq t_1 \leq \dots \leq t_{n+1} = T\}$ of $[0,T]$ define $MV(X,\pi)$ by*

$$MV(X,\pi) := \sum_{i=0}^{n} \mathbb{E}[|\mathbb{E}(X_{t_{i+1}} - X_{t_i} \mid \mathcal{F}_{t_i})|].$$

*Call $X$ **quasimartingale** if it is in addition adapted RCLL and if $MV(X) := \sup_\pi MV(X,\pi) < \infty$.*

**Remark 3.2.**

1. If $\pi \subset \pi'$, then $\mathrm{MV}(X,\pi) \leq \mathrm{MV}(X,\pi')$ by triangle inequality, so $\pi \mapsto \mathrm{MV}(X,\pi)$ is increasing.
2. Obvious examples of quasimartingales are: martingales, super/submartingales, process of integrable variation and linear combinations.

**Proposition 3.1.** *Suppose $S$ is a good integrator and bounded. Then for all $\epsilon > 0$ there exists a stopping time $\rho \leq T$ with $\mathbb{P}(\rho = T) \geq 1 - 3\epsilon$ and such that $S^\rho$ is a quasimartingale.*

**Remark 3.3.** Consequentially, $S$ a good integrator and bounded is locally a quasimartingale.

In order to proof Proposition 3.1 we need a $L^2$-version of **Komlós theorem**:

**Lemma 3.1.**

*(a) If $(g_n)_{n\in\mathbb{N}}$ is bounded in $L^2$, i.e. $\sup_{n\in\mathbb{N}} \|g_n\|_{L^2} < \infty$, then there exists a sequence $h_n \in conv(g_n, g_{n+1}, \dots) =: C_n$ (convex combination) such that $h_n \to h$ $\mathbb{P}$-a.s. and in $L^2$ for some $h \in L^2$.*

*(b) If $(f_n)_{n\in\mathbb{N}}$ is UI, then there exists a sequence $h_n \in conv(f_n, f_{n+1}, \dots)$ such that $h_n \to h$ in $L^1$ for some $h \in L^1$.*

*Proof of a).* For all $n \in \mathbb{N}, h_n \in C_n$ we have

$$\|h_n\|^2 \overset{\text{Jensen}}{\leq} \sup_{m \geq n} \|g_m\|^2 \leq \sup_{m \in \mathbb{N}} \|g_m\|^2 < \infty.$$

Next, define $A_n := \inf_{h \in C_n} \|h\|^2$ which is $\nearrow$ in $n$ because $C_n \searrow$ in $n$, and so

$$A_n \nearrow A := \sup_{n \in \mathbb{N}} \inf_{h \in C_n} \|h\|^2 \leq \sup_{m \in \mathbb{N}} \|g_m\|^2 < \infty.$$

For each $n$, choose $h_n \in C_n$ with

$$\|h_n\|^2 \leq A_n + \frac{1}{n} \leq A + \frac{1}{n}.$$

For $\epsilon > 0$, take $n$ large enough so that $A_n \geq A - \epsilon$. Then for $k, m \geq n$, $h_k \in C_k \subset C_n, h_m \in C_m \subset C_n$, so that $\frac{1}{2}(h_k + h_m) \in C_n$ (as a convex combination of elements of $C_n$) and hence (by definition of $A_n = \inf_{h \in C_n} \|h\|^2$)

$$\frac{1}{4}\|h_k + h_m\|^2 \geq A_n \geq A - \epsilon.$$

So for $k, m \geq n$, we get

$$\begin{aligned}\|h_k - h_m\|^2 &= 2\|h_k\|^2 + 2\|h_m\|^2 - \|h_k + h_m\|^2 \\ &\leq 4\left(A_n + \frac{1}{n}\right) - 4(A - \epsilon) \leq 4\left(A + \frac{1}{n}\right) - 4(A - \epsilon) = 4\left(\frac{1}{n} + \epsilon\right).\end{aligned}$$

So $(h_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in $L^2$ (which is complete as any $L^p$ space) hence converges in $L^2$ to some limit $h \in L^2$, and we also get $\mathbb{P}$-a.s. convergence along a subsequence.

Proof of b) Beiglböck/Siorpaes. $\square$

*Proof of Proposition 3.1.* For any bounded RCLL $X$

$$\lim_{t \searrow s} \mathbb{E}(X_t - X_s \mid \mathcal{F}_s) = 0.$$

Let us denote the $n$-th dyadic partition by $D_n := 2^{-n}T\mathbb{N}_0 \cap [0, T]$. Then for any partition $\pi$ of $[0, T]$, we can approximate $\mathrm{MV}(X, \pi)$ by $\mathrm{MV}(X, D_n)$ and because $\mathrm{MV}(X) = \lim_{n \to \infty} \mathrm{MV}(X, D_n)$ it is enough to look at $X$ along $D_n$.

1) $S$ is a good integrator, so $\forall \delta > 0, \forall \epsilon > 0, \exists \eta_0$ such that if $\|H\|_\infty \leq \eta \leq \eta_0$ implies $\mathbb{P}(|I_S(H)| \geq \delta) \leq \epsilon$. Fix $\delta > 0$ and choose $c := \delta/\eta_0 + 2\|S\|_\infty$. Then for all $H \in \mathbf{b}\mathcal{E}$ with $\|H\|_\infty \leq 1$, we get

$$\mathbb{P}(|I_S(H)| \geq c - 2\|S\|_\infty) \leq \mathbb{P}(|I_S(H\eta_0)| \geq \delta) \leq \epsilon.$$

For each $n \in \mathbb{N}$, let us define

$$\begin{aligned} H^n &:= \sum_{t_i \in D_n} 1_{(t_i, t_{i+1}]} \mathrm{sign}(\mathbb{E}[S_{t_{i+1}} - S_{t_i} \mid \mathcal{F}_{t_i}]) \in \mathbf{b}\mathcal{E} \\ H \cdot S &:= \text{stochastic integral process of } H \in \mathbf{b}\mathcal{E}. \\ \rho_n &:= \inf\{t_i \in D_n : H^n \cdot S_{t_i} \geq c - 2\|S\|_\infty\} \wedge T. \end{aligned}$$

Now $H^n \in \mathbf{b}\mathcal{E}$ has $\|H^n\|_\infty \leq 1$ and $I_S(H^n 1_{(\!(0,\rho_n]\!]}) = (H^n \cdot S)_{\rho_n} \geq c - 2\|S\|_\infty$ on $\{\rho_n < T\}$. This implies that $\mathbb{P}(\rho_n < T) \leq \epsilon$ so that $\mathbb{P}(\rho_n = T) \geq 1 - \epsilon$. Moreover, jumps of $S$ as well of $H^n \cdot S$ are bounded by $2\|S\|_\infty$ so that definition of $\rho_n$ gives $(H^n \cdot S)_{\rho_n} \leq c$. Using the definition of $H^n$ yields

$$\begin{aligned} c \geq \mathbb{E}((H^n \cdot S)_{\rho_n}) &= \sum_{t_i \in D_n} \mathbb{E}(1_{\{t_i < \rho_n\}} |\mathbb{E}(S_{t_{i+1}} - S_{t_i} \mid \mathcal{F}_{t_i})|) \\ &=: \mathrm{MV}(S^{\rho_n +}, D_n) \geq \mathrm{MV}(S^{\rho_n}, D_n) \end{aligned} \tag{*}$$

The last inequality ($\geq$) requires an argument:

$$\mathrm{MV}(S^\tau, D_n) - \mathrm{MV}(S^{\tau+}, D_n) = \sum_{t_i \in D_n} \mathbb{E}(|\mathbb{E}(S^\tau_{t_{i+1}} - S^\tau_{t_i} \mid \mathcal{F}_{t_i})| - 1_{\{t_i < \tau\}} |\mathbb{E}(S_{t_{i+1}} - S_{t_i} \mid \mathcal{F}_{t_i})|)$$

and

$$\mathbb{E}(S^\tau_{t_{i+1}} - S^\tau_{t_i} \mid \mathcal{F}_{t_i}) = 1_{\{t_i < \tau\}} (\mathbb{E}(S_{t_{i+1}} - S_{t_i} \mid \mathcal{F}_{t_i}) + \mathbb{E}(S_{t_{i+1} \wedge \tau} - S_{t_i} \mid \mathcal{F}_{t_i}))$$

so

$$\mathrm{MV}(S^\tau, D_n) - \mathrm{MV}(S^{\tau+}, D_n) \leq \sum_{t_i \in D_n} \mathbb{E}(1_{\{t_i < \tau\}} |\mathbb{E}(S_{t_{i+1} \wedge \tau} - S_{t_i} \mid \mathcal{F}_{t_i})|) \leq 2\|S\|_\infty$$

as there is at most one $i$ which gives a non zero summand.

**Problem**: $\rho_n$ still depends on $n$ !

2) Now we get rid of dependence of $\rho_n$ on $n$ via using Lemma 3.1 a):
For $g_n := 1_{\{\rho_n = T\}}$ to get convex $\mu^n_j, j = 1, \ldots, N^n$ such that that

$$h_n = \sum_{j=1}^{N^n} \mu^n_j g_{n_j}$$

(with $n_j \geq n$) converges to some $h$ $\mathbb{P}$-a.s.

Because $0 \leq h \leq 1$ and $\mathbb{E}(g_n) = \mathbb{P}(\rho_n = T) \geq 1 - \epsilon$ we get (by Lebesgue) $\mathbb{E}(h) \geq 1 - \epsilon$, we must have $\mathbb{P}(h < 2/3) < 3\epsilon$ (substitute $y = 1 - h$, then $\mathbb{E}(y) \leq \epsilon$ and by Markov's inequality $\mathbb{P}(h < 2/3) = \mathbb{P}(y > 1/3) \leq 3\mathbb{E}(y) \leq 3\epsilon$). So $1 - 3\epsilon < \mathbb{P}(\lim h_n = h \geq 2/3)$ and by **Egorov's theorem** (uniform convergence on large set) we must have $h_n \geq 1/2$ on $A$ for $n \geq n_0$ with $\mathbb{P}(A) \geq 1 - 3\epsilon$.

Now define

$$B_t^n := \sum_{j=1}^{N^n} \mu_j^n 1_{[\![0,\rho_{n_j}]\!]}(t).$$

Each $B^n$ is decreasing, adapted LC with $B_T^n = h_n$ because $1_{[\![0,\rho_{n_j}]\!]}(T) = 1_{\{\rho_{n_j}=T\}}$. Define stopping times $\sigma_n := \inf\{t \in [0,T] : B_t^n < 1/2\} \wedge T$, $\rho := \inf_{n \geq n_0} \sigma_n$. Then $B_\cdot^n \geq 1/2$ on $[\![0, \sigma_n]\!]$ and that for all $n \geq n_0$

$$1_{[\![0,\rho)\!)} \leq 1_{[\![0,\sigma_n)\!)} \leq 2B_\cdot^n = 2\sum_{j=1}^{N^n} \mu_j^n 1_{[\![0,\rho_{n_j})\!)} \tag{**}$$

Moreover, $\mathbb{P}(\rho = T) \geq 1 - 3\epsilon$ because $\{\rho = T\} \supset A$, indeed, $\rho < T$ implies that for some $n \geq n_0$ we have $B_t^n < 1/2$ for some $t \in [0,T]$ and then $h_n = B_T^n \leq B_t^n < 1/2$ and we are in the complement of $A$.

Now we compute

$$\begin{aligned}
\mathrm{MV}(S^\rho, D_n) - 2\|S\|_\infty &\overset{*}{\leq} \mathrm{MV}(S^{\rho+}, D_n) = \sum_{t_i \in D_n} \mathbb{E}(1_{\{t_i<\rho\}}|\mathbb{E}(S_{t_{i+1}} - S_{t_i} \mid \mathcal{F}_{t_i})|) \\
&\overset{**}{\leq} 2\sum_{j=1}^{N^n} \sum_{t_i \in D_n} \mu_j^n \mathbb{E}(1_{\{t_i<\rho_{n_j}\}}|\mathbb{E}(S_{t_{i+1}} - S_{t_i} \mid \mathcal{F}_{t_i})|) = 2\sum_{j=1}^{N^n} \mu_j^n \mathrm{MV}(S^{\rho_{n_j}+}, \underbrace{D_n}_{\subset D_{n_j}}) \\
&\leq 2\sum_{j=1}^{N^n} \mu_j^n \mathrm{MV}(S^{\rho_{n_j}+}, D_{n_j}) \leq 2c.
\end{aligned}$$

So $\mathrm{MV}(S^\rho) = \lim_{n\to\infty} \mathrm{MV}(S^\rho, D_n) \leq 2(c + \|S\|_\infty) < \infty$, we conclude that $\rho$ does the job. $\square$

We've mentioned that, loosely speaking, quasimartingales are stochastic analogues of FV functions, which are differences of two increasing functions. Next result is stochastic analogue/version of said result:

**Theorem 3.1** (Rao)**.** *Every quasimartingale, can be written as a difference of two RCLL, nonnegative supermartingales.*

*Proof.* For dyadic rationals $s \in D_n$, define

$$Y_s^n := \mathbb{E}\Big[\sum_{\substack{t_i \in D_n \\ t_i \geq s}} (\mathbb{E}[X_{t_i} - X_{t_{i+1}} \mid \mathcal{F}_{t_i}])^+ \mid \mathcal{F}_s\Big]$$
$$Z_s^n := \mathbb{E}\Big[\sum_{\substack{t_i \in D_n \\ t_i \geq s}} (\mathbb{E}[X_{t_i} - X_{t_{i+1}} \mid \mathcal{F}_{t_i}])^- \mid \mathcal{F}_s\Big]$$

Then $Y^n \geq 0, Z^n \geq 0$ and by telescopic sum $Y_s^n - Z_s^n = X_s - \mathbb{E}(X_T \mid \mathcal{F}_s)$ for $s \in D_n$, and both $Y^n, Z^n$ are along $D_n$ supermartingales. Moreover for $v \geq r \geq u$, we have

$$\begin{aligned}\mathbb{E}(X_v - X_u \mid \mathcal{F}_u)^{\pm} &= 2\mathbb{E}\left(\frac{1}{2}\mathbb{E}(X_v - X_r \mid \mathcal{F}_r) + \frac{1}{2}\mathbb{E}(X_r - X_u \mid \mathcal{F}_u) \mid \mathcal{F}_u\right)^{\pm} \\ &\overset{\pm \text{ convex}}{\leq} \mathbb{E}[(\mathbb{E}[X_v - X_r \mid \mathcal{F}_r])^{\pm} + (\mathbb{E}[X_r - X_u \mid \mathcal{F}_u])^{\pm} \mid \mathcal{F}_u].\end{aligned}$$

Use this with $v = t_{i+1} \in D_n$, $u = t_i \in D_n$ and $r = \frac{1}{2}(t_{i+1} + t_i) \in D_{n+1}$, sum up and condition on $\mathcal{F}_s$ to see that both $n \mapsto Y_s^n$ and $n \mapsto Z_s^n$ are increasing $\mathbb{P}$-a.s. So define the $\mathbb{P}$-a.s. limits

$$Y_s := \lim_{n\to\infty} Y_s^n + \mathbb{E}(X_T^+ \mid \mathcal{F}_s)$$
$$Z_s := \lim_{n\to\infty} Z_s^n + \mathbb{E}(X_T^- \mid \mathcal{F}_s)$$

on any $s \in \cup_{n\in\mathbb{N}} D_n =: D$. Since $\mathbb{E}(Y_s^n + Z_s^n) \leq \mathrm{MV}(X, D_n) \leq MV(X) < \infty$, we can conclude by monotone convergence that $Y_s, Z_s \in L^1$ for $s \in D$. Finally, by Fatou, both $Y, Z$ are nonnegative supermartingales on the index set $D = \cup_{n\in\mathbb{N}} D_n$ and $X_s = Y_s - Z_s$ for $s \in D$.

Standard argument, using that $\mathbb{F}$ is RC, allows us to extend $Y, Z$ to RCLL nonnegative supermartingales on $[0, T]$, see DM VI 1-4. □

**Definition 3.6.** *An adapted RCLL process $X = (X_t)_{0\le t\le T}$ is said to be of* ***class (D)*** *if the family $\{X_\tau : \tau \le T$ is a stopping time$\}$ is UI.*

**Theorem 3.2** (Doob-Meyer)**.** *Any supermartingale $X$ of class (D) has a (unique) decomposition $X - X_0 = M - A$ into RCLL martingale $M$ null at 0 and RCLL $A$ null at 0 which is increasing, integrable and predictable.*

*Proof.* **Uniqueness** (not really needed later): Any FV local martingale which is continuous must be constant (see BMSC), and any local martingale which is predictable must be continuous. Indeed, to see the latter, we make use of the predictable stopping theorem, i.e., let $L$ be a local martingale and $\tau$ a predictable stopping time (i.e. $\{(\omega, t) \in \Omega\times\mathbb{R}_+ : \tau(\omega) > t\} \in \mathcal{P}$), then we have $\mathbb{E}(\Delta L_\tau \mid \mathcal{F}_{\tau-}) = 0$.

For $X$ predictable and $\tau$ a stopping time, $\Delta X_\tau$ is $\mathcal{F}_{\tau-}$-measurable. So if $L \in \mathcal{M}_{\text{loc}}$ is predictable, then jump times $\tau$ of $L$ are predictable and so

$$\Delta L_\tau = \mathbb{E}(\Delta L_\tau \mid \mathcal{F}_{\tau-}) = 0$$

and we conclude that $L$ is continuous.

**Existence**: 1) For any $n \in \mathbb{N}$, view $X$ along $D_n$ as discrete-time supermartingale and write its classic Doob-decomposition as $X - X_0 = M^n - A^n$ on $D_n$, with

$$\begin{aligned} A^n_{t_{i+1}} - A^n_{t_i} &= \mathbb{E}(X_{t_i} - X_{t_{i+1}} \mid \mathcal{F}_{t_i}),\ t_i \in D_n \\ M^n_t &= X_t - X_0 + A^n_t,\ t \in D_n \end{aligned}$$

We want to construct $M_T$ as "limit" and so we first argue that $(M^n_T)_{n\in\mathbb{N}}$ is UI, in order to then use Lemma 3.1 b). By replacing $X_t$ by $X_t - \mathbb{E}(X_T \mid \mathcal{F}_t)$, we can assume that $X \ge 0$ and $X_T = 0$.

So $M^n_T = A^n_T - X_0$, and for any stopping time $\tau$ along $D_n$, using stopping theorem in finite discrete time gives $X_\tau = \mathbb{E}(A^n_T \mid \mathcal{F}_\tau) - A^n_\tau$.

For $c > 0$, define along $D_n$ stopping times $\tau_n(c) := \inf\{t_i \in D_n : A^n_{t_{i+1}} > c\} \wedge T$ so that $A^n_{\tau_n(c)} \le c$, hence $X_{\tau_n(c)} \ge \mathbb{E}(A^n_T \mid \mathcal{F}_{\tau_n(c)}) - c$ and $\{\tau_n(c) < T\} = \{A^n_T > c\}$. So we get

$$\mathbb{E}(A^n_T 1_{\{A^n_T > c\}}) = \mathbb{E}(\mathbb{E}(A^n_T \mid \mathcal{F}_{\tau_n(c)})1_{\{\tau_n(c)<T\}}) \le \mathbb{E}(X_{\tau_n(c)}1_{\{\tau_n(c)<T\}}) + c\mathbb{P}(\tau_n(c) < T)$$

But $\{\tau_n(c) < T\} \subset \{\tau_n(\frac{c}{2}) < T\}$ and thus

$$\mathbb{E}(X_{\tau_n(\frac{c}{2})} 1_{\{\tau_n(\frac{c}{2})<T\}}) = \mathbb{E}[(A_T^n - \overbrace{A^n_{\tau_n(\frac{c}{2})}}^{\leq c/2}) 1 \underbrace{\{\tau_n(c/2) < T\}}_{\supset\{\tau_n(c)<T\}}]$$
$$\geq \mathbb{E}[(\overbrace{A_T^n}^{\geq c} - A^n_{\tau_n(\frac{c}{2})}) 1_{\{\tau_n(c)<T\}}) \geq \frac{c}{2}\mathbb{P}(\tau_n(c) < T).$$

This yields (after mult. by 2) that $c\mathbb{P}(\tau_n(c) < T) \leq 2\mathbb{E}(X_{\tau_n(\frac{c}{2})} 1_{\{\tau_n(\frac{c}{2})<T\}})$ hence,

$$\mathbb{E}(A_T^n 1_{\{A_T^n > c\}}) \leq \mathbb{E}(X_{\tau_n(c)} 1_{\{\tau_n(c)<T\}}) + 2\mathbb{E}(X_{\tau_n(\frac{c}{2})} 1_{\{\tau_n(\frac{c}{2})<T\}}) \tag{*}$$

But

$$\mathbb{P}(\tau_n(c) < T) = \mathbb{P}(A_T^n > c) \overset{\text{Markov}}{\leq} \frac{\mathbb{E}(A_T^n)}{c} = \frac{\mathbb{E}(M_T^n + X_0)}{c} = \frac{\mathbb{E}(X_0)}{c} \xrightarrow{c\to\infty} 0$$

uniformly in $n \in \mathbb{N}$ (since RHS does not depend on $n$). Because $X$ is by assumption of class (D) we get

$$\lim_{c\to\infty} \sup_{n\in\mathbb{N}} \mathbb{E}(A_T^n 1_{\{A_T^n > c\}}) = 0$$

i.e. $(A_T^n)_{n\in\mathbb{N}}$ is UI, and then so is $(M_T^n)_{n\in\mathbb{N}}$ because $M_T^n = A_T^n - X_0$.

2) Now we use Lemma 3.1 b) to get "limit". More precisely, we get convex weights $\lambda_j^n, j = 1, \ldots, N^n$ and $n_j \geq n$ such that

$$L_T^n := \sum_{j=1}^{N^n} \lambda_j^n M_T^{n_j} \to M_T \text{ in } L^1$$

for some RV $M_T \in L^1$. Define associated RCLL martingales $L_t^n := \mathbb{E}(L_T^n \mid \mathcal{F}_t), M_t := \mathbb{E}(M_T \mid \mathcal{F}_t)$, $0 \leq t \leq T$. Then $L_t^n \to M_t$ in $L^1$ for all $t$, by Jensen. Extend each $A^n$ to a process on $[0, T]$ by piecewise constant LCRL interpolation along $D_n$ and set

$$B^n := \sum_{j=1}^{N^n} \lambda_j^n A^{n_j}, \text{ with same weights as above.}$$

Then $A := M - X + X_0$ is RCLL and for all $t \in D$, $B_t^n = L_t^n - X_t + X_0 \to M_t - X_t + X_0 = A_t$ in $L^1$. Along a subsequence, we get $\mathbb{P}$-a.s. convergence, simultaneously for all $t \in D$, and so $A$ is $\mathbb{P}$-a.s. increasing on $D$ like all the $A^n, B^n$. But $A$ is RC, and so $A$ has $\mathbb{P}$-a.s. increasing trajectories on $[0, T]$.

3) It remains to show that $A$ is (indistinguishable from) a predictable process.

All the $A^n, B^n$ are predictable, because they are LC adapted, and so it is enough to show that

$$A_t(\omega) = \limsup_{n\to\infty} B_t^n(\omega), \text{ for all } t \in [0,T] \ \mathbb{P}\text{-a.s.} \tag{**}$$

If $f_n, f : [0,T] \to \mathbb{R}$ are increasing with $f$ RC and $f_n(t) \to f(t)$ for all $t \in D$, then $\limsup_{n\to\infty} f_n(t) \le f(t)$ for all $t \in [0,T]$ and $\lim_{n\to\infty} f_n(t) = f(t)$ if $f$ is continuous at $t$.

So (**) can only fail at discontinuous (jump) times of $A$, and because $A$ is adapted RCLL we can write

$$\{\Delta A_n \neq 0\} \subset \bigcup_{n=1}^{\infty} [\![\tau_n]\!]$$

for a sequence of stopping times. So (**) will follow if we show that

$$\limsup_{n\to\infty} B_\tau^n(\omega) = A_\tau(\omega) \ \mathbb{P}\text{-a.s. for every stopping time } \tau \tag{***}$$

We know from the prev. page that $\limsup_{n\to\infty} B_\tau^n \le A_\tau$ $\mathbb{P}$-a.s. and $B_\tau^n \le B_t^n \to A_t$ in $L^1$. So we get

$$\liminf_{n\to\infty} \mathbb{E}(A_\tau^n) \overset{\text{conv. comb}}{\le} \limsup_{n\to\infty} \mathbb{E}(B_\tau^n) \overset{\text{Fatou}}{\le} \mathbb{E}(\limsup_{n\to\infty} B_\tau^n) \le \mathbb{E}(A_\tau)$$

thus (***) will follow if we can show that $\lim_{n\to\infty} \mathbb{E}(A_\tau^n) = \mathbb{E}(A_\tau)$.

Define $\sigma_n := \inf\{t \in D_n : t \ge \tau\}$ so that $\sigma_n \searrow \tau$ and $A_{\sigma_n}^n = A_\tau^n$ by piecewise LCRL-construction of $A^n$. So we get

$$\mathbb{E}(A_\tau^n) = \mathbb{E}(A_{\sigma_n}^n) = \mathbb{E}(M_{\sigma_n}^n - X_{\sigma_n} + X_0)$$
$$= \mathbb{E}(X_0 - \underbrace{X_{\sigma_n}}_{\to X_\tau \text{ a.s. by RC}}) \xrightarrow[n\to\infty]{X \text{ of class (D)}} \mathbb{E}(X_0 - X_\tau) = \mathbb{E}(M_\tau - X_\tau + X_0) = \mathbb{E}(A_\tau)$$

□

Now we can show:

**Theorem 3.3.** *Every good integrator is a semimartingale.*

*Proof.* For any adapted RCLL $S$

$$J^1_\cdot := \sum_{0<s\leq\cdot} \Delta S_s 1_{\{|\Delta S_s|>1\}}$$

is adapted, RCLL and of FV, hence a good integrator and a semimartingale. Next, $S - J^1$ is adapted RCLL with bounded jumps, hence locally bounded, and like $J^1$ a good integrator if $S$ is.

By Proposition 3.1, $S - J^1$ is locally a quasimartingale and hence by Theorem 3.1 locally difference of two supermartingales. But every supermartingale is locally of class (D), indeed let $X = (X_t)_{0\leq t\leq T}$ be a supermartingale, let $\tau_n := \inf\{t \in [0,T] : |X_t| > n\} \wedge T$, this gives for every stopping time $\sigma \leq T$,

$$|X^{\tau_n}_\sigma| \leq n + |X_{\tau_n}| \in L^1 \tag{*}$$

which is true because, $-X$ and then $|X|$ is a submartingale and thus

$$|X_{\tau_n}| \overset{\text{stopp. thm}}{\leq} \mathbb{E}(|X_T| \mid \mathcal{F}_{\tau_n}) \in L^1.$$

So by (*) we conclude that $X^{\tau_n}$ is locally of class (D). So $S - J^1$ is by Theorem 3.2 locally difference of martingale and difference of two predictable, increasing processes, thus $S - J^1$ is locally sum of martingale and a FV process, hence locally a semimartingale.

The same is then true for $S = S - J^1 + J^1$, and we are done because local semimartingales are semimartingales (see exercise below). □

**Exercise 3.1.** *Local semimartingales are semimartingales.*

**Remark 3.4.** Above proof shows us a bit more: any local-bounded semimartingale $S$ can be written as $S - S_0 = M + A$ with $M \in \mathcal{M}_{0,\text{loc}}$ and $A \in FV_0$ is even **predictable**. As in the proof of Theorem 3.2, one can argue that this decomposition is in fact unique, we call this the **canonical decomposition of special semimartingales**. For general semimartingales decomposition is not unique and $J$ is not predictable.

For converse of Theorem 3.3 we need an auxiliary result. For $Y$ RCLL, write $Y_t^* := \sup_{0\le s\le t} |Y_s|$, where $t \in [0,T]$ and for $H \in \mathbf{b}\mathcal{E}$, write $H \cdot Y$ for (elementary) stochastic integral process (i.e. $I_Y(H) = H \cdot Y_T$).

**Lemma 3.2.** *For any martingale $L$, $H \in \boldsymbol{b}\mathcal{E}$, $c > 0$ we have*

$$c\mathbb{P}(|I_L(H)| \ge c) \overset{(1)}{\le} c\mathbb{P}((H \cdot L)_T^* \ge c) \overset{(2)}{\le} 34\|H\|_{L^\infty}\|L_T\|_{L^1}$$

**Theorem 3.4.** *Every semimartingale is a good integrator.*

*Proof.* Sums of good integrators and local good integrators are good integrators, so it is enough to show that martingales and FV processes are good integrators.

For martingale $L$ and $H \in \mathbf{b}\mathcal{E}$

$$\mathbb{P}(|I_L(H)| \ge \delta) \overset{\text{L 3.2}}{\le} \frac{34}{\delta}\|H\|_{L^\infty}\|L_T\|_{L^1} \to 0, \text{ if } \|H\|_{L^\infty} \to 0.$$

For FV process $A$ and $H \in \mathbf{b}\mathcal{E}$,

$$|I_A(H)| = \left|\int_0^T H_s dA_s\right| \le \|H\|_{L^\infty} \underbrace{|A|_T}_{\text{Var of } A \text{ on } [0,T]} \to 0, \text{ even } \mathbb{P}\text{-a.s. as } \|H\|_{L^\infty} \to 0.$$

□

**Remark 3.5.** A variant of the above argument for martingales may have appeared in BMSC for square integrable martingales:

$$\|I_M(H)\|_{L^2} \overset{\text{Itô Isom.}}{=} \|H\|_{L^2(M)} = \left(\mathbb{E}\left(\int_0^T H_u^2 d\langle M\rangle_u\right)\right)^{1/2} \le \|H\|_{L^\infty}\|M_T\|_{L^2}.$$

*Proof of Lemma 3.2.* Part (1) is clear. For proving Theorem 3.4 we did not use (2), but only that LHS $\le$ RHS. That can be proved by viewing $L$ as a martingale in discrete time, which makes things simpler. However (2) will be needed later, and this cannot be reduced to discrete-time arguments.

1) Suppose first that $L \ge 0$. Then $Z := L \wedge c$ is a bounded, nonnegative supermartingale (since $\wedge$ is convex) and so Theorem 3.2 yields $Z = M - A$ with $M$ a martingale with $M_0 = Z_0$ and $A$ predictable, increasing integrable RCLL with $A_0 = 0$. Then $M = Z + A \ge 0$ and we claim that $M$ is square-integrable. Since $M_T^2 \le 2(Z_T^2 + A_T^2) \le 2c^2 + A_T^2$ (equivalent to $(a-b)^2 \ge 0$) and $M$ is a martingale it suffices, by Doobs $L^2$-inequality, to show that $A_T^2$ is integrable.

Look at

$$A_T^2 \overset{\text{Itô}}{=} 2\int_0^T A_{s-}dA_s + \sum_{0<s\le T}(\Delta A_s)^2 = \int_0^T A_{s-}dA_s + \int_0^T A_s dA_s \le 2\int_0^T A_s dA_s.$$

In the same way, using $A^n := A \wedge n$, we get

$$A_T A_T^n = \int_0^T A_{s-}dA_s^n + \int_0^T A_s^n dA_s,$$

so we get

$$\begin{aligned}
\int_0^T A_s^n dA_s &= \int_0^T (A_T - A_{s-})dA_s^n \\
&= \int_0^T (M_T - M_{s-})dA_s^n - \int_0^T (\underbrace{Z_T}_{\ge 0} - \underbrace{Z_{s-}}_{\le c})d\underbrace{A_s^n}_{\le A_s} \\
&\le \underbrace{\int_0^T (M_T - M_{s-})dA_s^n}_{=:Y_T} + cA_T \qquad (*)
\end{aligned}$$

Now $M \ge 0$ is a martingale and $A^n$ is increasing, predictable, null at 0, so we have:

**Exercise 3.2.** *Show that:*

$$\mathbb{E}(M_T A_T^n) = \mathbb{E}\left(\int_0^T M_{s-}dA_s^n\right).$$

Since $A^n$ is bounded, so both expressions in the above are finite, and hence we can take differences to obtain $\mathbb{E}(Y_T) = 0$ and then

$$\mathbb{E}\left(\int_0^T A_s^n dA_s\right) \overset{(*)}{\le} c\mathbb{E}(A_T).$$

But $A^n \nearrow A$, so $\int_0^T A_s^n dA_s \nearrow \int_0^T A_s dA_s$ and therefore

$$\mathbb{E}(A_T^2) \le 2\mathbb{E}\left(\int_0^T A_s dA_s\right) \overset{\text{mon. conv}}{\le} 2c\mathbb{E}(A_T) < \infty.$$

So $A_T \in L^2$ and $M$ is square integrable. Moreover, we can compute

$$\mathbb{E}(A_T) = \mathbb{E}(M_T - \underbrace{Z_T}_{\ge 0}) \le \mathbb{E}(M_T) = \mathbb{E}(M_0) = \mathbb{E}(Z_0) \le \mathbb{E}(L_0) = \mathbb{E}(L_T),$$

and then

$$\mathbb{E}(M_T^2) \le 2(\mathbb{E}(Z_T^2) + \mathbb{E}(A_T^2)) \le 2(c\mathbb{E}(L_T) + 2c\mathbb{E}(L_T)) = 6c\mathbb{E}(L_T).$$

2) Now take $H \in \mathbf{b}\mathcal{E}$ and w.l.o.g. $\|H\|_\infty \leq 1$. Then use that on $\{L_T^* \leq c\}$, we have $Z = L$ and so

$$\mathbb{P}((H \cdot L)_T^* \geq c) \leq \mathbb{P}(L_T^* > c) + \mathbb{P}((H \cdot Z)_T^* \geq c)$$

and because $L$ is a martingale, we get

$$\mathbb{P}(L_T^* > c) \overset{\text{Doob}}{\leq} \frac{1}{c}\mathbb{E}(L_T).$$

Next, $Z = M - A$ and $\|H\|_\infty \leq 1$ implies $H \cdot Z = H \cdot M - H \cdot A \leq H \cdot M + A$ and the latter is a submartingale. So we get

$$\begin{aligned}\mathbb{P}((H \cdot Z)_T^* \geq c) \leq \mathbb{P}((H \cdot M + A)_T^*)^2 \geq c^2) &\overset{\text{Doob}}{\leq} \frac{1}{c^2}\mathbb{E}((H \cdot M_T + A_T)^2) \\ &\leq \frac{2}{c^2}(\mathbb{E}((H \cdot M_T)^2) + \mathbb{E}(A_T^2)).\end{aligned}$$

Now $M$ is square-integrable, so Itô's isometry gives

$$\mathbb{E}((H \cdot M_T)^2) \leq \|H\|_\infty^2 \| E(M_T^2) \leq \mathbb{E}(M_T^2).$$

Bringing this all together, we find

$$\begin{aligned}\mathbb{P}((H \cdot L)_T^* \geq c) &\leq \frac{1}{c}\mathbb{E}(L_T) + \frac{2}{c^2}(\underbrace{\mathbb{E}(M_T^2)}_{\leq 6c\mathbb{E}(L_T)} + \underbrace{\mathbb{E}(A_T^2)}_{\leq 2c\mathbb{E}(A_T) \leq 2c\mathbb{E}(L_T)}) \\ &\leq \frac{1}{c}\mathbb{E}(L_T)17.\end{aligned}$$

3) Any martingale can be written as difference of two nonnegative martingales (**Krickeberg**). This gives an extra factor 2 and replaces $\mathbb{E}(L_T)$ by $\|L_T\|_{L^1}$. $\square$

## 3.1 Recap

We provide a quick summary of this section. We defined stochastic integrals for the class of integrands $H \in \mathbf{b}\mathcal{E}$. That is, for **any** $\mathbb{R}^d$-valued process $S = (S_t)_{0\leq t\leq T}$ we defined the stochastic integral as the map $I_S : \mathbf{b}\mathcal{E} \to L^0$ by $I_S(H) := H \cdot S_T = \sum_{i=0}^n h_i^{\mathrm{tr}}(S_{\tau_{i+1}} - S_{\tau_i})$. We then impose on $S$ the condition of being a **good integrator**, that is for $(H^n)_{n\in\mathbb{N}}, H \in \mathbf{b}\mathcal{E}$ we want that if $H^n \to H$ uniformly (in $(\omega, t)$), then $I_S(H^n) \to I_S(H)$ in probability. This condition has the (economic) interpretation that small changes in our self financing strategy $\varphi = (v_0, H)$ also only has small impact on the final wealth given by $v_0 + I_S(H)$. It was then our main goal in this section to establish the **Bichteler-Dellacherie theorem** which states that good integrators and semimartingales are the same thing. To prove this important result we proceeded as follows:

- We introduced the terminology of **quasimartingales**, i.e. processes whose variation on average is finite.
- We showed that $S$ bounded good integrators are locally quasimartingales.
  - Proof of the statement required a $L^2$-version of **Komlós theorem**.
- The **Theorem of Rao** establishes that every quasimartingale, can be written as a difference of two RCLL, nonnegative supermartingales.
- The general **Doob-Meyer theorem** shows that every supermartingale $X$ of class $(D)$ can be uniquely decomposed as $X - X_0 = M - A$ where $M$ is a RCLL martingale, null at 0 and $A$ is RCLL, increasing, integrable and predictable, null at 0.
- Theorem 3.3 states that every good integrator is a semimartingale.
  - Since local semimartingales are semimartingales we can work locally.
  - $J^1_\cdot := \sum_{0<s\leq\cdot} \Delta S_s 1_{\{|\Delta S_s|>1\}}$ is good integrator and semimartingale. $S - J^1$ has bounded jumps, hence locally bounded and is good integrator whenever $S$ is.
  - $S - J$ is then locally a quasimartingale and thus by Rao locally the difference of two supermartingales.
  - Every supermartingale is locally of class $(D)$. Hence by Doob-Meyer can be composed locally as difference of martingale and difference of two predictable, increasing processes. Thus $S - J^1$ is locally sum of martingale and a FV process, hence locally a semimartingale. Then so is $S = S - J^1 + J^1$.
- Theorem 3.4 then establishes that every semimartingale is a good integrator.

# 4 General stochastic integration

**Second main goal**: extend stochastic integration from $\mathbf{b}\mathcal{E}$ to a larger class of integrands $H$.

**Setup**: $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}),\ T \in (0, \infty),\ \mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ with usual conditions.

**Recall**:

1. $L^0$ is metric space for convergence in probability, metrized by

$$d_{L^0}(X, Y) = \mathbb{E}(1 \wedge |X - Y|).$$

2. We also use that for $Z \ge 0$

- $\mathbb{P}(Z > \delta) \le \frac{1}{\delta}\mathbb{E}(1 \wedge Z)$ for $0 < \delta \le 1$.
- $\mathbb{P}(Z > m) \le \mathbb{E}(1 \wedge Z)$ for $m \ge 1$.
- $\mathbb{E}(1 \wedge Z) \le \mathbb{P}(Z > c) + c$ (useful for small $c$).

Introduce spaces $\mathbb{L}$ and $\mathbb{D}$ of $\mathbb{R}^d$-valued adapted LCRL and RCLL processes respectively, and define

$$d(X^1, X^2) := \mathbb{E}(1 \wedge (X^1 - X^2)^*_T) = \mathbb{E}(1 \wedge \sup_{0 \le t \le T} |X^1_t - X^2_t|).$$

This $d$ is a metric, it metrices uniform (in $t$) convergence in probability. One can check that:

**Exercise 4.1.** *Both $(\mathbb{L}, d)$ and $(\mathbb{D}, d)$ are complete metric spaces. (Uses the fact that $\mathbb{F}$ is complete).*

Recall the space $\mathbf{b}\mathcal{E}$ of simple predictable processes. Add to this all processes of the form $H = h_0 1_{\{0\}}$ with $h_0 \in L^\infty(\mathcal{F}_0, \mathbb{R}^d)$ and call the resulting space $\mathbf{b}\mathcal{E}_0$ and define for $H$ as above $H \cdot X := h_0^{\mathrm{tr}} X_0$. Then define

$$\begin{aligned} d_E(X^1, X^2) &:= \sup_{\substack{H \in \mathbf{b}\mathcal{E}_0 \\ \|H\|_\infty \le 1}} \mathbb{E}(1 \wedge (H \cdot (X^1 - X^2))_T) \\ &\le d'_E(X^1, X^2) := \sup_{\substack{H \in \mathbf{b}\mathcal{E}_0 \\ \|H\|_\infty \le 1}} \mathbb{E}(1 \wedge (H \cdot (X^1 - X^2)^*_T)) \end{aligned}$$

Both $d_E$ and $d'_E$ are metrics and $d'_E \ge d$. It is sometimes easier to work with $d_E$, and we will now show that $d_E$ still controls $d$ and $d_E, d'_E$ are equivalent.

**Lemma 4.1.** *If $(X^n)_{n\in\mathbb{N}} \subset \mathbb{D}$ satisfies $d_E(X^n, X) \to 0$, then also $d(X^n, X) \to 0$.*

*Proof.* For $Y \in \mathbb{D}$ and $\delta > 0$ set $\tau := \inf\{t \in [0,T] : |Y_t| \geq \delta\} \wedge T$. Note that on $\{\tau < T\}$, we have $|Y_\tau| \geq \delta$ by RC and so $\{Y_T^* \geq \delta\} \subset \{|Y_\tau| \geq \delta\}$. But $H := 1_{[\![0,\tau]\!]} \in \mathbf{b}\mathcal{E}_0$ has $\|H\|_\infty \leq 1$ and $H \cdot Y_T = Y_\tau$. Thus we get

$$\mathbb{P}(Y_T^* \geq \delta) \leq \mathbb{P}(|Y_\tau| \geq \delta) \leq \frac{1}{\delta}\mathbb{E}(1 \wedge |Y_\tau|) = \frac{1}{\delta}\mathbb{E}(1 \wedge |H \cdot Y_T|) \leq \frac{1}{\delta} d_E(Y, 0).$$

Apply this to $Y := X^n - X$ and use that

$$d(Y, 0) \leq \mathbb{P}(Y_T^* \geq \delta) + \delta$$

this gives the result. □

Next auxiliary result is used to connect $d_E$ and $d'_E$:

**Lemma 4.2.** *For $(X^n)_{n\in\mathbb{N}} \subset \mathbb{D}$, $X \in \mathbb{D}$, we have $d'_E(X^n, X) \to 0$ if and only if $(H^n \cdot (X^n - X))_T \to 0$ in $L^0$ (i.e. in Probability) for every sequence $(H^n)_{n\in\mathbb{N}} \subset \boldsymbol{b}\mathcal{E}_0$, with $\|H^n\|_\infty \leq 1$.*

*Proof.* " $\Longrightarrow$ " Take $(H^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{E}_0$, set $Y := X^n - X$ and write

$$\mathbb{P}(|H^n \cdot Y_T| \geq \delta) \leq \frac{1}{\delta}\mathbb{E}(1 \wedge |H^n \cdot Y_T|) \leq \frac{1}{\delta} d_E(Y, 0) \leq \frac{1}{\delta} d'_E(Y, 0).$$

"$\Longleftarrow$" If $d'_E(X^n, X) \not\to 0$, then we can find $\delta_0 > 0$ and a sequence $(H^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{E}_0$, with $\|H^n\|_\infty \leq 1$ such that

$$\mathbb{P}((H^n \cdot (X^n - X))_T^* \geq \delta_0) \geq \delta_0, \text{ for all } n \in \mathbb{N}.$$

Define $\tau_n := \inf\{t \in [0,T] : |H^n \cdot (X^n - X)_t| > \delta_0\} \wedge T$, and $\tilde{H}^n := H^n 1_{[\![0,\tau_n]\!]}$. Then $(\tilde{H}^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{E}_0$ with $\|\tilde{H}^n\|_\infty \leq 1$ and $\tilde{H}^n \cdot (X^n - X)_T \not\to 0$ in $L^0$ by construction. □

**Corollary 4.1.** *For $(X^n)_{n\in\mathbb{N}} \subset \mathbb{D}$, $X \in \mathbb{D}$, we have $d'_E(X^n, X) \to 0$ if and only if $d_E(X^n, X) \to 0$.*

*Proof.* " $\Longrightarrow$ " This is clear from $d'_E \geq d_E$.
"$\Longleftarrow$" As in the proof of Lemma 4.2, we have for $\delta \in (0, 1]$ and any $H^n \in \mathbf{b}\mathcal{E}_0$ with $\|H^n\|_\infty \leq 1$ that

$$\mathbb{P}(|H^n \cdot Y_T| \geq \delta) \leq \frac{1}{\delta} d_E(Y, 0).$$

Using this for $Y := X^n - X$ shows that $d_E(X^n, X) \to 0$ implies $(H^n \cdot (X^n - X)_T)_{n\in\mathbb{N}} \to 0$ in $L^0$ for every sequence $(H^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{E}_0$ with $\|H^n\|_\infty \leq 1$. But then by Lemma 4.2 this implies that $d'_E(X^n, X) \to 0$. □

Denote by $\mathcal{S} \subset \mathbb{D}$ the space of $\mathbb{R}^d$-valued semimartingales.

**Theorem 4.1.** *With each of metric $d_E, d'_E$, $\mathcal{S}$ is complete (and a topological vector space).*

*Proof.* $\mathcal{S}$ equipped with $d_E$ or $d'_E$ is a metric space (and a topological vector space). If $(X^n)_{n\in\mathbb{N}} \subset \mathcal{S} \subset \mathbb{D}$ is Cauchy for $d_E$ or $d'_E \geq d_E$, then $(X^n)_{n\in\mathbb{N}} \subset \mathbb{D}$ is also Cauchy for $d$ by Lemma 4.1. Because $(\mathbb{D}, d)$ is complete, there exists $X \in \mathbb{D}$ with $d(X^n, X) \to 0$.

It remains to show that $X \in \mathcal{S}$ and $X^n \to X$ for $d_E$ or $d'_E$ respectively, as these two are equivalent by Corollary 4.1.

First we argue $d_E(X^m, X) \to 0$ as $m \to \infty$. Since $d(X^n, X) \to 0$ as $n \to \infty$ implies $X^n_\tau \to X_\tau$ in $L^0$ for any stopping time $\tau$ and hence $H \cdot X^n_T \to H \cdot X_T$ in $L^0$ for any $H \in \mathbf{b}\mathcal{E}_0$ with $\|H\|_\infty \leq 1$. So for $H \in \mathbf{b}\mathcal{E}_0$ with $\|H\|_\infty \leq 1$ we have

$$\begin{aligned}\mathbb{E}(1 \wedge |H \cdot (X^m - X)_T|) &= \lim_{n\to\infty} \mathbb{E}(1 \wedge |H \cdot (X^m - X^n)_T|)\\ &\leq \limsup_{n\to\infty} d_E(X^m, X^n) \xrightarrow{m\to\infty} 0,\end{aligned}$$

because $(X^n)_{n\in\mathbb{N}}$ is $d_E$-Cauchy. Taking the sup over all $H \in \mathbf{b}\mathcal{E}_0$ with $\|H\|_\infty \leq 1$ we get

$$d_E(X^m, X) \leq \limsup_{n\to\infty} d_E(X^m, X^n) \to 0 \text{ as } m \to \infty.$$

To show that even $X \in \mathcal{S}$, we know that it is equivalent to show that $X$ is a good integrator, i.e. $I_X(\mathbf{b}\mathcal{E}, \|\cdot\|_\infty) \to L^0$ is continuous. This is equivalent to show that

**Exercise 4.2.** *$X$ is a good integrator if and only if $\mathcal{X}_{(1)} := \{H \cdot X_T = I_X(H),\ H \in \boldsymbol{b}\mathcal{E}, \|H\|_\infty \leq 1\}$ is bounded in $L^0$, i.e.*

$$\lim_{m\to\infty} \sup_{Y\in\mathcal{X}_{(1)}} \mathbb{P}(|Y| \geq m) = 0$$

To prove that this indeed holds, write

$$\begin{aligned}\mathbb{P}(|I_X(H)| \geq 2m) &\leq \mathbb{P}(|I_{X^n}(H)| \geq m) + \mathbb{P}(|(H \cdot (X - X^n))_T| \geq m)\\ &\leq \mathbb{P}(|I_{X^n}(H)| \geq m) + d_E(X^n, X).\end{aligned}$$

The second term is smaller than $\epsilon$ for large enough $n$ because $X^n \to X$ w.r.t. $d_E$. Moreover, for fixed $n \in \mathbb{N}$, $X^n$ is a semimartingale, so that $\mathcal{X}^n_{(1)}$ is bounded in $L^0$.

So the first term is also smaller than $\epsilon$ for $m$ large, uniformly over $H \in \mathbf{b}\mathcal{E}$ with $\|H\|_\infty \leq 1$. So we get

$$\sup_{\substack{H\in\mathbf{b}\mathcal{E}\\ \|H\|_\infty\leq 1}} \mathbb{P}(|I_X(H)| \geq 2m) \xrightarrow{m\to\infty} 0.$$

Thus $\mathcal{X}_{(1)}$ is bounded in $L^0$ (bounded in Probability), hence $X \in \mathcal{S}$ and the proof is complete. $\square$

We already know that $(\mathbb{L}, d)$ is complete.

**Exercise 4.3.** *Any $H \in \mathbb{L}$ is adapted LC (by def. of $\mathbb{L}$), hence locally bounded*

Moreover, any bounded adapted process can be approximated uniformly in $(\omega, t)$ by processes in $\mathbf{b}\mathcal{E}_0$ (If $H_0 = 0$ for $H$ bounded adapted LC, the approximands are in $\mathbf{b}\mathcal{E}$; this is w.l.o.g. if $H$ is to be an integrand). For details to this statement compare with model solution of **Exercise** 4.1.

So $\mathbf{b}\mathcal{E}_0$ is dense in $\mathbb{L}$ for $d$, and this allows us for each semimartingale $S$ to extend stochastic integration $I_S$ from $(\mathbf{b}\mathcal{E}, \|\cdot\|_\infty)$ to $(\mathbb{L}, d)$ and resulting mapping $I_S : \mathbb{L} \to L^0$, $H \mapsto I_S(H) = H \cdot S_T$ is continuous with respect to $d$, i.e. $d(H^n, H) \to 0$ with $H^n, H \in \mathbb{L}$ implies that

$$I_S(H^n) = H^n \cdot S_T \to H \cdot S_T = I_S(H) \text{ in probability.}$$

We can define a stochastic integral process via

$$H \cdot S_t := I_S(H 1_{[\![0,t]\!]}) \text{ for } H \in \mathbb{L}.$$

This however does not give any path properties of $t \mapsto H \cdot S_t$. Our next result gives much more:

**Theorem 4.2.** *For every semimartingale $S$, the mapping $I_S : \mathbf{b}\mathcal{E} \to L^0$ can be extended to a continuous mapping $J_S : (\mathbb{L}, d) \to (\mathcal{S}, d'_E)$.*

**Remark 4.1.** Two statements:

- For $H \in \mathbb{L}$, stochastic integral process $H \cdot S := J_S(H)$ is well defined and a semimartingale (in particular adapted RCLL, hence we get path properties from the above theorem).
- By the continuity of $J_S$, we have if $(H^n) \subset \mathbb{L}, H \in \mathbb{L}$ satisfy $d(H^n, H) \to 0$ then $H^n \cdot S \to H \cdot S$ for $d'_E$.

*Proof.* As $\mathbf{b}\mathcal{E}_0$ is dense in $\mathbb{L}$ for $d$, any $H \in \mathbb{L}$ admits $(H^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{E}_0$ with $d(H^n, H) \to 0$. Then $(H^n)_{n\in\mathbb{N}}$ is Cauchy for $d$ (as its convergent w.r.t. $d$), and $X^n := H^n \cdot S = J_S(H^n)$ is in $\mathcal{S}$. If we show that $(X^n)_{n\in\mathbb{N}}$ is Cauchy for $d'_E$, then Theorem 4.1 implies that there exists $X \in \mathcal{S}$ with $d'_E(X^n, X) \to 0$ and then we define $J_S(H) := X =: H \cdot S$. That this $J_S$ is continuous for $d$ and $d'_E$ follows by construction.

So, it only remains to show that $X^n = H^n \cdot S$, $n \in \mathbb{N}$ is Cauchy for $d'_E$.

If $S = A$ is of FV, then this follows from classical results (pathwise) for Lebesgue-Stieltjes integrals. So let us suppose that $S = M \in \mathcal{M}_{0,\mathrm{loc}}$.

**Exercise 4.4.** *By localisation, we can assume w.l.o.g. that $M \in \mathcal{M}_{0,loc}$ is a (true) martingale.*

It is enough to show that $(H^k) \subset \mathbf{b}\mathcal{E}$ with $d(H^k, 0) \to 0$ implies that $d'_E(H^k \cdot M, 0) \to 0$. Take $K \in \mathbf{b}\mathcal{E}_0$ with $\|K\|_\infty \le 1$ and write

$$H^k = H^k 1_{\{(H^k)^* \le b\}} + H^k 1_{\{(H^k)^* > b\}} =: H'_k + H''_k \in \mathbf{b}\mathcal{E}_0.$$

Then look at $K \cdot (H^k \cdot M) = (KH^k) \cdot M$ and compute

$$\begin{aligned}\mathbb{P}((K \cdot (H^k \cdot M))^*_T \ge c) &\le \mathbb{P}((H^k)^*_T > b) + \mathbb{P}((KH'_k) \cdot M)^*_T \ge c) \\ &\overset{L3.2}{\le} \mathbb{P}((H^k)^*_T > b) + \frac{34}{c}\|H'_k\|_\infty \|M_T\|_{L^1}.\end{aligned}$$

So we get

$$\begin{aligned}d'_E(H^k \cdot M, 0) &= \sup_{\substack{K\in\mathbf{b}\mathcal{E}_0\\ \|K\|_\infty\le 1}} \mathbb{E}(1 \wedge (K \cdot (H^k \cdot M))^*_T) \\ &\le c + \mathbb{P}((H^k)^*_T > b) + \frac{34b}{c}\|M_T\|_{L^1}.\end{aligned}$$

Due to $d(H^k, 0) \to 0$, the second term 2 goes to 0 as $k \to \infty$ for any $b$. The terms 1 and 3 can be made arbitrary small by first making $b$, and then $c$ sufficiently small. So $d'_E(H^k \cdot M, 0) \to 0$. $\square$

To extend class of integrands beyond $\mathbb{L}$, we need a different approximation concept and hence a new idea. One concept needed is known from BMSC, it is redone here for the sake of completeness.

For a semimartingale $S$, process $S_-$ is adapted LC (hence in $\mathbb{L}$) and so $\int S_- dS = S_- \cdot S$ is well defined. We can approximate $S_-$ in $(\mathbb{L}, d)$ by

$$\sum_{\tau_i} 1_{(\!(\tau_i, \tau_{i+1}]\!]} \text{ for } 0 \leq \tau_0 \leq \cdots \leq \tau_{I^n} \leq T$$

to obtain

$$\int S_- dS = \lim_{n\to\infty} \sum_{i=0}^{I^n - 1} S_{\tau_i}(S^{\tau_{i+1}} - S^{\tau_i}) \qquad \text{(for } d'_E\text{)}$$

and in particular, this convergence holds uniformly over $t$, in probability (often abbreviated ucp).

**Definition 4.1.** *For semimartingale $S$, define* ***optional quadratic variation***

$$[S] := S^2 - S_0^2 - 2\int S_- dS = S^2 - S_0^2 - 2J_S(S_-).$$

**Remark 4.2.** By definition, $[S]$ is adapted, RCLL, null at 0.

**Lemma 4.3.** *$[S]$ is $\mathbb{P}$-a.s. increasing in $t$.*

*Proof.* Take dyadic partition $D_n = 2^{-n}T\mathbb{N}_0 \cap [0, T]$ and write

$$\int S_- dS = \lim_{n\to\infty} \sum_{t_i \in D_n} S_{t_i}(S^{t_{i+1}} - S^{t_i}).$$

Compute:

$$\begin{aligned} S^2 - S_0^2 - 2\sum S_{t_i}(S^{t_{i+1}} - S^{t_i}) &\overset{\text{tel. sum}}{=} \sum \left[((S^{t_{i+1}})^2 - (S^{t_i})^2) - 2S_{t_i}(S^{t_{i+1}} - S^{t_i})\right] \\ &= \sum_{t_i \in D_n} (S^{t_{i+1}} - S^{t_i})^2 =: V^n \end{aligned}$$

and conclude that $[S] = \lim_{n\to\infty} V^n$ ucp. **But** each $V^n$ is increasing on $D_n$, and so the limit is increasing on $D := \cup_{n\in\mathbb{N}} D_n$ because $D_n \subset D_{n+1}$. But $D$ is dense in $[0, T]$, so $[S]$ is increasing on $D$ and $[S]$ is RC(LL), hence $[S]$ is increasing on $[0, T]$ $\mathbb{P}$-a.s. $\square$

**Definition 4.2.** $\mathcal{H}_0^1 := \{M \in \mathcal{M}_{0,loc} : M_T^* = \sup_{0 \le s \le T} |M_s| \in L^1\}$. *If we identity indistinguishable processes, then* $\mathcal{H}_0^1$ *is a Banach space with* $\|M\|_{\mathcal{H}_0^1} := \|M_T^*\|_{L^1}$.

**Remark 4.3.** Clearly, each $M \in \mathcal{H}_0^1$ is a (true) martingale and UI (even of class (D)). See model solutions to the next Exercise if these statements are not clear.

**Exercise 4.5** (Important). $\mathcal{M}_{0,loc} = \mathcal{H}_{0,loc}^1$ *i.e. every local martingale null at* 0 *is locally in* $\mathcal{H}_0^1$.

The next result is crucial ingredient for the construction of stochastic integrals with respect to local martingales.

**Theorem 4.3** (Davis inequality). *There exists constants* $0 < c < C < \infty$ *such that for all* $M \in \mathcal{M}_{0,loc}$

$$c\mathbb{E}(([M]_T)^{1/2}) \le \mathbb{E}(M_T^*) \le C\mathbb{E}(([M]_T)^{1/2}).$$

**Remark 4.4.** Davis inequality gives in particular that the maximal and the quadratic norms on $\mathcal{H}_0^1$ are equivalent.

*Proof.* We can discuss a proof in two stages:
1) Show general results from result for discrete time martingale.
2) Elementary argument for discrete time.

2) (See Beiglböck/Siorpaes (2015)). Define

$$x_n^* := \max_{k=0,1,\ldots,n} |x_k|,$$
$$[x]_n := x_0^2 + \sum_{k=1}^{n} (x_k - x_{k-1})^2,$$
$$(h \cdot x)_n := \sum_{k=1}^{n} h_{k-1}(x_k - x_{k-1}).$$

Then prove that for the special choice $h_k := x_k / \sqrt{[x]_k + (x_k^*)^2}$ (bounded by one) we have:

$$([x]_N)^{1/2} \le 3x_N^* - (h \cdot x)_N,$$
$$x_N^* \le 6([x]_N)^{1/2} + 2(h \cdot x)_N.$$

(The statement is elementary analysis, but the proof needs work). Then apply this result to discrete-time martingale $M$ with $x_k := M_{t_k}(\omega)$, then $H_{t_k}(\omega) := h_{k-1}$ is bounded, predictable, and so stochastic integral $(H \cdot M)_T(\omega) = (h \cdot x)_N$ has expectation 0.

1) First, both $M^*$ and $[M]$ are increasing, so we can consider the stopped martingale process $M^{\tau_n}$ (later pass to the limit by monotone convergence).

Next, $\Delta[M] = (\Delta M)^2$ shows that $([M])^{1/2}$ is locally integrable iff $M$ is locally integrable (or $M^*$), and so, w.l.o.g., we can assume that all terms in the inequality of Theorem 4.3 are finite (after localisation). Now $M$, after localisation, is a martingale in $\mathcal{H}_0^1$ and

$$[M] = \lim_{n\to\infty} Y^n = \lim_{n\to\infty} \sum_i (M^{t_{i+1}} - M^{t_i})^2 \text{ ucp.}$$

hence, along a subsequence $\mathbb{P}$-a.s. uniformly over $[0,T]$. So $Y := \sup_{n\in\mathbb{N}} Y^n$ is RCLL with jumps

$$|\Delta Y_t| \le \sup_{n\in\mathbb{N}} |\Delta Y_t^n| \le 4(M_T^*)^2$$

and so $\sup_{0\le t\le T} |\Delta Y_t^{1/2}| \in L^1$. Use discrete time result along $D_n$ to get

$$\begin{tikzcd}[column sep=small] c\mathbb{E}((Y_T^n)^{1/2}) \arrow[d, "\text{dom.}"', "\text{conv.}"] & \le & \mathbb{E}(\sup\limits_{t_i\in D_n} |M_{t_i}|) \arrow[d, "\text{mon.}"', "\text{conv.}"] & \le & C\mathbb{E}((Y_T^n)^{1/2}) \arrow[d, "\text{dom.}"', "\text{conv.}"] \\ c\mathbb{E}([M]_T^{1/2}) & & \mathbb{E}(M_T^*) & & C\mathbb{E}([M]_T^{1/2}) \end{tikzcd}$$

□

Because $\mathcal{M}_{0,\text{loc}} = \mathcal{H}^1_{0,\text{loc}}$, we get as a direct consequence of Davis inequality

**Corollary 4.2.** *If $M \in \mathcal{M}_{0,loc}$, then $[M]^{1/2}$ is locally integrable.*

**Definition 4.3.** *For $M \in \mathcal{M}_{0,loc}$, define $\mathcal{L}^1(M)$ to be the space consisting of all predictable $H$ such that*

$$\|H\|_{L^1(M)} := \mathbb{E}\left[\left(\int_0^T H_s^2 d[M]_s\right)^{1/2}\right] < \infty.$$

*Identifying $H, H'$ if $\|H - H'\|_{L^1(M)} = 0$, yields space $L^1(M)$.*

**Remark 4.5.**

1. Above definition is for $d = 1$. For $d > 1$, $[M]$ is process with values in positive semidefinite $d \times d$-matrices, and

$$\|H\|_{L^1} = \mathbb{E}\left[\left(\int_0^T H_s^{\mathrm{tr}} d[M]_s H_s\right)^{1/2}\right].$$

2. If $M \in \mathcal{H}_0^1$, then $[M]^{1/2}$ is integrable and then $L^1(M) \supset \mathbf{b}\mathcal{P} :=$ {space of bounded predictable processes}.

3. If $A$ is of FV, say $|A|$, define $\mathrm{L}_{\mathrm{var}}(A)$ as space of predictable $H$ with $\int_0^T |H_s| d|A|_s < \infty$ $\mathbb{P}$-a.s. Notice that $\mathrm{L}_{\mathrm{var}}(A) \supset \mathbf{b}\mathcal{P}_{\mathrm{loc}}$.

**Lemma 4.4.** *If $M \in \mathcal{M}_{0,loc}$, then $\boldsymbol{b}\mathcal{E} \cap L^1(M)$ is dense in $L^1(M)$ for $\|\cdot\|_{L^1(M)}$.*

*Proof.* By Corollary 4.2 we can take $\tau_m \nearrow T$ stationarily such that $([M]^{\tau_m})^{1/2} = ([M^{\tau_m}])^{1/2}$ is integrable. If $H \in L^1(M)$, then

$$H^m := H 1_{[\![0,\tau_m]\!]} \to H \text{ in } L^1(M)$$

and $H^m \in L^1(M^{\tau_m})$. So w.l.o.g. assume $M \in \mathcal{H}_0^1$ and $H \in L^1(M)$. Then $H^k := H 1_{\{|H| \leq k\}} \in \mathbf{b}\mathcal{P}$ and $H^k \to H$ in $L^1(M)$ (by dominated convergence).

Now suppose $M \in \mathcal{H}_0^1$ and $H \in \mathbf{b}\mathcal{P} \subset L^1(M)$. Then we can find $H^j = \sum_k \lambda_k^j 1_{D_j^k}$ with $\lambda_k^j \in \mathbb{R}$, $D_j^k \in \mathcal{P}$ such that $\|H^j - H\|_\infty \to 0$ (and hence also in $L^1(M)$). Note however, that $H^j \notin \mathbf{b}\mathcal{E}$ because $D_j^k \in \mathcal{P}$ does not need to be rectangular.

But any $H$ of the form $H\lambda 1_D$ for $\lambda \in \mathbb{R}, D \in \mathcal{P}$ can be approximated in $L^1(M)$ by a sequence $(H^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{E}$ with $\|H^n\|_\infty \leq \|H\|_\infty = |\lambda|$.

**Exercise 4.6.** *Show that the claim in the previous phrase holds true by using a monotone class argument.*

Putting all this together gives the result. □

We could now use Lemma 4.4 and completeness of $\mathcal{H}_0^1$ to extend stochastic integrals for a local martingale, to $H \in L^1_{\text{loc}}(M)$ and in particular to $\mathbf{b}\mathcal{P}_{\text{loc}}$. We want, however, a stronger continuity property and hence proceed differently.

**Lemma 4.5.** *Let $(M^n)_{n\in\mathbb{N}}$, $Y$ be martingales with $|\Delta M^n_\tau| \leq |\Delta Y_\tau|$ for all $n \in \mathbb{N}$ and for all stopping times $\tau$. If $\mathbb{E}(1 \wedge [M^n]_T) \to 0$ i.e. $[M^n]_T \to 0$ in $L^0$, then $d'_E(M^n, 0) \to 0$.*

*Proof.* Take $(H^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{E}_0$ with $\|H^n\|_\infty \leq 1$ and set $X^n := H^n \cdot M^n$. Then $[X^n]_T \leq [M^n]_T \to 0$ in $L^0$, so that along a subsequence:

$$\mathbb{P}([X^{n_k}]_T \geq 2^{-k}) \leq 2^{-k}, \text{ for all } k \in \mathbb{N}.$$

By Borel-Cantelli $A := \sum_{k=1}^\infty [X^{n_k}]$ is $\mathbb{P}$-a.s. finite-valued and $[X^{n_k}]_T \to 0$ $\mathbb{P}$-a.s. for $k \to \infty$. Let $\tau_m := \inf\{t \in [0,T] : A_t \geq m \text{ or } |Y_t| \geq m\} \wedge T$ to get for all $k$

$$[X^{n_k}]_{\tau_m} \leq A_{\tau_m -} + (\Delta X^{n_k}_{\tau_m})^2 \leq m + (\Delta Y_{\tau_m})^2 \leq m + ((m+1)Y_{\tau_m}))^2.$$

So we obtain $\sup_{k\in\mathbb{N}}([X^{n_k}]_{\tau_m})^{1/2} \in L^1$. Use now Theorem 4.3 and Lebesgue to get

$$\mathbb{E}((X^{n_k})^*_{\tau_m}) \leq C\mathbb{E}([X^{n_k}]^{1/2}_{\tau_m}) \to 0 \text{ as } k \to \infty$$

and of course $\mathbb{P}(\tau_m = T) \to 1$ as $m \to \infty$. This implies that $(X^{n_k})^*_T = (H^{n_k} \cdot M^{n_k})^*_T \to 0$ in $L^0$ as $k \to \infty$. So every subsequence of $(H^n \cdot M^n)_{n\in\mathbb{N}}$ has a further subsequence which converges, uniformly on $[0,T]$, to 0 in $L^0$. Thus the original sequence $(H^n \cdot M^n)_{n\in\mathbb{N}}$ already satisfies this convergence, and so we get

$$\sup_{\substack{H\in\mathbf{b}\mathcal{E}_0 \\ \|H\|_\infty \leq 1}} \mathbb{P}((H \cdot M^n)^*_T \geq \epsilon) \xrightarrow{n\to\infty} 0, \text{ for all } \epsilon > 0.$$

(If this fails, find sequence $(H^n)_{n\in\mathbb{N}}$ violating the above). We conclude that $d'_E(M^n, 0) \to 0$ as $n \to \infty$. $\square$

**Corollary 4.3.** *Let $M$ be a martingale and $(H^n)_{n\in\mathbb{N}} \subset \boldsymbol{b\mathcal{E}}$ with $\|H^n\|_\infty \leq 1$. If $H^n \to 0$ pointwise, then $d'_E(H^n \cdot M, 0) \to 0$.*

**Remark 4.6.** The Corollary above gives a "nice and strong" convergence with respect to a "rich" topology.

*Proof.* Let $M^n := H^n \cdot M$ (which is a martingale) and $Y := M$ to get $|\Delta M^n_\tau| = |H^n_\tau \Delta M_\tau| \leq |\Delta Y_\tau|$ for all $n \in \mathbb{N}$ and for all stopping times $\tau$. Moreover,

$$[M^n]_T = \underbrace{\int_0^T \underbrace{(H^n_s)^2}_{\to 0, \text{ bdd. by } 1} \; d[M]_s}_{\text{bdd. by } [M]_T < \infty \ \mathbb{P}\text{-a.s.}} \to 0 \ \mathbb{P}\text{-a.s.}$$

hence also in $L^0$. So Lemma 4.5 gives the desired conclusion. $\square$

**Theorem 4.4.** *For any martingale $M \in \mathcal{H}_0^1$, the mapping $I_M(\boldsymbol{b\mathcal{E}}, \|\cdot\|_\infty) \to (\mathcal{S}, d'_E)$ admits a unique linear extension to $J_M : \boldsymbol{b\mathcal{P}} \to \mathcal{S}$ such that $H \cdot M := J_M(H)$ satisfies $[H \cdot M] = \int H^2 d[M]$ and $\Delta(H \cdot M) = H\Delta M$ and with dominated convergence property, i.e. that $(H^k) \subset \boldsymbol{b\mathcal{P}}$ with $\|H^k\|_\infty \leq 1$ and $H^k \to 0$ pointwise implies $d'_E(H^k \cdot M, 0) \to 0$. Moreover, $J_M : (\boldsymbol{b\mathcal{P}}, \|\cdot\|_\infty) \to (\mathcal{S}, d'_E)$ is continuous.*

*Proof.* Take $H \in \mathbf{b}\mathcal{P} \subset L^1(M)$ and use Lemma 4.4 to get $(H^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{E}$ with $\|H^n\|_\infty \leq \|H\|_\infty$ and $H^n \to H$ in $L^1(M)$-a.e. Then $H^n - H^m \to 0$ as $m, n \to \infty$ and hence $H^n - H^{m_n} \to 0$ as $n \to \infty$ for any $m_n \geq n$. So $d'_E((H^n - H^{m_n})\cdot M, 0) \to 0$ as by Corollary 4.3 and so $d'_E(H^n \cdot M, H^m \cdot M) \to 0$ as $m, n \to \infty$. So $(H^n \cdot M)_{n\in\mathbb{N}}$ is Cauchy in $\mathcal{S}$ for $d'_E$, and so by Theorem 4.1 there exists limit in $\mathcal{S}$ for $d'_E$ and we define $H \cdot M := d'_E \lim_{n\to\infty} H^n \cdot M$.

Linearity, uniqueness are clear, and

$$[H^n \cdot M] = \int (H^n) d[M],\ \Delta(H^n \cdot M) = H^n \Delta M,\ \text{for all } n \in \mathbb{N}$$

are easy because $H^n \in \mathbf{b}\mathcal{E}$. The result for $H$ then follows via limits. Dominated convergence property follows from Corollary 4.3 (plus its proof) for $M^k := H^k \cdot M$ and continuity for $(\mathbf{b}\mathcal{P}, \|\cdot\|_\infty)$ is then clear. $\square$

Next we extend from $\mathcal{H}_0^1$-integrators to semimartingales $S$. If $S = A$ is of FV, we can define $H \cdot A = \int H_s dA_s$, for $H \in \mathbf{b}\mathcal{P} \subset L_{\text{var}}(A)$ pathwise as usual Lebesgue-Stieltjes integral and check properties from known results.

So we only need to argue for $S = M \in \mathcal{M}_{0,\text{loc}} = \mathcal{H}^1_{0,\text{loc}}$ and this is easy - use the following exercise:

**Exercise 4.7.** *$M^{\tau_m} \in \mathcal{H}_0^1$ for $\tau_m \nearrow T$ stationarily.*

One more localisation allows us to extend from $\mathbf{b}\mathcal{P}$ to $\mathbf{b}\mathcal{P}_{\text{loc}}$, and so $H \cdot S = \int H dS$ is well-defined and a semimartingale for any $S \in \mathcal{S}$ and any $H \in \mathbf{b}\mathcal{P}_{\text{loc}}$.

Now define on $\mathcal{S}$ two metrics

$$\begin{aligned}
\widetilde{d}_E(S^1, S^2) &:= \sup_{\substack{H\in\mathbf{b}\mathcal{P}\\ \|H\|_\infty\leq 1}} \mathbb{E}(1 \wedge |H \cdot (S^1 - S^2)|_T) \geq d_E(S^1, S^2),\\
\widetilde{d}'_E(S^1, S^2) &:= \sup_{\substack{H\in\mathbf{b}\mathcal{P}\\ \|H\|_\infty\leq 1}} \mathbb{E}(1 \wedge (H \cdot (S^1 - S^2))^*_T)\\
&\geq \max(d'_E(S^1, S`2), \widetilde{d}_E(S^1, S^2))
\end{aligned}$$

Note that $\mathbf{b}\mathcal{P} \supset \mathbf{b}\mathcal{E}_0$.

One can argue as for Corollary 4.1 that $\widetilde{d}_E$ and $\widetilde{d}'_E$ are equivalent. Moreover, they are also equivalent to $d_E, d'_E$ (note that we already know that $d_E$ and $d'_E$ are equivalent). Indeed, $d'_E \leq \widetilde{d}'_E$ and for any $H \in \mathbf{b}\mathcal{P}$ with $\|H\|_\infty \leq 1$, take $(H^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{E}$ with $\|H^n\|_\infty \leq 1$ and $H^n \to H$ a.e. Then write

$$\begin{aligned} \mathbb{E}(1 \wedge |(H \cdot (S^1 - S^2))_T|) &= d_{L^0}(H \cdot S^1_T, H \cdot S^2_T) \\ \leq d_{L^0}(H \cdot S^1_T, H^n \cdot S^1_T) + d_{L^0}(H^n \cdot (S^1_T - S^2_T), 0) &+ d_{L^0}(H^n \cdot S^1_T, H \cdot S^2_T) \\ \leq d'_E((H - H^n) \cdot S^1_T) + d'_E(S^1, S^2) &+ d'_E((H - H^n) \cdot S^2). \end{aligned}$$

Terms 1 and 3 converge to 0 as $n \to \infty$ by the dominated convergence property (of Theorem 4.4) and so taking sup over all $H \in \mathbf{b}\mathcal{P}$, $\|H\|_\infty \leq 1$ gives $\widetilde{d}_E \leq d'_E$. This is enough.

**Terminology**: $\widetilde{d}'_E$ ("strongest") is **Emery metric** on $\mathcal{S}$ and corresponding (metric) topology is **Emery topology**.

We summarize our knowledge in the next theorem:

**Theorem 4.5.** *For every semimartingale $S$ and every $H \in \boldsymbol{b}\mathcal{P}$ stochastic integral $H \cdot S = \int H dS$ is well-defined and a semimartingale. Moreover the mapping*

$$\begin{cases} (\boldsymbol{b}\mathcal{P} \times \mathcal{S}, \|\cdot\|_\infty \times \widetilde{d}'_E) & \longrightarrow (\mathcal{S}, \widetilde{d}'_E) \\ (H, S) & \longmapsto H \cdot S \end{cases}$$

*is continuous.*

*Proof.* Take $(H^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{P}$ with $\|H^n - H\|_\infty \to 0$ and $(S^n) \subset \mathcal{S}$ with $\widetilde{d}'_E(S^n, S) \to 0$. Then

$$\begin{aligned} \widetilde{d}'_E(H^n \cdot S^n, H \cdot S) &\leq \widetilde{d}'_E(H^n \cdot S^n, H^n \cdot S) + \widetilde{d}'_E(H^n \cdot S, H \cdot S) \\ &= \sup_{\substack{K \in \mathbf{b}\mathcal{P} \\ \|K\|_\infty \leq 1}} \mathbb{E}(1 \wedge ((KH^n) \cdot (S^n - S))^*_T) + \widetilde{d}'_E((H^n - H) \cdot S, 0). \end{aligned}$$

The second term vanishes by dominated convergence property, and the first term goes to 0 by definition of $\widetilde{d}'_E$ because $d'_E(S^n - S, 0) \to 0$ as $n \to \infty$ (uses $\|H^n\|_\infty \leq 1$). ☐

In Theorem 4.4 we have $H \cdot M$ for $M \in \mathcal{H}_0^1$ and $H \in \mathbf{b}\mathcal{P}$. **But**: construction as limit in $\mathcal{S}$, so that

$$H \cdot M = \int H dM \text{ is in } \mathcal{S}.$$

Is $H \cdot M$ a martingale? More subtle than it looks: If $M$ is a martingale and $H$ is bounded, predictable, then $H \cdot M$ is a semimartingale - but it can fail to be a martingale **(unlike in discrete time! striking difference here!)**. Counterexamples in Herdegem/Herrimann, in continuous time. Nevertheless, we have:

**Proposition 4.1.** *If $M \in \mathcal{H}_0^1$ and $H \in \boldsymbol{b}\mathcal{P}$, then $H \cdot M \in \mathcal{H}_0^1$.*

*Proof.* As in Lemma 4.4, take $(H^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{P}$ with $\|H^n\|_\infty \leq \|H\|_\infty$ and $H^n \to H$ $L^1(M)$ a.e. Then by Theorem 4.4, $H^n \cdot M \to H \cdot M$ for $d'_E$. But better: by Theorem 4.3 (Davis inequality), $[M]_T^{1/2} \in L^1$ and so by Lebesgue dominated convergence

$$\mathbb{E}\left[\left(\int_0^T (H_s^n - H_s)^2 d[M]_s\right)^{1/2}\right] \xrightarrow{n\to\infty} 0.$$

Again by Theorem 4.3 $(H^n \cdot M)_{n\in\mathbb{N}}$ is Cauchy for $\|\cdot\|_{\mathcal{H}_0^1}$, and so the limit is again $\mathcal{H}_0^1$. $\square$

Now comes the final extension of integrands (most general case).

**Definition 4.4.** *Fix a semimartingale $S$. A predictable process $H$ is $S$-**integrable**, denoted as $H \in \mathcal{L}(S)$, if $(H^n \cdot S)_{n\in\mathbb{N}}$ is Cauchy in Emery topology (or for $d'_E$), where $H^n := H1_{\{|H|\leq n\}} \in \boldsymbol{b}\mathcal{P}$. Limit in $\mathcal{S}$ is called $H \cdot S$.*

**Remark 4.7.** If $H, H'$ are in $\mathcal{L}(S)$ with $|H| \leq |H'|$, then $\widetilde{d'_E}(H \cdot S, 0) \leq \widetilde{d'_E}(H' \cdot S, 0)$. Indeed, $\{H \neq 0\} \subset \{H' \neq 0\}$ implies for any $J \in \mathbf{b}\mathcal{P}$ with $\|J\|_\infty \leq 1$,

$$J \cdot (H \cdot S) = (JH) \cdot S = (JH1_{\{H\neq 0\}}) \cdot S = \left(J\frac{H}{H'}1_{\{H\neq 0\}}H'\right) \cdot S = K \cdot (H' \cdot S),$$

where $K := JH/H'1_{\{H\neq 0\}} \in \mathbf{b}\mathcal{P}$ with $\|K\|_\infty \leq 1$ due to $|H| \leq |H'|$. So definition of $\widetilde{d'_E}$ gives the desired result.

**Remark 4.8.** Some arguments above and in what follows assume $d = 1$. If $H$ and $S$ are $\mathbb{R}^d$-valued, then $H \cdot S$ is $\mathbb{R}$-valued, but e.g. $[S]$ is $\mathbb{R}^{d\times d}$-valued and $[H \cdot S] = \int H^{\mathrm{tr}} d[S] H$. Moreover, $H \cdot S$ can be well defined even if not all $H^i \cdot S^i$ are. For details, see Cherny/Shiryaev.

The next Proposition gives a characterisation of $\mathcal{L}(S)$.

**Proposition 4.2.** *Suppose $S$ is a semimartingale and $H$ is predictable. Then $H \in \mathcal{L}(S)$ iff for all sequences $(K^n)_{n\in\mathbb{N}} \subset \boldsymbol{b}\mathcal{P}$ with $|K^n| \leq |H|$ and $K^n \to 0$ pointwise, we have $\widetilde{d'_E}(K^n \cdot S, 0) \to 0$.*

**Remark 4.9.** This includes a dominated convergence property on $\mathbf{b}\mathcal{P}$ with bound in $\mathcal{L}(S) \supset_{\neq} \mathbf{b}\mathcal{P}$.

*Proof.* " $\Longrightarrow$ " Let $H^n := H1_{\{|H|\leq n\}}$. Then $(H^n \cdot S)_{n\in\mathbb{N}}$ is Cauchy for $\widetilde{d'_E}$ and so

$$\widetilde{d'_E}((H1_{\{|H|>n\}}) \cdot S, 0) = \widetilde{d'_E}(H \cdot S, H^n \cdot S) \to 0.$$

In consequence, by remark 4.7, $\widetilde{d'_E}((K^n1_{\{|H|>n\}}) \cdot S, 0) \to 0$. Take $m \neq n$ and write

$$(K^n1_{\{|H|\leq n\}}) \cdot S = (K^n1_{\{|H|\leq n\}}) \cdot S + (K^n1_{\{m<|H|\leq n\}}) \cdot S.$$

Recall $|K^n| \leq |H|$. For fixed $m$, the first term converges to 0 for $\widetilde{d'_E}$ by usual dominated convergence property for $\mathbf{b}\mathcal{P}$. Next,

$$\widetilde{d'_E}(2.\text{ term}, 0) \overset{\text{Rem 4.7}}{\leq} \widetilde{d'_E}(H1_{\{m<|H|\leq n\}}) \cdot S, 0) = \widetilde{d'_E}(H^n \cdot S, H^m \cdot S) \xrightarrow{n,m\to\infty} 0$$

by the Cauchy property. So we get $\widetilde{d'_E}((K^n1_{\{|H|\leq n\}}) \cdot S, 0) \to 0$ and hence $\widetilde{d'_E}(K^n \cdot S, 0) \to 0$.

"$\Longleftarrow$" For any $m_n \geq n$, define $K^n := H^{m_n} - H^n = H1_{\{n<|H|\leq m_n\}}$. So $|K^n| \leq |H|$ and $K^n \to 0$. So

$$\widetilde{d'_E}(H^{m_n} \cdot S, H^n \cdot S) = \widetilde{d'_E}(K^n \cdot S, 0) \xrightarrow{n\to\infty} 0$$

and as $m_n \geq n$ was arbitrary, we get $\widetilde{d'_E}(H^m \cdot S, H^n \cdot S) \to 0$ as $m, n \to \infty$. This is the Cauchy property for $(H^n \cdot S)_{n\in\mathbb{N}}$ and so $H \in \mathcal{L}(S)$. $\square$

Later, we need for the space of all stochastic integrals $\{H\cdot S : H \in \mathcal{L}(S)\}$ closedness result: Fix $S$ and define $d_S(H, H') := \widetilde{d'_E}(H \cdot S, H' \cdot S)$ for $H, H' \in \mathcal{L}(S)$, so integrands are close for $d_S$ iff their stochastic integrals are close for Emery metric.

Identify $H, H'$ if $d_S(H, H') = 0$ (gives equivalence classes) resulting space $L(S)$ is then metric space with metric $d_S$.

**Theorem 4.6** (Mémin). *Let $S = (S_t)_{0\le t\le T}$ be a $\mathbb{R}^d$-valued semimartingale. Then $(L(S), d_S)$ is a complete metric space. Equivalently, space of all stochastic integrals of $S$*

$$\left\{X = \int H dS = H \cdot S, H \in L(S)\right\}$$

*is closed for Emery topology (i.e. for $\widetilde{d}'_E$).*

**Remark 4.10.** We have seen in IMF that in **finite discrete** time, the space of all **final values** of stochastic integrals

$$G_T(\Theta) = \{G_T(\vartheta) : \vartheta \in \Theta\} = \left\{\int_0^T \vartheta_u dS_u : \vartheta \ \mathbb{R}^d\text{-valued predictable}\right\} \subset L^0$$

is closed in $L^0$. So we can ask: Is in general $\{H \cdot S_T : H \in L(S)\}$ closed in $L^0$? This is **not true** - in general, unlike in finite discrete time, a control on **final value** $H \cdot S_T$ does not give any control on the process $H \cdot S$, and so we cannot work.

Before we prove Theorem 4.6 we need an auxiliary result:

**Lemma 4.6.** *Suppose $(\gamma_k)_{k\in\mathbb{N}} \subset \boldsymbol{b}\mathcal{P}$ satisfy $\sum_{k=1}^\infty d_S(\gamma_k, 0) < \infty$ and define $[0, +\infty]$-valued process $G := \sum_{k=1}^\infty |\gamma_k|$. Then*

1. *For each $t \in [0, T]$, $\{\int_0^t H_u dS_u : H \in \boldsymbol{b}\mathcal{P}, |H| \le G\}$ is bounded in $L^0$.*
2. *If $(K^n)_{n\in\mathbb{N}} \subset \boldsymbol{b}\mathcal{P}$ satisfies $K^n \to 0$ pointwise and $|K^n| \le G$, then $\widetilde{d}'_E(K^n \cdot S, 0) \to 0$.*
   - *Does not follow from Proposition 4.2 because maybe $G \notin \mathcal{L}(S)$.*
3. *For any $H \in \boldsymbol{b}\mathcal{P}$, $\int H 1_{\{G=\infty\}} dS = 0$.*

*Proof.* We start with general considerations. Take $K \in \mathbf{b}\mathcal{P}$ with $|K| \le G$, set

$$G^m := \sum_{k=1}^m |\gamma_k| \nearrow G \text{ and } K^m := (K \wedge G^m) \vee (-G^m).$$

Then $|K^m| \le |K| \in \mathbf{b}\mathcal{P}$ and $K^m \to K$ pointwise, so that $d_S(\lambda K^m, 0) \to d_S(\lambda K, 0)$ as $m \to \infty$ for any $\lambda \ge 0$, by Theorem 4.4 (dominated convergence property). Moreoever, $|K^m| \le \sum_{k=1}^m (|K| \wedge |\gamma^k|)$ implies for fixed $\lambda > 0$, that

$$d_S(\lambda K, 0) = \lim_{m\to 0} d_S(\lambda K^m, 0) \le \lim_{m\to\infty} \sum_{k=1}^m d_S(\lambda(|K| \wedge |\gamma_k|), 0) = \sum_{k=1}^m d_S(\lambda(|K| \wedge |\gamma_k|, 0).$$

We can take $(\lambda_n) \subset [0,1]$ and $(K^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{P}$ to get analogous estimate for $d_S(\lambda_n K^n, 0)$. If $\lambda_n \to 0$ or $K^n \to 0$ pointwise, then for each $k \in \mathbb{N}$ we have $\lim_{n\to\infty} d_S(\lambda_n(|K^n| \wedge |\gamma_k|), 0) = 0$ by Theorem 4.4.

Moreover, $d_S(\lambda_n(|K^n| \wedge |\gamma_k|, 0) \leq d_S(\gamma_k, 0)$ which is summable over $k \in \mathbb{N}$ by assumption, so applying Lebesgue dominated convergence to the sum above yields

$$\lim_{n\to\infty} d_S(\lambda_n K^n, 0) \leq \lim_{n\to\infty} \sum_{k=1}^{\infty} d_S(\lambda(|K^n| \wedge |\lambda_k|, 0)$$

$$\overset{\text{Lebesgue}}{=} \sum_{k=0}^{\infty} \lim_{n\to\infty} d_S(\lambda_n(|K^n| \wedge |\lambda_k|), 0) = 0, \text{ if } \begin{cases} \lambda_n \xrightarrow{\text{pointwise}} 0 \\ \text{or} \\ K^n \xrightarrow{\text{pointwise}} 0 \end{cases}$$

We will now use the above derived general result, in order to conclude 1)-3).

1) Take any $(K^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{P}$ with $|K^n| \leq G$ and $\lambda_n \to 0$. Then $\lambda_n \int_0^t K_u^n dS_u \to 0$ in $L^0$, for each $t \in \mathbb{R}_+$

**Exercise 4.8.** *The setting as in 1) above ensures that*

$$\left\{ \int_0^t K_u dS_u : K \in \boldsymbol{b}\mathcal{P}, |K| \leq G \right\} \textit{ is bounded in } L^0.$$

2) Take $\lambda_n \equiv 1$. Then by definition

$$\widetilde{d}'_E(K^n \cdot S, 0) = d_S(K^n, 0) \to 0.$$

3) $G$ is predictable, and so $H1_{\{G=\infty\}} \in \mathbf{b}\mathcal{P}$ if $H \in \mathbf{b}\mathcal{P}$. For $c_n \to 0$, $H^n := c_n H 1_{\{G=\infty\}}$ satisfies $|H^n| \leq |G|$ because $G = \infty$ on $\{G = \infty\}$. By 1) we get

$$\left( c_n \int_0^t H_u 1_{\{G=\infty\}} dS_u \right)_{n\in\mathbb{N}} \text{ is bounded in } L^0.$$

This is only possible if $\int_0^t H_u 1_{\{G=\infty\}} dS_u = 0$ and $t$ was arbitrary. $\square$

*Proof of Theorem 4.6.* Equivalence is clear from the definition of $d_S$ from $\widetilde{d}'_E$ and the fact that $(\mathcal{S}, \widetilde{d}'_E)$ is complete.

Take $(\widetilde{H}^n)_{n\in\mathbb{N}} \subset L(S)$ which is Cauchy for $d_S$. We need to show that there exists $\widetilde{H} \in L(S)$ with $d_S(\widetilde{H}^n, \widetilde{H}) \to 0$. Approximate each $\widetilde{H}^n$ by $H^n \in \mathbf{b}\mathcal{P}$ (use definition of $L(S)$ and $d_S$) in $d_S$. Choose subsequence, again called $(H^n)_{n\in\mathbb{N}}$, with $d_S(H^n, H^{n-1}) \leq 2^{-n}$ (using Cauchy property plus $H^n \approx \widetilde{H}^n$). Then it is enough to find some $H \in L(S)$ with $H^n \cdot S \to H \cdot S$ for $\widetilde{d}'_E$.

Define $\gamma_n := H^{n+1} - H^n \in \mathbf{b}\mathcal{P}$ and $G := \sum_{n=1}^\infty |\gamma_n|$. For any $(K^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{P}$ with $|K^n| \leq G1_{\{G<\infty\}}$ and $K^n \to 0$, we get $\widetilde{d}'_E(K^n \cdot S, 0) \to 0$ by Lemma 4.6, 2). So $G1_{\{G<\infty\}} \in L(S)$ by Proposition 4.2 on $\{G < \infty\}$.

$\overline{H} := \sum_{n=1}^\infty \gamma_n = \sum_{n=1}^\infty (H^{n+1} - H^n)$ is well-defined and $\overline{H}1_{\{G<\infty\}}$ is predictable with $|\overline{H}1_{\{G<\infty\}}| \leq G$. So we can first use (for any test $(K^n)_{n\in\mathbb{N}}$) Lemma 4.6 2) and then again Proposition 4.2 to obtain that $\overline{H}1_{\{G<\infty\}}$ is in $L(S)$. So is then $H := (\overline{H} + H^1)1_{\{G<\infty\}}$ because $H^1 \in \mathbf{b}\mathcal{P} \subset L(S)$.

But $H^1 + \overline{H} - H^n = \sum_{k=n}^\infty \gamma_k$ implies that

$$d_S(H, H^n 1_{\{G<\infty\}}) \leq \sum_{k=n}^\infty d_S(\gamma_k, 0) \xrightarrow{n\to\infty} 0,$$

because $\sum_{k=1}^\infty d_S(\gamma_k, 0) < \infty$. Finally, $H^n \in \mathbf{b}\mathcal{P}$, so by Lemma 4.6 3)

$$\int H^n 1_{\{G<\infty\}} dS = H^n \cdot S$$

and thus

$$\widetilde{d}'_E(H \cdot S, H^n \cdot S) = d_S(H, H^n) \xrightarrow{n\to\infty} 0$$

and so we have found our $H$. $\square$

## 4.1 Recap

We provide again a short summary of this rather technical chapter. Our (second) main goal was to extend stochastic integration beyond $\mathbf{b}\mathcal{E}$, i.e. to a larger class of integrands $H$. To this extend we introduced the space $\mathbb{L}$ resp. $\mathbb{D}$ of $\mathbb{R}^d$-valued adapted LCRL resp RCLL processes and introduced the metric $d(X^1, X^2) := \mathbb{E}(1 \wedge (X^1 - X^2)^*_T)$. We've seen that both $(\mathbb{L}, d)$ and $(\mathbb{D}, d)$ are **complete** metric spaces (Exercise). We've introduced the space $\mathbf{b}\mathcal{E}_0$, which is $\mathbf{b}\mathcal{E}$ but all the spaces of the form $H = h_0 1_{\{0\}}$ with $h_0 \in L^\infty(\mathcal{F}_0, \mathbb{R}^d)$ added to it, then defined the metrics

$$\begin{aligned} d_E(X^1, X^2) &:= \sup_{H \in \mathbf{b}\mathcal{E}_0,\ \|H\|_\infty \le 1} \mathbb{E}(1 \wedge (H \cdot (X^1 - X^2))_T) \\ &\le d'_E(X^1, X^2) := \sup_{H \in \mathbf{b}\mathcal{E}_0,\ \|H\|_\infty \le 1} \mathbb{E}(1 \wedge (H \cdot (X^1 - X^2)^*_T)). \end{aligned}$$

We notice that $d'_E \ge d$ and then showed that $d_E$ controls $d$ (in the sense that convergences of $(X^n)_{n \in \mathbb{N}} \subset \mathbb{D}$ w.r.t $d_E$ implies convergence of the same sequence w.r.t $d$) and then in fact even showed that $d'_E$ and $d_E$ (as defined above) are equivalent. We have then seen that $\mathcal{S} \subset \mathbb{D}$, the space of $\mathbb{R}^d$-valued semimartingales, is a complete metric space with respect to $d_E, d'_E$.

Our first **major result** states that for every semimartingale $S$, the mapping $I_S : \mathbf{b}\mathcal{E} \to L^0$ can be extended to a continuous mapping $J_S : (\mathbb{L}, d) \to (\mathcal{S}, d'_E)$. Notice that with this (strong) result, for each $H \in \mathbb{L}$ (integrand), the stochastic integral $H \cdot S := J_S(H)$ is well defined and a semimartingale (in $\mathcal{S}$), in particular adapted RCLL and hence we get path properties. Moreover, by the continuity of the mapping $J_S$, we obtain that $H^n \to H$ w.r.t. $d$ for $(H^n) \subset \mathbb{L}, H \in \mathbb{L}$, implies $H^n \cdot S \to H \cdot S$ w.r.t $d'_E$.

Next, we want to extend class of integrands beyond $\mathbb{L}$, for this we need a different approximation concept and hence a new idea. The concept is already known from BMSC. For a semimartingale $S$, we define the **optional quadratic variation** as

$$[S] := S^2 - S_0^2 - 2\int S_- dS = S^2 - S_0^2 - 2J_S(S_-).$$

We notice that since $S$ is semimartingale (in particular RCLL), then $S_-$ is adapted and LCRL (hence in $\mathbb{L}$) and so $J_S(S_-) = S_- \cdot S$ is well defined. Also, by definition, $[S]$ is adapted, RCLL, null at 0 and it can be shown to be $\mathbb{P}$-a.s. increasing in $t$. **Davis inequality**, is then a crucial ingredient for the construction of stochastic integrals w.r.t. local martingales, it states that there exists constants $0 < c < C < \infty$ such that for all $M \in \mathcal{M}_{0,\mathrm{loc}}$ we have

$$c\mathbb{E}(([M]_T)^{1/2}) \le \mathbb{E}(M_T^*) \le C\mathbb{E}(([M]_T)^{1/2}).$$

For $M \in \mathcal{M}_{0,\mathrm{loc}}$, we define $\mathcal{L}^1(M)$ to be the space consistng of all predictable $H$ such that

$$\|H\|_{L^1(M)} := \mathbb{E}\left[\left(\int_0^T H_s^2 d[M]_s\right)^{1/2}\right] < \infty.$$

By identification this then yields the space $L^1(M)$. We also introduce the space $\mathcal{H}_0^1 := \{M \in \mathcal{M}_{0,\mathrm{loc}} : M_T^* = \sup_{0\le s\le T} |M_s| \in L^1\}$. For which we have (all Exercises) that every $M \in \mathcal{H}_0^1$ is a (true) martingale and even of class (D) and also we have $\mathcal{M}_{0,\mathrm{loc}} = \mathcal{H}_{0,\mathrm{loc}}^1$ (important). The last statement means that $M \in \mathcal{M}_{0,\mathrm{loc}}$ is locally in $\mathcal{H}_0^1$ and therefore (according to the last statements) $M \in \mathcal{M}_{0,\mathrm{loc}}$ is after localisation a (true) martingale. After a couple of auxiliary results (which involve the use of the optional quadratic variation) we then arrive at the major theorem:

**Theorem 4.4**: For any martingale $M \in \mathcal{H}_0^1$, the mapping $I_M(\mathbf{b}\mathcal{E}, \|\cdot\|_\infty) \to (\mathcal{S}, d_E')$ admits a unique linear extension to $J_M : \mathbf{b}\mathcal{P} \to \mathcal{S}$ such that $H\cdot M := J_M(H)$ satisfies $[H\cdot M] = \int H^2 d[M]$ and $\Delta(H\cdot M) = H\Delta M$ and with dominated convergence property, i.e. that $(H^k) \subset \mathbf{b}\mathcal{P}$ with $\|H^k\|_\infty \le 1$ and $H^k \to 0$ pointwise implies $d_E'(H^k\cdot M, 0) \to 0$. Moreover, $J_M : (\mathbf{b}\mathcal{P}, \|\cdot\|_\infty) \to (\mathcal{S}, d_E')$ is continuous.

One more localisation allows us to extend from $\mathbf{b}\mathcal{P}$ to $\mathbf{b}\mathcal{P}_{\mathrm{loc}}$, and so $H\cdot S = \int H dS$ is well-defined and a semimartingale for any $S \in \mathcal{S}$ and any $H \in \mathbf{b}\mathcal{P}_{\mathrm{loc}}$. We remark that in the setup of the theorem above, i.e. for $M \in \mathcal{H}_0^1$ and $H \in \mathbf{b}\mathcal{P}$, we have by construction as limit in $\mathcal{S}$ that $H\cdot M \in \mathcal{S}$ is a semimartingale. However, we've seen in class that it is actually a (true) martingale, in fact we have $H\cdot M \in \mathcal{H}_0^1$.

Next we introduce on $\mathcal{S}$ (space of semimartingales in $\mathbb{R}^d$) two metrics

$$\widetilde{d}_E(S^1, S^2) := \sup_{H\in\mathbf{b}\mathcal{P},\ \|H\|_\infty\le 1} \mathbb{E}(1 \wedge |H\cdot(S^1-S^2)|_T) \ge d_E(S^1, S^2),$$
$$\widetilde{d}'_E(S^1, S^2) := \sup_{H\in\mathbf{b}\mathcal{P},\ \|H\|_\infty\le 1} \mathbb{E}(1 \wedge (H\cdot(S^1-S^2))_T^*)$$

One can again argue that these two metrics are equivalent. Moreover, they are also equivalent to $d_E, d_E'$. The "strongest" metric $\widetilde{d}'_E$ is called the **Emery metric** on $\mathcal{S}$ and the corresponding (metric) topology is called **Emery topology**. These notions allow us to introduce the final extension of integrands (most general case). We fix a semimartingale $S$ and define a predictable process $H$ (i.e. $H \in \mathcal{P}$) to be $S$-**integrable**, denoted as $H \in \mathcal{L}(S)$, if $(H^n\cdot S)_{n\in\mathbb{N}}$ is Cauchy in **Emery topology** (or for $d_E'$), where $H^n := H1_{\{|H|\le n\}} \in \mathbf{b}\mathcal{P}$. The Limit in $\mathcal{S}$ (recall that $\mathcal{S}$ is complete) is called $H\cdot S$. Moreover we have characterized $\mathcal{L}(S)$; we've seen that $H \in \mathcal{L}(S)$ iff for all sequence $(K^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{P}$ with $|K^n| \le |H|$ and $K^n \to 0$ pointwise, we have $K^n\cdot S \to 0$ w.r.t. $\widetilde{d}'_E$ (i.e. Emery metric).

As a final result in this section, we discussed a result that we will need later in the course. Later (to discuss absence of arbitrage) we need for the space of all stochastic integrals $\{H \cdot S : H \in \mathcal{L}(S)\}$ a closedness result.

For a fixed semimartingale $S \in \mathcal{S}$ and $H, H' \in \mathcal{L}(S)$ we define

$$d_S(H, H') = \widetilde{d}'_E(H \cdot S, H' \cdot S).$$

So integrands are close for $d_S$ iff their stochastic integrals are close for Emery metric. Identifying $H, H'$ if $d_S(H, H') = 0$ (gives equivalence classes) resulting space $L(S)$ is then a metric space with metric $d_S$.

The **Theorem of Mémin** then establishes that $(L(S), d_S)$ is a complete metric space. Equivalently, space of all stochastic integrals of $S$, $\{H \cdot S : H \in L(S)\}$ is closed for Emery toplogy (i.e. w.r.t. $\widetilde{d}'_E$)

# 5 The fundamental theorem of asset pricing (FTAP)

**Goal**: Precise formulation for equivalence between "absence of arbitrage" and "existence of an EMM" in general continuous time model.

**Setup**: $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ satisfying the usual conditions. Moreover, $B = 1$, $S$ adapted RCLL $\mathbb{R}^d$-valued and also assume $S$ is a **semimartingale**.

Recall space $L(S)$ of $\mathbb{R}^d$-valued predictable $S$-integrable $\vartheta$, define $G(\vartheta) := \int \vartheta dS = \vartheta \cdot S$, and call $\vartheta$ $a$-**admissible** for $a \ge 0$ if $G_.(\vartheta) \ge -a$ $\mathbb{P}$-a.s. Call $\Theta^a$ all $a$-admissible $\vartheta \in L(S)$, $\Theta_{\text{adm}} = \cup_{a \ge 0} \Theta^a$ and introduce

$$\mathcal{G}^a := G_T(\Theta^a) = \left\{ G_T(\vartheta) = \int_0^T \vartheta_u dS_u : \vartheta \in \Theta^a \right\}$$
$$\mathcal{G}_{\text{adm}} := G_T(\Theta_{\text{adm}}) = \bigcup_{a \ge 0} \mathcal{G}^a.$$

**Interpretation**: any $g \in \mathcal{G}^a$ represents the final wealth of some self-financing trading strategy $\varphi \stackrel{\wedge}{=} (0, \vartheta)$ starting from 0 initial wealth and with wealth bounded from below by $-a$.

**No arbitrage**: (NA) is defined to be the condition $\mathcal{G}_{\text{adm}} \cap L^0_+ = \{0\}$.

Recall that $Q \approx \mathbb{P}$ is an ELMM for $S$ if $S \in \mathcal{M}_{\text{loc}}(Q)$. The next Lemma we've already discussed in section 2, we now have the tools to rigorously prove it.

**Lemma 5.1.** *Suppose $S$ is adapted RCLL. If $S$ admits an ELMM, say $Q$, then $S$ satisfies (NA), i.e. $G_T(\Theta_{adm}) \cap L^0_+ = \{0\}$.*

For a complete proof, we need an auxiliary result:

**Proposition 5.1** (Ansel-Stricker)**.** *Suppose $M$ is a local martingale and $\vartheta \in L(M)$. If $\int \vartheta dM \ge -b$ for some $b \ge 0$, then also $\int \vartheta dM$ is a local martingale and a supermartingale.*

We first prove a more general result, due to De Donno/Pratelli.

**Lemma 5.2.** *Suppose $X$ is adapted RCLL. Let $(\gamma_n)_{n \in \mathbb{N}} \subset L^1$ and $(\tau_n)_{n \in \mathbb{N}}$ a localising sequence (increasing to $T$ stationarily) be such that $X^{\tau_n} \ge \gamma_n$ for all $n \in \mathbb{N}$. Let $(M^n)_{n \in \mathbb{N}}$ be a sequence of martingales with $d(M^n, X) \to 0$ as $n \to \infty$ (i.e. $M^n \to X$ uniformly in $t \in [0,T]$ in $L^0$) and $(\Delta M^n_\sigma)^\pm \le (\Delta X_\sigma)^\pm$ for all $n \in \mathbb{N}$ and for all stopping times $\sigma$. Then $X$ is a local martingale.*

*Proof.* **Idea**: Construct a good sequence $(\sigma_m)_{m\in\mathbb{N}}$ of stopping times, then $M^n_{t\wedge\sigma_m} \xrightarrow{n\to\infty} X_{t\wedge\sigma_m}$ in $L^0$ for all $m \in \mathbb{N}$, argue that convergence even takes place in $L^1$ and that $X_{t\wedge\sigma_m} \in L^1$. Then $X^{\sigma_m}$ is a martingale, i.e. $X$ is a local martingale.

Assume $M^n_0 \equiv X_0 = 0$. Define stopping times

$$\rho_n := \inf\{t \in [0,T] : X_t > n \text{ or } M^n_t > X_t + 1 \text{ or } M^n_t < X_t - 1\} \wedge T.$$

By passing to a subsequence, we may assume that $\sum_{n=1}^{\infty} \mathbb{P}(\rho_n < T) < \infty$. Then $\mathbb{P}(\inf_{n\geq m} \rho_n = T) \xrightarrow{m\to\infty} 1$ because

$$\mathbb{P}(\inf_{n\geq m} \rho_n = T) = 1 - \mathbb{P}\left(\bigcup_{n\geq m} \{\rho_n < T\}\right) \geq 1 - \underbrace{\sum_{n=m}^{\infty} \mathbb{P}(\rho_n < T)}_{\xrightarrow{m\to\infty} 0} \xrightarrow{m\to\infty} 1.$$

Thus $\sigma_m := \tau_m \wedge \inf_{n\geq m} \rho_n \nearrow T$ stationarily as $m \to \infty$.

**Claim**: $X^{\sigma_m}$ is a martingale.

To see this, start with

$$(\Delta M^n_{t\wedge\sigma_m})^- \leq (\Delta X_{t\wedge\sigma_m})^- \overset{X^{\tau_m}\geq\gamma_m, X_{\rho_m^-}\leq m}{\leq} m - \gamma_m.$$

Moreover, for $n \geq m$ and $t < \sigma_m$, we have $M^n_{t\wedge\sigma_m} = M^n_t \geq X_t - 1 \geq \gamma - 1$, and so for all $n \geq m$ and all $t \in \mathbb{R}^+_0$, $M^n_{t\wedge\sigma_m} \geq \gamma_m - 1 - (m - \gamma_m) \in L^1$. So we obtain

$$\mathbb{E}(X_{t\wedge\sigma_m}) = \mathbb{E}(\lim_{n\to\infty} M^n_{t\wedge\sigma_m}) \overset{\text{Fatou}}{\leq} \liminf_{n\to\infty} \underbrace{\mathbb{E}(M^n_{t\wedge\sigma_m})}_{=0,\ (\text{mart.})} = 0.$$

In particular we get

$$\mathbb{E}(X^+_{t\wedge\sigma_m} - X^-_{t\wedge\sigma_m}) = \mathbb{E}(X_{t\wedge\sigma_m}) \leq 0, \text{ i.e. } \mathbb{E}(X^+_{t\wedge\sigma_m}) \leq \mathbb{E}(X^-_{t\wedge\sigma_m})$$

and $X^-_{t\wedge\sigma_m} \in L^1$ because $X_{t\wedge\sigma_m} \geq \gamma_m \in L^1$ (use if $f \geq a$ then $-f^- \geq a$ i.e. $f^- \leq -a$). Hence $\mathbb{E}(X^+_{t\wedge\sigma_m}) \leq \mathbb{E}(X^-_{t\wedge\sigma_m}) < \infty$ and so $X_{t\wedge\sigma_m} \in L^1$ for all $t$. Then also $\Delta X_{t\wedge\sigma_m} \in L^1$ because $X_{(t\wedge\sigma_m)^-} \leq m$.
By an analog argument, we get

$$M^n_{t\wedge\sigma_m} \leq m + 1 + (\Delta M^n_{t\wedge\sigma_m})^+ \leq m + 1 + |\Delta X_{t\wedge\sigma_m}| \in L^1, \text{ for all } n \in \mathbb{N}.$$

Moreover, $M^n_{t\wedge\sigma_m} \xrightarrow{n\to\infty} X_{t\wedge\sigma_m}$ in $L^0$, hence in $L^1$ and so $X^{\sigma_m}$ is a martingale because all $(M^n)^{\sigma_m}$ are. $\square$

*Proof of Proposition 5.1.* Since $\mathcal{M}_{0,\text{loc}} = \mathcal{H}^1_{0,\text{loc}}$ allows us by stopping to assume that $M \in \mathcal{H}^1$. Take $\vartheta \in L(M)$ and $\vartheta^n_. := \vartheta.1_{\{|\vartheta \leq n\}}$ and set $M^n := \int \vartheta^n dM$. By definition, $M^n \to \int \vartheta dM =: X$ for Emery metric $\widetilde{d}'_E \geq d$, and each $M^n \in \mathcal{H}^1$ by Proposition 4.1. Because $X \geq -b$, all assumptions of Lemma 5.2 are satisfied and so $X = \int \vartheta dM$ is a local martingale. Moreover, as $X \geq -b$ it is also a supermartingale by Fatou. $\square$

**Remark 5.1.** From the proof above: conclusion still holds if instead of $\int \vartheta dM \geq -b$, we only have that all large jumps of $\int \vartheta dM$ (say $\geq C$) have lower bounded in $L^1$.

We now prove Lemma 5.1

*Proof of Lemma 5.1.* We first argue that $S$ is a $\mathbb{P}$-semimartingale: Indeed, by Theorem 3.3 and Theorem 3.4 semimartingales are the same as good integrators and like $L^0$, this concept is the same for $\mathbb{P}$ and for any $R \approx \mathbb{P}$. Now $S \in \mathcal{M}_{\text{loc}}(Q)$, so $S$ is a $Q$-semimartingale and as $Q \approx \mathbb{P}$, $S$ is also a $\mathbb{P}$-semimartingale. So $L(S)$, $\Theta_{\text{adm}}$ (NA) all make sense under $\mathbb{P}$.

Suppose $\vartheta \in \Theta_{\text{adm}}$. Then $\int \vartheta dS \geq$ constant a.s. and $S \in \mathcal{M}_{\text{loc}}(Q)$; so $\int \vartheta dS$ is a $Q$-local martingale and $Q$-supermartingale by Proposition 5.1 (Ansel-Stricker). Hence $\mathbb{E}_Q(G_T(\vartheta)) \leq 0$, if also $G_T(\vartheta) \in L^0_+$, i.e. nonnegative a.s., we must have that $G_T(\vartheta) = 0$ a.s. and so we get (NA). $\square$

So, the existence of a ELMM $\implies$ (NA), and we have seen from counterexample that $\not\Longleftarrow$. So we need to strengthen (NA).

**Recall**: $\mathcal{C}^\infty := (G_T(\Theta_{\text{adm}}) - L^0_+) \cap L^\infty$ is set of all bounded payoffs which can be dominated/superreplicated by final wealth of some admissible self-financing strategy with 0 initial wealth.

**Exercise 5.1.** *(NA) $\mathcal{G}_T(\Theta_{adm}) \cap L^0_+ = \{0\}$ is equivalent to (NA) $\mathcal{C}^\infty \cap L^\infty_+ = \{0\}$.*

**Definition 5.1.** *A semimartingale $S = (S_t)_{0 \leq t \leq T}$ satisfies **no free lunch with vanishing risk** if*

- *(NFLVR): $\overline{\mathcal{C}^\infty}^{L^\infty} \cap L^\infty_+ = \{0\}$.*

*where $\overline{\cdot}^{L^\infty}$ is (norm) closure in $L^\infty$.*

**Remark 5.2.**

- By definition, (NFLVR) $\implies$ (NA). Moreover, (NFLVR) forbids not only, like (NA), generating profits directly from 0, but also asymptotically (i.e. as limits in $L^\infty$).

- Later, we want to use Kreps-Yan theorem, for $p = \infty$. Then we need to use on $L^\infty$ weak* topology $\sigma(L^\infty, L^1)$ and we have $\overline{\cdot}^{L^\infty} \subset \overline{\cdot}^{\sigma(L^\infty,L^1)}$. We also need in Kreps-Yan theorem closedness for $\sigma(L^\infty, L^1)$. On economic grounds, using in (NFLVR) $\overline{\cdot}^{L^\infty}$ is much more natural than it would be using $\overline{\cdot}^{\sigma(L^\infty,L^1)}$ (NFL). This is at the root of later technical difficulties.

**Proposition 5.2.** *For a semimartingale $S = (S_t)_{0\le t\le T}$ the following are equivalent:*

1. *$S$ satisfies (NFLVR)*

2. *Any sequence $g_n = G_T(\vartheta^n)$ in $\mathcal{G}_{adm}$ for $n \in \mathbb{N}$ with $G_T^-(\vartheta^n) \to 0$ in $L^\infty$ converges to $0$ in $L^0$.*

   - *That is, if losses go to $0$ uniformly, then wealth or gains go to $0$ in $L^0$.*

3. *$S$ satisfies (NA) plus* ***no unbounded profit with bounded risk****, (NUPBR): $\mathcal{G}^1 = G_T(\Theta^1) = \{G_T(\vartheta) : \vartheta \text{ is } 1\text{-admissible}\}$ is bounded in $L^0$ (i.e. $\sup_{g\in\mathcal{G}^1} \mathbb{P}(|g| \ge n) \to 0$ as $n \to \infty$).*

**Remark 5.3.** In brief [important]: (NFLVR)=(NA)+(NUPBR).

*Proof.* "2) $\implies$ 3)": Any $f \in \mathcal{C}^\infty \cap L^\infty_+$ has form $0 \le f = g - Y$ with $g \in G_T(\Theta_{\text{adm}})$, $Y \ge 0$. So $g \ge 0$ and by taking $g_n \equiv g$ in 2) gives $g \equiv 0$, so we have (NA).

If (NUPBR) fails, then $\mathcal{G}^1$ is not bounded in $L^0$ and so there exists $\gamma_n \nearrow +\infty$ with $\mathbb{P}(g_n \ge \gamma_n) \ge \delta > 0$ for some sequence $(g_n)_{n\in\mathbb{N}} \subset \mathcal{G}^1$. But $g_n = G_T(\vartheta^n)$ with $\vartheta^n$ 1-admissible means that

$$\widetilde{g}_n := \frac{1}{\gamma_n} g_n$$

is in $\mathcal{G}_{\text{adm}}$ with $\widetilde{g}_n^- \to 0$ in $L^\infty$ because $\|\widetilde{g}_n^-\|_{L^\infty} \le \frac{1}{\gamma_n}$ (using $g_n \ge -1 \implies g_n^- \le 1$). But we also have $\mathbb{P}(\widetilde{g}_n \ge 1) = \mathbb{P}(g_n \ge \gamma_n) \not\to 0$, so $\widetilde{g}_n \not\to 0$ in $L^0$ and this contradicts 2). Hence (NUPBR) holds as well.

In order to continue the proof we need an auxiliary result from probability theory which is extremely useful

**Lemma 5.3** (Komlós)**.** *For any sequence $(X_n)_{n\in\mathbb{N}} \subset L^0_+$, there exists convex combinations $\widetilde{X}_n \in conv(X_n, X_{n+1}, \dots)$, $n \in \mathbb{N}$ such that $(\widetilde{X}_n)_{n\in\mathbb{N}}$ converges $\mathbb{P}$-a.s. to some $\widetilde{X}_\infty$ taking values in $[0, +\infty]$. Moreover, if $\mathbb{P}(X_n \geq \alpha) \geq \delta > 0$ for some $\alpha > 0$, then $\mathbb{P}(\widetilde{X}_\infty > 0) > 0$. If conv$(X_1, X_2, \dots)$ is bounded in $L^0$, then $\widetilde{X}_\infty < \infty$ $\mathbb{P}$-a.s.*

For a proof of Lemma 5.3 see Appendix B.

"3) $\Longrightarrow$ 1)": Suppose (NA) holds but (NFLVR) fails, then we claim that (NUPBR) fails as well and the result follows. To see this: (NFLVR) fails means there exists a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{C}^\infty$ and $f \in L^\infty_+ \setminus \{0\}$ with $\|f - f_n\|_\infty \leq \frac{1}{n}$.

Now $f_n \leq g_n$ for some $g_n \in \mathcal{G}_{\text{adm}}$ (by Definition of $\mathcal{C}^\infty$), so $\|g_n^-\|_\infty \leq \|f_n^-\|_\infty \leq \frac{1}{n}$, and due to (NA), $g_n = G_T(\vartheta^n)$ with $\vartheta^n \in \Theta_{\text{adm}}$ and $g_n \geq -\frac{1}{n}$ implies that $G.(\vartheta^n) \geq -\frac{1}{n}$.

**Exercise 5.2.** *Verify that the statement in the previous paragraph holds.*

This means that $n\vartheta^n$ is 1-admissible and so $ng_n \in \mathcal{G}^1$. Using Lemma 5.3 gives $\widetilde{g}_n \in \text{conv}(g_n, g_{n+1}, \dots)$ with $\widetilde{g}_n \to \widetilde{g}_\infty$ $\mathbb{P}$-a.s., and $g_n \geq f_n$ gives $\widetilde{g}_\infty \geq f$ so that $\mathbb{P}(\widetilde{g}_\infty > 0) > 0$. But now $(n\widetilde{g}_n)_{n\in\mathbb{N}} \subset \mathcal{G}^1$ has $\frac{1}{n} n\widetilde{g}_n = \widetilde{g}_n$ not converging to 0 in $L^0$. So $\mathcal{G}^1$ is not bounded in $L^0$ and thus (NUPBR) fails.

"1) $\Longrightarrow$ 2)": Take $(g_n)_{n\in\mathbb{N}} \subset \mathcal{G}_{\text{adm}}$ with $g_n^- \to 0$ in $L^\infty$ and suppose $g_n \not\to 0$ in $L^0$. Then along a subsequence, again called $(g_n)_{n\in\mathbb{N}}$, we get $\mathbb{P}(g_n \geq \alpha) \geq \delta > 0$ for some $\alpha > 0$.

Look at $f_n := g_n \wedge 1 = g_n - (g_n - g_n \wedge 1) \in \mathcal{G}_{\text{adm}} - L^0_+$. Then $f_n^- = g_n^- \to 0$ in $L^0$ so that we can assume $f_n \geq -a$ for some $a > 0$ (and $a$ can be arbitrarily small). So $f_n \in L^\infty$, hence $f_n \in \mathcal{C}^\infty$.

Take $\widetilde{f}_n \in \text{conv}(f_n, f_{n+1}, \dots) \subset \mathcal{C}^\infty$ (subset since $\mathcal{C}^\infty$ is convex set itself) with $\widetilde{f}_n \to \widetilde{f}_\infty$ $\mathbb{P}$-a.s. and also $\widetilde{f}_\infty \geq -a$ $\mathbb{P}$-a.s. and $\mathbb{P}(\widetilde{f}_\infty > 0) =: \beta > 0$. As we can take $a$ to be arbitrarily small, this gives $\widetilde{f}_\infty \in L^\infty_+ \setminus \{0\}$.

Now $\widetilde{f}_n \to \widetilde{f}_\infty$ $\mathbb{P}$-a.s. and $(\widetilde{f}_n)_{n\in\mathbb{N}} \subset \mathcal{C}^\infty$. By **Egorov's theorem**, $\mathbb{P}$-a.s. convergence implies uniform convergence (i.e. in $L^\infty$) on subsets of large measure. So there exists $B \in \mathcal{F}$ with $\mathbb{P}(B) \geq 1 - \frac{\beta}{2}$ and $\widetilde{f}_n 1_B \to \widetilde{f}_\infty 1_B$ in $L^\infty$. But then $\widetilde{f}_n 1_B - \widetilde{f}_n^- 1_{B^c} = \widetilde{f}_n - \widetilde{f}_n^+ 1_{B^c}$ is a sequence in $\mathcal{C}^\infty$ which converges in $L^\infty$ because

$f_n^- \to 0$ in $L^\infty$ like $g_n^- = f_n^-$. So $\widetilde{f}_\infty 1_B \in \overline{C^\infty}^{L^\infty} \cap L^\infty_+$, but

$$\mathbb{P}(\widetilde{f}_\infty 1_B > 0) \geq \mathbb{P}(\widetilde{f}_\infty > 0) - \mathbb{P}(B^c) \geq \beta - \frac{\beta}{2} = \frac{\beta}{2} > 0$$

and this contradicts (NFLVR). So we are done. □

Main results needs one more concept.

**Definition 5.2.** *An adapted $\mathbb{R}^d$-valued RCLL process $X = (X_t)_{0\leq t\leq T}$ is called a ($\mathbb{P}$-)$\sigma$-**martingale** if $X - X_0 = \int \Psi dM$ for an $\mathbb{R}^d$-valued local ($\mathbb{P}$-) martingale $M$ and a one-dimensional integrand $\Psi \in L(M)$ (meaning $\Psi \in L(M^i)$ for all $i = 1, \dots, d$) with $\Psi > 0$.*

**Remark 5.4.** Loosely speaking, $dX = \Psi dM$ with $\Psi > 0$, so $dX$ and $dM$ go in the same direction.

**Definition 5.3.** ***Equivalent $\sigma$-martingale measure*** *(E$\sigma$M) for $S$ is a probability measure $Q \approx \mathbb{P}$ (on $\mathcal{F}_T$) such that $S$ is a $Q$-$\sigma$-martingale.*

**Definition 5.4.** ***Equivalent separating measure*** *(ESM) for $S$ is a probability measure $Q \approx \mathbb{P}$ (on $\mathcal{F}_T$) with $\mathbb{E}_Q(G_T(\vartheta)) \leq 0$ for all $\vartheta \in \Theta_{adm}$.*

**Remark 5.5.**

1. We have the following chain of implications: Martingale $\Longrightarrow$ local martingale $\Longrightarrow$ $\sigma$-martingale (take $\Psi \equiv 1$), converses are not true in general.

2. $X$ $\sigma$-martingale and $X - X_0 \geq$ const, means that $X$ is a local martingale by Ansel-Stricker (Proposition 5.1); useful for $X \geq 0$ and $X_0$ non random. More generally, using Lemma 5.2 a $\sigma$-martingale which is locally bounded from below (e.g. continuous) is a local martingale.

3. If $S$ is a $Q$-$\sigma$-martingale and $\vartheta \in \Theta_{\text{adm}}$. Then

$$G(\vartheta) = \int \vartheta dS = \int \Psi\vartheta dM \geq -a$$

is a local $Q$-martingale and $Q$-supermartingale by Ansel-Stricker.

   - So EMM $\Longrightarrow$ ELMM $\Longrightarrow$ E$\sigma$MM $\Longrightarrow$ ESM.

**Exercise 5.3.** *Conversely: if $S$ is (locally) bounded, then ESM $\Longrightarrow$ E(L)MM.*

**Example 5.1** (Emery). *Let $\tau \sim$Exp$(1)$ and $Z$ independent of $\tau$ taking values $\pm 1$ with probability $1/2$ each. Set $M_t := Z1_{\{t \geq \tau\}}$ for $t \geq 0$ and use $\mathbb{F} := \mathbb{F}^M$. Then $M_t - M_s$ for $t > s$ is different from $0$ iff $s < \tau \leq t$. This implies that for $u \leq s$ and on $\{\tau > s\}$ we have $M_u = Z1_{\{u \geq \tau\}} = 0$, so $\mathcal{F}_s^M \cap \{\tau > s\} = \{A \cap \{\tau > s\} : A \in \mathcal{F}_s^M\}$ is trivial and so anything $\mathcal{F}_s^M$-measurable must be constant almost surely on $\{\tau > s\}$. Therefore, for $A \in \mathcal{F}_s^M$ and $t > s$ we have*

$$\mathbb{E}((M_t - M_s)1_A) = \mathbb{E}(1_{\{s<\tau\leq t\}} Z 1_A) \overset{Z \text{ indep. } \tau}{=} const \cdot \underbrace{\mathbb{E}(Z)}_{=0} \mathbb{P}(s < \tau \leq T) = 0,$$

*So $M$ is a martingale, and of course it is of FV.*

*Define $\Psi(t) := \frac{1}{t}$, this is predictable (it is non-random) and positive on $(0, \infty)$, and it is also in $L(M)$. The martingale $M$ is constant except for single jump at $\tau$, and so for any $t > 0$*

$$\int_0^t \Psi(u) dM_u = \Psi(\tau) \Delta M_\tau 1_{\{t \geq \tau\}} = \frac{Z}{\tau} 1_{\{t \geq \tau\}}.$$

*Of course $\int \Psi dM$ is a $\sigma$-martingale.*

***Claim**: $\int \Psi dM$ is **not** a local martingale!*

***Problem**: $(\Psi \cdot M)_\sigma \notin L^1$ for all stopping times $\sigma \not\equiv 0$, so one cannot make $\Psi \cdot M$ integrable by any localisation and thus it cannot be a local martingale.*

*Indeed, if $\sigma$ is any stopping time w.r.t. $\mathbb{F}^M$, then on $\{\sigma < \tau\}$, $\sigma$ must be constant (similar argument as above) and if $\sigma \not\equiv 0$, then $\sigma \geq \tau$ on $\{\tau \leq \epsilon\}$ for some $\epsilon > 0$. So*

$$|(\Psi \cdot M)_\sigma| = \left| \int_0^\sigma \Psi(u) dM_u \right| = \left| \frac{Z}{\tau} 1_{\{\sigma \geq \tau\}} \right| \overset{Z \in \{\pm 1\}}{=} \frac{1}{\tau} 1_{\{\sigma \geq \tau\}} \geq \frac{1}{\tau} 1_{\{\tau \leq \epsilon\}} \notin L^1,$$

*because*

$$\mathbb{E}\left( \frac{1}{\tau} 1_{\{\tau \leq \epsilon\}} \right) \overset{\tau \sim Exp(1)}{=} \int_0^\epsilon \frac{1}{u} e^{-u} du = +\infty.$$

*This ends the example.*

We now come to our main result (a kind of converse to Lemma 5.1):

**Theorem 5.1** (FTAP, Delbaen-Schachermayer)**.** *Suppose $S = (S_t)_{0 \le t \le T}$ is a $\mathbb{R}^d$-valued semimartingale. Then the following are equivalent:*

1. *$S$ satisfies (NFLVR).*
2. *$S$ admits an equivalent separating measure.*
3. *$S$ admits an equivalent $\sigma$-martingale measure.*

**Outline of the proof**:

**Exercise 5.4.** *"3) $\Longrightarrow$ 1) is easy and goes essentially like the proof of Lemma 5.1.*

"1) $\Longrightarrow$ 2)" is conceptually similar to proof of 2.2 (DMW; Dalang-Morton-Willinger). We want to use Kreps-Yan Theorem for $p = \infty$, and so first step (which is hard) shows that if we have (NFLVR), then $\mathcal{C}^\infty \subset L^\infty$ is weak* closed in $L^\infty$. This is both difficult and remarkable, see later.

Second step applies Kreps-Yan to $\mathcal{C}^\infty$ to get $Q \approx \mathbb{P}$ on $\mathcal{F}_T$ with $\mathbb{E}_Q(Y) \le 0$, for all $Y \in \mathcal{C}^\infty$. We want $\mathbb{E}_Q(g) \le 0$ for all $g = g_T(\vartheta) \in \mathcal{G}_{\mathrm{adm}}$. For that, take $g \in \mathcal{G}_{\mathrm{adm}}$ and write

$$L^\infty \ni g \wedge n = g - (g - g \wedge n) \in \mathcal{G}_{\mathrm{adm}} - L^0_+$$

to see that $g \wedge n \in \mathcal{C}^\infty$. So $\mathbb{E}_Q(g \wedge n) \le 0$ for all $n \in \mathbb{N}$ and $g \ge$ constant, so that Fatou gives $\mathbb{E}_Q(g) \le 0$ for all $g \in \mathcal{G}_{\mathrm{adm}}$. Thus $Q$ is an ESM.

"2) $\Longrightarrow$ 3)" If $S$ is locally bounded, any ESM is an ELMM, hence an E$\sigma$MM, as remarked (Remark 5.5) But if $S$ is unbounded, an ESM need not be an E$\sigma$MM, and we can even have $\mathcal{G}_{\mathrm{adm}} = \{0\}$. So $\mathbb{E}_Q(g) \le 0$ for all $g \in \mathcal{G}_{\mathrm{adm}}$ does not give a lot of information.

**But**, one can show that the set of all E$\sigma$MM is dense in set of ESM if the latter is not the empty set. This is quite technical, uses semimartingale characterisation. See Delbaen-Schachermayer Section 8.3. (outline) and Sections 14.3 and 14.4 (details).

(End of outline)

Key step in proof "1) $\implies$ 2)" is

**Theorem 5.2.** *If $S$ satisfies (NFLVR), then $\mathcal{C}^\infty = (\mathcal{G}_{adm} - L^0_+) \cap L^\infty$ is weak*-closed in $L^\infty$.*

**Remark 5.6.** Obviously we have the following inclusions

$$\mathcal{C}^\infty \subset \overline{\mathcal{C}^\infty}^{L^\infty} \subset \overline{\mathcal{C}^\infty}^{\sigma(L^\infty, L^1)},$$

and as a consequence of Theorem 5.2 we get equality throughout.

Outline of main steps proving Theorem 5.2.

**Notation**: Write $\mathcal{C}^0_{\mathrm{adm}} := \mathcal{G}_{\mathrm{adm}} - L^0_+$, so that $\mathcal{C}^\infty = \mathcal{C}^0_{\mathrm{adm}} \cap L^\infty$.

**Step 1**: A result from functional analysis says that a convex set $C \subset L^\infty$ is weak*-closed iff for any uniformly bounded sequence in $C$ which converges $\mathbb{P}$-a.s, the limit is still in $C$ (Uses the Krein-Smulian theorem).

**So**: it is enough to show that any uniformly bounded sequence in $\mathcal{C}^0_{\mathrm{adm}}$ which converges $\mathbb{P}$-a.s. has its limit still in $\mathcal{C}^0_{\mathrm{adm}}$.

**Definition 5.5.** *$A \subset L^0$ is said to be **Fatou-closed** if any sequence $A$ which is uniformly bounded from below and $\mathbb{P}$-a.s. convergent has its limit still in $A$. If $A$ is a cone, it is enough to check this for the uniform lower bound $-1$.*

**Step 2**: **Goal**: Show that $\mathcal{C}^0_{\mathrm{adm}}$ is Fatou-closed. We know that $\mathcal{C}^0_{\mathrm{adm}}$ is a convex cone. Take $(f_n)_{n\in\mathbb{N}} \subset \mathcal{C}^0_{\mathrm{adm}}$ with $f_n \geq -1$ and $f_n \to f$ $\mathbb{P}$-a.s. Then $-1 \leq f_n \leq g_n$ with $g_n = G_T(\vartheta^n)$ where $\vartheta^n \in \Theta_{\mathrm{adm}}$.

We have (NFLVR), hence by Proposition 5.2 also (NA) and so $G_.(\vartheta^n) \geq -\mathrm{const.}$ plus we know that $G_T(\vartheta^n) \geq -1$, so using (NA) implies $G_.(\vartheta^n) \geq -1$, so that $\vartheta^n \in \Theta^1$ and $g_n \in \mathcal{G}^1$. Use Lemma 5.3 (Komlos) to get $\widetilde{g}_n \in \mathrm{conv}(g_n, g_{n+1}, \dots) \subset \mathcal{G}^1$ with $\widetilde{g}_n \to \widetilde{g}_\infty$ $\mathbb{P}$-a.s. and of course $\widetilde{g}_\infty \geq -1$. Moreover, $g_n \geq f_n$ and $f_n \to f$ $\mathbb{P}$-a.s. implies that also $\widetilde{g}_\infty \geq f$. So we conclude that

$$\widetilde{g}_\infty \in \mathcal{D}_f := \{g \in L^0 : g \geq f\} \cap \overline{\mathcal{G}^1}^{L^0}.$$

**If** $\mathcal{G}^1$ were closed in $L^0$, or **if** we could show $\widetilde{g}_\infty = G_T(\vartheta)$ for some $\vartheta \in \Theta_{\mathrm{adm}}$ then we should have

$$f = G_T(\vartheta) - (\widetilde{g}_\infty - f) \in \mathcal{G}_{\mathrm{adm}} - L^0_+ = \mathcal{C}^0_{\mathrm{adm}}.$$

**But**, above properties are not true in general. We only have $\widetilde{g}_n = G_T(\widetilde{\vartheta}^n)$ with $\widetilde{\vartheta}^n \in \Theta_{\mathrm{adm}}$ and $\widetilde{g}_n \to \widetilde{g}_\infty$ $\mathbb{P}$-a.s., and this gives no information on asymptotics of $(\widetilde{\vartheta}^n)_{n\in\mathbb{N}}$ or $G_.(\widetilde{\vartheta}^n)$. So we need to do more work.

**Definition 5.6.** *An element $a \in A \subset L^0$ is said to be* ***maximal in*** *$A$ if $h \in A$ and $h \geq a$ ($\mathbb{P}$-a.s.), then $h = a$ ($\mathbb{P}$-a.s.).*

**Step 3**: We have $\mathcal{D}_f = \{g \in L^0 : g \geq f\} \cap \overline{\mathcal{G}^1}^{L^0} \neq \emptyset$. Moreover, by Proposition 5.2, (NFLVR) implies (NUPBR) i.e. $\mathcal{G}^1$ is bounded in $L^0$.

**Exercise 5.5.** *Show that if $\mathcal{G}^1$ is bounded in $L^0$, then $\overline{\mathcal{G}^1}^{L^0}$ is also bounded in $L^0$.*

So by the above exercise $\mathcal{D}_f$ is also bounded in $L^0$. Moreover, $\mathcal{D}_f$ is also closed in $L^0$. Indeed, suppose $h_n \to h$ in $L^0$ with $(h_n)_{n\in\mathbb{N}} \subset \mathcal{D}_f$. Take subsequence $h_{n_k} \to h$ $\mathbb{P}$-a.s. as $k \to \infty$ (convergence in probability implies convergence on a subsequence almost surely). Such $h_{n_k} \in \overline{\mathcal{G}^1}^{L^0}$ has form $h_{n_k} = \lim_{m\to\infty} g_{k_m}$ $\mathbb{P}$-a.s. with $g_{k_m} \in \mathcal{G}^1$. Use diagonal argument to get $h = \lim_{l\to\infty} g'_l$ $\mathbb{P}$-a.s. with $g'_l \in \mathcal{G}^1$, so $h \in \overline{\mathcal{G}^1}$ and of course $h \geq f$. Thus $h \in \mathcal{D}_f$ and $\mathcal{D}_f$ is closed in $L^0$.

**But now**: every closed bounded $\emptyset \neq A \subset L^0$ has a maximal element by Zorn's lemma: if $B \subset A$ is totally ordered, then esssup$A$ is majorant for $B$ and finite $\mathbb{P}$-a.s. because $A$ is bounded in $L^0$, and it is in $A$ because $A$ is closed in $L^0$. Thus Zorn's lemma applies.

Now take maximal element $h_0 \in \mathcal{D}_f$. Then $h_0 \geq f$ and $h_0 = \lim_{n\to\infty} G_T(\vartheta^n)$ in $L^0$ for sequence $(\vartheta^n)_{n\in\mathbb{N}}$ of 1-admissible integrands, i.e. $(\vartheta^n)_{n\in\mathbb{N}} \subset \Theta^1$. Then $f = h_0 - (h_0 - f)$ yields $f \in \mathcal{C}^0_{\mathrm{adm}}$, **if** we can show that $h_0 \in \mathcal{G}_{\mathrm{adm}}$ i.e. $h_0 = G_T(\vartheta)$ for some $\vartheta \in \Theta_{\mathrm{adm}}$.

**Note**: In contrast to $\widetilde{g}_\infty$ from before, $h_0$ is also **maximal** in $\mathcal{D}_f$.

**Step 4**: If $h_0$ is maximal in $\mathcal{D}_f$ and $h_0 = \lim_{n\to\infty} G_T(\vartheta^n)$ in $L^0$ with $(\vartheta^n)_{n\in\mathbb{N}} \subset \Theta^1$, then $G^*_T(\vartheta^n) = \sup_{0\leq t\leq T} |G_t(\vartheta^n)|$, $n \in \mathbb{N}$ is Cauchy in $L^0$. Therefore $(G_\cdot(\vartheta^n))_{n\in\mathbb{N}} \subset \mathbb{D}$ is Cauchy for $d$ (see Section 4) and hence has limit for $d$.

**Indirect argument**: Let us denote $\widetilde{Y}^*_t := \sup_{0\leq s\leq t} Y_s$. Suppose there are $i_k, j_k \to \infty$ for $k \to \infty$ with $\mathbb{P}(G^*_T(\vartheta^{i_k} - \vartheta^{j_k}) > \alpha) \geq \delta > 0$ for some $\alpha > 0$. Then $\tau_k := \inf\{t \in [0,T] : G_t(\vartheta^{i_k} - \vartheta^{j_k}) > \alpha\} \wedge T$. has $\mathbb{P}(\tau_k > T) \geq \delta$ and let us consider $\widetilde{\vartheta}^k := \vartheta^{i_k} 1_{[\![0,\tau_k]\!]} + \vartheta^{j_k} 1_{(\!(\tau_k,T]\!]}$, then we have $\widetilde{\vartheta}^k \in \Theta^1$. Indeed, $G_\cdot(\widetilde{\vartheta}^k) = G_\cdot(\vartheta^{i_k}) \geq -1$ on $[\![0,\tau_k]\!]$ and for $t > \tau_k$ we have

$$
\begin{aligned}
G_t(\widetilde{\vartheta}^k) &= G_{\tau_k}(\vartheta^{i_k}) + (G_t(\vartheta^{j_k}) - G_{\tau_k}(\vartheta^{j_k})) \\
&= \underbrace{G_t(\vartheta^{j_k})}_{\geq -1} + \underbrace{(G_{\tau_k}(\vartheta^{i_k}) - G_{\tau_k}(\vartheta^{j_k}))}_{\geq \alpha > 0}.
\end{aligned}
$$

So $\widetilde{\vartheta}^k \in \Theta^1$ as claimed.

Moreover,

$$G_T(\widetilde{\vartheta}^*) = 1_{\{\tau_k=T\}}G_T(\vartheta^{i_k}) + 1_{\{\tau_k<T\}}G_T(\vartheta^{j_k}) + \xi_k$$

with $\xi_k = 1_{\{\tau_k<T\}}(G_{\tau_k}(\vartheta^{i_k}) - G_{\tau_k}(\vartheta^{j_k})) \geq 0$ having $\mathbb{P}(\xi_k \geq \alpha) \geq \delta > 0$ for all $k \in \mathbb{N}$. Then use Lemma 5.3 (Komlos) plus convexity of $\mathcal{G}^1$ to get in $\mathcal{D}_f$ an element $h_0+\eta$ with $\eta = \lim_{k\to 0}\widetilde{\xi}_k$ in $L^0_+ \setminus \{0\}$. So $h_0+\eta$ is in $\mathcal{D}_f$ and $h_0+\eta \geq h_0$ and $h_0+\eta \neq h_0$ which contradicts the maximality of $h_0$.

To finish the proof, we need a recent result by Cuchiero/Teichmann ($\approx$ 2016):

**Theorem 5.3** (Cuchiero/Teichmann)**.** *Suppose $S$ satisfies (NUPBR). Assume $(\vartheta^n)_{n\in\mathbb{N}} \subset \Theta^1$ is such that $(G.(\vartheta^n))_{n\in\mathbb{N}}$ converges for $d$ to some process $X \in \mathbb{D}$ such that $X_T$ is maximal in $\overline{\mathcal{G}}^{L^0}$. Then $(G.(\vartheta^n))_{n\in\mathbb{N}} \to X$ for Emery metric $\widetilde{d'_E}$. As a consequence, $X = G(\vartheta)$ for some $\vartheta \in \Theta^1$ and so $X_T \in \mathcal{G}^1$.*

**How does that imply Theorem 5.2**? We need $h_0 \in \mathcal{G}^1$.

From Step 4, $(G(\vartheta^n))_{n\in\mathbb{N}} \subset \mathbb{D}$ with $(\vartheta^n)_{n\in\mathbb{N}} \subset \Theta^1$ is Cauchy for $d$, hence convergent for $d$ to some $X \in \mathbb{D}$. Moreover, $X_T = \lim_{n\to\infty} G_T(\vartheta^n) = h_0 \geq f$ and is maximal in $\overline{\mathcal{G}^1}^{L^0}$. By Theorem 5.3, then $G(\vartheta^n) \to X$ for $\widetilde{d'_E}$. But by Theorem 4.6 (Mémin) space of all Stochastic Integrals of $S$ is closed for $\widetilde{d'_E}$, so $X = G(\vartheta)$ for some $\vartheta \in L(S)$.

But $(\vartheta^n)_{n\in\mathbb{N}} \subset \Theta^1$ gives $G.(\vartheta^n) \geq -1$, and $G.(\vartheta^n) \to X. = G.(\vartheta)$ for $d$, i.e. uniformly in $t$ in probability. Hence also $G.(\vartheta) \geq -1$, meaning $\vartheta \in \Theta^1$, and so $h_0 = X_T = G_T(\vartheta)$ is in $\mathcal{G}^1$. This is exactly what we need to conclude that $\mathcal{C}^0_{\text{adm}}$ is Fatou-closed. $\square$

*Proof of Theorem 5.3.* Also needs more work, omitted here. The main steps can be summarized as follows:

1) Show that (NUPBR) implies that $(\mathcal{G}(\vartheta^n))_{n\in\mathbb{N}}$ satisfies property called (P-UT); this is a kind of boundedness for Emery topology and can be used to control convergence of all parts from decomposition of $\mathcal{G}(\vartheta^n)$ except FV parts.

2) Use maximality to show that also convergence of FV parts holds. $\square$

# 6 No-arbitrage properties in some model classes

## 6.1 (NUPBR) and related results

**Goal**: Study what can be said about models that satisfy (NUPBR).

**Recall**: $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space with filtration $\mathbb{F}$ satisfying the usual conditions, $S = (S_t)_{0 \le t \le T}$ is an $\mathbb{R}^d$-valued semimartingale.

$$\Theta^1 := \left\{ \vartheta \in L(S) : G.(\vartheta) = \vartheta \cdot S = \int \vartheta dS \ge -1 \right\}$$

define

$$\mathcal{X}^1 := 1 + G(\Theta^n) = \left\{ X = 1 + \int \vartheta dS \ge 0 : \vartheta \in L(S) \right\}.$$

This is the space of all wealth processes of self-financing strategies with initial wealth 1 and nonnegative running wealth. This also gives $\mathcal{X}_T^1 = 1 + G_T(\Theta^1) = 1 + \mathcal{G}^1$. Also recall that (NUPBR) means that $\mathcal{G}^1$ is bounded in $L^0$. We also need

$$\mathcal{X}_{++}^1 := \{X \in \mathcal{X}^1 : X > 0 \text{ and } X_- > 0\}$$

Also recall that by definition $\mathcal{G}^1 = G_T(\Theta^1) = \mathcal{H}_T^1 - 1$ is bounded in $L^0$ iff $S$ satisfies (NUPBR).

**Definition 6.1.** ***Equivalent $\sigma$-martingale density*** *(E$\sigma$MD) for $S$ is a local martingale $Z > 0$ with $Z_0 = 1$ such that $ZS$ is a ($\mathbb{P}$-) $\sigma$-martingale. If $ZS$ is even a local martingale we say that $Z$ is ELMD.*

**Remark 6.1.**

1. Suppose $Q \approx \mathbb{P}$ is E$\sigma$MM or ELMM for $S$. Denote by $Z = Z^{Q;\mathbb{P}}$ density process of $Q$ with respect to $\mathbb{P}$. Then $Z > 0$ is ($\mathbb{P}$-) martingale and by Bayes, $ZS$ is ($\mathbb{P}$-) $\sigma$-(respectively local) martingale. We have $Z_0 = 1$ iff $Q = \mathbb{P}$ on $\mathcal{F}_0$ (e.g. if $\mathcal{F}_0$ is trivial). **So**: E$\sigma$MD or ELMD generalises E$\sigma$MM respectively ELMM because $Z$ only needs to be a local martingale.

2. We have the following

   **Exercise 6.1.** *Any ELMD is of course an E$\sigma$MD. If $S$ is also continuous, the converse holds as well.*

**Definition 6.2.** *A* ***Numéraire portfolio*** *is an element $X^{np} \in \mathcal{X}^1_{++}$ such that $X/X^{np}$ is a $\mathbb{P}$-supermartingale for all $X \in \mathcal{X}^1_{++}$.*

**Intuition**: A supermartingale is decreasing on average, so $X/X^{\mathrm{np}}$ supermartingale means loosely that $X^{\mathrm{np}}$ has "better performance" then any other $X \in \mathcal{X}^1_{++}$.

**Exercise 6.2** (Easy). *If $X^{np}$ exists, it is* ***unique****.*

**Remark 6.2.** Suppose $Z$ is E$\sigma$MD or ELMD for $S$. Take any $\vartheta \in L(S)$ and compute $ZG(\vartheta)$. This is then a Stochastic integral of a local ($\mathbb{P}$-) martingale. So if $\vartheta \in \Theta_{\mathrm{adm}}$, then $ZG(\vartheta)$ has local integrable lower bound and is therefore by Lemma 5.2 a $\mathbb{P}$-supermartingale. A $Z$ with this property (plus $Z_0 = 1,\ Z > 0$ and $Z$ local martingale) is sometimes called **(equivalent/strict) supermartingale deflator** for $S$ (or better for $G(\Theta_{\mathrm{adm}})$). **If** in addition $1/Z$ is in $\mathcal{X}^1_{++}$, then $1/Z := X^{\mathrm{np}}$ is the numéraire portfolio.

We now give some general connections:

**Theorem 6.1.** *For a $\mathbb{R}^d$-valued semimartingale $S = (S_t)_{0 \le t \le T}$ are equivalent:*

1. *$S$ satisfies (NUPBR).*
2. *There exists an E$\sigma$MD for $S$.*
3. *There exists a numéraire portfolio $X^{np}$ for $S$.*

*Proof.* "1) $\iff$ 2)" Takaoka, Theorem 2.6.
"2) $\iff$ 3)" Karatzas/Kardaras Theorem 4.12. □

Let us focus on the case where $S$ is continuous.

## 6.2 The continuous case

In this section, consider the case where $S = (S_t)_{0 \le t \le T}$ is an $\mathbb{R}^d$-valued adapted **continuous process**.

**Proposition 6.1.** *Suppose $S$ is continuous. If $S$ admits an E$\sigma$MD $Z$ [and in particular, if $S$ satisfies (NFLVR)], then $S$ satisfies the* ***structure condition*** *(SC) i.e., $S$ is a continuous semimartingale which can be written as $S = S_0 + M + A$ with $M \in \mathcal{M}^c_{0,loc}$ and $A \in cFV_0$ of the form $A = \int d\langle M\rangle\lambda$ with a predictable $\mathbb{R}^d$-valued $\lambda \in L^2_{loc}(M)$ (i.e., $\int \lambda^{tr} d\langle M\rangle\lambda$ is finite-valued).*

**Remark 6.3.** Loosely speaking: $A \ll \langle M\rangle$ with $\lambda = \frac{dA}{d\langle M\rangle} \in L^2_{\mathrm{loc}}(M)$.

*Proof.* $S = \frac{1}{Z}(ZS)$ and $ZS$ is $\sigma$-martingale, hence a semimartingale and $\frac{1}{Z}$ is a semimartingale by Itô's formula. More carefully: $Z > 0$ is local martingale, hence supermartingale, and so $Z_- > 0$ by minimum principle (BMSC). Moreover, can also argue that $\sigma$-martingale $ZS$ is actually local martingale (by Lemma 5.2 because $S$ is continuous) hence a semimartingale. As $Z_- > 0$, we can use Itô's formula and conclude that $1/Z$ is a semimartingale.

Write $S = S_0 + M + A$, $M \in \mathcal{M}^c_{0,\mathrm{loc}}$, $A \in \mathrm{cFV}_0$. Now look at local martingale $ZS$. Write $Z = \mathcal{E}(N)$ with

$$N := \int \frac{1}{Z_-} dZ \in \mathcal{M}_{0,\mathrm{loc}}.$$

Use generalised version of (Galtchonk-) Kunita-Watanabe decomposition to get

$$N = \int \vartheta dM + L$$

with $\vartheta \in L^2_{\mathrm{loc}}(M)$, $L \in \mathcal{M}_{0,\mathrm{loc}}$ and strongly orthogonal to $M$. (Detail: $N = N^c + N^d$ with $N^c, N^d \in \mathcal{M}_{0,\mathrm{loc}}$ and $N^c$ continuous and $N^d \perp$ to all continuous local martingales).

$L \perp M$ is equivalent (by the product rule) to $[L, M] \in \mathcal{M}_{0,\mathrm{loc}}$ and because $M$ is continuous, $[L, M] = \langle L, M\rangle$ and so $[L, M] = \langle L, M\rangle \equiv 0$ by local martingale property. Now compute:

$$\begin{aligned}\underbrace{ZS - Z_0 S_0}_{\text{local mart.}} &\overset{\text{Itô}}{=} \int Z_- dS + \int S_- dZ + [Z, S] \\ &= \left(\int Z_- dM + \int S_- dZ\right) + \left(\int Z_- dA + \langle Z, S\rangle\right).\end{aligned}$$

Next,

$$\langle Z, S\rangle = \langle \int Z_- dN, M\rangle = \int Z_- d \underbrace{\langle N, M\rangle}_{=\int d\langle M\rangle \nu} = \int Z_- d\langle M\rangle \nu.$$

So $\int Z_-(dA + d\langle M\rangle\nu)$ is local martingale and predictable, hence continuous (predictable Optional Stopping Theorem), and of FV, which gives that $\int Z_-(dA + d\langle M\rangle\nu) \equiv 0$. So $A + \int d\langle M\rangle\nu \equiv \text{const} = 0$ and so we can take $\lambda := -\nu$ to get the claim. □

Corollary of preceding proof is a parametrization of all E$\sigma$MDs(=ELMDs) for a given **continuous** $S$. In view of Proposition 6.1 we can assume w.l.o.g. that $S$ satisfies (SC), since this is necessary for existence of E$\sigma$MD or ELMD.

**Corollary 6.1.** *Suppose $S$ is continuous and satisfies (SC). Then:*

1. *A local martingale $Z > 0$ with $Z_0 = 1$ is E$\sigma$MD and ELMD for $S$ iff it has form $Z = \mathcal{E}(-\int \lambda dM)\mathcal{E}(L)$ for some $L \in \mathcal{M}_{0,loc}$ strongly orthogonal to $M$ with $\Delta L > -1$. In particular, $\hat{Z} := \mathcal{E}(-\int \lambda dM)$ is always an E$\sigma$MD and ELMD.*

2. *$Q \approx \mathbb{P}$ on $\mathcal{F}_T$ is an E$\sigma$MM and ELMM for $S$ iff its density process $Z := Z^{Q;\mathbb{P}}$ has form $Z = Z_0\hat{Z}\mathcal{E}(L)$ with $L$ as in 1) and such that product is a true $\mathbb{P}$-martingale, strictly positive on $[0, T]$.*

*Proof.* $S$ continuous implies that E$\sigma$MD=ELMD and E$\sigma$MM=ELMM and 2) easily follows from 1).

1) Proof of Proposition 6.1 implies that any E$\sigma$MD $Z$ for $S$ has form $Z = \mathcal{E}(N) = \mathcal{E}(-\int \lambda dM + L)$ and this equals $\mathcal{E}(-\int \lambda dM)\mathcal{E}(L)$ by Yor's formula because $[L, M] = \langle L, M\rangle \equiv 0$ due to $L \perp M$. So $\mathcal{E}(L) > 0$ because $Z > 0$ and $\hat{Z} = \mathcal{E}(-\int \lambda dM) > 0$ (because $M$ is continuous) and because $\Delta(\mathcal{E}(L)) = \Delta(\int \mathcal{E}(L)_- dL) = \mathcal{E}(L)_-\Delta L$, we get (because $\Delta(\mathcal{E}(L)) = \mathcal{E}(L) - \mathcal{E}(L)_-$), that

$$\underbrace{\mathcal{E}(L)}_{>0} = \underbrace{\mathcal{E}(L)_-}_{>0}\underbrace{(1 + \Delta L)}_{\Longrightarrow >0}$$

and so $\Delta L > -1$. This gives " $\Longrightarrow$ ".

For "$\Longleftarrow$" write $Z = \hat{Z}\mathcal{E}(L)$, then $Z = \mathcal{E}(-\int \lambda dM + L)$ by Yor, and then do same computation as in proof of Proposition 6.1 to find

$$ZS - Z_0S_0 = \int Z_- dM + \int S_- dZ + 0$$

so that $ZS$ is a local martingale, i.e. $Z$ is an ELMD. $\square$

**Names and terminology**: If $S$ is continuous with (SC), so $S = S_0 + M + \int d\langle M\rangle\lambda$. The process $\hat{Z} := \mathcal{E}(-\int \lambda dM)$ is called **minimal ELMD** for $S$. If it is a true martingale, strictly positive, then corresponding ELMM $\hat{\mathbb{P}}$ with $\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} := \hat{Z}_T$ is called **minimal ELMM** for $S$ (or shortly **minimal martingale measure**).

For continuous $S$, $\hat{\mathbb{P}}$ as above, meaning that

$$\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} := \mathcal{E}\left(-\int \lambda dM\right)_T$$

is natural candidate for ELMM (if it exists, as probability measure) and then it also has other useful properties. If $S$ has jumps but still satisfies (SC) (e.g. with $M \in \mathcal{M}^2_{0,\mathrm{loc}}$) one must be more careful.

The process $\hat{K} := \int \lambda^{\mathrm{tr}} d\langle M\rangle\lambda = \langle \int \lambda dM\rangle$ is often called **mean-variance trade-off** (MVT) process of $S$. Why? For $d = 1$ we have $A = \int \lambda d\langle M\rangle$ and then

$$\hat{K} = \int \lambda^2 d\langle M\rangle = \int \left(\frac{dA}{d\langle M\rangle}\right)^2 d\langle M\rangle = \int \frac{(dA)^2}{d\langle M\rangle} = \int \frac{(\mathbb{E}[dS_t \mid \mathcal{F}_t])^2}{\mathrm{Var}[dS_t \mid \mathcal{F}_{t-}]}$$

which explains the name.

**Remark 6.4.**

1. In particular continuous adapted process admits an E$\sigma$MD/ELMD iff $S$ satisfies (SC).

2. If $\hat{Z} > 0$ is a true $\mathbb{P}$-martingale and $Q$ is any E$\sigma$MM/ELMM, then in representation 2) from Corollary 6.1 process $Z_0\mathcal{E}(L)$ is $\hat{\mathbb{P}}$-martingale.

**A large model class: Itô processes**:

Start with $\mathbb{R}^n$-valued BM $W$ on general $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ (filtration satisfying the usual conditions). Consider **undiscounted prices**: $\widetilde{S}^0 = \widetilde{B}$ and $\widetilde{S} = (\widetilde{S}^i)_{i=1,\dots,d}$ given by

$$
\begin{aligned}
d\widetilde{B}_t &= \widetilde{B}_t r_t dt, \quad \widetilde{B}_0 = 1 \\
d\widetilde{S}^i_t &= \widetilde{S}^i_t \mu^i_t dt + \widetilde{S}^i_t \sum_{j=1}^n \sigma^{ij}_t dW^j_t, \quad \widetilde{S}^i_0 = s^i_0 > 0.
\end{aligned}
$$

with predictable, suitable integrable coefficients $r, \mu^j, \sigma^{ij}$.

**Discounted prices**: $S^i = \widetilde{S}^i / \widetilde{B}$ satisfy

$$
dS^i_t = S^i_t b_t dt + S^i_t \sum_{j=1}^n \sigma^{ij}_t dW^j_t, \quad S^i_0 = s_0 > 0,
$$

with $b^i = \mu^i - r$. More compactly:

$$
dS_t = \operatorname{diag}(S_t)(b_t dt + \sigma_t dW_t)
$$

with $b$ $\mathbb{R}^d$-valued, $\sigma$ $\mathbb{R}^{d\times n}$-valued, $b = \mu - r1^d$ where $1^d$ denotes the $d$-dimensional vector saturated by 1's.

**Remark 6.5.** If $S$ satisfies (SC) and if $\mathbb{F} = \mathbb{F}^W$ and $S > 0$, then (perhaps up to some integrability) $S$ must be an Itô process.

Write $S = S_0 + M + A$, then

$$
\begin{aligned}
dA_t &= \operatorname{diag}(S_t) b_t dt, \\
dM_t &= \operatorname{diag}(S_t) \sigma_t dW_t.
\end{aligned}
$$

Hence,

$$
d\langle M \rangle_t = \operatorname{diag}(S_t) \sigma_t \sigma_t^{\mathrm{tr}} \operatorname{diag}(S_t) dt.
$$

Assume $d \le n$ (that is, more sources of uncertainty than risky assets) and $\operatorname{rank}(\sigma_t) \equiv d$ (i.e. full rank), so that $\sigma_t \sigma_t^{\mathrm{tr}}$ is invertible for all $t$. Define (the $\mathbb{R}^n$-valued)

$$
\overline{\lambda}_t := \sigma_t^{\mathrm{tr}} (\sigma_t \sigma_t^{\mathrm{tr}})^{-1} b_t
$$

in order to see that

$$dS_t = \mathrm{diag}(S_t)\sigma_t(\overline{\lambda}_t dt + dW_t),$$

and

$$\begin{aligned} dA_t &= \mathrm{diag}(S_t)\sigma_t\overline{\lambda}_t dt \\ &= \mathrm{diag}(S_t)(\sigma_t\sigma_t^{\mathrm{tr}})\mathrm{diag}(S_t) \times \mathrm{diag}(S_t)^{-1}(\sigma_t\sigma_t^{\mathrm{tr}})^{-1}b_t dt = d\langle M\rangle_t\lambda_t \end{aligned}$$

with $\mathbb{R}^d$-valued process

$$\begin{aligned} \lambda_t &:= \mathrm{diag}(S_t)^{-1}(\sigma_t\sigma_t^{\mathrm{tr}})^{-1}b_t \\ &= \mathrm{diag}(S_t)^{-1}(\sigma_t\sigma_t^{\mathrm{tr}})^{-1}(\mu_t - r_t 1^d). \end{aligned}$$

Note that

$$\int \lambda dM = \int b_t^{\mathrm{tr}}(\sigma_t\sigma_t^{\mathrm{tr}})^{-1}\sigma_t dW_t = \int \overline{\lambda} dW$$

and therefore

$$\hat{K} = \int \lambda_t^{\mathrm{tr}} d\langle M\rangle_t\lambda_t = \langle \int \lambda dM\rangle = \langle \int \overline{\lambda} dW\rangle = \int |\overline{\lambda}_t|^2 dt.$$

So $S$ satisfies (SC) iff this process is finite-valued. In that case, E$\sigma$MDs/ELMDs (they are the same since we're in the continuous case) $Z$ for $S$ are parametrized as

$$Z = Z_0\hat{Z}\mathcal{E}(L) = Z_0\mathcal{E}\left(-\int \overline{\lambda} dW\right)\mathcal{E}(L)$$

with $L \in \mathcal{M}_{0,\mathrm{loc}}$ strongly orthogonal to $M$, i.e. to $\int \sigma dW$, and $\Delta L > -1$.

So far, $\mathbb{F}$ has been general. If $\mathbb{F}$ is ($\mathbb{P}$-augmentation of raw filtration) generated by $W$, $\mathbb{F} = \mathbb{F}^W$, then we can say more.

**Lemma 6.1.** *Suppose Itô process-model has $\mathbb{F} = \mathbb{F}^W$, $M = \int diag(S)\sigma dW$ and $L \in \mathcal{M}_{0,loc}$. Then we have always $\Delta L > -1$, and $L$ is strongly orthogonal to $M$ iff $L = \int \nu dW$ with $\nu \in L^2_{loc}(W)$ and $\sigma\nu \equiv 0$. As a consequence, E$\sigma$MDs/ELMDs are parametrized by*

$$Z = \hat{Z}\mathcal{E}\left(\int \nu dW\right) = \mathcal{E}\left(-\int (\overline{\lambda} - \nu) dW\right)$$

*with $\overline{\lambda} = \sigma^{tr}(\sigma\sigma^{tr})^{-1}b$ and $\sigma\nu \equiv 0$.*

*Proof.* $\mathcal{F}_0$ is trivial so that $Z_0 = 1$. $\mathbb{F}^W$ has the representation property (Itô's representation theorem), so that any $L \in \mathcal{M}_{0,\mathrm{loc}}(\mathbb{F}^W)$ is continuous and of the form $L = \int \nu dW$ with $\nu$ ($\mathbb{R}^d$-valued) in $L^2_{\mathrm{loc}}(W)$. Moreover, $L \perp M$ iff $\langle L, M\rangle \equiv 0$, and

$$\langle L, M\rangle = \int \operatorname{diag}(S)\sigma\nu dt.$$

This gives the result. □

If we also suppose that $d = n$, then $\sigma$ is a $d \times d$-matrix and so it has full rank $d$ iff it is invertible or, equivalently, its kernel is just 0. But then the only E$\sigma$MD/ELMD for $S$ is

$$\hat{Z} = \mathcal{E}\left(-\int \lambda dM\right) = \mathcal{E}\left(-\int \overline{\lambda} dW\right) = \mathcal{E}\left(-\int \sigma^{-1} b dW\right).$$

So there is at most one candidate for E$\sigma$MM/(ELMM, namely $\hat{\mathbb{P}}$, it will be an ELMM iff $\hat{Z}$ is a true $\mathbb{P}$-martingale.

**Example 6.1** (Black-Scholes model)**.** *Itô process model with $d = 1 = n$, constant coefficients $r, \mu, \sigma$ and $\mathbb{F} = \mathbb{F}^W = \mathbb{F}^S$. So*

$$dS_t = S_t((\mu - r)dt + \sigma dW_t),$$

*the minimal ELMD is*

$$\hat{Z} = \mathcal{E}\left(-\int \frac{\mu - r}{\sigma} dW\right) = \mathcal{E}\left(-\frac{\mu - r}{\sigma} W\right)$$

*and hence a true martingale, and so $S$ admits $\hat{\mathbb{P}}$ as ELMM, hence it is arbitrage-free (satisfies (NFLVR)). In fact:*

$$dS_t = S_t\sigma \underbrace{\left(\frac{\mu - r}{\sigma} dt + dW_t\right)}_{=: d\widehat{W}_t}$$

*where $\widehat{W}$ is by Girsanov's theorem a $\hat{\mathbb{P}}$-BM. So $dS_t = S_t\sigma d\widehat{W}_t$ gives $S = s_0\mathcal{E}(\sigma\widehat{W})$ and hence $S$ is actually a $\hat{\mathbb{P}}$-martingale and consequently $\hat{\mathbb{P}}$ is even an EMM for $S$.*

Now return to the case of general $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and $\mathbb{R}^d$-valued adapted **continuous** process $S = (S_t)_{0\leq t\leq T}$. We know: $S$ admits E$\sigma$MD/ELMD iff $S$ satisfies (SC), and in that case $\hat{Z} = \mathcal{E}(-\int \lambda dM)$ is an E$\sigma$MD/ELMD. Using continuity of $S$ and $A = \int d\langle M\rangle\lambda$ by (SC) gives

$$\begin{aligned}
1/\hat{Z} = 1/\mathcal{E}\left(-\int \lambda dM\right) &= \exp\left(\int \lambda dM - \frac{1}{2}\langle\int \lambda dM\rangle\right)\\
&= \exp\Big(\int \lambda dM + \underbrace{\int \lambda^{\mathrm{tr}} d\langle M\rangle}_{=\int \lambda dA} - \frac{1}{2}\int \lambda^{\mathrm{tr}} d\underbrace{\langle M\rangle}_{=\langle S\rangle}\lambda\Big)\\
&= \exp\left(\int \lambda dS - \frac{1}{2}\langle\int \lambda dS\rangle\right) = \mathcal{E}\left(\int \lambda dS\right)\\
&= 1 + \int \mathcal{E}\left(\int \lambda dS\right)\lambda dS.
\end{aligned}$$

So $1/\hat{Z} = 1 + \int \vartheta dS$ for some $\vartheta \in L(S)$ and $1/\hat{Z} > 0$, $(1/\hat{Z})_- > 0$ (by continuity), so $1/\hat{Z} \in \mathcal{X}^1_{++}$.

**Corollary 6.2.** *If $S$ is continuous and satisfies (SC), then numéraire portfolio $X^{np}$ exists and equals $1/\hat{Z}$.*

*Proof.* $\hat{Z}$ is an E$\sigma$MD and $1/\hat{Z} \in \mathcal{X}^1_{++}$. Now see remark before Theorem 6.1. $\square$

Converse of Corollary 6.2 is:

**Proposition 6.2.** *Suppose $S$ is a continuous semimartingale and numéraire portfolio $X^{np}$ exists. Then $S$ satisfies (SC).*

*Proof.* **Quick argument**: By Theorem 6.1, existence of $X^{\mathrm{np}}$ implies existence of an E$\sigma$MD, and this implies (SC) by Proposition 6.1. But we did not prove Theorem 6.1 in general.

**Direct proof** (using the continuity of $S$): For ease of notation, work with $d = 1$. Start with $X \in \mathcal{X}^1_{++}$, hence $X$ is of the form $X = 1 + \vartheta \cdot S$ with $X > 0$ and $X_- > 0$. Now write

$$X = 1 + \vartheta \cdot S = 1 + \Big(X_- \underbrace{\frac{\vartheta}{X_-}}_{=:\pi}\Big) \cdot S = 1 + X_- \cdot (\pi \cdot S)$$

with $\pi = \frac{\vartheta}{X_-} \in L(S)$. So $X = \mathcal{E}(\pi \cdot S)$ (by unique solution to the SDE $dX = X_- d(\pi \cdot S)$) and for any other $\overline{X} \in \mathcal{X}^1_{++}$, we also have $\overline{X} = \mathcal{E}(\overline{\pi} \cdot S)$. [Intuition: $\vartheta$ is number of shares, so $\pi$ is a bit like fractions of wealth].

Now use continuity of $S = S_0 + M + A$ to compute

$$\begin{aligned}\frac{X}{\overline{X}} = \frac{\mathcal{E}(\pi \cdot S)}{\mathcal{E}(\overline{\pi} \cdot S)} &= \exp\left((\pi - \overline{\pi}) \cdot S - \frac{1}{2}\pi^2 \cdot \langle S \rangle + \frac{1}{2}\overline{\pi}^2 \cdot \langle S \rangle\right) \\ &= \exp\left((\pi - \overline{\pi}) \cdot M - \frac{1}{2}(\pi - \overline{\pi})^2 \cdot \langle M \rangle\right) \\ &\quad \times \exp\left((\pi - \overline{\pi}) \cdot A + \frac{1}{2}(\pi - \overline{\pi})^2 \cdot \langle M \rangle - \frac{1}{2}(\pi^2 - \overline{\pi}^2) \cdot \langle M \rangle\right).\end{aligned}$$

Factor 1 is $\mathcal{E}((\pi - \overline{\pi}) \cdot M)$ and hence a local martingale like $M$. Factor 2 is FV and predictable.

Now take $\overline{X} = X^{\mathrm{np}}$ which exists by assumption. Then $X/\overline{X} = X/X^{\mathrm{np}}$ is a supermartingale for all $X \in \mathcal{X}^1_{++}$. So above product is a supermartingale (and strictly positive) for any choice of $\pi = \frac{\vartheta}{X_-}$. But any strictly positive supermartingale (and with strictly positive left limits) has a unique multiplicative decomposition as product of a strictly positive local martingale and a predictable decreasing process. So factor 2 must always be decreasing for $\overline{X} = X^{\mathrm{np}}$ which means that

$$(\pi - \overline{\pi}) \cdot A + \underbrace{\frac{1}{2}(\pi - \overline{\pi})^2 \cdot \langle M \rangle - \frac{1}{2}(\pi^2 - \overline{\pi}^2) \cdot \langle M \rangle}_{= (-\pi\overline{\pi} + \overline{\pi}^2) \cdot \langle M \rangle = (\overline{\pi}(\overline{\pi} - \pi)) \cdot \langle M \rangle} = (\pi - \overline{\pi}) \cdot (A - \overline{\pi} \cdot \langle M \rangle)$$

must be decreasing for all $\pi$ (resulting from some $X \in \mathcal{X}^1_{++}$). But this implies that $A - \overline{\pi} \cdot \langle M \rangle \equiv 0$ (since $\pi - \overline{\pi}$ can be made both, positive and negative) or $A = \overline{\pi} \cdot \langle M \rangle$; moreover $\overline{\pi}$ is $L^2_{\mathrm{loc}}(M)$ by continuity of $M, S, A$ and so this means that we have (SC). $\square$

**Corollary 6.3.** *For a $\mathbb{R}^d$-valued* ***continuous*** *semimartingale $S = (S_t)_{0 \le t \le T}$ the following are equivalent:*

1. *$S$ admits E$\sigma$MD/ELMD.*
2. *$S$ satisfies (SC).*
3. *There exists a numéraire portfolio $X^{np}$.*
   - *In addition, we then have $X^{np} = 1/\hat{Z}$.*

To round this subsection off, we prove

**Lemma 6.2.** *Suppose $S$ is adapted $\mathbb{R}^d$-valued* ***RCLL*** *process. If $S$ admits an E$\sigma$MD, then $S$ is a semimartingale and satisfies (NUPBR). [In particular, this applies if $S$ is an E$\sigma$MM].*

**Remark 6.6.** Looking back at Theorem 6.1, Lemma 6.2 is actually "2) $\implies$ 1)" there. Converse is more difficult.

*Proof of Lemma 6.2.* Semimartingale property of $S = \frac{1}{Z}(ZS)$ for an E$\sigma$MD $Z$ follows as shown in Proposition 6.1. Moreover, as seen before Theorem 6.1 (or exercise), we know that $ZX$ is a supermartingale for any $X \in \mathcal{X}^1$, i.e. $X = 1 + \int \vartheta dS \geq 0$.

So we get

$$\mathbb{E}(Z_T X_T) \leq \mathbb{E}(\underbrace{Z_0}_{=1}\underbrace{X_0}_{=1}) = 1, \text{ for all } X \in \mathcal{X}^1$$

and this shows that

$$Z_T \mathcal{G}^1 = Z_T \mathcal{X}^1_T = \{Z_T g : g \in \mathcal{G}^1\}$$

is bounded in $L^1$, hence also in $L^0$ and because $Z_T > 0$ almost surely, also $\mathcal{G}^1$ is then bounded in $L^0$ as the exercise below establishes.

**Exercise 6.3.** *Show that since $Z_T \mathcal{G}^1$ is bounded in $L^0$ and $Z_T > 0$ a.s., also $\mathcal{G}^1$ is bounded in $L^0$.*

Since $\mathcal{G}^1$ is bounded in $L^0$, this is exactly (NUPBR). □

### 6.2.1 Recap

We had a more in-depth look at (NUPBR). We have seen (without proof) an advanced result from Takaoka/Schweizer namely Theorem 6.1. This result states that for a general (RCLL) semimartingale $S$; $S$ satisfies (NUPBR), there exists a E$\sigma$MD for $S$, there exists a **numéraire portfolio** $X^{\text{np}}$ for $S$, are all equivalent statements.

We then switch from the RCLL workframe to the **continuous** one. We shown that if $S$ admits an E$\sigma$MD, then $S$ must necessarily satisfy the **structure condition** (SC) i.e. it can be written as $S = S_0 + M + \int d\langle M \rangle \lambda$ with $M \in \mathcal{M}^c_{0,\text{loc}}$ and $\lambda$ is predictable $\lambda \in L^2_{\text{loc}}(M)$. We characterized the structure for a $Z > 0$ with $Z_0 = 1$ to be an E$\sigma$MD for a continuous $S$ that satisfies (SC) as $Z = \mathcal{E}(-\int \lambda dM)\mathcal{E}(L)$ (with some properties). We then arrived at that in the continuous setting $S$ admits E$\sigma$MD=ELMD, $S$ satisfies (SC), there exists a numéraire portfolion $X^{\text{np}}$ are all equivalent. Finally, we discussed "2) $\implies$ 1)" in the RCLL setting in Thm 6.1.

## 6.3 Lévy models

To model price processes with jumps, Lévy processes provide a large and fairly tractable class.

**Recall**: $\mathbb{R}^d$-valued Lévy process $L = (L_t)_{0\le t\le T}$ has independent stationary increments, is RCLL and starts at 0.

We then have the **Lévy-Khinchin representation**

$$\mathbb{E}(e^{iu^{\mathrm{tr}}L_t}) = e^{t\Psi(u)}, \text{ for } u \in \mathbb{R}^d$$

with the **characteristic exponent**

$$\Psi(u) = ib^{\mathrm{tr}}u - \frac{1}{2}u^{\mathrm{tr}}\Sigma u + \int_{\mathbb{R}^d}\left(e^{iu^{\mathrm{tr}}x} - 1 - iu^{\mathrm{tr}}1_{\{|x|\le 1\}}\right)\nu(dx)$$

for Lévy triplet $(b, \Sigma, \nu)$. This holds for all $iu \in \mathbb{C}^d$ with $u \in \mathbb{R}^d$ and extends, for $d = 1$ to $z \in \mathbb{C}$ with $0 \le \mathrm{Re} z < \alpha$ if $\mathbb{E}(e^{\alpha L_t}) < \infty$.

We now study what happens if we combine the Lévy property with some kind of martingale property. For ease of notation, set $\kappa(\cdot) := \Psi(-i)$ so that $\mathbb{E}(e^{z^{\mathrm{tr}}L_t}) = e^{t\kappa(z)}$.

**Proposition 6.3.** *Suppose $L$ is a Lévy process and $S = S_0 \exp(L)$ [taken component wise] with $S_0 \in \mathbb{R}^d_+$ fixed is a $\sigma$-martingale. Then $S$ is a true martingale. [This holds on $[0,T]$ or also on $[0,\infty)$, but not on $[0,\infty]$].*

*Proof.* Coordinates of $L$ are all Lévy processes (but not necessarily independent processes), it is enough to argue for $d = 1$. W.l.o.g. let $S_0 = 1$.

Now $S = e^L$ is a $\sigma$-martingale and it is nonnegative, thus it is a local martingale and a supermartingale by Ansel-Stricker. So for any $t \in [0,T]$ we have $\mathbb{E}(S_t) \le \mathbb{E}(S_0) = 1$ and so $S$ will be a martingale on $[0,T]$, if $\mathbb{E}(S_T) = 1$ (every supermartingale with constant expectation is a true martingale).

We have $\mathbb{E}(e^{L_1}) = \mathbb{E}(S_1) < \infty$ so that $c := \log \mathbb{E}(e^{L_1})$ is well-defined and because $L$ is Lévy, this implies that $\mathbb{E}(e^{L_t}) = e^{ct}$ for all $t$. Again using the Lévy property of $L$, this implies that for any $T$ that

$$(S_t e^{-tc})_{0\le t\le T} = (e^{L_t - tc})_{0\le t\le T}$$

is a martingale on $[0,T]$. So it is of class (D) on $[0,T]$, which gives that also $S$ itself is of class (D) on $[0,T]$ and so $S$ is actually a true martingale on $[0,T]$. $\square$

**Remark 6.7.** Again assume $d = 1$ and $e^L$ is a $\sigma$-martingale. Then $\mathbb{E}(e^{L_1}) < \infty$ and we get $\mathbb{E}(e^{zL_t}) = e^{t\kappa(z)}$ for all $z \in \mathbb{C}$ with $0 \leq \mathrm{Re}(z) < 1$ and in particular for $z \in 0, 1)$, where

$$\kappa(z) = \Psi(-iz) = bz + \frac{1}{2}\Sigma z^2 + \int_{\mathbb{R}} \left(e^{2x} - 1 - zx1_{\{|x|\leq 1\}}\right) \nu(dx).$$

One can show with a bit of analysis that this extends to $z = 1$, so we get that $\mathbb{E}(e^{L_t}) = e^{t\kappa(1)}$ and because $S = e^L$ is a martingale, this gives $\mathbb{E}(S_t) \equiv 1$ so that we can get

$$0 = \kappa(1) = b + \frac{1}{2}\Sigma + \int_{\mathbb{R}} \left(e^x - 1 - x1_{\{|x|\leq 1\}}\right) \nu(dx).$$

This is a the condition on the Lévy triplet $(b, \Sigma, \nu)$ for $S = e^L$ to be a martingale.

The arithmetic case has same result as Proposition 6.3 but the proof is more complicated because we cannot use Ansel-Stricker.

**Proposition 6.4.** *If $L$ is a Lévy process and a $\sigma$-martingale, then it is a true martingale.*

*Proof.* In a first step, one shows that for Lévy processes being $\sigma$-martingale is the same as being a local martingale. This holds because a Lévy process has deterministic seminmartingale characteristics, see Jacod/Shiryaev III 6.35 and III 6.39.

So assume $L$ is Lévy and local martingale; we want to show that $L$ is a true martingale. Define

$$J^L_\cdot := \sum_{0<s\leq \cdot} \Delta L_s 1_{\{|\Delta L_s|>1\}}.$$

Then $J^L$ is Lévy like $L$ and of FV. Moreover, we get $L = J^L + M + ct$, where $M$ is a true martingale with bounded jumps and $c \in \mathbb{R}$ [see BMSC]. So w.l.o.g. assume $M \equiv 0$ and $L = J^L + ct$ is Lévy and local martingale and of FV. This implies that $\Delta L$ is locally integrable and so the FV process $J^L$ has locally integrable variation.

Hence, by Doob-Meyer, $J^L$ has compensator $\widetilde{J}$, i.e. there exists $\widetilde{J}$ predictable and of locally integrable variation such that $J^L - \widetilde{J}$ is a local martingale. **But** $J^L$ is Lévy, so $\widetilde{J}$ is deterministic and linear in $t$, i.e., $\widetilde{H}_t = \gamma t$ for some $\gamma \in \mathbb{R}$ and this is given by

$$\gamma = \int_{\mathbb{R}} x1_{\{|x|>1\}}\nu(dx)$$

(and then actually $\gamma = -c$).

Next, $\widetilde{J}$ is non-random and of integrable variation, so that we get

$$\int_{\mathbb{R}^d} |x| 1_{\{|x|>1\}} \nu(dx) < \infty.$$

But $L_t = J_t^L + ct$, and so this means that $\mathbb{E}(|\Delta L_t|) = b < \infty$, thus the local martingale $J^L - \widetilde{J}$ is of integrable variation. So it is a true martingale (because it is of class (D)) and so is then $L$. This completes the proof. $\square$

# 7 Pricing and hedging by replication

Consider a financial market, arbitrage-free, and time $T$ payoff $H \in L^0_+(\mathcal{F}_T)$. Two basic questions we can ask:

- What is a reasonable **time-$t$ value** for $H$?
- after selling at $t$, how do we **manage resulting risk**?

## 7.1 Basic ideas and results

**Definition 7.1.** *Given $H \in L^0_+(\mathcal{F}_T)$, **replicating strategy for** $H$ is a self-financing, admissible $\varphi \triangleq (v_0, \vartheta)$ with $V_T(\varphi) = H$ $\mathbb{P}$-a.s. Then $H$ is said to be **replicable/attainable** by $\varphi$.*

**Fundamental idea (valuation/pricing by replication)**: If $H \in L^0_+(\mathcal{F}_T)$ is replicable by $\varphi$, time-$t$ value of $H$ must be $V_t(\varphi)$ to avoid arbitrage.

*"Proof".* Consider on $[t, T]$ two strategies:

1. buy $H$ at its time-$t$ price $\pi_t(H)$ and wait until $T$;
2. use self-financing strategy $\Psi \triangleq (V_t(\varphi), \vartheta)$.

Both i) and ii) have zero-cash flows on $(t, T)$. Indeed, i) by definition and ii) because its self-financing. Moreover, both have as $T$ payoff the value of $H$; i) again by definition and ii) because $V_T(\varphi) = H$.

So if $\pi_t(H) \neq V_t(\varphi)$ then we can buy cheaper and sell more expensive product to get profit, hence create arbitrage. □

**Problem**: using buy and sell creates difference of strategies, and even if they are admissible, difference can fail to be so. More **fundamentally**, the argument omits specifying which "meta-strategies" are allowed for creating arbitrage.

For easier/alternative **computation** of $V_t(\varphi)$, note that $H$ is attainable iff there exists $\varphi \triangleq (v_0, \vartheta)$ admissible with

$$H = V_T(\varphi) = v_0 + \int_0^T \vartheta_u dS_u \ \mathbb{P}\text{-a.s.}$$

and then $V_t\varphi) = v_0 + \int_0^t \vartheta_u dS_u$. So if $Q$ is an E$\sigma$MM for $S$ and $\vartheta$ is "nice enough", so that $\int \vartheta dS$ is a true $Q$-martingale (and also $H$ or $v_0$ are in $L^1(Q)$). Then

$$V_t(\varphi) = \mathbb{E}_Q(H \mid \mathcal{F}_t), \ 0 \leq t \leq T \tag{7.1}$$

**valuation/pricing by risk-neutral expectation**.

**Remark 7.1.**

1. Key argument is to transform $V_t(\varphi)$ by **riskless dynamic trading** on $[t,T]$ into $H$. In particular:
   (a) Key idea is **hedging by dynamic trading** getting valuation is side product.
   (b) $Q$ is purely **auxiliary tool**, in that sense, "$Q$-probabilites" are artificial.
   (c) LHS in (7.1) does not depend on $Q$, RHS does not depend on $\varphi$.
   (d) Assumption that $H$ is replicable is crucial!
2. **Basic structure** of (7.1) is **very simple**: discounted time-$t$ value is conditional expectation of discounted payoff under E$\sigma$MM $Q$. Simple recipe, often used, sometimes abused.
3. Precise mathematical formulation needs work.

For **finite discrete time**, results are easy to state: For the rest of this subsection, suppose $S = (S_k)_{k=0,1\ldots,T}$ is $\mathbb{R}^d$-valued adapted and $S^0 \equiv 1$. Denote by $\mathbb{P}_{e(,\mathrm{loc})}(S)$ set of all E(L)MMs and recall from Corollary 2.1 (Dalang-Morton-Willinger) that $S$ satisfies (NA) iff $\mathbb{P}_e(S) \neq \emptyset$ iff $\mathbb{P}_{e,\mathrm{loc}}(S) \neq \emptyset$. Recall notion of attainability/replicability for $H$ and call market $(S,\mathbb{F})$ **complete** if every $H \in L^0_+(\mathcal{F}_T)$ is attainable.

**Lemma 7.1.** *If $\mathcal{F}_0$ is trivial, then the following are equivalent:*

1. *$(S,\mathbb{F})$ is complete.*
2. *Every $H \in L^0_+(\mathcal{F}_T)$ admits representation*

$$H = H_0 + \int_0^T \vartheta_u dS_u \ \mathbb{P}\text{-}a.s.$$

*with $H_0 \in \mathbb{R}$ and $\vartheta$ $\mathbb{F}$-predictable ($S$-integrable) admissible.*

*Proof.* Just take $\varphi \triangleq (H_0,\vartheta)$. □

**So**: completeness means that up to constants, $S$ spans via stochastic integrals all $\mathcal{F}_T$-measurable random variables.

Next results are from IMF, we will only repeat the statements.

**Theorem 7.1.** *Suppose $\mathcal{F}_0$ is trivial and $S$ satisfies (NA). Then for $H \in L^0_+(\mathcal{F}_T)$ are equivalent:*

1. *$H$ is attainable.*
2. $\sup_{Q \in \mathbb{P}_{e,loc}(S)} \mathbb{E}_Q(H) < \infty$ *is attained in some $Q^* \in \mathbb{P}_{e,loc}(S)$ (sup=max).*
3. *The mapping $\mathbb{P}_{e,loc}(S) \to \mathbb{R}$, $Q \mapsto E_Q(H)$ is constant.*

**Theorem 7.2.** *Suppose $\mathcal{F}_0$ is trivial, $S$ satisfies (NA) and $\mathcal{F}_T = \mathcal{F}$. Then the following are equivalent:*

1. *$(S, \mathbb{F})$ is complete.*
2. *$\#\mathbb{P}_{e,loc}(S) = 1$, i.e. there exists **exactly one** ELMM for $S$.*

**Remark 7.2.** Theorem 7.2 follows easily from Theorem 7.1, Theorem 7.1 is more difficult, it needs so-called **optional decomposition theorem**.

In continuous time, there are analogues to Theorem 7.1 and Theorem 7.2; but formulation is more delicate and proofs are more technical. <u>See later</u>.

## 7.2 An Illustration: The Black-Scholes formula

Consider model with one bank account $\widetilde{B}_t = \widetilde{S}^0_t = e^{rt}$ and one stock

$$\widetilde{S}_t = \widetilde{S}^1_t = s_0 \exp\left(\sigma W_t + \left(\mu - \frac{1}{2}\sigma^2\right) t\right);$$

so discounted stock price $S = \widetilde{S}/\widetilde{B} = \widetilde{S}^1/\widetilde{S}^0$ satisfies SDE

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

For payoff, consider **call option** $\widetilde{H} = (\widetilde{S}_T - \widetilde{K})^+$ with discounted payoff $H = \widetilde{H}/\widetilde{B}_T = (S_T - K)^+$ with $K = \widetilde{K}/\widetilde{B}_T = \widetilde{K}e^{-rT}$. What is the value of call option at time $t \leq T$?

Assume $\mathbb{F} = \mathbb{F}^W$ is generated by $W$ (and augmented by $\mathbb{P}$-nullsets from $\mathcal{F}^W_T$) by Lemma 6.1 and subsequent example, only one candidate for E$\sigma$MD/ELMD and for density of an E$\sigma$MM/ELMM

$$\widehat{Z}_t = \mathcal{E}\left(-\frac{\mu - r}{\sigma} W\right)_t = \exp\left(-\frac{\mu - r}{\sigma} W_t - \frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 t\right)$$

This $\widehat{Z} > 0$ is a true $\mathbb{P}$-martingale; so $\frac{d\widehat{P}}{d\mathbb{P}} := \widehat{Z}_T$ gives a probability measure $\widehat{P} \approx \mathbb{P}$ on $\mathcal{F}_T$, and also

$$dS_t = S_t\sigma\left(\frac{\mu - r}{\sigma}dt + dW_t\right) = S_t\sigma d\widehat{W}_t$$

where $\widehat{W}_t := W_t + \frac{\mu-r}{\sigma}t$, $0 \le t \le T$, is by Girsanov a $\widehat{P}$-BM. So $S = S_0\mathcal{E}(\sigma\widehat{W})$ is even a true $\widehat{P}$-martingale, so that $S$ admits a unique EMM $\widehat{P}$, on $[0,T]$. In Analogy to Theorem 7.2 we **suspect** that model is complete and $H$ is attainable/replicable. So we can **compute**

$$\begin{aligned}
\widehat{V}_t &:= \widehat{\mathbb{E}}(H \mid \mathcal{F}_t) = \widehat{\mathbb{E}}((S_T - K)^+ \mid \mathcal{F}_t) \\
&= \widehat{\mathbb{E}}\Big[\Big(\underbrace{S_t}_{\mathcal{F}_t\text{-meas.}} \cdot \underbrace{e^{\sigma(\widehat{W}_T - \widehat{W}_t) - \frac{1}{2}\sigma^2(T-t)}}_{\text{indep. of } \mathcal{F}_t} - K\Big)^+ \mid \mathcal{F}_t\Big] \\
&= \widehat{\mathbb{E}}\left[\left(ae^{bZ-c} - d\right)^+\right]\Big|_{\substack{a=S_t, Z=\frac{\widehat{W}_T - \widehat{W}_t}{\sqrt{T-t}} \sim \mathcal{N}(0,1) \text{ under } \widehat{P} \\ b=\sigma\sqrt{T-t},\ c=\frac{1}{2}b^2,\ d=K}} =: \widehat{v}(t, S_t),
\end{aligned}$$

where function $\widehat{v}(t,x)$ can be computed explicitly. For undiscounted value, natural guess is then $\widetilde{V}_t = \widehat{V}_t\widetilde{B}_t =: \widetilde{v}(t, \widetilde{S}_t)$ and working out computations gives

$$\widetilde{v}(t, \widetilde{S}_t) = \widetilde{S}_t\Phi(d_1) - \widetilde{K}e^{-r(T-t)}\Phi(d_2) \tag{7.2}$$

with

$$d_{1,2} = \frac{\log\dfrac{\widetilde{S}_t}{\widetilde{K}e^{-r(T_t}} \pm \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}} \tag{7.3}$$

and

$$\Phi(z) = \widehat{P}(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}dx$$

is the cdf of an $\mathcal{N}(0,1)$-distributed random variable. Formula in (7.2), (7.3) is **Black-Scholes formula** for price of call option, its **derivation** was awarded with a Nobel price in 1997.

To justify that $\widetilde{V}_t = \widetilde{v}(t, \widetilde{S}_t)$ as reasonable time-$t$ value, we still need to check whether $H$ is attainable (in a good sense).

One way uses Itô's representation theorem: as $\mathbb{F} = \mathbb{F}^W = \mathbb{F}^{\widehat{W}}$, any $H \in L^1(\widehat{P}, \mathcal{F}_T^{\widehat{W}})$ has unique representation as

$$H = \widehat{\mathbb{E}}(H) + \int_0^T \Psi_u d\widehat{W}_u = \widehat{\mathbb{E}}(H) + \int_0^T \vartheta_u dS_u,$$

where $\int \Psi d\widehat{W} = \int \vartheta dS$ is $\widehat{P}$-martingale; this uses $dS = S\sigma d\widehat{W}$, so $\vartheta_u := \Psi_u/(\sigma S_u)$. Moreover, if $H \geq 0$ (as e.g. call), then

$$\int \vartheta dS \geq -\widehat{\mathbb{E}}(H)$$

is admissible integrand and strategy $\varphi \overset{\wedge}{=} (\widehat{\mathbb{E}}(H), \vartheta)$ is self-financing, admissible.

Furthermore,

$$V_T(\varphi) = \widehat{\mathbb{E}}(H) + \int_0^T \vartheta_u dS_u = H \text{ a.s.}$$

so that $H$ is final value of an admissible self-financing strategy whose wealth process $V(\varphi) = \widehat{\mathbb{E}}(H) + \int \vartheta dS$ is $\widehat{P}$-martingale. So any $H \in L_1^+(\mathcal{F}_T, \widehat{P})$ is attainable in the above sense with a nice strategy, and so we have a kind of completeness. Finally, note that for call option, we have $0 \leq H = (S_t - K)^+ \leq S_T \in L^1(\widehat{P})$.

**Alternative argument**: that call option is attainable even works without specifying $\mathbb{F}$. Start with function $\widetilde{v}(t,x)$ from (7.2), (7.3) and check by computations (elementary, but laborious) that $\widetilde{v}$ satisfies PDE

$$\frac{\partial \widetilde{v}}{\partial t} + rx\frac{\partial \widetilde{v}}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 v^2}{\partial x^2} - r\widetilde{v} = 0,$$
$$\widetilde{v}(T,x) = (x-K)^+.$$

Now $\widetilde{S}_t = \widetilde{B}_t S_t = e^{rt} S_t$ satisfies $d\widetilde{S}_t = \widetilde{S}_t(rdt + \sigma d\widehat{W}_t)$; so Itô's formula gives $d\widetilde{V}_t = d\widetilde{v}(t, \widetilde{S}_t) = \ldots dt + \ldots d\widehat{W}_t$. Working out the calculations and using from above PDE that $\widetilde{v}_t + rx\widetilde{v}_x + \frac{1}{2}\sigma^2 x^2 \widetilde{v}_{xx} = r\widetilde{v}$ yields

$$d\widetilde{V}_t = \frac{\partial \widetilde{v}}{\partial x}(t, \widetilde{S}_t)\sigma \widetilde{S}_t d\widehat{W}_t + r\widetilde{V}_t dt = \frac{\partial \widetilde{v}}{\partial x} d\widetilde{S}_t + \left( r\widetilde{V}_t - \frac{\partial \widetilde{v}}{\partial x}(t, \widetilde{S}_t) r \widetilde{S}_t \right) dt.$$

We can rewrite this more nicely by setting

$$\vartheta_t := \frac{\partial \widetilde{v}}{\partial x}(t, \widetilde{S}_t),\ \varphi_t^0 := \frac{1}{\widetilde{B}_t}\left( \widetilde{v}(t, \widetilde{S}_t) - \widetilde{S}_t \frac{\partial \widetilde{v}}{\partial x}(t, \widetilde{S}_t) \right)$$

to get

$$d\widetilde{V}_t = \vartheta_t d\widetilde{S}_t + \varphi_t^0 d\widetilde{B}_t \text{ and } \widetilde{V}_t = \widetilde{v}(t, \widetilde{S}_t) = \vartheta_t \widetilde{S}_t + \varphi_t^0 d\widetilde{B}_t.$$

Now integrating from $t$ to $T$ and using $\widetilde{v}(T, \widetilde{S}_T)(\widetilde{S}_T - \widetilde{K})^+ = \widetilde{K}$ finally gives

$$(\widetilde{S}_T - \widetilde{K})^+ = \widetilde{V}_T = \widetilde{V}_t + \int_t^T d\widetilde{V}_t = \widetilde{v}(t, \widetilde{S}_t) + \int_t^T (\vartheta_u d\widetilde{S}_u + \varphi_u^0 d\widetilde{B}_u).$$

**So**: $\varphi = (\varphi^0, \vartheta)$ is self-financing stragey for undiscounted prices $(\widetilde{B}, \widetilde{S})$, it is admissible because $\widetilde{V}_t = \widetilde{v}(t, \widetilde{S}_t) \geq 0$, and it transforms $\widetilde{v}(t, \widetilde{S}_t)$ on $[t, T]$ into payoff $(\widetilde{S}_T - \widetilde{K})^+$. **Hence**: guess produces explicit replicating strategy, and call is attainable.

# 8 Superreplication and the optimal decomposition

**Basic question**: given a general ("not attainable") payoff, how to hedge and price this?

**Standard setup**: $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space with filtration satisfying the usual conditions. Moreover, $T < \infty$, $S^0 \equiv 1$, $S$ is an $\mathbb{R}^d$-valued semimartingale satisfying (NFLVR); so set $\mathbb{P}_{e,\sigma}$ of E$\sigma$MMs for $S$ is nonempty.

Fix payoff $H \in L^0_+(\mathcal{F}_T)$. If $H$ is "not attainable", then there **exists no** admissible self-financing $\varphi$ with $V_T(\varphi) = H$ a.s. How to hedge/price $H$ at time $t \leq T$?

**Idea**: look at strategies producing at least $H$ at $T$ and find the cheapest among these.

**Definition 8.1.** *The **superreplication price** of $H \in L^0_+(\mathcal{F}_T)$ is defined as*

$$\begin{aligned}\pi^s(H) :&= \inf\left\{v_0 \in \mathbb{R} : v_0 + \int_0^T \vartheta_u dS_u \geq H \text{ a.s. for a } \vartheta \in \Theta_{adm}\right\}\\ &= \inf\{v_0 \in \mathbb{R} : H - v_0 \in \mathcal{G}_{adm} - L^0_+ = \mathcal{C}^0_{adm}\}.\end{aligned}$$

**Intuition**: can sell $H$ at 0 for $\pi^s(H)$ without incurring any risk, because $\varphi \triangleq (\pi^s(H), \vartheta^s)$ produces at least $H$ at $T$ in an admissible self-financing way, and is also competitive. Price is from seller's perspective (notice the supscript $s$).

**Warning**: infimum is perhaps not attained, so we do not know whether $\vartheta^s$ for $v_0 := \pi^s(H)$ exists!

**Lemma 8.1.** *Assume (NFLVR) or (equivalently) $\mathbb{P}_{e,\sigma} \neq \emptyset$. For any (payoff) $H \in L^0_+(\mathcal{F}_T)$, we have*

$$\pi^s(H) \geq \sup_{Q \in \mathbb{P}_{e,\sigma}} \mathbb{E}_Q(H).$$

*Proof.* By convention $\inf = +\infty$ and thus the inequality is trivial in that case. So w.l.o.g., there exists $v_0 \in \mathbb{R}$ and $\vartheta \in \Theta_{\text{adm}}$ with $v_0 + G_T(\vartheta) \geq H$ a.s. For any $Q \in \mathbb{P}_{e,\sigma}$, $G(\vartheta) = \int \vartheta dS$ is under $Q$ a supermartingale (usual Ansel-Stricker argument) and so

$$\mathbb{E}_Q(H) \leq v_0 + \mathbb{E}_Q(G_T(\vartheta)) \leq v_0,$$

take sup over $Q$, inf over $v_0$ to conclude. $\square$

**Goals**: prove that we have equality in Lemma 8.1, and that inf for $\pi^s(H)$ is attained.

Fix $H \in L^0_+(\mathcal{F}_T)$ and define adapted "process" with values in $[0, +\infty]$

$$U_t := \operatorname*{ess\,sup}_{Q \in \mathbb{P}_{e\sigma}} \mathbb{E}_Q(H \mid \mathcal{F}_t),\ 0 \le t \le T.$$

If $\mathcal{F}_0$ is trivial, then $U_0 = \sup_{Q \in \mathbb{P}_{e,\sigma}} \mathbb{E}_Q(H)$.

**Proposition 8.1.** *Assume (NFLVR) or $\mathbb{P}_{e,\sigma} \neq \emptyset$ and $H \in L^0_+(\mathcal{F}_T)$.*
*If $\sup_{Q \in \mathbb{P}_{e,\sigma}} \mathbb{E}_Q(H) < \infty$, then $U$ is a $Q$-supermartingale, (simultaneously) for all $Q \in \mathbb{P}_{e,\sigma}$.*

*Proof.* $U$ is adapted and nonnegative. We fix $Q \in \mathbb{P}_{e,\sigma}$ and $Q$-supermartingale property saying that $\mathbb{E}_Q(U_t \mid \mathcal{F}_s) \le U_s$ for $s \le t$. Then $Q$-integrability follows as well by taking expectations and using that $\sup_{Q \in \mathbb{P}_{e,\sigma}} \mathbb{E}_Q(H) < \infty$.

For each $t \in [0, T]$, introduce set

$$\begin{aligned}\mathcal{Z}_t :&= \{Z : Z \text{ is density process w.r.t. } Q \text{ of some } R \in \mathbb{P}_{e,\sigma}, \text{ and } Z_s = 1 \text{ for } s \le t\} \\ &= \{Z : Z \text{ is density process w.r.t. } Q \text{ of some } R \in \mathbb{P}_{e,\sigma} \text{ with } R = Q \text{ on } \mathcal{F}_t\}.\end{aligned}$$

Then $\mathcal{Z}_t$ is nonempty because $Z \equiv 1 \in \mathcal{Z}_t$, for $R := Q$. and $\mathcal{Z}_t \subset \mathcal{Z}_s$ for $s \le t$. Moreover, we claim that

$$\mathcal{Z}_t = \left\{ \frac{Z^R_{t\vee\cdot}}{Z^R_t} : Z^R \text{ is density process w.r.t. } Q \text{ of some } R \in \mathbb{P}_{e,\sigma} \right\}.$$

⌜ "$\subset$" For $Z = Z^R$ with $R \in \mathbb{P}_{e,\sigma}$, we have $Z_s = 1$ for $s \le t$ and hence $Z_t = 1$; so

$$Z_\cdot = 1_{\{\cdot \le t\}} + Z_\cdot 1_{\{\cdot > t\}} = \frac{Z^R_{t\vee\cdot}}{Z^R_t}.$$

"$\supset$" Take $R \in \mathbb{P}_{e,\sigma}$ with density process $Z^R$ w.r.t. $Q$, and define $Z_\cdot := Z^R_{t\vee\cdot}/Z^R_t$. Then $Z > 0$, $Z_S = 1$ for $s \le t$ and $Z_\cdot$ is a $Q$-martingale like $Z^R$. So we can use $Z$ as a density process w.r.t. $Q$ of some $R' \approx Q \approx \mathbb{P}$ on $\mathcal{F}_T$. We know that $Q \in \mathbb{P}_{e,\sigma}$, $R \in \mathbb{P}_{e,\sigma}$; so by Bayes, $S$ is $Q$-$\sigma$-martingale and $Z^R S$ is $Q$-$\sigma$-martingale, and then so is

$$Z_\cdot S_\cdot = S_\cdot 1_{\{\cdot \le t\}} + \frac{Z^R_\cdot S_\cdot}{Z^R_t} 1_{\{\cdot > t\}}.$$

So $S$ is an $R'$-$\sigma$-martingale by Bayes, and so $R' \in \mathbb{P}_{e,\sigma}$. ⌟

Now we look at

$$
\begin{aligned}
U_t = \operatorname*{ess\,sup}_{R\in\mathbb{P}_{e,\sigma}} \mathbb{E}_R(H \mid \mathcal{F}_t) &\overset{\text{Bayes rule}}{=} \operatorname*{ess\,sup}_{R\in\mathbb{P}_{e,\sigma}} \mathbb{E}_Q\left(\frac{HZ_T^R}{Z_t^R} \mid \mathcal{F}_t\right) \\
&\overset{\text{above}}{=} \operatorname*{ess\,sup}_{Z\in\mathcal{Z}_t} \mathbb{E}_Q(HZ_T \mid \mathcal{F}_t) =: \operatorname*{ess\,sup}_{Z\in\mathcal{Z}_t} \Gamma_t(Z).
\end{aligned}
$$

But the family $\{\Gamma_t(Z) : Z \in \mathcal{Z}_t\}$ is directed upward: indeed, consider

**Exercise 8.1.** *If $Z, Z' \in \mathcal{Z}_t$ and $A \in \mathcal{F}_t$, then $Z1_A + Z'1_{A^c}$ is again in $\mathcal{Z}_t$.*

and so for $A := \{\Gamma_t(Z) \geq \Gamma_t(Z')\} \in \mathcal{F}_t$, we get $Z'' := Z1_A + Z'+_{A^c} \in \mathcal{Z}_t$ and

$$
\begin{aligned}
\max(\Gamma_t(Z), \Gamma_t(Z')) &= 1_A\Gamma_t(Z) + 1_{A^c}\Gamma_t(Z') \\
&= \mathbb{E}_Q(H(\underbrace{1_A Z_T + 1_{A^c} Z'_T}_{Z''_T}) \mid \mathcal{F}_t) = \Gamma_t(Z'').
\end{aligned}
$$

Hence, for each $t$, there exists a sequence $(Z^{(n)})_{n\in\mathbb{N}} \subset \mathcal{Z}_t$ with

$$
U_t = \nearrow\text{-}\lim_{n\to\infty} \Gamma_t(Z^{(n)}) = \nearrow\text{-}\lim_{n\to\infty} \mathbb{E}_Q(HZ_T^{(n)} \mid \mathcal{F}_s).
$$

Consider under $Q$ on $\mathcal{F}_s$ to get

$$
\begin{aligned}
\mathbb{E}_Q(U_t \mid \mathcal{F}_s) &\overset{\text{mon. int.}}{=} \nearrow\text{-}\lim_{n\to\infty} \mathbb{E}_Q[\mathbb{E}_Q(HZ_T^{(n)} \mid \mathcal{F}_t) \mid \mathcal{F}_s] \\
&\overset{Z^{(n)}\in\mathcal{Z}_t\subset\mathcal{Z}_s}{\leq} \operatorname*{ess\,sup}_{Z\in\mathcal{Z}_s} \mathbb{E}_Q(HZ_T \mid \mathcal{F}_s) = U_s.
\end{aligned}
$$

This is $Q$-supermartingale property. □

Using standard argument, using $\mathbb{F}$ is RC, can also show that $U$ admits RCLL version; see Krankov and D/M. We choose this version and call it again $U$.

One **concrete example** of process which is $Q$-supermartingale, simultaneously for all $Q \in \mathbb{P}_{e,\sigma}$, looks as follows. Start with $x \in \mathbb{R}$, $\vartheta \in L(S)$, $C$ adapted increasing (RCLL) with $C_0 = 0$. Define

$$
V^{x,\vartheta,C} := x + \int \vartheta dS - C.
$$

Interpret $(x, \vartheta, C)$ as **generalized strategy**: $x$ is initial wealth, $\vartheta$ describes numbers of shares in $S$ held, and $C$ is total (cumulative) amount spent for consumption on $[0, t]$.

Then wealth evolution is

$$x + \int \vartheta dS - C = V^{x,\vartheta,C}.$$

Note that $C$ is nonnegative and $V^{x,\vartheta,C} + C = x + \int \vartheta dS$, so if $V^{x,\vartheta,C}$ is bounded below, then $\vartheta \in \Theta_{\text{adm}}$. Whenever $\vartheta$ is admissible, then $x + G(\vartheta) = x + \int \vartheta dS$ is a $Q$-supermartingale, for all $Q \in \mathbb{P}_{e,\sigma}$ (Ansel-Stricker). If in addition $V^{x,\vartheta,C}$ is bounded below, then

$$0 \leq C = x - V^{x,\vartheta,C} + \int \vartheta dS \leq \text{const} + \int \vartheta dS$$

shows that $C$ must be $Q$-integrable, for all $Q \in \mathbb{P}_{e,\sigma}$, and then $V^{x,\vartheta,C}$ is also a $Q$-supermartingale, for all $Q \in \mathbb{P}_{e,\sigma}$.

The next result establishes that this **concrete example** is the **only** example.

**Theorem 8.1** (Optional decomposition, Kramkov). *Suppose $S$ satisfies (NFLVR) or $\mathbb{P}_{e,\sigma} \neq \emptyset$. If a nonnegative process $U$ is a $Q$-supermartingale, for all $Q \in \mathbb{P}_{e,\sigma}$ and if $U_0$ is bounded, then there exists $\vartheta \in \Theta_{adm}$ and $C$ adapted increasing (RCLL) null at 0 such that*

$$U = U_0 + \int \vartheta dS - C.$$

*Moreover, if $U_0 \in \mathbb{R}$, e.g. if $\mathcal{F}_0$ is trivial, then $U = V^{U_0,\vartheta,C}$.*

An immediate consequence of the optional decomposition theorem is that we get equality in Lemma 8.1 More precisely, we get:

**Theorem 8.2** (hedging duality, first version). *Assume (NFLVR) or $\mathbb{P}_{e,\sigma} \neq \emptyset$ and $\mathcal{F}_0$ is trivial. Then for any $H \in L^0_+(\mathcal{F}_T)$, we have*

$$\pi^s(H) = \inf\{v_0 \in \mathbb{R} : H - v_0 \in \mathcal{G}_{adm} - L^0_+\} = \sup_{Q \in \mathbb{P}_{e,\sigma}} \mathbb{E}_Q(H).$$

*Moreover, if the RHS is finite, then the infimum is attained, i.e. a minimum.*

*Proof.* "$\geq$" was shown in Lemma 8.1. For "$\leq$" define $U$ as RCLL version of

$$U_t = \operatorname*{ess\,sup}_{Q \in \mathbb{P}_{e,\sigma}} \mathbb{E}_Q(H \mid \mathcal{F}_t), \ 0 \leq t \leq T$$

and note that $U$ is nonnegative with $U_T = H$ a.s. Moreover, as $\mathcal{F}_0$ is trivial, we have $U_0 = \sup_{Q \in \mathbb{P}_{e,\sigma}} \mathbb{E}_Q(H)$. Without loss of generality (for "$\leq$") assume $U_0 < \infty$.

By Proposition 8.1 and Theorem 8.1 (optional decomposition) we then get $\vartheta \in \Theta_{\mathrm{adm}}$ and $C$ increasing, null at 0 with

$$U = U_0 + \int \vartheta dS - C.$$

So $C_T \geq 0$ and $H - U_0 = U_T - U_0 = G_T(\vartheta) - C_T \in \mathcal{G}_{\mathrm{adm}} - L^0_+$. This shows that $U_0 \geq \pi^s(H)$, hence equality, and $U_0$ attains inf. $\square$

**Comment on optional decomposition theorem**: It can be seen as a uniform version of Doob-Meyer:

- $U$ is supermartingale, simultaneously for all $Q \in \mathbb{P}_{e,\sigma}$, thus by Doob Meyer we get $U = U_0 + M^Q - A^Q$ for all $Q$.
- here: $\int \vartheta dS$ is a local $Q$-martingale (Ansel-Stricker) simultaneously for all $Q \in \mathbb{P}_{e,\sigma}$, and $C$ increasing, does not depend on $Q$.
- **note**: $A^Q$ is predictable, $C$ in general is not (only optional).

*Proof of Theorem 8.1.* Too difficult in general; see Kramkov, Föllmer/Kalomov. **Here** instead: we use different argument under extra assumptions that $\mathbb{F}$ is **continuous** (i.e. all local martingales (w.r.t. $\mathbb{P}$ or equivalent, w.r.t. any $R \approx \mathbb{P}$) in $\mathbb{F}$ are continuous). Standard example: $\mathbb{F} = \mathbb{F}^W$ (and e.g. Itô process model)!

**Outline of main steps**:

(a) parametrize E$\sigma$MMs $Q$ via density process $Z = Z_0 \hat{Z} \mathcal{E}(N)$ as in section 6.2.

(b) Decompose $U$ under $Q$ via Doob-Meyer and Kunita-Watanabe as

$$U - U_0 = \hat{L} - \hat{B} = \hat{\vartheta} dS + \hat{N} - \hat{B}.$$

(c) use Girsanov to express $\hat{N}$ under $Q \in \mathbb{P}_{e,\sigma}$; then use $Q$-supermatingale property of $U$ to get more information about $\hat{N}$.

(d) deduce that $\hat{N} \equiv 0$.

First, by (NFLVR), $S$ is a $\sigma$-martingale under some $Q \approx \mathbb{P}$ on $\mathcal{F}_T$, hence Stochastic Integral of some local $Q$-martingale, which is continuous, because $\mathbb{F}$ is continuous. So $S$ is continuous, and write $S - S_0 = M + A$ with $M \in \mathcal{M}^c_{0,\mathrm{loc}}(\mathbb{P})$ and, by Proposition 6.1, because $\mathbb{P}_{e,\sigma} \neq \emptyset$, we have $A = \int d\langle M \rangle \lambda$ with $\lambda$ predictable in $L^2_{\mathrm{loc}}(M)$. This is the structure condition (SC).

Define minimal E$\sigma$MD $\hat{Z} = \mathcal{E}(-\int \lambda dM)$ so that (again by continuity of $S$) $\hat{Z}S \in \mathcal{M}_{\text{loc}}(\mathbb{P})$. For **simplicity**, assume that $\hat{Z}$ is a true $\mathbb{P}$-martingale. Then $d\hat{P} := \hat{Z}_T d\mathbb{P}$ defines $\hat{P} \approx \mathbb{P}$ on $\mathcal{F}_T$, density process is $\hat{Z}$, and $S$ is by Bayes a local $\hat{P}$-martingale. $\hat{P}$ is minimal E$\sigma$MM.

Now take any $Q \in \mathbb{P}_{e,\sigma}$ and write its density process $Z$ (which is an E$\sigma$MD for $S$) w.r.t. $\mathbb{P}$ as $Z = Z_0 \hat{Z}\mathcal{E}(N)$ (see section 6.2) with $N \in \mathcal{M}_{0,\text{loc}}(\mathbb{P})$ and $N \perp M$ under $\mathbb{P}$. As $\mathbb{F}$ is continuous, $N \in \mathcal{M}^c_{0,\text{loc}}(\mathbb{P})$ and $N \perp M$ is equivalent to $\langle N, M\rangle \equiv 0 \equiv \langle N, S\rangle$. Moreover, $\hat{Z}Z_0\mathcal{E}(N) = Z$ is strictly positive $\mathbb{P}$-martingale so that (by Bayes) $Z_0\mathcal{E}(N) > 0$ is $\hat{P}$-martingale and in fact it is a density process of $Q$ w.r.t. $\hat{P}$.

Now $U$ is a $\hat{P}$-supermartingale and so Doob-Meyer under $\hat{P}$ gives $U - U_0 = \hat{L} - \hat{B}$ with $\hat{L} \in \mathcal{M}_{0,\text{loc}}(\hat{P})$ and $\hat{B}$ predictable, increasing with $\hat{B}_0 = 0$. As $S$ is continuous, we can use Galtchouck-Kunita-Watanabe to write $\hat{L} = \int \hat{\vartheta} dS + \hat{N}$ with $\hat{\vartheta} \in L(S)$ and $\hat{N} \in \mathcal{M}_{0,\text{loc}}(\hat{P})$ with $\hat{N} \perp S$ under $\hat{P}$. By continuity, this means $0 \equiv \langle \hat{N}, S\rangle = \langle \hat{N}, M\rangle$. But $\hat{Z} = \mathcal{E}(-\int \lambda dM) = 1 - \int \hat{Z}\lambda dM$, and so we also get $\langle \hat{N}, \hat{Z}\rangle \equiv 0$.

Next, using Itô and continuity of $\hat{Z}$, we get

$$\frac{1}{\hat{Z}} = \ldots d\hat{Z} + (\text{cont. FV})$$

and then also $\langle \hat{N}, 1/\hat{Z}\rangle \equiv 0$. But $\hat{Z}$ is $\mathbb{P}$-martingale, so (by Bayes) $\frac{1}{\hat{Z}}$ is $\hat{P}$-martingale and so we get that $\hat{P}$-local martingale $\hat{N}$ and $\frac{1}{\hat{Z}}$ are $\hat{P}$-orthogonal. Hence $\hat{N}\frac{1}{\hat{Z}} \in \mathcal{M}_{\text{loc}}(\hat{P})$ which means (by Bayes) that $\hat{N} \in \mathcal{M}_{0,\text{loc}}(\mathbb{P})$. In summary:

$$U - U_0 = \int \hat{\vartheta} dS - \hat{B} + \hat{N} \tag{7.1}$$

with $\hat{N} \in \mathcal{M}_{0,\text{loc}}(\mathbb{P})$ and $\hat{N} \perp M$ under $\mathbb{P}$.

Now take any $Q \in \mathbb{P}_{e,\sigma}$ and write its density process w.r.t. $\mathbb{P}$ as $Z = \hat{Z}Z_0\mathcal{E}(N)$. Use Girsanov from $\hat{P}$ to $Q$ and that density process of $Q$ w.r.t. $\hat{P}$ is $Z_0\mathcal{E}(N)$ to write $\hat{N} \in \mathcal{M}_{0,\text{loc}}(\hat{P})$ as $\hat{N} = L^Q + \langle \hat{N}, N\rangle$ with $L^Q \in \mathcal{M}_{0,\text{loc}}(Q)$. So we get, under $Q$,

$$U - U_0 = \underbrace{\left(\int \hat{\vartheta} dS + L^Q\right)}_{\in \mathcal{M}_{0,\text{loc}}(Q)} - \underbrace{\left(\hat{B} - \langle \hat{N}, N\rangle\right)}_{c\text{FV}_0} \tag{7.2}$$

But $U - U_0$ is a $Q$-supermartingale, so it has a unique DOob-Meyer decomposition and hence term 2 in (7.2) must be the predictable increasing part from Doob-Meyer decomposition. So in (7.2) the term

$$\hat{B} - \langle \hat{N}, N \rangle \tag{7.3}$$

must be increasing and this holds for any $N$ coming from a $Q \in \mathbb{P}_{e\sigma}$. By localisation, then, (7.3) must hold for all $N \in \mathcal{M}_{0,\text{loc}}(\mathbb{P})$ with $N \perp M$ under $\mathbb{P}$. In particular, we can take $N := \alpha\hat{N}$ for any $\alpha > 0$ and so $\hat{B} - \alpha\langle\hat{N}\rangle$ is increasing for all $\alpha > 0$.

Hence we must have $\langle\hat{N}\rangle \equiv 0$, hence $\hat{N} \equiv \hat{N}_0 = 0$ and so $\hat{N} \equiv 0$ and (7.1) reads $U - U_0 = \int \hat{\vartheta} dS - \hat{B}$. Finally, $\hat{\vartheta} \in L(S)$ is in $\Theta_{\text{adm}}$ because

$$\int \hat{\vartheta} dS = \underbrace{U}_{\geq 0} - \underbrace{U_0}_{\text{bounded}} + \underbrace{\hat{B}}_{\geq 0} \geq -\text{const.}$$

□

**Remark 8.1.** Thanks to $\mathbb{F}$ being continuous, $C := \hat{B}$ is continuous, hence predictable, in general this is not true.

**Comments (on Theorem 8.2)**: This **superreplication approach** is general conceptually nice, and has good structural properties. Moreover, it also comes up in other situations, as auxiliary tool.

- As approach to pricing/hedging, very extreme because **totally-risk-averse**: $\pi^s(H)$ allows seller of $H$ to find hedging strategy whose outcome dominates $H$ a.s.; so all risk goes to buyer (and is paid by him).
- Nice price for seller, but it may happen that bounded $H$ has $\pi^s(H) = \|H\|_{L^\infty}$, or call has $\pi^s((S_T - K)^1) = S_0$. Clearly, only few buyers (if any) will pay that.
- Can similarly define buyer price $\pi^b(H)$ and prove that $\pi^b(H) = -\pi^s(-H) = \inf_{Q \in \mathbb{P}_{e,\sigma}} \mathbb{E}_Q(H)$.
- Major drawback: using admissible strategy requires $H \geq 0$ (or $H$ bounded from below) for seller price, $H$ bounded above for buyer price and hence only **bounded** payoffs can be studied for both parties.
- **Even worse**: can have (NFLVR) with $\mathcal{G}_{\text{adm}} = \{0\}$ then

  $$\{v_0 \in \mathbb{R} : H - v_0 \in \mathcal{G}_{\text{adm}} - L^0_+\} = \{v_0 \in \mathbb{R} : H \leq v_0\} \neq \emptyset,$$

  iff $H$ is bounded from above. This needs new ideas.

**Remark 8.2.** Can use optional decomposition or hedging duality to characterize payoffs which are "attainable". For that, call $H \in L^0_+(\mathcal{F}_T)$ **attainable** if $H = H_0 + G_T(\vartheta)$ for some $(H_0, \vartheta) \in R \times \Theta_{\mathrm{adm}}$ with the extra property that $G.(\vartheta)$ is true $P^*$-martingale for some $P^* \in \mathbb{P}_{e,\sigma}$.

**Exercise 8.2.** *Show that $H$ is attainable (in the above sense) iff*

$$\sup_{Q \in \mathbb{P}_{e,\sigma}} \mathbb{E}_Q(H) < \infty,$$

*and is attained in some $Q^* \in \mathbb{P}_{e,\sigma}$, i.e. sup is max.*

# 9 Duality for unbounded $S$

**Goal**: extend hedging duality from Theorem 8.2 to possibly **unbounded payoffs** and settings which may have $\mathcal{G}_{\text{adm}} = \{0\}$.

**Key idea**: allow **random** lower bounds for $G.(\vartheta)$.

**Recall**: if $S$ satisfies (NFLVR), then there exists E$\sigma$MM $Q$ for $S$, i.e.

$$S - S_0 = \int \Psi dM$$

with $M \in \mathcal{M}_{0,\text{loc}}(Q)$ and $\Psi > 0$ is in $L(M)$.

**Exercise 9.1.** *By changing $M$ and $\Psi$ we can even achieve that $M \in \mathcal{H}_0^1(Q)$, i.e. $M_T^* = \sup_{0 \le t \le T} |M_t| \in L^1(Q)$.*

So the one-dimensional process $\zeta = 1/\psi$ is then strictly positive and has the property that

$$\sup_{0 \le t \le T} \left| \int_0^t \zeta_u dS_u \right| = (\zeta \cdot S)_T^* = M_T^* \in L^1(Q).$$

If $S$ is locally bounded ( and $Q$ an ELMM for $S$), then we can even choose

$$\zeta = \sum_{n=1}^{\infty} \gamma_n 1_{(\!(\tau_{n-1}, \tau_n]\!]}$$

to get $\zeta > 0$ with $(\zeta \cdot S)_T^* \le$const.

The next definition introduces "good random lower bounds".

**Definition 9.1.** *A **feasible weight function** is a random variable $w \ge 1$ satisfying:*

1. *$(\zeta \cdot S)_T^* \le w$ $\mathbb{P}$-a.s. for some one dimensional $\zeta > 0$ in $L(S)$.*
2. *$\mathbb{E}_Q(w) < \infty$ for some $Q \in \mathbb{P}_{e,\sigma}$.*

**Remark 9.1.**

1. Good lower bound for $G.(\vartheta)$ should be large enough to allow many strategies, but not so large that we also allow e.g. doubling (or other extreme) strategies.
2. If $S$ satisfies (NFLVR) or $\mathbb{P}_{e,\sigma} \neq \emptyset$, then feasible weight functions exists.
3. W.l.o.g. can impose $w$ to be $\mathcal{F}_T$-measurable.
4. If $S$ satisfies (NFLVR) and is locally bounded, then every constant $\ge 1$ is a feasible weight function.

**Notation**: For any random variable $w \geq 1$, set $\mathbb{P}^w_{e,\sigma} := \{Q \in \mathbb{P}_{e,\sigma} : \mathbb{E}_Q(w) < \infty\}$. If $S$ satisfies (NFLVR) and if $w$ is a feasible weight function (which exists thanks to remark 2 above) then $\mathbb{P}^w_{e,\sigma} \neq \emptyset$.

**Definition 9.2.** *Let $w \geq 1$ be a random variable with $\mathbb{E}_Q(w) < \infty$ for some $Q \in \mathbb{P}_{e,\sigma}$. For a constant $a \geq 0$, call $\vartheta \in L(S)$ $(a, w)$-**admissible** if*

$$G_t(\vartheta) \geq -a\mathbb{E}_Q(w \mid \mathcal{F}_t) \ \mathbb{P}\text{-a.s. for all } t \in [0, T] \text{ and for all } Q \in \mathbb{P}_{e,\sigma}.$$

*Then write $\vartheta \in \Theta^a_w$ and set $\Theta_w := \bigcup_{a \geq 0} \Theta^a_w$ ($w$-admissible $\vartheta \in L(S)$). Also set*

$$\mathcal{G}_w := G_T(\Theta_w) = \{G_T(\vartheta) : \vartheta \text{ is } w\text{-admissible}\}.$$

**Remark 9.2.** For $w \equiv 1$, $w$-admissible reduces to admissible. For clarity, write $(1, 1)$-admissible, for what was called 1-admissible before, i.e. if we have $G_.(\vartheta) \geq -1$. Also if $w \equiv 1$, we have $\mathcal{G}_{\text{adm}} = \mathcal{G}_1$ (Note difference between $\mathcal{G}_1$ and $\mathcal{G}^1$).

**Lemma 9.1.** *Assume (NFLVR) and let $w \geq 1$ be a random variable satisfying $\mathbb{E}_Q(w) < \infty$ for some $Q \in \mathbb{P}_{e,\sigma}$. For any $w$-admissible $\vartheta \in L(S)$, the stochastic integral process $G(\vartheta) = \int \vartheta dS$ is then a $Q$-supermartingale. This gives:*

1. *If $g \in \mathcal{G}_w$ and $Q \in \mathbb{P}^w_{e,\sigma}$, then $\mathbb{E}_Q(g) \leq 0$.*
2. *If $g \in \mathcal{G}_{adm}$ and $Q \in \mathbb{P}_{e,\sigma}$, then $\mathbb{E}_Q(g) \leq 0$.*
3. *If $f \in \mathcal{G}_w - L^0_+$ and $Q \in \mathbb{P}^w_{e,\sigma}$, then $\mathbb{E}_Q(f) \leq 0$.*
4. *Suppose $g \in \mathcal{G}_w$, so that $g = G_T(\vartheta)$ for some $w$-admissible $\vartheta$, and fix $Q \in \mathbb{P}^w_{e,\sigma}$. If $\widetilde{w} \geq 1$ is another random variable with $\mathbb{E}_Q(\widetilde{w}) < \infty$ and $g \geq -\widetilde{w}$ a.s., then $\vartheta$ is $(1, \widetilde{w})$-admissible, and in particular $g \in \mathcal{G}_{\widetilde{w}}$.*

*Proof.* $Q \in \mathbb{P}_{e,\sigma}$ gives $G(\vartheta) = \int \vartheta dS = \int \vartheta \Psi dM$ for some $M \in \mathcal{M}_{0,\text{loc}}(Q)$. Moreover, $\vartheta \in \Theta_w$ gives $G(\vartheta) \geq -a\mathbb{E}_Q(w \mid \mathbb{F})$ so that $G(\vartheta)$ is bounded below by a $Q$-martingale. Now apply Ansel-Stricker plus subsequent comment (lower bound with suitable integrability is enough) to get that $G(\vartheta)$ is a $Q$-supermartingale. The rest is then easy:

1. $g = G_T(\vartheta)$ with $\vartheta \in \Theta_w$ gives $\mathbb{E}_Q(g) = \mathbb{E}_Q(G_T(\vartheta)) \overset{\text{superm.}}{\leq} \mathbb{E}_Q(G_0(\vartheta)) = 0$.
2. special case of 1) with $w \equiv 1$.
3. follows from 1) because $f \leq g$ and $g \in \mathcal{G}_w$. Note that $g = G_T(\vartheta) \geq -a\mathbb{E}_Q(w \mid \mathcal{F}_T) = -aw$ (because $w$ is w.l.o.g. $\mathcal{F}_T$-measurable) for some $a \geq 0$ and so $g^- \in L^1(Q)$, and hence $g \in L^1(Q)$, hence $f^+ \in L^1(Q)$.
4. $G_t(\vartheta) \overset{\text{supm.}}{\geq} \mathbb{E}_Q(G_T(\vartheta) \mid \mathcal{F}_t) = \mathbb{E}_Q(g \mid \mathcal{F}_t) \geq -\mathbb{E}_Q(\widetilde{w} \mid \mathcal{F}_t)$ a.s. $\forall t \in [0, T]$.

$\square$

**Remark 9.3.** Part 4) of Lemma 9.1 is extension of result that under (NA), any admissible $\vartheta$ with $G_T(\vartheta) \geq -b$ is automatically even $(b, 1)$-admissible. However, proof technique is different.

**Goal**: extend hedging duality (Theorem 8.2) to possibly unbounded payoffs.

**Notation**: Recall from Section 5 that one crucial consequence of (NFLVR) was Thereom 5.2 stating that $\mathcal{C}^\infty = \mathcal{C}^\infty_{\text{adm}} = (\mathcal{G}_{\text{adm}} - L^0_+) \cap L^\infty$ is weak*-closed in $L^\infty$. Underlying that was result that $\mathcal{G}^1_{\text{adm}} - L^0_+$ is Fatou-closed under (NFLVR), where

$$\begin{aligned}\mathcal{G}^1_{\text{adm}} &= \{g = G_T(\vartheta) : \vartheta \text{ is } (1,1)\text{-admissible}\} \\ &= \{g = G_T(\vartheta) : \vartheta \in L(S) \text{ has } G_.(\vartheta) \geq -1\}.\end{aligned}$$

But (as pointed out before), for general $S$, we might have $\mathcal{G}_{\text{adm}} = \{0\}$, and so superreplication with admissible strategies is limit for unbounded payoffs. Hence we need larger sets $\widetilde{\mathcal{G}}, \widetilde{\mathcal{C}}$.

For that, start with $w \geq 1$ satisfying $\mathbb{E}_Q(w) < \infty$ for some $Q \in \mathbb{P}_{e,\sigma}$ and consider

$$\begin{aligned}\mathcal{G}^1_w &:= \{g = G_T(\vartheta) : \vartheta \text{ is } (1,w)\text{-admissible}\} \\ &\;= \{g = G_T(\vartheta) : \vartheta \in L(S) \text{ has } G_t(\vartheta) \geq -\mathbb{E}_Q(w \mid \mathcal{F}_t) \text{ a.s. } \forall t,\ \forall Q \in \mathbb{P}^w_{e,\sigma}\}. \\ \mathcal{G}_w &= \{g = G_T(\vartheta) : \vartheta \text{ is } w\text{-admissible}\}. \\ \mathcal{C}^\infty_w &:= \frac{1}{w}(\mathcal{G}_w - L^0_+) \cap L^\infty.\end{aligned}$$

These are $w$-analogues of $\mathcal{G}^1_{\text{adm}},\ \mathcal{G}_{\text{adm}},\ \mathcal{C}^\infty_{\text{adm}}$.

**Interpretation**: elements of $\mathcal{C}^\infty_{\text{adm}}$ are payoffs which are bounded, $w$-discounted and (after discounting) can be superreplicated by final value of a $w$-admissible self-financing strategy $\varphi \mathrel{\hat{=}} (0, \vartheta)$ with zero initial wealth.

First (and most difficult) result is analogue of Fatou-closedness of $\mathcal{G}^1_{\text{adm}} - L^0_+$.

**Theorem 9.1.** *Assume (NFLVR) and take random variable $w \geq 1$ with $\mathbb{E}_Q(w) < \infty$ for some $Q \in \mathbb{P}_{e,\sigma}$. Then convex cone $\mathcal{G}^1_w - L^0_+$ is closed in $L^0$.*

*Proof.* Hard; uses compactness result for sequences of martingales. See D/S, Chapter 15 and in particular Corollary 15.4.11. □

Theorem 9.1 has many nice consequences:

**Corollary 9.1.** *Suppose (NFLVR) and let $w \geq 1$ satisfies $\mathbb{E}_Q(w) < \infty$ for some $Q \in \mathbb{P}_{e,\sigma}$. Then $\mathcal{C}_w^\infty = \frac{1}{w}(\mathcal{G}_w - L_+^0) \cap L^\infty$ is weak*-closed in $L^\infty$. (w-analogue of Theorem 5.2).*

*Proof.* $\mathcal{C}_w^\infty \subset L^\infty$ is convex cone, so can use the same criteria for weak*-closedness as in proof of Theorem 5.2. Take sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{C}_\infty^w$ uniformly bounded by 1 and converging to $f$ $\mathbb{P}$-a.s., then show that $f \in \mathcal{C}_\infty^w$.

$|f_n| \leq 1$ for all $n \in \mathbb{N}$ implies that $|f| \leq 1$, so $f \in L^\infty$. Each $f_n \in \mathcal{C}_w^\infty = \frac{1}{w}(\mathcal{G}_w - L_+^0)$ admits some $g_n \in \mathcal{G}_w$ with $g_n \geq wf_n \geq -w$ as $|f_n| \leq 1$. Write $g_n = G_T(\vartheta^n)$ with $\vartheta \in \Theta_w$. Use part 4) of Lemma 9.1 with $\widetilde{w} = w$, to conclude that $\vartheta^n$ is even $(1, w)$-admissible so that $g_n \in \mathcal{G}_w^1$ and $wf_n \in \mathcal{G}_w^1 - L_+^0$.

But now $wf_n \to wf$ a.s. and hence in $L^0$, and thus $wf \in \mathcal{G}_w^1 - L_+^1$ by Theorem 9.1. Put differently, $f \in L^\infty$ and

$$f \in \frac{1}{w}(\mathcal{G}_w^1 - L_+^0) \subset \frac{1}{w}(\mathcal{G}_w - L_+^0)$$

or $f \in \mathcal{C}_w^\infty$ as desired. $\square$

To attain a duality result, we now want to use the **bipolar theorem** (see appendix)applied to the convex cone $\mathcal{C}_w^\infty$.

Identify probability measures $R \ll \mathbb{P}$ with its density $\frac{dR}{d\mathbb{P}} \in L^1(\mathbb{P})$, and so view $\mathbb{P}_{e,\sigma}$ etc as subset of $L^1(\mathbb{P})$. We can use the shorthand

$$w\mathbb{P}_{e,\sigma} \triangleq \left\{ wY : Y = \frac{dQ}{d\mathbb{P}} \text{ for some } Q \in \mathbb{P}_{e,\sigma}^w \right\}.$$

Note that $\mathbb{E}_\mathbb{P}(wY) = \mathbb{E}_Q(w) < \infty$ for all such $Y$ by definition of $\mathbb{P}_{e,\sigma}^w$; hence $w\mathbb{P}_{e,\sigma} \subset L^1(\mathbb{P})$. Finally, set $\mathbb{P}_{a,\sigma} := \{Q \ll \mathbb{P} \text{ (on } \mathcal{F}_T) : S \text{ is } Q\text{-}\sigma\text{-martingale}\}$.

**Exercise 9.2.** *Show that (with above identification via density), $\mathbb{P}_{e,\sigma}$ is $L^1(\mathbb{P})$-dense in $\mathbb{P}_{a,\sigma}$, provided that $\mathbb{P}_{e,\sigma} \neq \emptyset$.*

Analogously, $\mathbb{P}_{e,\sigma}^w$ is $L^1(\mathbb{P})$-dense in $\mathbb{P}_{a,\sigma}^w$ if $\mathbb{P}_{e,\sigma}^w \neq \emptyset$, and then also $w\mathbb{P}_{e,\sigma}^w$ is $L^1(\mathbb{P})$-dense in $w\mathbb{P}_{a,\sigma}^w$.

**Theorem 9.2.** *Suppose $S$ satisfies (NFLVR) and $w$ is a feasible weight function. Then*

$$\textit{polar of } \mathcal{C}_w^\infty \textit{ in } L^1(\mathbb{P}) = \textit{cone}(w\mathbb{P}_{e,\sigma}^w). \tag{9.1}$$

*Hence we obtain for any $f \in L^\infty$ that*

$$f \in \mathcal{C}_w^\infty \iff \mathbb{E}_Q(wf) \leq 0, \textit{ for all } Q \in \mathbb{P}_{e,\sigma}^w. \tag{9.2}$$

This is a generalised version of hedging duality, still for bounded payoffs. Relaxing this, however, is not hard:

**Corollary 9.2.** *Suppose (NFLVR) and let $w$ be a feasible weight function. Then for any $f \geq -w$, we have*

$$f \in \mathcal{G}_w - L^0_+ \iff \mathbb{E}_Q(f) \leq 0, \text{ for all } Q \in \mathbb{P}^w_{e,\sigma}.$$

*Proof.* " $\implies$ " This is just Lemma 9.1, part 3).
"$\impliedby$" If $f \geq -w$, then $f_n := \frac{f \wedge n}{w}$ is bounded from below and also from above as $w \geq 1$. So we get

$$\mathbb{E}_Q\left(w \frac{f \wedge n}{w}\right) \leq \mathbb{E}_Q(f) \leq 0, \text{ for all } Q \in \mathbb{P}^w_{e,\sigma}$$

which implies by Theorem 9.2, (9.2) that $\frac{f \wedge n}{w} \subset \mathcal{C}^\infty_w \subset \frac{1}{w}(\mathcal{G}_w - L^0_+)$ or $f \wedge n \in \mathcal{G}_w - L^0_+$, for all $n \in \mathbb{N}$. So we get sequence $(g_n)_{n\in\mathbb{N}} \subset \mathcal{G}_w$ with $g_n \geq w f_n = f \wedge n \geq -w$. Now use part 4) of Lemma 9.1 to conclude that $g_n = G_T(\vartheta^n)$ with $\vartheta^n$ even $(1, w)$-admissible, so $(g_n)_{n\in\mathbb{N}} \subset \mathcal{G}^1_w$ and then even $f \wedge n \in \mathcal{G}^1_w - L^0_+$.

Because $f \wedge n \to f$ a.s. and also $\mathcal{G}^1_w - L^0_+$ is closed in $L^0$ by Theorem 9.1, we get $f \in \mathcal{G}^1_w - L^0_+ \subset \mathcal{G}_w - L^0_+$ as desired. $\square$

Corollary 9.2 extends basic hedging duality from Theorem 8.2 to general $S$ and unbounded payoffs. We discuss this more in detail in section 10.

**Remark 9.4.** Suppose $S$ satisfies (NFLVR) and take $w \equiv 1$. If $S$ is locally bounded, we have seen that this $w$ is then a feasible weight function, and then also $\mathbb{P}^w_{e,\sigma} = \mathbb{P}_{e,\sigma}$. We can also take any $w \equiv$const. So for locally bounded $S$, Corollary 9.2 says that any payoff $f$, bounded from below, is superreplicable with an admissible strategy iff $\sup_{Q\in\mathbb{P}_{e,\sigma}} \mathbb{E}_Q(f) \leq 0$.

*Proof of Theorem 9.2.* 1) (9.2) follows from (9.1): Indeed, we know that $w\mathbb{P}^w_{e,\sigma}$ is $L^1(\mathbb{P})$-dense in $w\mathbb{P}^w_{a,\sigma}$. So RHS says that $f$ is in polar of cone $(w, \mathbb{P}^w_{a,\sigma})$. On LHS, $\mathcal{C}^\infty_w$ is convex cone with vertex at 0 and weak*-closed by Corollary 9.1, so that $\mathcal{C}^\infty_w = (\mathcal{C}^\infty_w)^{oo} = ((\mathcal{C}^\infty_w)^o)^o$ by bipolar theorem (see appendix). SO LHS says that $f$ is in polar of $(\mathcal{C}^\infty_w)^o$, and so (9.2) follows from (9.1).

2) Proof of (9.1): "$\supset$" Take any $Q \in \mathbb{P}^w_{e,\sigma}$ and $f \in \mathcal{C}^\infty_w$ so that $wf \in \mathcal{G}_w - L^0_+$. Use Lemma 9.1 part 3) to get $\mathbb{E}_Q(wf) \leq 0$, i.e. $\mathbb{E}_\mathbb{P}(fw\frac{dQ}{d\mathbb{P}}) \leq 0$ for all $f \in \mathcal{C}^\infty_w$. By density, this extends to $Q \in \mathbb{P}^w_{a,\sigma}$ and so polar of $\mathcal{C}^\infty_w \subset w\mathbb{P}^w_{a,\sigma} \subset \text{cone}(w\mathbb{P}^w_{a,\sigma})$.

"$\subset$" Take any $Y \in (\mathcal{C}_w^\infty)^o$; then $Y \in L^1(\mathbb{P})$ and $\mathbb{E}_\mathbb{P}(Yf) \leq 0$, for all $f \in \mathcal{C}_w^\infty$. Now $w \geq 1$ so that $\frac{Y}{w}$ is also in $L^1(\mathbb{P})$. For any $h \in \mathcal{G}_w - L_+^0$ such that $\frac{h}{w}$ is bounded, i.e. $\frac{h}{w} \in \mathcal{C}_w^\infty$, we get

$$\mathbb{E}_\mathbb{P}\left(\frac{Y}{w}h\right) = \mathbb{E}_\mathbb{P}\left(Y\frac{h}{w}\right) \leq 0.$$

Take $h := -1_{\{Y<0\}}$ and note that $\frac{h}{w}$ is bounded as $w \geq 1$. This implies that $Y \geq 0$ $\mathbb{P}$-a.s. and so $\frac{dQ}{d\mathbb{P}} := \text{const}\frac{Y}{w}$ defines a probability measure $Q \ll \mathbb{P}$ with $\mathbb{E}_Q(w) = \text{const}\mathbb{E}_\mathbb{P}(Y) < \infty$ and also $\mathbb{E}_Q(h) \leq 0$ for any $h \in \mathcal{G}_w - L_+^0$ with $\frac{h}{w}$ bounded. We argue below that $Q$ is therefore in $\mathbb{P}_{a,\sigma}^w$ and so $Y = w\frac{Y}{w} = \text{const}.w\frac{dQ}{d\mathbb{P}} \in \text{cone}(w\mathbb{P}_{a,\sigma}^w)$ which gives "$\subset$".

**Need to show**: $S$ is a $Q$-$\sigma$-martingale.
$w$ is a feasible weight function, so there exists $\zeta \in L(S)$, $\zeta > 0$ with $(\zeta \cdot S)_T^* \leq w$. Take $s \leq t$, $A \in \mathcal{F}_s$ and set

$$h_\pm^i := \pm 1_A \int_s^t \zeta_u dS_u^i.$$

These are in $\mathcal{G}_w - L_+^0$ with $\frac{h_\pm^i}{w}$ bounded, hence $\mathbb{E}_Q(h_\pm^i) \leq 0$. This implies

$$\mathbb{E}_Q\left(1_A \int_s^t \zeta_u dS_u^i\right) = 0.$$

Moreover, $\mathbb{E}_Q(w) < \infty$, and so we get that $M := \int \zeta dS$ is a $Q$-martingale. So:

$$S - S_0 = \int \frac{1}{\zeta} dM$$

shows that $S$ is a $Q$-$\sigma$-martingale. $\square$

# 10 Superreplication, pricing, and hedgeable payoffs

**Goal**: study hedging and pricing of a given payoff, and characterize those for which seller and buyer price agree.

**Setup as usual**: $S^0 \equiv 1$, $S$ $\mathbb{R}^d$-valued semimartingale on $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ on $[0, T]$, and (NFLVR) for $S$ or $\mathbb{P}_{e,\sigma} \neq \emptyset$. **Fix feasible weight function** $w$.

**Note**: most concepts and results depend on choice of $w$, even if notation does not show this.

**Definition 10.1.** *Take $f \in L^0(\mathcal{F}_T)$. If $f \geq -w$, call $f$ ($w$-)* ***superreplicable at price*** *$a \in \mathbb{R}$ if $f - a \in \mathcal{G}_w - L^0_+$. Define $\Gamma_+ = \Gamma_+(f) := \{a \in \mathbb{R} : f - a \in \mathcal{G}_w - L^0_+\}$ and call $\alpha = \alpha(f) := \inf \Gamma_+$* ***the superreplication*** *or* ***seller*** *or* ***ask price*** *of $f$ (w.r.t. $w$). In same way, if $f \leq w$, call $f$* ***subreplicable at price*** *$b \in \mathbb{R}$ if $-f + b \in \mathcal{G}_w - L^0_+$, and set $\Gamma_- := \{b \in \mathbb{R} : -f + b \in \mathcal{G}_w - L^0_+\}$ and call $\beta = \sup \Gamma_-$* ***subreplication/buyer/bid price*** *of $f$.*

**Intuition**: Suppose we get $a$. Let us use $w$-admissible strategy $\varphi \stackrel{\wedge}{=} (0, \vartheta)$ to generate $g = G_T(\vartheta)$ from 0 in wealth. Then $f - a \leq g$ means that total result $a + g$ is enough to cover payout $f$; so we can sell $f$ at $a$ without risk. If we spend $b$ and use some $\vartheta$ to generate $g = G_T(\vartheta)$, then $-f + b \leq g$ means that at $T$, total expense $b - g \leq f$ is less than buying $f$ at $T$ outright; so should be willing to buy $f$ at time 0 for $b$. Taking inf/sup then gives competitive prices.

**Remark 10.1.**

1. By definition, $\Gamma_+$ and $\Gamma_-$ are both intervals, and $\beta(f) = -\alpha(-f)$. So results for $\alpha$ translate directly into results for $\beta$.

2. Intuition suggests $\alpha \geq \beta$. See later for proof.

First result says that if $\alpha, \beta$ are finite, then they are attained; this is analogous to Theorem 8.2.

**Theorem 10.1.** *Suppose $S$ satisfies (NFLVR) and $w$ is a feasible weight function. If $f \geq -w$ and $\alpha(f) = \inf \Gamma_+(f)$ is in $\mathbb{R}$, then $\alpha \in \Gamma_+$, i.e. $\Gamma_+ = [\alpha, \infty)$ is closed and there exists $g \in \mathcal{G}_w$ with $\alpha + g \geq f$, i.e. $f$ is superreplicable at $\alpha$. If $f \leq w$ and $\beta(f) \in \mathbb{R}$, then $\beta \in \Gamma_-$ and $\Gamma_- = (-\infty, \beta]$.*

*Proof.* Only argue for $\alpha$. Next, by the definitions, it is enough to show that $\Gamma_+$ is closed. So take $(a_n)_{n \in \mathbb{N}} \subset \Gamma_+$ decreasing to $\alpha$ and for each $n$ some $g_n \in \mathcal{G}_w$

with $f \leq g_n + a_n$. Then $(a_n)_{n\in\mathbb{N}}$ is bounded by some $A > 0$, and for $f \geq -w$ and $w \geq 1$ yields $g_n \geq f - a_n \geq -w(1 + A) =: -w'$. This gives $\mathcal{G}_{w'} = \mathcal{G}_w$ and $\mathbb{E}_Q(w') = (1 + A)\mathbb{E}_Q(w) < \infty$ for all $Q \in \mathbb{P}^w_{e,\sigma}$.

Next use Lemma 9.1, part 4) to get that $g_n = \mathcal{G}_T(\vartheta^n)$ with $\vartheta^n$ $(1, w')$-admissible, i.e. $g_n \in \mathcal{G}^1_{w'} \subset \mathcal{G}^1_{w'} - L^0_+$. Moreover, $g_n = f - a_n \to f - \alpha$ a.s., hence in $L^0$. By Theorem 9.1 $\mathcal{G}^1_{w'} - L^0_+$ is closed in $L^0$ and so $f - \alpha \in \mathcal{G}^1_{w'} - L^0_+ \subset \mathcal{G}_{w'} - L^0_+ = \mathcal{G}_w - L^0_+$. This means that $\alpha \in \Gamma_+$ and so we are done. $\square$

**Remark 10.2.** Above proof looks very simple, whereas proof of Theorem 8.2 used optional decomposition. This is misleading: key argument behind 9.1 (whose proof we skipped) also underlies optional decomposition. See D/S, Section 15.5.

Suppose that $|f| \leq w$ and $\alpha, \beta$ are both finite. Then we expect $\beta \leq \alpha$ (so that $\Gamma_+ \cap \Gamma_-$ contains at most one point). This is indeed true: For any $c \in \mathbb{R}$ with $f - c \in \mathcal{G}_w - L^0_+$, we get from Lemma 9.1 part 3) that $\mathbb{E}_Q(f - c) \leq 0$, for all $Q \in \mathbb{P}^w_{e,\sigma} \neq \emptyset$ (by (NFLVR)). Similarly, if $-f + c \in \mathcal{G}_w - L^0_+$, then $\mathbb{E}_Q(-f + c) \leq 0$. So by Corollary 9.2, we have both $f - \alpha, -f + \beta$ are in $\mathcal{G}_w - L^0_+$, and so

$$\beta \leq \inf_{Q\in\mathbb{P}^w_{e,\sigma}} \mathbb{E}_Q(f) \leq \sup_{Q\in\mathbb{P}^w_{e,\sigma}} \mathbb{E}_Q(f) \leq \alpha. \tag{$\bigstar$}$$

Essentially, this is the same argument as for Lemma 8.1.

**Theorem 10.2** (Hedging duality, general form)**.** *Suppose $S$ satisfies (NFLVR) and $w$ is a feasible weight function. If $f \geq -w$, then*

$$\alpha = \inf\{a \in \mathbb{R} : f - a \in \mathcal{G}_w - L^0_+\} = \sup_{Q\in\mathbb{P}^w_{e,\sigma}} \mathbb{E}_Q(f), \tag{9.1}$$

*and if these expressions are finite, infimum is attained as minimum.*
*If $f \leq w$, then*

$$\beta = \sup\{b \in \mathbb{R} : -f + b \in \mathcal{G}_w - L^0_+\} = \inf_{Q\in\mathbb{P}^w_{e,\sigma}} \mathbb{E}_Q(f),$$

*and if these expressions are finite, supremum is attained as maximum.*

*Proof.* Only argue for $\alpha$, by discussion (see ($\bigstar$)) above and by Theorem 10.1, we only need to show that $\alpha \leq \sup_{Q\in\mathbb{P}^w_{e,\sigma}} \mathbb{E}_Q(f)$, and so w.l.o.g. we can assume the RHS to be finite.

Take any $a < \alpha = \inf \Gamma_+$; then by definition, $f - a \notin \mathcal{G}_w - L^0_+$. But we have $f \geq -w$ and $w \geq 1$, so $f - a \geq -w(1 + |a|) =: -w'$ and $w'$ is a feasible weight function like $w$, and $\mathcal{G}_{w'} = \mathcal{G}_w$. So $f - a \geq -w'$ is not in $\mathcal{G}_{w'} - L^0_+$, and thus Corollary 9.2 implies that $\mathbb{E}_{\widetilde{Q}}(f-a) > 0$ for some $\widetilde{Q} \in \mathbb{P}^{w'}_{e,\sigma} = \mathbb{P}^w_{e,\sigma}$. So we get

$$a < \mathbb{E}_{\widetilde{Q}}(f) \leq \sup_{Q \in \mathbb{P}^w_{e,\sigma}} \mathbb{E}_Q(f) < \infty,$$

and as $a < \alpha$ was arbitrary, we get $\alpha \leq \sup_{Q \in \mathbb{P}^w_{e,\sigma}} \mathbb{E}_Q(f)$. This is enough. $\square$

By above results, any payoff $|f| \leq w$ should have a **price within interval**

$$[\beta, \alpha] = \left[ \inf_{Q \in \mathbb{P}^w_{e,\sigma}} \mathbb{E}_Q(f), \sup_{Q \in \mathbb{P}^w_{e,\sigma}} \mathbb{E}_Q(f) \right].$$

This gives arbitrage-free price bounds. For $\beta = \alpha$, there is a unique price, and we study later what this implies about $f$. For $\beta < \alpha$, we have

**Lemma 10.1.** *Suppose $S$ satisfies (NFLVR) and $w$ is a feasible weight function. Let $f$ be a* ***bounded*** *payoff with $|f| \leq w$ (no serious restriction). Then if $-\infty < \beta(f) < \alpha(f) < \infty$ we have:*

1. *$f$ is strictly superreplicable at price $\alpha$, i.e. there exists $g \in \mathcal{G}_w$ with $f \leq \alpha + g$ $\mathbb{P}$-a.s. and $\mathbb{P}(f < \alpha + g) > 0$.*
2. *$f$ is strictly subreplicable at price $\beta$.*

**Intuition**: We cannot superreplicate $f$ for less than $\alpha$ - but there is some surplus at $T$.

*Proof.* By symmetry, we only argue 1). By Theorem 10.1 $\alpha \in \Gamma_+$ and so there exists $g_0 \in \mathcal{G}_w$ with $f \leq \alpha + g_0$. Write $g_0 = G_T(\vartheta^0)$ with $\vartheta^0$ $w$-admissible and define

$$\tau := \inf\{t \in [0,T] : G_t(\vartheta^0) \geq 1 + |\alpha| + \|f\|_\infty\}.$$

By RC, $G_\tau(\vartheta^0) \geq 1 + |\alpha| + \|f\|_\infty \geq \max(0, 1 + f - \alpha)$ on $\{\tau \leq T\}$. Moreover, set $\vartheta := \vartheta^0 1_{(\!(0,\tau]\!]} \in L(S)$ to get

$$G_t(\vartheta) = G_t(\vartheta^0) 1_{\{t<\tau\}} + G_\tau(\vartheta^0) 1_{\{t \geq \tau\}}. \tag{$\star$}$$

We know that $\vartheta^0$ is $w$-admissible, and $G_\tau(\vartheta^0) \geq 0 \geq -a\mathbb{E}_Q(w \mid \mathbb{F})$, for $a > 0$ and $Q \in \mathbb{P}^w_{e,\sigma}$, and so also $\vartheta$ is $w$-admissible. Next,

$$\begin{aligned} \alpha + G_T(\vartheta) &\overset{(\star)}{=} (\alpha + g_0) 1_{\{T<\tau\}} + (\alpha + G_\tau(\vartheta^0)) 1_{\{T \geq \tau\}} \\ &\geq f 1_{\{T<\tau\}} + (1+f) 1_{\{T \geq \tau\}} = f + 1_{\{T \geq \tau\}}. \end{aligned}$$

So $g := G_T(\vartheta) \in \mathcal{G}_w$ superreplicates $f$ from initial wealth $\alpha$, and this is already strict if $\mathbb{P}(T \geq \tau) > 0$.

What happens if $\mathbb{P}(T \geq \tau) = 0$, i.e. if $\tau > T$ $\mathbb{P}$-a.s.?

In this case, we have $G_{\cdot}(\vartheta^0)$ bounded above (by $1 + |\alpha| + \|f\|_\infty$); and because $\vartheta^0$ is $w$-admissible, it is also for each $Q \in \mathbb{P}^w_{e,\sigma}$ bounded below by some constant multiple of $Q$-martingale $\mathbb{E}_Q(w \mid \mathbb{F})$. So $G_{\cdot}(\vartheta^0)$ is a local $Q$-martingale (by Ansel-Stricker) and then a true $Q$-martingale (since it is bounded), for any $Q \in \mathbb{P}^w_{e,\sigma}$. So we get

$$\mathbb{E}_Q(g_0) = \mathbb{E}_Q(G_T(\vartheta^0)) = 0, \text{ for all } Q \in \mathbb{P}^w_{e,\sigma}$$

and we also have $f \leq \alpha + g_0$ a.s. If we had equality ("=") a.s., we should get $\mathbb{E}_Q(f) = \alpha$ for all $Q \in \mathbb{P}^w_{e,\sigma}$, which gives $\alpha = \beta$, a contradiction. So also in case $\tau > T$ $\mathbb{P}$-a.s., we get $\mathbb{P}(f < \alpha + g_0) > 0$ and we are done. $\square$

**Remark 10.3.**

1. Proof crucially exploits that $f$ is bounded. If only $|f| \leq w$, it is not clear what happens.
2. Assumption $|f| \leq w$ is harmless if $f$ is bounded; indeed, take any feasible weight function $w_0$ and set $w := w_0(\|f\|_\infty \vee 1)$ to get feasible weight function $w$ with $|f| \leq w$ and $\mathcal{G}_w = \mathcal{G}_{w_0}$. So such $w$ always exists.

Now we study those payoffs $f$ for which $\alpha = \beta$ i.e. buyer and seller price agree. Let us first discuss how one might define **attainability** of a given payoff.

**Recall**: We called $f$ superreplicable at price $a$ if $f - a \in \mathcal{G}_w - L^0_+$, i.e. $f \leq a + g$ for some $g \in \mathcal{G}_w$. So we might want to call $f$ attainable at price $c$ if $f = c + g$ for some $g \in \mathcal{G}_w$. Then one could try to define seller price also as inf over such $c$. But there are two problems:

1. It may happen that (even if $\Gamma_+ \neq \emptyset$), getting $f = c + g$ a.s. is impossible (even if $f \leq c + g$ a.s. works). So trying to have exact replication is too restrictive.
2. It could happen that $f = c + g$ a.s. for some $g \in \mathcal{G}_w$, but also $f \leq c + g'$ a.s. for some other $g' \in \mathcal{G}_w$ with $\mathbb{P}(c + g' > f) > 0$. So $\vartheta$ from $G_T(\vartheta) = g$ would exactly replicate $f$ at cost $c$, but strategy $\vartheta'$ from $G_T(\vartheta') = g'$ achieves strictly more than $\vartheta$ at some initial cost $c$. Translated back to $f$, this means that for amount $c$, we can afford payoff $c + g = f$; but this payoff $f$ is not a good choice because we could also afford, at some price $c$, the better payoff $f' = c + g'$.

In order to avoid issue 2) as described on the above, we recall the following definition of maximality:

**Definition 10.2.** *For a subset $A \subset L^0$, an element $a \in A$ is* ***maximal in*** *$A$ if for $a' \in A$ with $a' \geq a$ a.s. implies $a' = a$ a.s. (In particular, $f$ as in issue 2) is not maximal in $c + \mathcal{G}_w$)*

**Definition 10.3.** *Let $w$ be a feasible weight function. Then $f \in L^0$ is $w$-hedgeable if $f = c + g$ a.s. with $c \in \mathbb{R}$ and $g \in \mathcal{G}_w$ is maximal in $\mathcal{G}_w$.*

Loosely speaking, non-maximal $g \in \mathcal{G}_w$ means **stupid trading** - one could do strictly better at same cost.

**Remark 10.4.** Having non-maximal $g \in \mathcal{G}_w$ does **not imply** that there is arbitrage! Indeed: Suppose $g, g' \in \mathcal{G}_w$ with $g' \geq g$ a.s. and $\mathbb{P}(g' > g) > 0$. Would like to buy $g'$ and sell $g$ to generate from 0 initial wealth final payoff $g' - g \in L^0_+ \setminus \{0\}$. **But**: these "assets" need not be for sale; and writing $g = G_T(\vartheta)$, $g' = G_T(\vartheta')$ and then using $\varphi \overset{\wedge}{=} (v_0 = 0, \vartheta' - \vartheta)$ has $V_0(\varphi) = v_0 = 0$ and $V_T(\varphi) = G_T(\vartheta' - \vartheta) = g' - g \in L^0_+ \setminus \{0\}$, but $\vartheta' - \vartheta$ doesn't have to be $w$-admissible!

Back to our question: what can we say about $f$ if $\alpha(f) = \beta(f)$?

**Proposition 10.1.** *Suppose (NFLVR) and $w$ is a feasible weight function. Let $f$ be payoff with $|f| \leq w$. If $\alpha = \beta$ and both are in $\mathbb{R}$, then $f$ is $w$-hedgeable.*

*Proof.* By Theorem 10.1 $\alpha \in \Gamma_+$ and $\beta \in \Gamma_-$ so that $f - \alpha$ and $-f + \beta = -(f - \alpha)$ are both in $\mathcal{G}_w - L^0_+$. So there exists $g_1, g_2 \in \mathcal{G}_w$ with $f \leq \alpha + g_1$ and $-f \leq -\alpha + g_2$. Adding these up gives $g_1 + g_2 \geq 0$ a.s. but Lemma 9.1 part 1) also gives $\mathbb{E}_Q(g_1 + g_2) \leq 0$ for any $Q \in \mathbb{P}^w_{e,\sigma}$ so that we get $g_1 + g_2 = 0$ a.s. or $g_2 = -g_1$ a.s. But then $\alpha + g_1 = \alpha - g_2 \leq f \leq \alpha + g_1$ a.s. and thus $f = \alpha + g_1$ a.s.

It remains to show that $g_1$ is maximal in $\mathcal{G}_w$. By Theorem 10.2 the assumption $\alpha = \beta$ means that

$$\sup_{Q \in \mathbb{P}^w_{e,\sigma}} \mathbb{E}_Q(f) = \inf_{Q \in \mathbb{P}^w_{e,\sigma}} \mathbb{E}_Q(f)$$

and thus the mapping $Q \mapsto \mathbb{E}_Q(f)$ is finite-valued (because $\alpha = \beta \in \mathbb{R}$) and constant, with value $\alpha$. But $f = \alpha + g_1$ a.s. and so we must have $\mathbb{E}_Q(g_1) = 0$ for all $Q \in \mathbb{P}^w_{e,\sigma}$ and in particular

$$\mathbb{E}_Q(g_1) = 0, \text{ for some } Q \in \mathbb{P}^w_{e,\sigma}. \tag{10.2}$$

But (10.2) already implies that $g_1$ is maximal in $\mathcal{G}_w$. Indeed: if $g' \in \mathcal{G}_w$ with $g' \geq g_1$ a.s. then we get

$$0 \overset{(10.2)}{=} \mathbb{E}_Q(g_1) \leq \mathbb{E}_Q(g') \overset{\text{L 9.1,1}}{\leq} 0$$

and therefore $g' = g_1$ a.s. ($Q$ and $\mathbb{P}$ a.s.). □

Proof of Proposition 10.1 shows that $f$ is $w$-hedgeable if it has **same expectation** under all $Q \in \mathbb{P}^w_{e,\sigma}$. This is analogous to discrete-time result, and also intuitive: if $f$ has **unique price**, then it can be hedged.

Converse looks even more plausible: if $f$ is $w$-hedgeable, then $f = c + g$ a.s. and $g \in \mathcal{G}_w$ is maximal; so price of $f$ should be $c$, and -as in discrete time- should be computable as $c = \mathbb{E}_Q(f)$ for any $Q \in \mathbb{P}^w_{e,\sigma}$. Surprisingly, this is not true! If $f$ is bounded, this is correct (see below); but in general, there can be $g \in \mathcal{G}_w$ with $\mathbb{E}_Q(g) = 0$ for some $Q \in \mathbb{P}^w_{e,\sigma}$, but $\mathbb{E}_{Q'}(g) < 0$ for some other $Q' \in \mathbb{P}^w_{e,\sigma}$. See next example.

Correct extension of Theorem 7.1 to general case is as follows:

**Theorem 10.3** (Characterisation of $w$-hedgeable payoffs)**.** *Suppose (NFLVR) and $w$ is a feasible weight function. For any payoff $f \geq -w$ the following are equivalent:*

*1. $f$ is $w$-hedgeable.*

*2. $f = c + g$ for some $c \in \mathbb{R}$ and some $g \in \mathcal{G}_w$ satisfying $\mathbb{E}_{Q^*}(g) = 0$ for some $Q^* \in \mathbb{P}^w_{e,\sigma}$.*

*2'. $f = c + G_T(\vartheta)$ for some $c \in \mathbb{R}$ and $\vartheta \in \Theta_w$ such that $G_.(\vartheta)$ is a $Q^*$-martingale on $[0, T]$ for some $Q^* \in \mathbb{P}^w_{e,\sigma}$.*

*3. $\sup_{Q \in \mathbb{P}^w_{e,\sigma}} \mathbb{E}_Q(f)$ is finite and attained in some $Q^{**} \in \mathbb{P}^w_{e,\sigma}$.*

**Remark 10.5.** If $f$ is $w$-hedgeable and **bounded**, then in $2')$ stochastic integral process $G_.(\vartheta)$ is $Q^*$-martingale whose final value $f - c$ is bounded. But then $G_.(\vartheta)$ is bounded, and so it is $Q$-martingale under all $Q \in \mathbb{P}^w_{e,\sigma}$. In turn, this gives $\mathbb{E}_Q(g) = \mathbb{E}_Q(G_T(\vartheta)) = 0$ for all $Q \in \mathbb{P}^w_{e,\sigma}$ and so $Q \mapsto \mathbb{E}_Q(f) \equiv c$ is constant over $\mathbb{P}^w_{e,\sigma}$ - exactly as in discrete time!

*Proof.* "2) $\iff$ 2')": For any $g \in \mathcal{G}_w$ with $g = G_T(\vartheta)$ and $\vartheta \in \Theta_w$, $G_.(\vartheta)$ is by Lemma 9.1 under any $Q \in \mathbb{P}^w_{e,\sigma}$ a $Q$-supermartingale, null at 0.
So $0 = \mathbb{E}_{Q^*}(g) = \mathbb{E}_{Q^*}(G_T(\vartheta))$ is equivalent to $G_.(\vartheta)$ being a $Q^*$-martingale.

"2) $\implies$ 3)": For any $Q \in \mathbb{P}^w_{e,\sigma}$, part 1) of Lemma 9.1 gives $\mathbb{E}_Q(f) = c + \mathbb{E}_Q(g) \leq c$; so if we have equality for $Q^*$, then $Q^{**} := Q^*$ attains the supremum (and that is also finite).

"3) $\implies$ 1)": Set $c := \sup_{Q \in \mathbb{P}^w_{e,\sigma}} \mathbb{E}_Q(f) = \mathbb{E}_{Q^*}(f)$. This is finite, by assumption, and so Theorem 10.2 and Theorem 10.1 gives that $c \in \Gamma_+$ so that $f \leq c + g$

for some $g \in \mathcal{G}_w$. By Lemma 9.1 part 1) we get

$$c = \mathbb{E}_{Q^*}(f) \overset{f \le c+g}{\le} c + \mathbb{E}_{Q^*}(g) \overset{\text{L 9.1}}{\le} c$$

and so we must have $f = c+g$ a.s. and $\mathbb{E}_{Q^*}(g) = 0$. This implies that $g$ is maximal in $\mathcal{G}_w$ (see argument after (10.2)) and so 1) follows.

"1) $\implies$ 2)": We need to show existence of some $Q^* \in \mathbb{P}^w_{e,\sigma}$ with $\mathbb{E}_{Q^*}(g) = 0$, assuming that $g \in \mathcal{G}_w$ is maximal in $\mathcal{G}_w$.

**Idea**: use variant of Kreps-Yan Theorem (see Appendix) for $p = \infty$. But $g$ is not bounded...
**Goal**: find another $w'$ such that $\frac{g}{w} \in L^\infty$ and apply Kreps-Yan theorem to $\frac{g}{w'}$.

First, $\mathbb{E}_Q(g) \le 0$ by Lemma 9.1 and so $g^+ \in L^1(Q)$, for any $Q \in \mathbb{P}^w_{e,\sigma}$. Define $w' := w + g^+$. This is again a feasible weight function and $\mathbb{P}^w_{e,\sigma} = \mathbb{P}^{w'}_{e,\sigma}$. Next, $\mathcal{G}_w \subset \mathcal{G}_{w'}$ and so we do not know if $g$ is also maximal in $\mathcal{G}_{w'}$. But it is: Indeed, if $\widetilde{g} \ge g$ and $\widetilde{g} \in \mathcal{G}_{w'}$, then $\widetilde{g} \ge g \ge -aw$ (as $g \in \mathcal{G}_w$) and so part 4) of Lemma 9.1 implies that actually $\widetilde{g} \in \mathcal{G}_{aw} = \mathcal{G}_w$. But now maximality of $g$ in $\mathcal{G}_w$ gives $\widetilde{g} = g$ a.s. and so $g \in \mathcal{G}_w \subset \mathcal{G}_{w'}$ is even maximal in $\mathcal{G}_{w'}$ as claimed.

**Now**: $g \ge -aw$ for some $a \ge 0$, as $g \in \mathcal{G}_w$, and $g \le g^+$. So $\frac{g}{w'} = \frac{g}{w+g^+}$ is bounded. Moreover, we claim that

$$\left(\mathcal{C}_{w'} - \frac{g}{w'}\right) \bigcap L^\infty_+ = \{0\},$$

because $g$ is maximal in $\mathcal{G}_{w'}$.

⌈ If $h \in \mathcal{C}_w - \frac{g}{w'}$, then $h = \frac{1}{w'}(\widetilde{g} - \widetilde{Y}) - \frac{g}{w'}$ with $\widetilde{g} \in \mathcal{G}_{w'}, \widetilde{Y} \ge 0$. If also $h \ge 0$, then we get $\widetilde{g} = hw' + \widetilde{Y} + g \ge g$. By maximality of $g$ in $\mathcal{G}_{w'}$, this yields $\widetilde{g} = g$ a.s., hence $hw' + \widetilde{Y} = 0$, and this forces $\widetilde{Y} = 0$ and $h = 0$.
This proves the claim above. ⌋

Next, $\mathcal{C}_{w'}$ is weak* closed in $L^\infty$ by Corollary 9.1. So we can apply modified Kreps-Yan Theorem (Appendix) and get some $Q_0^* \approx \mathbb{P}$ with $\mathbb{E}_{Q_0^*}(Y) \le 0$, for all $Y \in \mathcal{C}_{w'}$, and $\mathbb{E}_{Q_0^*}(\frac{g}{w'}) = 0$. Define probability measure $Q^* \approx \mathbb{P}$ by $\frac{dQ^*}{dQ_0^*} := \text{const}\frac{1}{w'}$. Then $\mathbb{E}_{Q^*}(w') = \text{const} < \infty$, $\mathbb{E}_{Q^*}(g) = 0$ and $\mathbb{E}_{Q^*}(Yw') \le 0$ for all $Y \in \mathcal{C}_{w'}$.

But now, for any $\widetilde{Y} \in \mathcal{G}_{w'} - L^0_+$ such that $\frac{\widetilde{Y}}{w'}$ is bounded, $\widetilde{Y}/w'$ is in $\mathcal{C}_{w'}$, by definition. So $\mathbb{E}_{Q^*}(\widetilde{Y}) \le 0$ for all such $\widetilde{Y}$. Now last step of proof of Theorem 9.2 shows that therefore $Q^* \in \mathbb{P}_{e,\sigma}$, and so also $Q^* \in \mathbb{P}^{w'}_{e,\sigma} = \mathbb{P}^w_{e,\sigma}$. That is 1) $\implies$ 2). □

**Example 10.1.** *For unbounded $f$, 3) in Theorem 10.3 cannot be improved to*

*3'. $Q \mapsto \mathbb{E}_Q(f)$ is constant over $Q \in \mathbb{P}^w_{e,\sigma}$*

*We construct model and $g \in \mathcal{G}_w$ with $\mathbb{E}_{Q'}(g) = 0$ for some $Q' \in \mathbb{P}^w_{e,\sigma}$ (and so $g$ is maximal in $\mathcal{G}_w$, see after (10.2)), but $\mathbb{E}_Q(g) < 0$ for some other $Q \in \mathbb{P}^w_{e,\sigma}$. **So**: a complete analogy to the discrete-time result of Theorem 7.1 is impossible.*

*For simplicity, work on $[0, \infty)$ or $[0, \infty]$. Let $W, W'$ be independent BMs under $\mathbb{P}$, define $X_t := \mathcal{E}(W)_t = \exp(W_t - \frac{1}{2}t)$ and $Y := \mathcal{E}(W')$ and*

$$\begin{aligned}\sigma &:= \inf\{t \geq 0 : X_t = 1/2\},\\ \tau &:= \inf\{t \geq 0 : Y_t = 2\}.\end{aligned}$$

*By LLN for BM, $X_t \to 0$ a.s. as $t \to \infty$ so that $\sigma < \infty$ $\mathbb{P}$-a.s. Also $Y_t \to 0$ as $t \to \infty$, but $\mathbb{P}(\tau = \infty) = 1/2$. **Indeed**: $Y_\tau = 2$ on $\{\tau < \infty\}$, $Y_\tau = 0$ on $\{\tau = \infty\}$ and $Y^\tau$ is bounded martingale. So stopping theorem gives*

$$1 = \mathbb{E}(Y_0) = \mathbb{E}(Y_\tau) = 2\mathbb{P}(\tau < \infty).$$

*Now define $S := X^{\tau\wedge\sigma}$ so that $X \in \mathcal{M}_{loc}(\mathbb{P})$ with $S_0 = 1$. Then $S_\infty = X_{\tau\wedge\sigma} = X_\sigma 1_{\{\tau=\infty\}} + X_{\tau\wedge\sigma}1_{\{\tau<\infty\}}$ and $X_\sigma = 1/2$. But $W, W'$ are independent; so $X$ and $\sigma$ are independent of $\tau$ and $\mathbb{E}(X_{\sigma\wedge t}) = \mathbb{E}(X_0) = 1$ for any fixed $t$. So we can compute:*

$$\begin{aligned}\mathbb{E}(S_\infty) &= \frac{1}{2}\mathbb{P}(\tau = \infty) + \mathbb{E}[\underbrace{\mathbb{E}(X_{\sigma\wedge\tau} \mid \tau)}_{\overset{indep.}{=} \underbrace{\mathbb{E}(X_{\sigma\wedge t})}_{=1}|_{t=\tau}} 1_{\{\tau<\infty\}}]\\ &= \frac{1}{2}\underbrace{\mathbb{P}(\tau = \infty)}_{=1/2} + 1\underbrace{\mathbb{P}(\tau < \infty)}_{=1/2} = \frac{3}{4} < 1.\end{aligned}$$

*So the payoff $g := S_\infty - S_0 = \int_0^\infty 1dS$ is in $\mathcal{G}_1$ because $\int_0^\cdot 1dS = S - S_0 \geq -1$, and $Q := \mathbb{P} \in \mathbb{P}^1_{e,loc}$, but $\mathbb{E}_Q(g) < 0$.*

*To find $Q' \in \mathbb{P}^1_{e,loc}$ with $\mathbb{E}_{Q'}(g) = 0$ (which means that $S$ is a **true** $Q'$-martingale) we first define $Z := Y^{\tau\wedge\sigma}$. This is a martingale on $[0, \infty)$ and bounded (by 2) hence of class (D), and so we can define $Q' \ll \mathbb{P}$ by $\frac{dQ'}{d\mathbb{P}} := Z_\infty = Y_{\tau\wedge\sigma}$; but $\sigma < \infty$ $\mathbb{P}$-a.s. and $Y > 0$ on $[0, \infty)$ so that $Z_\infty > 0$ a.s. and so $Q' \approx \mathbb{P}$.*

***Claim**: $S$ is a $Q'$-martingale on $[0, \infty]$ (which then implies $\mathbb{E}_{Q'}(g) = 0$). By Bayes, it is equivalent to argue that $ZS$ is $\mathbb{P}$-martingale on $[0, \infty]$. But*

$$SZ = (XY)^{\tau\wedge\sigma} = (\mathcal{E}(W)\mathcal{E}(W'))^{\tau\wedge\sigma} = (\mathcal{E}(W + W'))^{\tau\wedge\sigma}$$

*is a nonnegative martingale on* $[0,\infty)$ *and hence a supermartingale on* $[0,\infty)$ *as well as on* $[0,\infty]$. ***So***: *We only need to argue that* $\mathbb{E}(S_\infty Z_\infty) = 1$.

***Now***: $Y^\tau$ *is* $\mathbb{P}$-*UI martingale so that stopping theorem gives* $Y_{\tau\wedge\sigma} = \mathbb{E}(Y_\tau \mid \mathcal{F}_{\tau\wedge\sigma})$, *and* $Y_\tau = 0$ *on* $\{\tau = \infty\}$, $Y_\tau = 2$ *on* $\{\tau < \infty\}$. *Moreover,* $\mathbb{E}(X_{\tau\wedge\sigma} 1_{\{\tau<\infty\}}) = 1/2$ *as seen before.* ***Hence***:

$$\mathbb{E}(S_\infty Z_\infty) = \mathbb{E}(X_{\tau\wedge\sigma} Y_{\tau\wedge\sigma}) \overset{\mathbb{E}(Y_\tau|\mathcal{F}_{\tau\wedge\sigma})=Y_{\tau\wedge\sigma}}{=} \mathbb{E}(X_{\tau\wedge\sigma} Y_\tau) = \mathbb{E}(X_{\tau\wedge\sigma} 2 \cdot 1_{\{\tau<\infty\}}) = 2 \cdot \frac{1}{2} = 1.$$

*This concludes the example.*

# 11 Utility maximisation I: The primal problem

**Goal**: study basic problem of **optimal portfolio choice** with preferences given by **expected utility** from terminal wealth.

**Setup**: standard model of financial market with $T < \infty$, $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ (satisfying the usual conditions), bank account $S^0 \equiv 1$ and discounted asset prices given by $\mathbb{R}^d$-valued semimartingale $S = (S_t)_{0\leq t\leq T}$. Assume $\mathcal{F}_0$ is trivial (i.e. know everything at beginning) and impose absence of arbitrage via $\mathbb{P}_{e,\sigma} \neq \emptyset$.

Fix **initial wealth** $x > 0$ and consider self financing strategy $\varphi \stackrel{\wedge}{=} (x, \vartheta)$ with $\vartheta \in L(S)$ predictable $\mathbb{R}^d$-valued. Resulting **wealth process** is $V(\varphi) = V(x,\vartheta) = x + G(\vartheta) = x + \int \vartheta dS$ and we want $V(\varphi) \geq 0$; so strategy $\varphi \stackrel{\wedge}{=} (x,\vartheta)$ is 0-admissible and integrand $\vartheta$ is $x$-admissible, $\vartheta \in \Theta^x_{\mathrm{adm}}$.

**Goal**: find 0-admissible strategy $\varphi \stackrel{\wedge}{=} (x,\vartheta)$ to **maximise expected utility** $\mathbb{E}[U(V_T(x,\vartheta))]$ **from terminal wealth** (over all $\vartheta \in \Theta^x_{\mathrm{adm}}$) where $U$ is a utility function.

**Notation**: for $x > 0$, set

$$\begin{aligned}\mathcal{V}(x) &:= \left\{ V = V(x,\vartheta) = x + \int \vartheta dS : \vartheta \in L(S) \text{ and } V \geq 0 \right\} \\ &= \left\{ x + \int \vartheta dS : \vartheta \in \Theta^x_{\mathrm{adm}} \right\}.\end{aligned}$$

**Primal optimisation problem**: find $u(x) := \sup_{V\in\mathcal{V}(x)} \mathbb{E}(U(V_T))$ for $x > 0$, where $U$ is a utility function (on $(0,\infty)$).

**Definition 11.1.** *A* ***utility function*** *(on $(0,\infty)$) is a mapping $U : (0,\infty) \to \mathbb{R}$, strictly increasing, strictly concave, in $C^1$ and satisfying* ***Inada conditions****:*

$$\begin{aligned} U'(0) &:= \lim_{x\searrow 0} U'(x) = +\infty, \\ U'(\infty) &:= \lim_{x\to\infty} U'(x) = 0. \end{aligned}$$

**Interpretation**: $U$ quantifies subjective preferences: monetary amount $z$ is given a subjective utility/level of happiness $U(z)$. Increasing means *"more is better" (people are greedy)* and concave means *"an extra dollar means more increase of happiness for a beggar than for a billionaire"*. Then **indirect utility** $u(x)$ is maximal expected utility one can achieve from initial wealth $x$ via trading with 0-admissible strategy $\varphi \stackrel{\wedge}{=} (x,\vartheta)$.

**Remark 11.1.**

0. Having $\varphi$ 0-admissible matches with dom$U = (0,\infty)$. Having dom$U = (-a,\infty)$ with $a \geq 0$ is easily treated via translation. Allowing dom$U = \mathbb{R}$ becomes more complicated for defining "allowed" strategies, see e.g. Biagini/Frittelli (2005,2008).

1. A priori, $U$ is defined only on $(0,\infty)$. But $U(0) := \lim_{x\searrow 0} U(x)$ exists by monotonicity and is in $[-\infty,+\infty)$ and so $U(V_T)$ is well defined on $[-\infty,+\infty)$.

2. For any random variable $f \geq 0$, set $\mathbb{E}(U(f)) := -\infty$ if $U(f) \notin L^1(\mathbb{P})$. This is harmless because we maximize and $\vartheta \equiv 0$ gives $u(x) \geq U(x) > -\infty$.

3. If $S$ is not locally bounded, it may happen that $\Theta^x_{\mathrm{adm}} = \{0\}$ and $\mathcal{V}(x) = \{x\}$. Then problem is trivial with $u(x) = U(x)$.

4. Consider the following:

   **Exercise 11.1.** *If $S$ allows arbitrage and $U$ is unbounded above, then $u \equiv +\infty$.*

   So in general, problem only makes sense in arbitrage-free market.

**Questions**:

(a) Does there **exist** optimal strategy $\vartheta^* \in \Theta^x_{\mathrm{adm}}$, i.e.

$$\mathbb{E}[U(V_T(x,\vartheta^*))] = u(x) = \sup_{\vartheta\in\Theta^x_{\mathrm{adm}}} \mathbb{E}[U(V_T(x,\vartheta))] \ ?$$

If yes, how to find or describe it?

(b) How does optimal expected utility $u(x)$ depend on $x$?

**Exercise 11.2** (Easy)**.** *If $U$ is increasing and concave (not necessarily strictly) then so is $u$. If also $u(x_0) < \infty$ for some $x_0 > 0$, then $u(x) < \infty$ for all $x > 0$. We can get stronger properties with (one) extra condition on $U$; see later.*

To make our problem more tractable, generalise:

$$\begin{aligned}\mathcal{C}(x) &:= \{f \in L^0_+(\mathcal{F}_T) : f \leq V_T \text{ for some } V \in \mathcal{V}(x)\}\\ &= \{x + G_T(\vartheta) - Y \geq 0 : \vartheta \in \Theta^x_{\mathrm{adm}}, Y \geq 0\ \mathcal{F}_T\text{-measurable}\}\\ &= (x + G_T(\Theta^x_{\mathrm{adm}}) - L^0_+) \cap L^0_+(\mathcal{F}_T)\\ &= (x + \mathcal{G}^x_{\mathrm{adm}} - L^0_+) \cap L^0_+(\mathcal{F}_T) :\end{aligned}$$

space of nonnegative time-$T$ payoffs one can superreplicate from initial wealth $x$ via 0-admissible self-financing strategy.

Of course, $\mathcal{C}(x) \supset \{V_T : V \in \mathcal{V}(x)\}$ and we have

$$u(x) := \sup_{V \in \mathcal{V}(x)} \mathbb{E}(U(V_T)) = \sup_{f \in \mathcal{C}(x)} \mathbb{E}(U(f)). \tag{$\spadesuit$}$$

$\lceil$ "$\leq$" is clear from the inclusion above. Conversely, if $f \in \mathcal{C}(x)$, then $f \leq V_T$ for some $V \in \mathcal{V}(x)$; and as $U$ is increasing we have $\mathbb{E}(U(f)) \leq \mathbb{E}(U(V_T)) \leq u(x)$ $\rfloor$

**Note**: If $f^* \in \mathcal{C}(x)$ is optimal, then $f^* \leq V_T(x, \vartheta^*)$ for some $\vartheta^* \in \Theta^x_{\text{adm}}$, and then $\vartheta^*$ is optimal for original problem (follows by the equality demonstrated above).

The set $\mathcal{C}(x)$ is **convex** and **solid**, i.e. $f \in \mathcal{C}(x)$ and $0 \leq f' \leq f$ implies $f' \in \mathcal{C}(x)$. We also have:

**Lemma 11.1.** *Assume (NFLVR) and $\mathcal{F}_0$ is trivial. Then*

$$\mathcal{C}(x) = \{f \in L^0_+(\mathcal{F}_T) : \mathbb{E}_Q(f) \leq x, \textit{ for all } Q \in \mathbb{P}_{e,\sigma}\}.$$

*Proof.* This is a variation of the hedging duality in Theorem 8.2.
For "$\subset$", note that $G(\vartheta) = \int \vartheta dS$ is for any $\vartheta \in \Theta^x_{\text{adm}}$ and any $Q \in \mathbb{P}_{e,\sigma}$ a $Q$-supermartingale (Ansel-Stricker). So $f \leq x + G_T(\vartheta)$ gives

$$\mathbb{E}_Q(f) \leq x + \mathbb{E}_Q(G_T(\vartheta)) \leq x.$$

For "$\supset$", define $U$ as RCLL version of $U_t := \operatorname{ess\,sup}_{Q \in \mathbb{P}_{e,\sigma}} \mathbb{E}_Q(f \mid \mathcal{F}_t)$ for $0 \leq t \leq T$. Note that $U_T = f$ as $f$ is $\mathcal{F}_T$-measurable, and $U_0 = \sup_{Q \in \mathbb{P}_{e,\sigma}} \mathbb{E}_Q(f) \leq x$ as $\mathcal{F}_0$ is trivial. Then Proposition 8.1 and Theorem 8.1 imply that $U = U_0 + \int \vartheta dS - C$ with $\vartheta \in \Theta_{\text{adm}}$ and $C$ is increasing, null at 0. So

$$\int \vartheta dS = \underbrace{U}_{\geq 0} - \underbrace{U_0}_{\leq x} + \underbrace{C}_{\geq 0} \geq -x$$

hence $\vartheta \in \Theta^x_{\text{adm}}$ and

$$f = U_T = U_0 + G_T(\vartheta) - C_T \leq x + G_T(\vartheta), \ \vartheta \in \Theta^x_{\text{adm}}$$

so that $f \in \mathcal{C}(x)$. $\square$

In view of Lemma 11.1 (and the identity ($\spadesuit$) above), we can rewrite the primal problem as maximise $\mathbb{E}(U(f))$ over $f \in L_+(\mathcal{F}_T)$ subject to $\mathbb{E}_Q(f) \leq x$, for all $Q \in \mathbb{P}_{e,\sigma}$. But we cannot simply write Lagrange function, because number of constraints (the for all $Q \in \mathbb{P}_{e,\sigma}$ part) is in general infinite.

It turns out to be useful to generalise E$\sigma$MDs as well. Start with $Q \in \mathbb{P}_{e,\sigma}$, denote by $Z = Z^{Q;\mathbb{P}}$ its density process and take any $V \in \mathcal{V}(x)$, so $V = x + \int \vartheta dS$, $\vartheta \in \Theta^x_{\text{adm}}$. Then $Z > 0$ is a $\mathbb{P}$-martingale with $Z_0 = 1$ (as $\mathcal{F}_0$ is trivial) and by Bayes and Ansel-Stricker, $ZV$ is a $\mathbb{P}$-supermartingale. This motivates:

**Definition 11.2.** *For $z > 0$, we call $\mathcal{Z}(z)$ the family of all nonnegative $\mathbb{F}$-adapted RCLL process $Z = (Z_t)_{0 \le t \le T}$ with $Z_0 = z$ such that $ZV$ is a $\mathbb{P}$-supermartingale for all $V \in \mathcal{V}(1)$ (or, equivalently, for all $V \in \mathcal{V}(x)$, for all $x > 0$; note that $\mathcal{V}(x) = x\mathcal{V}(1)$.)*

Note that $\vartheta \equiv 0$ gives $V(1,0) = 1 \in \mathcal{V}(1)$; so each $Z \in \mathcal{Z}(z)$ is itself a $\mathbb{P}$-supermartingale. Next $Z^{Q;\mathbb{P}} \in \mathcal{Z}(1)$ for any $Q \in \mathbb{P}_{e,\sigma}$ and $\mathcal{Z}(z) = z\mathcal{Z}(1)$ for all $z > 0$; so $\mathcal{Z}(z) \neq \emptyset$ by (NFLVR).

In analogy to $\mathcal{C}(x)$, define for $z > 0$

$$\mathcal{D}(z) := \{h \in L^0_+ : h \le Z_T \text{ for some } Z \in \mathcal{Z}(z)\}.$$

Now take any $V \in \mathcal{V}(x),\ Z \in \mathcal{Z}(z)$. Then $ZV$ is $\mathbb{P}$-supermartingale starting at $zx$ so that $\mathbb{E}(Z_T V_T) \le zx$. Define

$$J(y) := \sup_{x>0}(U(x) - xy), \text{ for } y > 0.$$

Then $J$ is **decreasing** and **convex** as supremum of convex (actually affine) function $\mathcal{C}_x(y) = U(x) - xy$. Then $U(V_T) \le J(Z_T) + Z_T V_T$ and so $\mathbb{E}(U(V_T)) \le \mathbb{E}(J(Z_T)) + zx$. If we take supremum over $\mathcal{V}(x)$ on LHS, inf over $\mathcal{Z}(z)$ on the RHS, we get

$$u(x) \le j(z) + xz \text{ for all } x > 0,\ z > 0,$$

where we set $j(z) := \inf_{Z \in \mathcal{Z}(z)} \mathbb{E}(J(Z_T))$. This $j$ is called the **dual problem**.

**Note**: primal problem maximizes concave functional, dual problem minimizes convex functional.

In analogy to primal problem, we also have

**Exercise 11.3.**

$$j(z) := \inf_{Z \in \mathcal{Z}(z)} \mathbb{E}(J(Z_T)) = \inf_{h \in \mathcal{D}(z)} \mathbb{E}(J(h)).$$

Finally we also get $j(z) \ge \sup_{x>0}(u(x) - xz)$ for all $z > 0$ and $u(x) \le \inf_{z>0}(j(z) + zx)$ for all $x > 0$. We shall see later that under one extra assumption on $U$, we actually have equalities, and this will have many useful consequences.

The next result slightly extends Lemma 11.1.

**Lemma 11.2.** *Assume (NFLVR) and $\mathcal{F}_0$ is trivial. For any $f \in L^0_+(\mathcal{F}_T)$, we then have*

$$f \in \mathcal{C}(x) \iff \sup_{h \in \mathcal{D}(1)} \mathbb{E}(fh) \leq x.$$

*As a consequence, $\mathcal{C}(x)$ is closed in $L^0$.*

*Proof.* "$\Longleftarrow$" For any $Q \in \mathbb{P}_{e,\sigma}$, $h := Z_T^{Q;\mathbb{P}} = \frac{dQ}{d\mathbb{P}}$ is in $\mathcal{D}(1)$, so we can use Lemma 11.1 to get $f \in \mathcal{C}(x)$.
" $\Longrightarrow$ " $f \leq V_T$ with $V \in \mathcal{V}(x)$, $h \leq Z_T$ with $Z \in \mathcal{Z}(1)$ gives $\mathbb{E}(fh) \leq \mathbb{E}(V_T Z_T) \leq x$ because $VZ$ is $\mathbb{P}$-supermartingale started at $x$.

Finally, $f_n \to f$ in $L^0$ for $f_n \in \mathcal{C}(x)$, thus w.l.o.g. $f_n \to f$ $\mathbb{P}$-a.s. (on a subsequence). Then we have

$$\mathbb{E}(fh) \overset{\text{Fatou}}{\leq} \mathbb{E}(f_n h) \leq x.$$

Hence $f \in \mathcal{C}(x)$ and $\mathcal{C}(x)$ is closed in $L^0$. □

# 12 Utility maximisation II: The dual problem and its use

**Goal**: prove existence of solution to the dual problem and show how this helps to tackle the primal problem.

Recall that for a utility function $U$ on $(0, \infty)$, we have defined

$$J(y) := \sup_{x>0}(U(x) - xy), \text{ for } y > 0$$

(convex, decreasing). This is the **Legendre transform** or convex conjugate of $-U(-\cdot)$, with convention $U(x) := \infty$ for $x < 0$. See Rockafellar, Chapter 12. We collect properties of $J$.

**Lemma 12.1.** *If $U$ is a utility function, then $J : (0, \infty) \to \mathbb{R}$ is strictly decreasing, strictly convex, in $C^1$ with $J'(0) = -\infty$, $J'(\infty) = 0$ as well as $J(0) = U(\infty)$, $J(\infty) = U(0)$. Moreover, we have conjugacy relation*

$$U(x) = \inf_{y>0}(J(y) + xy), \text{ for } x > 0.$$

*In addition $J' = -I := -(U')^{-1}$ is minus inverse of $U'$, and $J(y) = U(I(y)) - yI(y)$.*

**Example 12.1.** *Classical utilities on $(0, \infty)$ are $U(x) = \log x$ with $J(y) = -\log y - 1$, and $U(x) = \frac{1}{\gamma}x^\gamma$ with $\gamma < 1$, $\gamma \neq 0$, where $J(y) = \frac{1-\gamma}{\gamma}y^{\frac{\gamma}{\gamma-1}}$: **power**/CRRA utility. Note that for $\gamma < 0$, then $U \leq 0$ is bounded above; for $\gamma > 0$, then $U \geq 0$ is unbounded above. For $\gamma \to 0$ (By Bernoulli de L'Hospital) we get* log *utility.*

**Goal now**: show that **dual problem** of finding

$$j(z) := \inf_{h \in \mathcal{D}(z)} \mathbb{E}(J(h))$$

has a unique solution $h^*_z \in \mathcal{D}(z)$ if $j(z) < \infty$, i.e. function $F(h) := \mathbb{E}(J(h))$ attains infimum over over $\mathcal{D}(z)$ in a unique $h^*_z$.

**Classic**: continuous function on compact set has minimum. How? Approximate infimum along a sequence; use compactness to get converging subsequence whose limit is candidate; use continuity of $F$ to compute $F$ in limit and show its minimality.

In our settting, use Komlós lemma (Lemma 5.3) to produce candidate; then show that $\mathcal{D}(z)$ is convex and closed in $L^0$ so that it also contains candidate; then show that $F$ is convex and lower semicontinuous on $\mathcal{D}(z)$ to obtain that candidate is minimiser.

**Proposition 12.1.** *For any $z > 0$, the set $\mathcal{D}(z)$ is convex, solid and closed in $L^0$.*

*Proof.* Recall $\mathcal{D}(z) = \{h \in L^0_+ : h \leq Z_T \text{ for some } Z \in \mathcal{Z}(z)\}$. So solid is clear and convex holds because $\mathcal{Z}(z)$ is convex.

The proof for closedness is a bit more complicated than for Proposition IV 3.2. in IMF because we have time index $t \in [0,T]$. Take $(h_n)_{n\in\mathbb{N}}$ in $\mathcal{D}(z)$ converging to $L^0$ to some $h$, so $h \geq 0$ and $h < \infty \mathbb{P}$-a.s., i.e. $h \in L^0_+$. W.l.o.g. (take subsequence and call it again $h_n$), we can assume $h_n \to h$ $\mathbb{P}$-a.s. Take $Z^n \in \mathcal{Z}(z)$ with $h_n \leq Z^n_T$. As $h_n \geq 0$, $Z^n \geq 0$, we can use Lemma 5.3 (Komlós) plus diagonal argument to find convex combinations $\widetilde{h}_n$ and $\widetilde{Z}^n_r$, for all rational $r \in [0,T]$, converging all simultaneously, $\mathbb{P}$-a.s., to some $h_\infty$ and $Z^\infty_r$, respectively.

A priori, we might have values $+\infty$ for $h_\infty$; but $h_n \to h$ $\mathbb{P}$-a.s., $\widetilde{h}_n \in \operatorname{conv}(h_n, h_{n+1}, \dots)$, which gives $\widetilde{h}_\infty = h < \infty$ $\mathbb{P}$-a.s. Moreover, $h_n \leq Z^n_T$ $\mathbb{P}$-a.s., for all $n \in \mathbb{N}$, and we can use same convex combinations for $\widetilde{h}_n$ and all $\widetilde{Z}^n_r$; so also get $h_\infty = h \leq Z^\infty_T$.

It only remains to show that $Z^\infty_T \leq Z_T$ for some $Z \in \mathcal{Z}(z)$. Each $\widetilde{Z}^n$ is in $\mathcal{Z}(z)$ as $\mathcal{Z}(z)$ is convex, and so $\widetilde{Z}^n_0 = z$, for all $n \in \mathbb{N}$, hence $Z^\infty_0 = z$, and each $\widetilde{Z}^n V$ is $\mathbb{P}$-supermartingale for each $V \in \mathcal{V}(1)$, and nonnegative. So take rational $r \leq s$ and compute

$$\mathbb{E}(Z^\infty_s V_S \mid \mathcal{F}_r) \overset{\text{Fatou}}{\leq} \liminf_{n\to\infty} \mathbb{E}(\widetilde{Z}^n_s V_s \mid \mathcal{F}_r) \overset{\text{supermg.}}{\leq} \liminf_{nto\infty} \widetilde{Z}^n_r V_r = Z^\infty_r V_r,$$

and so $Z^\infty V$ and also $Z^\infty$ itself (take $V \equiv 1$) are $\mathbb{P}$-supermartingale on $[0,T] \cap \mathbb{Q}$. But now standard construction in martingale theory (see DM, Theorem IV.2, uses $\mathbb{F}$ is RC) gives existence of an RCLL $\mathbb{P}$-supermartingale $Z = (Z_t)_{0\leq t\leq T}$ on $[0,T]$ with $Z_r \leq Z^\infty_r$ for all rational $r$, indeed, can take

$$Z_t := \lim_{r \searrow\searrow t} Z^\infty_r,$$

and so we get $Z_0 = Z^\infty_0 = z$ and $Z_T = Z^\infty_T$. As above, using construction of $Z$ from $Z^\infty$ and Fatou to conclude that $ZV$ is $\mathbb{P}$-supermaringale on $[o,T]$, for all $V \in \mathcal{V}(1)$. So $Z \in \mathcal{Z}(z)$ and $h \leq Z_T$ is in $\mathcal{D}(z)$. $\square$

**Proposition 12.2.** *For each $z > 0$, we have*

1. *$\mathcal{D}(z)$ is bounded in $L^1(\mathbb{P})$, and family $\{(J(h))^- : h \in \mathcal{D}(z)\}$ is $\mathbb{P}$-UI.*
2. *$F : L^0_+(\mathbb{P}) \to [-\infty,\infty]$, $h \mapsto F(h) := \mathbb{E}(J(h))$ is lower-semicontinuous on $\mathcal{D}(z)$: if $h_n \to h$ in $L^0$ with $h_n, h \in \mathcal{D}(z)$, then $F(h) \leq \liminf_{n\to\infty} F(h_n)$.*

*Proof.* See IMF, Proposition IV 3.3. or KS Lemma 3.2. $\square$

**Theorem 12.1.** *Suppose $\mathbb{P}_{e,\sigma} \neq \emptyset$ so that each $\mathcal{D}(z) \neq \emptyset$. For each $z > 0$ with $j(z) < \infty$, there exists a unique solution $h_z^* \in \mathcal{D}(z)$ to the dual problem, i.e.*

$$\inf_{h \in \mathcal{D}(z)} \mathbb{E}(J(h)) = j(z) = \mathbb{E}(J(h_z^*)).$$

*Proof.* Uniqueness is clear from strict convexity of $J$. For existence, take $(h_n)_{n\in\mathbb{N}} \subset \mathcal{D}(z)$ with $F(h_n) = \mathbb{E}(J(h_n)) \searrow j(z) < \infty$. All $h_n$ are nonnegative, so Lemma 5.3 gives $\widetilde{h}_n \in \operatorname{conv}(h_n, h_{n+1}, \dots)$ converging to some $h$ $\mathbb{P}$-a.s. A priori, might have $h = +\infty$ with positive probability. By Proposition 12.2, $\mathcal{D}(z)$ is bounded in $L^1(\mathbb{P})$, hence also in $L^0$, and so Lemma 5.3 implies that $h < \infty$ $\mathbb{P}$-a.s; so $h \in L^0_+$, $\widetilde{h}_n \to h$ $\mathbb{P}$-a.s. and hence in $L^0$.

But $\mathcal{D}(z)$ is convex, so $(\widetilde{h}_n)_{n\in\mathbb{N}} \subset \mathcal{D}(z)$, and closed in $L^0$ (by Proposition 12.1) so that also $h \in \mathcal{D}(z)$. By Lemma 12.1, $J$ is convex, so

$$F(\widetilde{h}_n) = \mathbb{E}(J(\widetilde{h}_n)) \leq \sup_{k\geq n} \mathbb{E}(J(h_n)) = \sup_{k\geq n} F(h_k) = F(h_n),$$

as $n \mapsto F(h_n)$ is decreasing. So $\widetilde{h}_n \in \mathcal{D}(z)$ gives

$$j(z) \leq F(\widetilde{h}_n) \leq F(h_n) \searrow j(z),$$

and hence

$$\mathbb{E}(J(h)) = F(h) \leq \liminf_{n\to\infty} F(\widetilde{h}_n) = j(z),$$

and so $h_z^* := h$ is optimal. $\square$

Convex analysis gives extra properties of $j$.

**Corollary 12.1.** *The function $j$ is decreasing, strictly convex on $\{z > 0 : j(z) < \infty\}$ and continuous on interior of $\{j < \infty\}$ (which is an interval from some $a$ to $+\infty$).*

*Proof.* See IMF C IV 3.5. or KS Lemma 3.3. $\square$

Now let us see how to use the dual problem in order to tackle the primal. Fix $x > 0$ and take $f \in \mathcal{C}(x)$ so that $f \leq V_T$ for some $V \in \mathcal{V}(x)$. Take $z > 0$ and $h \in \mathcal{D}(z)$ so that $h \leq Z_T$ for some $Z \in \mathcal{Z}(z)$. Then

$$\mathbb{E}(fh) \leq \mathbb{E}(V_T Z_T) \overset{VZ \text{ supermg}}{\leq} zx.$$

By definition of $J$, $U(f) \leq J(h) + fh$, and so $\mathbb{E}(U(f)) \leq \mathbb{E}(J(h)) + zx$. Sup over $f \in \mathcal{C}(x)$ gives $u(x)$; inf over $h \in \mathcal{D}(z)$ gives on RHS $j(z) + zx$. But each side also provides a bound for other side, and so we should get optima by making bounds sharp via equality.

So aim for equalities everywhere. Maximiser for $J(y) = \sup_{x>0}(U(x) - xy)$ is given via $U'(x) = y$ or $x = I(y)$ so that, as in Lemma 12.1, $J(y) = U(I(y)) - yI(y)$. So let us choose, for $h \in \mathcal{D}(z)$ $f := I(h)$ to get $U(f) = J(h) + fh$. To get $\mathbb{E}(fh) = xz$, we then want next $\mathbb{E}(hI(h)) = xz$. If we have that, then we get for $h \in \mathcal{D}(z)$,

$$\mathbb{E}[U(I(h))] = \mathbb{E}(J(h)) + xz \geq j(z) + zx \geq \inf_{z>0}(j(z) + zx) \overset{\text{Sec. 11}}{\geq} u(x).$$

First inequality becomes equality for $h := h_z^*$. The second inequality becomes equality if $z$ is a minimiser for $j/z) + zx$, i.e., b¡ solving $j'(z) = -x$.

**Reverse engineering** suggests **recipe**:

1. Start with $x > 0$ and define $z = z_x$ via $-j'(z) = x$.
2. Solve dual problem for $z$ to get $h_z^*$. Then define $f_x^* := I(h_z^*) = I(h_{z_x}^*)$.
3. Show that $\mathbb{E}(h_{z_x}^* I(h_{z_x}^*)) = xz_x$.
4. Show that $f_x^* \in \mathcal{C}(x)$.

If all that can be done, then we obtain

$$\begin{aligned} u(x) &\overset{4)}{\geq} \mathbb{E}(U(f_x^*)) = \mathbb{E}[U(I(h_{z_x}^*))] \overset{J(y)=U(I(y))-yI(y)}{=} \mathbb{E}(J(h_{z_x}^*)) + \mathbb{E}(h_{z_x}^* I(h_{z_x}^*)) \\ &\overset{2),3)}{=} j(z_x) + xz_x \geq \inf_{z>0}(j(z) + xz) \overset{\text{Sect. 11}}{\geq} u(x) = \sup_{f \in \mathcal{C}(x)} \mathbb{E}(U(f)). \end{aligned}$$

This shows that $f_x^*$ is optimal and also that $u(x) = \inf_{z>0}(j(z) + xz)$. If we can make this work for all $x > 0$, then we get

$$j(z_x) \overset{2)}{=} \mathbb{E}(J(h_{z_x}^*)) = \mathbb{E}[U(I(h_{z_x}^*)) - h_{z_x}^* I(h_{z_x}^*)] \overset{2),3)}{\mathbb{E}} (U(f_x^*)) - xz_x \overset{4)}{=} u(x) - xz_x.$$

We also know from Section 11 that for all $z > 0$, $j(z) \geq \sup_{x>0}(u(x) - xz)$. If the range of the achieved $z_x$ is $(0, \infty)$, then we also get $j(z) = \sup_{x>0}(u(x) - xz)$ for all $z > 0$. This means that value functions $u, j$ satisfy same conjugacy as original $U, J$.

So far, just ideas and wishes. Does it really work? The answer to that question is yes, under an extra condition (on $U$), and after quite a bit of auxiliary technical results.

# 13 Utility maximisation III: Auxiliary results

**Goal**: provide auxiliary results in order to show that recipe in 4 steps actually works.

**Standing assumptions in this section**: $\mathbb{P}_{e,\sigma} \neq \emptyset$, $U$ utility function on $(0,\infty)$ and $u(x_0) < \infty$ for some $x_0 > 0$.

To make 1) work, need $j$ to be smooth. As $j(z) = \mathbb{E}(J(h_z^*))$, start by looking at $h_z^*$.

**Lemma 13.1.** *Mapping $(0,\infty) \to L_+^0$, $z \mapsto h_z^*$ is continuous on interior of $\{j < \infty\}$ : If $z > 0$, $j(z) < \infty$ and $z \in int(\{j < \infty\})$ and if $z_n > 0$, $j(z_n) < \infty$, $z_n \to z$, then $h_{z_n}^* \to h_z^*$ in $L^0$.*

*Proof.* See IMF, L IV 5.1. or KS, L 3.6. □

Subsequent results need one extra condition on $U$.

**Definition 13.1.** *$U$ has **reasonable asymptotic elasticity at** $+\infty$; $RAE_{+\infty}(U)$, if*

$$AE_{+\infty}(U) := \limsup_{x\to\infty} \frac{xU'(x)}{U(x)} < 1.$$

**Intuition**: we look at ratio $\frac{U'(x)}{U(x)/x}$. **Marginal utility** $U'(x) = U'(x)1 \overset{\text{Taylor}}{\approx} U(x+1) - U(x)$ **measures increase of utility** as wealth increases from $x$ to $x+1$. On the other hand, if $U(0) = 0$, then

$$\frac{U(x)}{x} = \frac{1}{x}(U(x) - U(0)) = \frac{1}{x}\sum_{j=1}^{x}(U(j) - U(j-1))$$

measures **average increase of utility** as wealth increases successively from 0 to $x$.

**Exercise 13.1.** *$U$ is concave, so $U(j) - U(j-1) \geq U'(x)$ for $x \in (j-1, j)$ so that we always get $AE_{+\infty}(U) \leq 1$.*

Having equality would mean that for large wealth, marginal and average utility behave in same way, or that $U$ is almost linear for large $x$. From exonomic viewpoint, this looks unreasonable.

**Example 13.1.**

- *For* $U(x) = \log x$ *has*
$$\frac{xU'(x)}{U(x)} = \frac{1}{\log x},$$
*so* $AE_{+\infty}(U) = 0$.

- *For* $U(x) = \frac{1}{\gamma}x^\gamma$ *where* $\gamma < 1,\ \gamma \neq 0$, *we get*
$$\frac{xU'(x)}{U(x)} = \gamma,$$
*so* $AE_{+\infty}(U) = \gamma < 1$.

- *If for instance* $U(x) = \frac{x}{\log x}$ *for large* $x$, *then*
$$\frac{xU'(x)}{U(x)} = \log x \frac{\log x - 1}{(\log x)^2} = 1 - \frac{1}{\log x}.$$
*So here we get* $AE_{+\infty}(U) = 1$, *and indeed, this* $U$ *is almost linear for large* $x$.

Condition $\mathrm{RAE}_{+\infty}(U)$ gives on $U$ an estimate of the form $U'(x) \leq \beta \frac{U(x)}{x}$ for large $x$ (and with $\beta < 1$). For conjugate $J$, this translates into a similar condition for small $y$.

**Lemma 13.2.** *Let* $U$ *be a utility function with conjugate* $J$. *If* $AE_{+\infty}(U) < 1$, *then there are* $y_0 > 0$ *and* $C \in (0, \infty)$ *with*
$$-J'(y) \leq C\frac{J(y)}{y}, \text{ for } 0 < y \leq y_0.$$

*Proof.* See IMF, L IV 5.2. or KS, L 6.3. □

**Lemma 13.3.** *If* $AE_{+\infty}(U) < 1$, *then mapping* $(0, \infty) \to \mathbb{R}$, $z \mapsto H(z) := \mathbb{E}(h_z^* I(h_z^*))$ *is continuous on interior of* $\{j < \infty\}$.

*Proof.* In view of Lemma 13.1, only need to check uniform integrability. This uses Proposition 12.2, Lemma 13.2, Corollary 12.1, Lemma 13.1. For details, see IMF L IV 5.3. or KS, L3.7. □

Next result is **key mathematical result**. Note: condition $\mathrm{RAE}_{+\infty}(U)$ is needed only for part of it.

**Theorem 13.1.** *Suppose $\mathbb{P}_{e,\sigma} \neq \emptyset$, $U$ is a utility function (with Inada conditions) and $u(x_0) < \infty$ for some $x_0 > 0$. Then we have conjugacy:*

$$j(z) = \sup_{x>0}(u(x) - xz) \text{ for all } z > 0, \tag{13.1}$$

*(we do not say that sup is finite), and therefore $j(z) < \infty$ for $z \geq z_0$ with some $z_0 \in (0, \infty)$. If $AE_{+\infty}(U) < 1$, then $j(z) < \infty$ for all $z > 0$.*

*Sketch of proof.* Because $J(y) = \sup_{x>0}(U(x) - xy)$, it looks plausible that we should have

$$\mathbb{E}(J(h)) = \sup_{f \in L^\infty_+} \mathbb{E}(U(f) - fh)$$

and so we get

$$j(z) = \inf_{h \in \mathcal{D}(z)} \mathbb{E}(J(h)) = \inf_{h \in \mathcal{D}(z)} \sup_{f \in L^\infty_+} \mathbb{E}(U(f) - fh).$$

We should like to interchange inf and sup, and this needs a **minimax theorem**. Then we should get

$$j(z) = \sup_{f \in L^\infty_+} \inf_{h \in \mathcal{D}(z)} \mathbb{E}(U(f) - fh).$$

We also know that, by Lemma 11.2

$$\inf_{h \in \mathcal{D}(z)} \mathbb{E}(-fh) = -\sup_{h \in \mathcal{D}(z)} \mathbb{E}(fh) > -\infty \text{ iff for some } x > 0,\ f \in \mathcal{C}(x).$$

So in effect, we should get

$$j(z) = \sup_{x>0} \sup_{f \in \mathcal{C}(x)} (\mathbb{E}[U(f)] - xz),$$

which is (13.1). □

⌈ Look at literature shows that for minimax theorems, one almost always needs **compactness** of one set at least. So we approximate $L^\infty_+$ by compact balls, use minimax result, and then pass to limit. In more detail: view $L^\infty$ as dual of $L^1$ and equip it with topology $\sigma(L^\infty, L^1)$. For each $n \in \mathbb{N}$, the ball $B_n := \{f \in L^\infty_+ : f \leq n\}$ is then weak*-compact (Alaoglu Theorem from Functional Analysis). Moreover, each $\mathcal{D}(z)$ is by Proposition 12.2 a convex subset of $L^1$. Next, mapping

$$B_n \times \mathcal{D}(z) \to \mathbb{R},\ (f, h) \mapsto \mathbb{E}(U(f) - fh)$$

is in $f$ concave (like $U$) and in $h$ linear, hence convex, and continuous. [...] So we can use minimax theorem (e.g. Aubin (1979), T 2.7.1.) to get

$$\sup_{f \in B_n} \inf_{h \in \mathcal{D}(z)} \mathbb{E}(U(f) - fh) = \inf_{h \in \mathcal{D}(z)} \sup_{f \in B_n} \mathbb{E}(U(f) - fh).$$

With a lot of work, one can now let $n \to \infty$ and eventually arrives at conjugacy (13.1). For details and rest of argument see IMF, T IV 5.4. or KS, L 3.4. and T 3.2. or L3.8. $\rfloor$

Next step is to derive **smoothness of** $j$.

**Lemma 13.4.** *Suppose $AE_{+\infty}(U) < \infty$. Then $j$ is in $C^1$ on $(0, \infty)$ with $j'$ strictly increasing with $-zj'(z) = \mathbb{E}(h_z^* I(h_z^*))$ for all $z > 0$.*

*Proof.* See IMF, L IV 3.5. or KS, L 3.8. □

Step 1) from recipe asks for solution $z$ of $-j'(z) = x$. As $j$ is continuous and strictly monotonic, uniqueness is clear. For existence, we need to know range of values of $j'$.

**Lemma 13.5.** *We always have $j'(\infty) = \lim_{z \to \infty} j'(z) = 0$. If $AE_{+\infty}(U) < 1$, then also $j'(0) = \lim_{z \searrow 0} j'(z) = -\infty$.*

*Proof.* See IMF, L IV 5.6. or KS, T 3.2. □

Last result in this section gives an extra optimality property of dual optimizer $h_z^*$.

**Lemma 13.6.** *For all $z > 0$ and $h \in \mathcal{D}(z)$, if $j(z) < \infty$, we have*

$$\mathbb{E}(hI(h_z^*)) \leq \mathbb{E}(h_z^* I(h_{z_x}^*)).$$

*Proof.* See IMF, LIV 6.1. or KS L 3.9. □

With all that, we are ready to solve primal problem.

# 14 Utility Maximisation IV: Solving the primal problem

**Goal**: solve problem of finding

$$u(x) := \sup_{f \in \mathcal{C}(x)} \mathbb{E}(U(f)), \text{ for } x > 0.$$

a) via duality; b) directly.

**In Section 12 we obtained the following recipe:**

1. Define $z = z_x$ via $-j'(z) = x > 0$. By Lemma 13.4 and Lemma 13.5, $-j' : (0, \infty) \to (0, \infty)$ is continuous (as $j \in C^1$), strictly decreasing and surjective. So 1) is uniquely for any $x > 0$ and produces $z_x > 0$.

2. Solve dual problem for $z_x > 0$ to get dual optimiser $h^*_{z_x} \in \mathcal{D}(z_x)$. Then set $f^*_x := I(h^*_{z_x})$. By Theorem 13.1, using $\mathrm{AE}_{+\infty}(U) < 1$, we have $j(z) < \infty$ for all $z > 0$; so dual optimiser $h^*_z$ exists, by Theorem 12.1, for any $z > 0$ and is unique.

3. Show that $\mathbb{E}(h^*_{z_x} I(h^*_{z_x})) = x z_x$. By Lemma 13.4, $\mathbb{E}(h^*_z I(h^*_z)) = -zj'(z)$ for any $z > 0$. Now plug in $z_x$ and use step 1).

4. Show that $f^*_x \in \mathcal{C}(x)$. For that, use Lemma 11.2 analogue $\mathbb{E}(hf^*_x) \le x$ for any $h \in \mathcal{D}(1)$. But $\mathcal{D}(z) = z\mathcal{D}(1)$ for any $z > 0$, so take any $h \in \mathcal{D}(1)$ and set $\widetilde{h} := z_x h \in \mathcal{D}(z_x)$. Then compute

$$\mathbb{E}(hf^*_x) = \frac{1}{z_x}\mathbb{E}(\widetilde{h} I(h^*_{z_x})) \overset{\text{L.13.6}}{\le} \frac{1}{z_x}\mathbb{E}[h^*_{z_x} I(h^*_{z_x})] \overset{3)}{=} x.$$

   Using Lemma 11.2, we obtain $f^*_x \in \mathcal{C}(x)$ if it is $\mathcal{F}_T$-measurable. This holds if $\mathcal{F} = \mathcal{F}_T$ (Better way to resolve this: define a priori $\mathcal{D}(z)$ as subset of $L^0(\mathcal{F}_T)$, so automatically, without $\mathcal{F} = \mathcal{F}_T$, $f$ is $\mathcal{F}_T$-measurable).

In summary we get main result:

**Theorem 14.1.** *Suppose $\mathcal{F}_0$ is trivial, $\mathcal{F} = \mathcal{F}_T$, $\mathbb{P}_{e,\sigma} \neq \emptyset$, $U$ is a utility function on $(0, \infty)$ (includes Inada conditions), $u(x_0) < \infty$ for some $x_0 > 0$ and $AE_{+\infty}(U) < 1$. For every $x > 0$, primal problem of maximising expected utility from final wealth has a unique solution $f^*_x \in \mathcal{C}(x)$, given by $f^*_x = I(h^*_{z_x})$, where $h^*_{z_x}$ is unique solution to dual problem for $z_x > 0$ which is defined by $-j'(z_x) = x$.*

*Proof.* By preceding discussion, construction is feasible and produces candidate $f_x^* \in \mathcal{C}(x)$. Uniqueness of solution is direct from strict concavity. Only need to argue optimality of $f_x^*$:

$$u(x) \geq \mathbb{E}[U(f_x^*)] \overset{2)}{=} \mathbb{E}[U(I(h_{z_x}^*)] = \mathbb{E}[J(h_{z_x}^*) + h_{z_x}^* I(h_{z_x}^*)] \overset{2),3)}{=} j(z_x) + xz_x$$
$$\geq \inf_{z>0}(j(z) + xz) \overset{\text{Sect. 11}}{\geq} u(x) = \sup_{f\in\mathcal{C}(x)} \mathbb{E}[U(f)].$$

This proves optimality of $f_x^*$. □

**Remark 14.1.**

1. From Theorem 13.1, we always get (even without AE)

   $$j(z) = \sup_{x>0}(u(x) - xz), \text{ for all } z > 0.$$

   (might be $+\infty$ without AE). Proof of Theorem 14.1 also gives $u(x) = \inf_{z>0}(j(z) + zx)$ for all $x > 0$.

2. How "necessary" is condition $\mathrm{AE}_{+\infty}(U) < 1$? Leaving all other conditions unchanged, it is essentially sharp in the following sense: If we have it, then utility maximisation problem is solvable for any reasonable model $S$. Conversely: If we do not have $\mathrm{AE}_{+\infty}(U) < 1$, one can construct reasonable model $S$ where utility maximisation problem is not solvable. See Kramkov/Schachermayer.

How about solving primal problem directly? Nice to know if all work on dual problem can be avoided, even if that gives a lot of extra properties. If one only wants **existence** of primal optimiser, one can do things, with a few extra assumptions.

Let us introduce extra conditions

$$U(x)^+ \leq k(1 + x^\beta) \text{ for } x > 0, \text{ with constants } k \in (0, \infty), \beta \in (0, 1). \qquad (U+)$$

This is clearly satisfied by power utility $U(x) = \frac{1}{\gamma}x^\gamma$ if $\gamma \in (0, 1)$; for $\gamma < 0$, $U$ is even bounded above.

**Lemma 14.1.** *Suppose that either $U(\infty) < \infty$, or that $U$ satisfies $(U+)$ and also that there exists an $E\sigma MM$ $\widetilde{Q}$ such that $(\frac{d\widetilde{Q}}{d\mathbb{P}})^{-1}$ has moments of all orders. Then $U^+(\mathcal{C}(x)) = \{(U(f))^+ : f \in \mathcal{C}(x)\}$ is $\mathbb{P}$-UI (uniformly integrable) for all $x > 0$.*

*Proof.* If $U(\infty) < \infty$, this is clear. Under $(U+)$, take $p > 1$ with $\beta p < 1$. Then use $(U+)$ to argue that $U^+(\mathcal{C}(x))$ is bounded in $L^p(\mathbb{P})$ (and hence $\mathbb{P}$-UI), as follows:

$$\mathbb{E}[((U(f))^+)^p] \overset{(U+)}{\leq} k^p \mathbb{E}[(1+f^\beta)^p] \leq A(p) + \beta(p)\mathbb{E}[f^{\beta p}]$$

and, setting $\widetilde{Z} := \frac{d\widetilde{Q}}{d\mathbb{P}}$, we get for $r := \frac{1}{\beta p} > 1$ (conjugate to $s$)

$$\mathbb{E}[f^{\beta p}] = \mathbb{E}[(f\widetilde{Z})^{\beta p}(\widetilde{Z})^{-\beta p}] \overset{\text{Hölder}}{\leq} (\mathbb{E}[f\widetilde{Z}])^{1/r}(\mathbb{E}[(\widetilde{Z})^{-s\beta p}])^{1/2} = (\underbrace{\mathbb{E}_{\widetilde{Q}}[f]}_{\leq x \ L11.1})^{\beta p}(\mathbb{E}[\widetilde{Z}^{-s\beta p}])^{1/s}$$

$$\leq x^{\beta p}\text{const.}(\widetilde{Q}, s, \beta, p) < \infty, \text{ uniformly over } f \in \mathcal{C}(x).$$

This is enough. $\square$

**Remark 14.2.** How do we verify the assumption?

**Exercise 14.1.** *Suppose $S$ is continuous and satisfies (S.C.) (structure condition); so $S = S_0 + M + \int d\langle M\rangle\lambda$. Look at MVT process $K = \int \lambda^{tr} d\langle M\rangle\lambda = \langle \int \lambda dM\rangle$. Define $\hat{Z} := \mathcal{E}(-\int \lambda dM) > 0$ and*

$$\frac{d\hat{P}}{d\mathbb{P}} := \hat{Z}_T.$$

*If $K$ is bounded (uniformly in $\omega$, , by a constant) then $\hat{P} \in \mathbb{P}_{e,loc}$ and*

$$\left(\frac{d\hat{P}}{d\mathbb{P}}\right)^{-1}$$

*has moments of all orders.*

**Proposition 14.1.** *Suppose $\mathcal{F}_0$ is trivial, $u(x_0) < \infty$ for some $x_0 > 0$ and either $U(\infty) < \infty$ and $\mathbb{P}_{e,\sigma} \neq \emptyset$, or that $U$ satisfies $(U+)$ and there exists an $E\sigma MM$ $\widetilde{Q}$ s.t. $(d\widetilde{Q}/d\mathbb{P})^{-1}$ has moments of all orders. For any $x > 0$, there exists then unique solution $f^*_x \in \mathcal{C}(x)$ to primal problem.*

*Proof.* Uniqueness is clear from **strict** concavity of $U$. Moreover, $u(x) < 0$ for all $x > 0$ follows from $U$ being concave and increasing (see exercise sheet 8). For existence, take $(f_n)_{n\in\mathbb{N}} \subset \mathcal{C}(x)$ with $\mathbb{E}[U(f_n)] \nearrow u(x) < \infty$, so this can be done. Now all $f_n \geq 0$; so by Lemma 5.3 (Komlos) there exists $\widetilde{f}_n \in \text{conv}(f_n, f_{n+1}, \dots)$ with $\widetilde{f}_n \to f_\infty$ $\mathbb{P}$-a.s. with $f_\infty$ for some $\mathcal{F}_T$-measurable $f_\infty$ with values in $[0, +\infty]$. But by Lemma 11.1, $\mathcal{C}(x)$ is bounded in $L^1(Q)$ for any $Q \in \mathbb{P}_{e,\sigma} \neq \emptyset$, hence also bounded in $L^0(Q)$, which is the same space for $\mathbb{P}$, i.e. $L^0(Q) = L^0(\mathbb{P})$ since $Q \approx \mathbb{P}$. Again by Lemma 5.3, we thus get $f_\infty < \infty$ $\mathbb{P}$-a.s., hence $f_\infty \in L^0_+(\mathcal{F}_T)$ and hence $\widetilde{f}_n \to f_\infty$ in $L^0$.

Now $\mathcal{C}(x)$ is convex, so $\widetilde{f}_n \in \mathcal{C}(x)$ for all $n \in \mathbb{N}$, and $\mathcal{C}(x)$ is closed in $L^0$, so that $f_\infty \in \mathcal{C}(x)$.

**Claim**: $f_x^* := f_\infty$ is optimal $\rightsquigarrow$ need to compute $\mathbb{E}[U(f_\infty)]$.

We have $U(f_n) \to U(f_\infty)$ $\mathbb{P}$-a.s. (since $U$ is concave, then continuous where it is finite). If $U(\infty) < \infty$, we can use Fatou limsup version to get

$$\mathbb{E}[U(f_\infty)] \geq \limsup_{n\to\infty} \mathbb{E}[U(\widetilde{f}_n)].$$

If we have $(U+)$, then by Lemma 14.1

$$\begin{aligned}\mathbb{E}[U^+(f_\infty)] &= \lim_{n\to\infty} \mathbb{E}[U^+(\widetilde{f}_n)],\\ \mathbb{E}[(U^-(f_\infty)] &\overset{U^-\geq 0,\text{ Fatou}}{\leq} \liminf_{n\to\infty} \mathbb{E}[U^-(\widetilde{f}_n)].\end{aligned}$$

So by taking differences we get again

$$\mathbb{E}[U(f_\infty)] \geq \limsup_{n\to\infty} \mathbb{E}[U(\widetilde{f}_n)].$$

But now write

$$\mathbb{E}[U(\widetilde{f}_n)] \geq \inf_{k\geq n} \mathbb{E}[U(f_k)] \overset{\mathbb{E}[U(f_k)]\nearrow u(x)}{=} \mathbb{E}[U(f_n)]$$

moreover, $n \mapsto \mathbb{E}[U(f_n)]$ is increasing, and so

$$u(x) \geq \mathbb{E}[U(f_\infty)] \geq \limsup_{n\to\infty} \mathbb{E}[U(\widetilde{f}_n)] \geq \limsup_{n\to\infty} \mathbb{E}[(U(f_n)] = u(x)$$

so $f_\infty$ is optimal. $\square$

Alternatively we have:

**Proposition 14.2.** *Suppose $\mathcal{F}_0$ is trivial, $u(x_0) < \infty$ for some $x_0 > 0$, $\mathbb{P}_{e,\sigma} \neq \emptyset$ and either $U(\infty) < \infty$ or ($U \geq 0$ plus $AE_{+\infty}(U) < 1$). For any $x > 0$, primal problem has unique solution $f_x^* \in \mathcal{C}(x)$.*

*Proof.* See IMF, T. IV. 7.2. and subsequent remark. $\square$

**Remark 14.3.** Example $U(x) = \log(x)$ does not satisfy assumption of Proposition 14.2, but does satisfy $(U+)$; so if there exists $\widetilde{Q}$, E$\sigma$MM, Proposition 14.1 applies.

# 15 Appendix

## 15.1 Appendix A: The Kreps–Yan theorem

This section contains an important separation theorem prov ed independently by D. Kreps and J.-A. Yan around the same time. It is a crucial ingredient for proving most versions of the fundamental theorem of asset pricing and also comes up in the Bichteler–Dellacherie characterisation of semimartingales as good integrators in stochastic analysis.

We begin by recalling some concepts and results from functional analysis. Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For $p \in [1, \infty)$, the dual of the space $L^p$ is $(L^p)* = L`q$ with $q$ conjugate to $p$, meaning that $\frac{1}{p} + \frac{1}{q} = 1$. This is not true for $p = \infty$. If we fix conjugate numbers $p, q$ both in $[1, \infty]$, the dual pairing between $L^p$ and $L^q$ is given by

$$(Y, Z) := \mathbb{E}(YZ), \text{ for } Y \in L^p, Z \in L^q.$$

• For $p \in [1, \infty)$, the *weak topology* on $L^p$, denoted by $\sigma(L^p, L^q)$, is the defined to be the coarsest topology on $L^p$ which makes all the linear functionals $Y \mapsto (Y, Z)$ continuous for all $Z \in L^q$. So a sequence $(Y_n)_{n\in\mathbb{N}} \subset L^p$ converges to $Y$ in $\sigma(L^p, L^q)$ if and only if

$$\lim_{n\to\infty} \mathbb{E}(Y_n Z) = \mathbb{E}(YZ) \text{ for each } Z \in L^q.$$

• For $p \in (1, \infty]$, the *weak* toplogy* on $L^p$, also denoted by $\sigma(L^p, L^q)$, vies $L^p$ as the dual of $L^q$ (which explains why we must take $p > 1$); it is the coarsest topology on $L^p$ which makes all the linear functions $Y \mapsto (Y, Z)$ continuous for all $Z \in L^q$.

It is clear from the above definitions that for $1 < p < \infty$, the weak and the weak* topology coincide. For $p = 1$, there is only the weak topology on $L^1$, with $Y_n \to Y$ in $\sigma(L^1, L^\infty)$ if and only if $\lim_{n\to\infty} \mathbb{E}(Y_n Z) = \mathbb{E}(YZ)$ for each $Z \in L^\infty$. For $p = \infty$, there is only the weak* topology on $L^\infty$, with $Z_n \to Z$ in $\sigma(L^\infty, L^1)$ if and only if $\lim_{n\to\infty} \mathbb{E}(YZ_n) = \mathbb{E}(YZ)$ for each $Y \in L^1$.

Finally, a convex subset of $L^p$, for $p \in [1, \infty)$, is weakly closed, i.e. closed for the weak topology $\sigma(L^p, L^q)$, if and only if it is (strongly) closed in $L^p$, i.e. for the norm-topology on $L^p$. Note that $p = \infty$ is again not allowed here.

After these preliminaries, we are now in a position to formulate and prove the announced separation result.

**Theorem 15.1** (Kreps/Yan). *Fix conjugate $p, q \in [1, \infty]$ and suppose that $C \subset L^p$ is a convex cone with $C \supset -L^p_+$ and $C \cap L^p_+ = \{0\}$. If $C$ is closed in $\sigma(L^p, L^q)$ (meaning that it is weak* closed if $p = \infty$), then there exists a probability measure $Q \approx \mathbb{P}$ with $\frac{dQ}{d\mathbb{P}} \in L^q$ and $\mathbb{E}_Q(Y) \leq 0$ for all $Y \in C$.*

*Proof.* The proof consists of a combination of a separation argument with an exhaustion argument and goes as follows.

1) For any fixed $x \in L^p_+ \setminus \{0\}$, the assumption gives $x \notin C$. The Hahn-Banach theorem thus allows us to strictly separate $x$ from $C$: i.e. there exists some $z_x \in L^q$ with $(x, z_x) > \alpha$ and $(Y, z_x) \leq \alpha$ for all $Y \in C$. Because $C$ is a cone, we may take $\alpha = 0$. Choosing $Y := -1_{\{z_x<0\}}$, which is in $C$ because $C \supset -L^p_+$, next gives $-\mathbb{E}(z_x 1_{\{z_x<0\}}) = (Y, z_x) \leq 0$ and therefore $z_x \geq 0$, and because the separation is strict, we must have $z_x \not\equiv 0$ to avoid $(x, z_x) = 0$. So we can, and we will, normalise $z_x$ to have $\mathbb{E}(z_x) = 1$, for each $x \in L^p_+ \setminus \{0\}$.

2) Now consider the family $\mathcal{G}$ of all sets $\Gamma_x = \{z_x > 0\} \in \mathcal{F}$, where $x$ runs through $L^p_+ \setminus \{0\}$. For any set $A \in \mathcal{F}$ with $\mathbb{P}(A) > 0$, we have $\mathbb{P}(A \cap \Gamma_x) > 0$ for some $\Gamma_x \in \mathcal{G}$; indeed, $1_A \in L^p_+ \setminus \{0\}$ and therefore we can take $x = I_A$ and use that

$$0 < \mathbb{E}(1_a z_{1_a}) = \mathbb{E}(1_A z_{1_A} 1_{\{z_{1_A}>0\}}) = \mathbb{E}(1_A z_{1_A} 1_{\Gamma_{1_A}})$$

to conclude that we must have $\mathbb{P}(A \cap \Gamma_{1_A}) > 0$. By the proceeding Lemma below, this implies that the family $\mathcal{G}$ contains a countable subfamily of sets whose union has probability 1. So there is a sequence $(x_n)_{n\in\mathbb{N}}$ in $L^p_+$ such that

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} \Gamma_{x_n}\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} \{z_{x_n} > 0\}\right) = 1.$$

Defining $z := \sum_{n=1}^{\infty} 2^{-n} z_{x_n}$ therefore yields a random variable $z > 0$ $\mathbb{P}$-a.s. which is in $L^q$ like all the $z_{x_n}$, and we also have

$$\mathbb{E}(Yz) = \sum_{n=1}^{\infty} 2^{-n} \mathbb{E}(Y z_{x_n}) \leq 0, \text{ for all } Y \in C.$$

Finally, monotone integration yields

$$\mathbb{E}(z) = \sum_{n=1}^{\infty} 2^{-n} \mathbb{E}(z_{x_n}) = 1$$

so that $dQ := z d\mathbb{P}$ gives the desires probability measure. □

The following abstract result provides the missing step in the proof of Kreps/Yan Theorem.

**Lemma 15.1.** *Let $\Lambda \neq \emptyset$ be an index family and $\mathcal{G} = (\Gamma_\lambda)_{\lambda\in\Lambda}$ a family of sets in $\mathcal{F}$ such that any set $A \in \mathcal{F}$ with $\mathbb{P}(A) > 0$ has a nontrivial intersection with some $\Gamma_\lambda \in \mathcal{G}$, meaning that $\mathbb{P}(A\cap\Gamma_\lambda) > 0$. Then there exists an at most countably subfamiliy $(\Gamma_{\lambda_n})_{n\in\mathbb{N}}$ of sets in $\mathcal{G}$ whose union has probability 1.*

*Proof.* Suppose first that $\mathcal{G}$ is closed under countable unions. Then $\sup_{\lambda\in\Lambda}\mathbb{P}(A_\lambda)$ is attained in some $\Gamma_{\lambda^*} \in \mathcal{G}$, because we can approximate the supremum along a sequence $(\Gamma_{\lambda_m})_{m\in\mathbb{N}}$ and take $\Gamma_{\lambda^*} := \cup_{m=1}^{\infty}\Gamma_{\lambda_m}$, which is in $\mathcal{G}$ by the above closedness assumption. If we had $\mathbb{P}(\Gamma^c_{\lambda^*}) > 0$, we could find a set $\Gamma_\lambda \in \mathcal{G}$ with $\mathbb{P}(\Gamma^c_{\lambda^*}\cap\Gamma_\lambda) > 0$ by the assumption on $\mathcal{G}$, and so we should get $\mathbb{P}(\Gamma_\lambda \cup \Gamma_{\lambda^*}) > \mathbb{P}(\Gamma_{\lambda^*})$, contradicting the maximality of $\Gamma_{\lambda^*}$. So $\Gamma_{\lambda^*}$ has probability 1 and we can take the family consisting of this single set.

In general, we consider the family $\mathcal{G}'$ formed by all countable unions of sets from $\mathcal{G}$; this family satisfies the same assumptions as $\mathcal{G}$. Applying the above argument to $\mathcal{G}'$ then gives the assertion. ∎

## 15.2 Appendix B: A Komlós-type lemma from probability theory

These notes provide a formulation and proof for an elementary lemma from probability theory which is extremely useful in many optimisation problems involving convexity. Recall that $L^0$ denotes the vector space of all (equivalence classes of, for the equivalence relation of equality $\mathbb{P}$-a.s.s) random variables on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and taking values in $\mathbb{R}$. For a sequence $(Y_n)_{n\in\mathbb{N}}$ in $L^0$, we denote for $m \in \mathbb{N}$ by $\mathrm{conv}(Y_m, Y_{m+1}, \dots)$ the set of all (finite) convex combinations of $(Y_k)_{k\geq m}$ i.e. all $Y$ of the form

$$Y = \sum_{k=m}^{\infty} \lambda_k Y_k$$

with the $\lambda_k \geq 0$ satisfying $\sum_{k=m}^{\infty} \lambda_k = 1$ and at most finally many $\lambda_k \neq 0$.

**Lemma 15.2.** *For any sequence $(Y_n)_{n\in\mathbb{N}}$ of nonnegative random variables, there exists a sequence $(\widetilde{Y}_n)_{n\in\mathbb{N}}$ with $\widetilde{Y}_n \in conv(Y_n, Y_{n+1}, \dots)$ for all $n$ and $\widetilde{Y}_n \to \widetilde{Y}_\infty$ $\mathbb{P}$-a.s. for some random variable $\widetilde{Y}_\infty$ taking values in $[0, +\infty]$. Moreover, if $\mathbb{P}(Y_n \geq \alpha) \geq \delta > 0$ for some $\alpha > 0$, then $\mathbb{P}(\widetilde{Y}_\infty > 0) > 0$. If $conv(Y_1, Y_2, \dots)$ is bounded in $L^0$, then $\widetilde{Y}_\infty < \infty$ $\mathbb{P}$-a.s.*

*Proof.* Set $C_n := \mathrm{conv}(Y_n, Y_{n+1}, \dots) \supset C_{n+1}$ so that the sequence $J_n := \inf_{Y\in C_n} \mathbb{E}(e^{-Y})$ increases to some $J \leq 1$. Take a sequence $(Y'_n)_{n\in\mathbb{N}}$ with $Y'_n \in C_n$ and $\mathbb{E}(e^{-Y'_n}) \leq J_n + \frac{1}{n}$ for all $n$. For $\epsilon > 0$, define the set

$$B_\epsilon := \left\{ (x, y) \in [0, \infty)^2 : |x - y| \geq \epsilon \text{ and } x \wedge y \leq \frac{1}{\epsilon} \right\}.$$

As the mapping $z \mapsto e^{-z}$ is convex, we always have

$$e^{-(x+y)/2} \leq \frac{1}{2}(e^{-x} + e^{-y}).$$

For $(x, y) \in B_\epsilon$, a calculation gives

$$e^{-(x+y)/2} - \frac{1}{2}(e^{-x} + e^{-y}) \leq -\delta, \text{ for some } \delta = \delta(\epsilon) > 0,$$

and therefore

$$e^{-(x+y)/2} \leq \frac{1}{2}(e^{-x} + e^{-y}) - \delta 1_{B_\epsilon}(x, y).$$

Choosing $x := Y'_m$ and $y := Y'_n$ yields for $n \neq m$ that

$$
\begin{aligned}
J_{n\wedge m} &\leq \mathbb{E}\left(e^{-(Y'_m+Y'_n)/2}\right) \\
&\leq \frac{1}{2}\left(\mathbb{E}[e^{-Y'_m}] + \mathbb{E}[e^{-Y'_n}]\right) - \delta\mathbb{P}[(Y'_m, Y'_n) \in B_\epsilon] \\
&\leq \frac{1}{2}\left(J_m + \frac{1}{m} + J_n + \frac{1}{n}\right) - \delta\mathbb{P}[(Y'_m, Y'_n) \in B_\epsilon],
\end{aligned}
$$

and so we obtain that

$$\lim_{n,m\to\infty} \mathbb{P}[(Y'_m, Y'_n) \in B_\epsilon] = 0.$$

Considering the separate cases $|x-y| < \epsilon$ or $x \wedge y > \frac{1}{\epsilon}$ or $(x, y) \in B_\epsilon$ leads to the estimate

$$|e^{-x} - e^{-y}| \leq \epsilon + 2e^{-1/\epsilon} + 2 \times 1_{B_\epsilon}(x, y).$$

This gives in turn that

$$\left|\mathbb{E}[e^{-Y'_m} - e^{-Y'_n}]\right| \leq \epsilon + 2e^{-1/\epsilon} + 2\mathbb{P}[(Y'_m, Y'_n) \in B_\epsilon]$$

so that $(e^{-Y'_n})_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^1(\mathbb{P})$ and hence convergent in $L^1(\mathbb{P})$. Therefore this sequence has a subsequence $(e^{-\widetilde{Y}_n})_{n\in\mathbb{N}}$ which converges $\mathbb{P}$-a.s., and then the sequence $(\widetilde{Y}_n)_{n\in\mathbb{N}}$ is also $\mathbb{P}$-a.s. convergent and has $\widetilde{Y}_n \in C_n$ like for $Y'_n$.

If $\mathbb{P}(Y_n \geq \alpha) \geq \delta > 0$, then $\mathbb{E}(e^{-Y_n}) \leq 1 - \delta + \delta e^{-\delta}$ and the same bound $< 1$ holds for any $\widetilde{Y}_n \in C_n$ by Jensen's inequality. So $\mathbb{E}(e^{-\widetilde{Y}_\infty}) \leq 1 - \delta + \delta e^{-\delta}$ by dominated convergence, and so $\mathbb{P}(\widetilde{Y}_\infty > 0) > 0$.

Finally,

**Exercise 15.1.** *If a set is bounded in $L^0$, all its accumulation points in $L^0$ are finite-valued $\mathbb{P}$-a.s.*

□

**Remark 15.1.** If one has extra properties for the original sequence $(Y_n)_{n\in\mathbb{N}}$, one can also say more about the limit $\widetilde{Y}_\infty$. For example, if all the $Y_n$ are bounded by some constant, the same is true for the $\widetilde{Y}_n$ and hence also for $\widetilde{Y}_\infty$.

## 15.3 Appendix C: Essential supremum and infimum

These notes briefly recall the definition and main properties of the essential supremum and infimum of a family of (possibly extended) real-valued random variables. We fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, an arbitrary index set $\Lambda \neq \emptyset$ and a family $(Y_\lambda)_{\lambda\in\Lambda}$ of (possibly extended) real-valeud random variables on $(\Omega, \mathcal{F}, \mathbb{P})$.

**Definition 15.1.** *A random variable $Z$ is called **essential supremum** of the family $(Y_\lambda)_{\lambda\in\Lambda}$ if*

*i) $Z \geq Y_\lambda$ $\mathbb{P}$-a.s. for each $\lambda \in \Lambda$.*

*ii) $Z \leq Z'$ $\mathbb{P}$-a.s. for each random variable $Z'$ satisfying $Z' \geq Y_\lambda$ $\mathbb{P}$-a.s. for each $\lambda \in \Lambda$.*

*We then write briefly $Z = \operatorname{ess\,sup}_{\lambda\in\Lambda} Y_\lambda$. The **essential infimum** is defined analogously by simply reversing all inequalities above.*

**Remark 15.2.**

1. If $\Lambda$ is countable, we can take the pointwise supremum $Z(\omega) := \sup_{\lambda\in\Lambda} Y_\lambda(\omega)$; this is measurable and thus a random variable. But if $\Lambda$ is uncountable, this no longer works; on the one hand, the pointwise supremum may fail to be measurable, and on the other hand, i) and ii) can also fail, as illustrated by the subsequent example.

2. By ii), an essential supremum is $\mathbb{P}$-a.s. unique, so we only have to prove its existence.

3. The subsequent results can of course also be formulated and proved (with obvious changes) for the essential infimum instead of supremum.

4. Since the definition and all the arguments below only involve the order structure of $\mathbb{R}$, but not the actual values of the random variables under consideration, everything works equally well if we allow the $Y_\lambda$ to take values in $[-\infty, +\infty]$.

**Example 15.1.** *Let $\Omega = [0, 1]$, $\mathbb{P} = \lambda =$ Lebesgue measure, $\Lambda = [0, 1]$ and $Y_\lambda(\omega) = 1_{\{\lambda\}}(\omega)$. Then*

$$\sup_{\lambda\in\Lambda} Y_\lambda(\omega) = 1, \text{ for each fixed } \omega,$$

*and so the pointwise supremum $\sup_{\lambda\in\Lambda} Y_\lambda \equiv 1$ is here measurable. But for every fixed $\lambda$, we also have $Y_\lambda = 0$ $\mathbb{P}$-a.s. and thus obviously*

$$\operatorname*{ess\,sup}_{\lambda\in\Lambda} Y_\lambda = 0 \ (\mathbb{P}\text{-a.s.})$$

**Proposition 15.1.** *For any family $(Y_\lambda)_{\lambda\in\Lambda}$ of (possibly extended) real-valued random variables,* $\operatorname{ess\,sup}_{\lambda\in\Lambda} Y_\lambda = Z$ *exists, and* $Z = \sup_{j\in J_0} Y_j$ *for some countable subset $J_0$ of $\Lambda$.*

*Proof.* Since the above definition only involves the order structure of $\mathbb{R}$, we may and do assume without loss of generality that all $Y_\lambda$ are bounded, uniformly in $\lambda$ and $\omega$. Set

$$c := \sup\left\{\mathbb{E}\left(\sup_{j\in J} Y_j\right) : J \subset \Lambda \text{ countable}\right\}$$

and choose a sequence $(J_n)_{n\in\mathbb{N}}$ of countable subset of $\Lambda$ such that

$$\lim_{n\to\infty} \mathbb{E}\left(\sup_{j\in J_n} Y_j\right) = c.$$

Then $J_0 := \cup_{n\in\mathbb{N}} J_n \subset \Lambda$ is countable, so $Z := \sup_{j\in J_0} Y_j$ is a random variable, and $\mathbb{E}(Z) = c$ by monotone integration. We claim that $Z$ does the job, and so we check the required properties.

ii) If $Z' \geq Y_\lambda$ $\mathbb{P}$-a.s. for each $\lambda \in \Lambda$, then also $\mathbb{P}(Z' \geq Y_j \text{ for all } j \in J_0) = 1$ because $J_0$ is countable, and thus $Z' \geq Z$ $\mathbb{P}$-a.s. by the definition of $Z$.

i) For each $\lambda \in \Lambda$, we have $Z \vee Y_\lambda = \max(Z, Y_\lambda) \geq Z$, and by the definitions of $c$ and $J_0$,

$$\mathbb{E}(Z \vee Y_\lambda) = \mathbb{E}\left(\sup_{j\in J_0\cup\{\lambda\}} Y_j\right) \leq c = \mathbb{E}(Z).$$

Hence $Z \vee Y_\lambda - Z \geq 0$ and $\mathbb{E}(Z \vee Y_\lambda - Z) \leq 0$; so we must have $Z \vee Y_\lambda = Z$ $\mathbb{P}$-a.s., and thus $Z \geq Y_\lambda$ $\mathbb{P}$-a.s. This holds for each $\lambda \in \Lambda$, and so $Z$ satisfies i). $\square$

**Corollary 15.1.** *Suppose that $(Y_\lambda)_{\lambda\in\Lambda}$ is directed upward, i.e., for each pair $\lambda, \lambda'$ in $\Lambda$, there is some $\mu \in \Lambda$ such that* $\max(Y_\lambda, Y_{\lambda'}) \leq Y_\mu$*; this holds in particular if the family $(Y_\lambda)_{\lambda\in\Lambda}$ is closed under taking maxima. Then there is a sequence $(j_n)_{n\in\mathbb{N}}$ in $\Lambda$ such that*

$$\operatorname*{ess\,sup}_{\lambda\in\Lambda} Y_\lambda = \nearrow - \lim_{n\to\infty} Y_{j_n} \ \mathbb{P}\text{-}a.s.,$$

*i.e., $Y_{j_n} \leq Y_{j_{n+1}}$ $\mathbb{P}$-a.s. for each $n$ and $Y_{j_n} \nearrow$* $\operatorname{ess\,sup}_{\lambda\in\Lambda} Y_\lambda$ *$\mathbb{P}$-a.s.*

*Proof.* Choose $J_0 = \{\lambda_n : n \in \mathbb{N}\} \subset \Lambda$ countable with $\operatorname{ess\,sup}_{\lambda\in\Lambda} Y_\lambda = \sup_{n\in\mathbb{N}} Y_{\lambda_n}$. Set $j_1 := \lambda_1$ and choose recursively an element $j_n$ of $\Lambda$ such that $\max(Y_{j_{n-1}}, Y_{\lambda_n}) \leq Y_{j_n}$. Then clearly

$$Y_{j_{n-1}} \leq Y_{j_n} \leq \operatorname*{ess\,sup}_{\lambda\in\Lambda} Y_\lambda \ \mathbb{P}\text{-a.s. for all } n \in \mathbb{N},$$

and induction yields

$$Y_{j_n} \geq \max_{k=1,\dots,n} Y_{\lambda_k},$$

so that

$$\operatorname*{ess\,sup}_{\lambda\in\Lambda} Y_\lambda \geq \nearrow - \lim_{n\to\infty} Y_{j_n} = \sup_{n\in\mathbb{N}} Y_{j_n} \geq \sup_{n\in\mathbb{N}} Y_{\lambda_n} = \operatorname*{ess\,sup}_{\lambda\in\Lambda} Y_\lambda.$$

This gives the assertion. □

## 15.4 Appendix D: The bipolar theorem

These notes provide a formulation of the bipolar theorem from functional analysis. We formulate the result here for the setting we need, which means that we use the dual pair $(L^\infty, L^1)$ with the duality pairing given by $(Z, Y) = \mathbb{E}(ZY)$ for $Z \in L^\infty$ and $Y \in L^1$.

**Definition 15.2.** *For a subset $C \subset L^\infty$, the* ***polar*** *of $C$ in $L^1$ is*

$$C^\circ := \{Y \in L^1 : (Z, Y) \leq 1 \textit{ for all } Z \in C\}$$

*In the same way, the polar in $L^\infty$ of $D \subset L^1$ is*

$$D^\circ := \{Z \in L^\infty : (Z, Y) \leq 1 \textit{ for all } Y \in D\}.$$

*The* ***bipolar*** *of $C \subset L^\infty$ is then the polar of $C^\circ$,*

$$C^{\circ\circ} := (C^\circ)^\circ \subset L^\infty.$$

It is easy to check that for any $D \subset L^1$, the polar $D^\circ$ is a convex set in $L^\infty$, that $0 \in D^\circ$ and that $D^\circ$ is $\sigma(L^\infty, L^1)$-closed, i.e. weak* closed in $L^\infty$. If $C \subset L^\infty$ is a cone with vertex at 0 (meaning that $\lambda C \subset C$ for all $\lambda > 0$) then we also have

$$C^\circ = \{Y \in L^1 : (Z, Y) \leq 0 \text{ for all } Z \in C\};$$

so $C^\circ$ is then also a cone with vertex at 0, and hence

$$C^{\circ\circ} = \{Z \in L^\infty : (Z, Y) \leq 0 \text{ for all } Y \in C^\circ\}.$$

**Theorem 15.2** (Bipolar theorem)**.** *For any $C \subset L^\infty$, its bipolar $C^{\circ\circ}$ is the $\sigma(L^\infty, L^1)$-closed convex hull of $C \cup \{0\}$, i.e., the smallest convex and weak* closed subset of $L^\infty$ containing $C$ and 0. In particular, if $C$ is a convex cone with vertex at 0, then $C^{\circ\circ}$ is the weak* closure of $C$; if in addition $C$ is weak* closed, then $C^{\circ\circ} = C$.*

*Proof.* See H. H. Schaefer (with M. P. Wolff) (1999), "Topological Vector Spaces", second edition, Springer, Theorem IV.1.5. □

**Remark 15.3.** While the above result looks simple, it is not quite straightforward. In fact, the argument for showing that the bipolar $C^{\circ\circ}$ is contained in the $\sigma(L^\infty, L^1)$-closed convex hull of $C \cup \{0\}$ uses the separation theorem for convex sets and is thus based on the Hahn-Banach theorem.

## 15.5 Appendix E: Bayes formula and Girsanov transformation

In this section, we briefly discuss the following basic question: How does a (continuous) $\mathbb{P}$-semi-martingale behave under a different probability measure $Q$, where typically $Q$ is absolutely continuous or even equivalent to $\mathbb{P}$? Since the path properties of the finite variation part of the semi-martingale do not change, it is enough to address this question for local $\mathbb{P}$-martingales. The results we discuss here are only meant to refresh knowledge acquired in any basic stochastic calculus class and can be found in more detail there.

We fix a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le\infty}, \mathbb{P})$ with filtration satisfying the usual conditions. Fix a probability measure $Q \ll \mathbb{P}$ and choose an RCLL version of the $\mathbb{P}$-martingale

$$Z_t := \mathbb{E}_{\mathbb{P}}\left[\frac{dQ}{d\mathbb{P}} \mid \mathcal{F}_t\right], \text{ for } 0 \le t \le \infty.$$

$Z$ is a $\mathbb{P}$-martingale on $[0,\infty]$, hence UI under $\mathbb{P}$, and $Z_t \ge 0$ $\mathbb{P}$-a.s. for each $t$ since $Q \ll \mathbb{P}$. We call $Z$ the **density process** (of $Q$ with respect to $PP$). If $Q \approx \mathbb{P}$, we even have $Z_t > 0$ $\mathbb{P}$-a.s. for each $t$.

**Remark 15.4.**

1. Both $Z$ and $Z_-$ are strictly positive $Q$-a.s., i.e., we have $Q$-a.s. $Z_t > 0$, simultaneously for all $t \ge 0$: the trajectories of $Z$ never reach 0, $Q$-a.s. (neither by a jump, nor by creeping down).

2. By the previous remark, $\frac{1}{Z}$ and $\frac{1}{Z_-}$ are well defined $Q$-a.s. Being adapted and left-continuous, $\frac{1}{Z_-}$ is also predictable and locally bounded (with respect to $Q$).

3. If $\mathbb{P} \approx Q$, all this also holds $\mathbb{P}$-a.s, and this allows us to write $Z = Z_0\mathcal{E}(L)$ for a unique $L \in \mathcal{M}_{0,\mathrm{loc}}(\mathbb{P})$; moreover, we have $EE_{\mathbb{P}}(Z_0) = \mathbb{E}_{\mathbb{P}}(\frac{dQ}{d\mathbb{P}}) = 1$. Indeed, setting

$$L := \int \frac{1}{Z_-} dZ$$

gives $dZ = Z_- dL$, whose unique solution (up to the initial value $Z_0$) is by definition the stochastic exponential, i.e. $Z = Z_0\mathcal{E}(L)$.

4. The same results hold on $[0,T]$ if we only assume $Q \ll \mathbb{P}$ on $\mathcal{F}_T$ or $Q \approx \mathbb{P}$ on $\mathcal{F}_T$, respectively.

The next result, known as Bayes formula, shows how to convert conditional expectations under $Q$ and $\mathbb{P}$.

**Proposition 15.2** (Bayes formula)**.** *Suppose that $Q \ll \mathbb{P}$ with density process $Z$. Then:*

1. *For any stopping times $\sigma \leq \tau$ and $U_\tau$ $\mathcal{F}_\tau$-measurable nonnegative or in $L^1(Q)$, we have the* ***Bayes formula***

$$\mathbb{E}_Q(U_\tau \mid \mathcal{F}_\sigma) = \frac{1}{Z_\sigma}\mathbb{E}_{\mathbb{P}}(Z_\tau U_\tau \mid \mathcal{F}_\sigma) \ Q\text{-a.s.}$$

2. *Suppose now that $Q \approx \mathbb{P}$. Then an $\mathbb{F}$-adapted process $Y$, null at $0$, is a (local) $Q$-martingale if and only if $ZY$ is a (local) $\mathbb{P}$-martingale.*

**Theorem 15.3** (Girsanov transformation)**.** *Take $Q \approx \mathbb{P}$ and assume that hte density process $Z$ is continuous. If $M \in \mathcal{M}^c_{0,loc}(\mathbb{P})$, then*

$$\widetilde{M} := M - \int \frac{1}{Z} d\langle Z, M\rangle = M - \langle L, M\rangle \text{ is in } \mathcal{M}^c_{0,loc}(Q),$$

*where we write $Z = Z_0\mathcal{E}(L)$. In particular, $M$ is a continuous $Q$-semimartingale with explicit $Q$-decomposition $M = \widetilde{M} + \widetilde{A}$, where $\widetilde{A} = M - \widetilde{M} = \langle L, M\rangle$.*

If $M = W$ is a Brownian motion under $\mathbb{P}$, we have a more precise result.

**Theorem 15.4.** *Suppose that $W$ is a $\mathbb{P}$-Brownian motion and $Q \approx \mathbb{P}$ has a density process of the form $Z = \mathcal{E}\left(\int b_s dW_s\right)$ for some predictable process $b$. Then $W$ is under $Q$ a Brownian motion with drift $b$, i.e.,*

$$W = \widetilde{W} + \int b_s ds$$

*for a $Q$-Brownian motion $\widetilde{W}$.*

## 15.6 Appendix F: The Kunita-Watanabe decomposition and Itô's representation theorem

Our goal in this section is to establish a decomposition of the martingale space $\mathcal{H}_0^2 = \{\text{RCLL martingales } M \text{ such that } \sup_{t\geq 0}\mathbb{E}(M_t^2) < \infty,\ M_0 = 0\}$ into one space which is spanned by a given element $M$, and a second space orthogonal to the first one. To that end, we first need notions of orthogonality. Recall that we have identified $\mathcal{H}^2$ with $L^2(\Omega, \mathcal{F}_\infty, \mathbb{P})$.

**Definition 15.3.** *Two elements $M, N$ of $\mathcal{H}_0^2$ are called **weakly orthogonal** if they satisfy $(M,N)_{\mathcal{H}^2} = \mathbb{E}(M_\infty N_\infty) = 0$, i.e., they are orthogonal in the sense of the Hilbert space $\mathcal{H}^2$. $M, N$ are called **strongly orthogonal**, written as $M \perp N$, if their product $MN$ is a martingale (which is UI since it is a martingale on the closed interval $[0,\infty]$). More generally $M, N \in \mathcal{M}_{0,loc}^c$ (or more generally $\mathcal{H}_{0,loc}^2$) are called **strongly orthogonal**, denoted $M \perp N$, if their product $MN$ is a local martingale.*

**Theorem 15.5** (Kunita-Watanabe decomposition)**.** *Fix $M \in \mathcal{M}_{0,loc}^c$. Then every $N \in \mathcal{M}_{0,loc}^c$ can be uniquely written as*

$$N = H \cdot M + L, \text{ for some } H \in L_{loc}^2(M) \text{ and some } L \in \mathcal{M}_{0,loc}^c \text{ with } L \perp M$$

*The integrand $H$ is a predictable density of $\langle M, N\rangle$ with respect to $\langle M\rangle$.*

According to the Kunita-Watanabe decomposition, any continuous local martingale in an arbitrary filtration can be decomposed into a stochastic integral with respect to a given continuous local martingale $M$ and a strongly orthogonal term $L$. This holds in particular if $M = W$ is a Brownian motion. If in addition the filtration is generated by $W$, we know very precisely how $L$ looks: it is zero. This is the message of Itô's representation theorem, which says that **in a Brownian filtration, every local martingale is a stochastic integral with respect to the underlying Brownian motion**.

**Theorem 15.6** (Itô's representation theorem)**.** *Every random variable $F \in L^1(\mathcal{F}_\infty^W, \mathbb{P})$, admits a unique representation*

$$F = \mathbb{E}(F) + \int_0^\infty H_s dW_s \ \mathbb{P}\text{-a.s.}$$

*with a process $H \in L_{loc}^2(W)$ such that $\int HdW$ is a martingale (on $[0,\infty]$). As a consequence, every local $(\mathbb{P};\mathbb{F}^W)$-martingale $N$ is of the form*

$$N = N_0 + \int HdW \text{ with some } H \in L_{loc}^2(W)$$

*and in particular has a continuous version.*

# 16 Solutions to Exercises

**Exercise** 1.1: Construct a similar arbitrage strategy as $1_{[\![0,\tau]\!]}$ from the lecture on a finite horizon and with a positive process.

*Proof.* We start with a geometric Brownian motion with $\mu = 0,\ \sigma = 1$, so that

$$\widetilde{S}_t = \exp\left(W_t - t/2\right).$$

Recall from BMSC that the law of large numbers for Brownian motion states that

$$\lim_{t\to\infty} \frac{W_t}{t} = 0 \ \mathbb{P}\text{-a.s.}$$

Hence we get (as is already known from BMSC) that

$$\widetilde{S}_t = \exp\left(t\left[\frac{W_t}{t} - \frac{1}{2}\right]\right) \to 0 \text{ as } t \to \infty,$$

since

$$\frac{W_t}{t} - \frac{1}{2} \xrightarrow{t\to\infty} -\frac{1}{2}, \text{ so that } t\left[\frac{W_t}{t} - \frac{1}{2}\right] \xrightarrow{t\to\infty} -\infty.$$

So the stopping time $\widetilde{\tau} := \inf\{t \geq 0 : \widetilde{S}_t = 1/2\}$ is a.s. finite. Set $\Psi(t) := \tan t$ and define $S_t := \widetilde{S}_{\Psi(t)}$ for $t \in [0, \pi/2)$ and $S_{\pi/2} = 0$ (which agrees with $\widetilde{S}_\infty = 0$). This yields again a continuous process $S$ and a stopping time $\tau := \Psi^{-1} \circ \widetilde{\tau} \in [0, \pi/2)$ a.s. Thus we can use the predictable, self-financing strategy $\varphi = (v_0 = 0, \vartheta)$ (i.e. with $v_0 = 0$) going short on $(\!(0, \tau]\!]$, i.e.

$$\vartheta_t := -1_{(\!(0,\tau]\!]}(t)$$

which is adapted and left-continuous, hence $S$-integrable ($S$ is a continuous local martingale). It follows that the value (wealth) process is given by

$$V(\varphi) = \int \vartheta dS = -(S^\tau - S_0).$$

In particular we end up with final value

$$V_{\pi/2}(\varphi) = S_0 - S^\tau_{\pi/2} = S_0 - S_\tau = 1 - \frac{1}{2} = \frac{1}{2}.$$

Thus we have $S > 0$ (because $\widetilde{S} > 0$) and $\varphi$ starts from 0, is self-financing and ends up with wealth 1/2, which is clearly an arbitrage. □

**Exercise** 2.1: Show that even if there exists an ELMM $Q$, $\mathrm{NA}_{\mathrm{det}}$ and hence also $\mathrm{NA}_{\mathrm{elem}}$ can fail. In particular, can you argue directly that $\vartheta$ is not admissible in your example?

*Proof.* Recall that in Exercise 1.1 (previous Exercise), we've constructed a process

$$S_t = \begin{cases} \exp(W_{\tan(t)} - \tan(t)/2), & \text{for } 0 \leq t < \pi/2 \\ 0, & \text{for } t = \pi/2. \end{cases}$$

Obviously, $S$ is a local martingale on $[0, \pi/2]$, but it is **not a martingale**. Indeed, assume for contradiction that $S$ is a martingale on $[0, \pi/2]$, then we must have $\mathbb{E}(S_t) = \mathbb{E}(S_0) = 1$, but clearly (by the law of large numbers for Brownian Motion as $\tan(\pi/2) = +\infty$) $\mathbb{E}(S_{\pi/2}) = 0$.

We now take $\vartheta_t = -1_{(0,\pi/2]}(t)$ which is a very simple integrand i.e. $\vartheta \in \mathbf{b}\mathcal{E}_{\mathrm{det}}$ and hence also $\vartheta \in \mathbf{b}\mathcal{E}$. Moreover we have

$$G_{\pi/2}(\vartheta) = -(S_{\pi/2} - S_0) = 1,$$

which gives an (deterministic and hence also elementary) arbitrage opportunity (money pump). That is we have shown that

$$G_T(\mathbf{b}\mathcal{E}_{\mathrm{det}}) \cap L^0_+ = G_{\pi/2}(\mathbf{b}\mathcal{E}_{\mathrm{det}}) \cap L^0_+ \neq \{0\}.$$

Hence if we take $Q = \mathbb{P}$ as our ELMM (under which $S$ is a local martingale but not a martingale) then $\mathrm{NA}_{\mathrm{det}}$ (and hence also $\mathrm{NA}_{\mathrm{elem}}$) can still fail.

Let us now see if we can argue directly that $\vartheta$ is not admissible in this setup. Assume for contradiction that $\vartheta$ is admissible i.e. we have that $G_t(\vartheta) \geq -a$ $\mathbb{P}$-a.s. for all $0 \leq t \leq T = \pi/2$ for some $a \geq 0$. But if $G(\vartheta)$ is uniformly bounded from below, then $S$ is bounded from above as

$$G(\vartheta) = -(S - S_0) = 1 - S \geq -a \implies S \leq a - 1.$$

But by the very definition of our process $S = (S_t)_{0 \leq t \leq \pi/2}$ we have that $S \geq 0$, which gives that $S$ is bounded and we know (easy exercise) from BMSC that every bounded local martingale is a (true) martingale. But as we've already argued above, $S$ is not a martingale. $\square$

**Exercise** 2.2: Show that if there exists an (elementary) arbitrage opportunity

$$\vartheta = \sum_{k=1}^{N} h_k 1_{(\!(\tau_{k-1},\tau_k]\!]} \in \mathbf{b}\mathcal{E},$$

then there also exists a "one-step buy-and-hold" arbitrage opportunity of the form $\vartheta^* = h1_{(\!(\sigma_0,\sigma_1]\!]} \in \mathbf{b}\mathcal{E}$.

*Proof.* Define

$$k^* := \min\{k \in 1,\dots,N\} : G_{\tau_k}(\vartheta) \in L^0_+ \setminus \{0\}\},$$

then $k^*$ is well-defined since we know that for the strategy $\varphi = (0,\vartheta)$ there exists an (elementary, i.e. $\vartheta \in \mathbf{b}\mathcal{E}$) arbitrage opportunity, moreover we observe that $k^*$ is deterministic. We set $\sigma_0 := \tau_{k^*-1}$ and $\sigma_1 := \tau_{k^*}$. Furthermore, we set

$$h := \begin{cases} h^{k^*} & \text{if } \mathbb{P}(G_{\tau_{k^*-1}}(\vartheta) = 0) = 1, \\ h^{k^*} 1_{\{G_{\tau_{k^*-1}}(\vartheta)<0\}} & \text{if } \mathbb{P}(G_{\tau_{k^*-1}}(\vartheta) = 0) < 1. \end{cases}$$

In words, we wait until $G_{\cdot}(\vartheta)$ first becomes genuinely positive and then use the single step of $\vartheta$ from the previous $\tau_{k-1}$ on the set where the previous gains were zero (first case) or genuinely negative (second case). We notice that $\mathbb{P}(G_{\tau_{k^*-1}}(\vartheta) < 0) > 0$ in the second case by the definition of $k^*$.

We claim that $\vartheta^* := h_{(\!(\sigma_0,\sigma_1]\!]} \in \mathbf{b}\mathcal{E}$ is an arbitrage opportunity. Indeed:

• In the first case, i.e. if $G_{\tau_{k^*-1}}(\vartheta) = 0$ almost surely:

$$G_T(\vartheta^*) = G_{\tau_{k^*}}(\vartheta) - G_{\tau_{k^*-1}}(\vartheta) = G_{\tau_{k^*}}(\vartheta) \in L^0_+ \setminus \{0\}.$$

• In the second case, i.e. if $\mathbb{P}(G_{\tau_{k^*-1}}(\vartheta) < 0) > 0$ which means the gain at time $\tau_{k^*-1}$ is genuinely negative with positive probability, we get:

$$\begin{aligned} G_T(\vartheta^*) &= \left(G_{\tau_{k^*}}(\vartheta) - G_{\tau_{k^*-1}}(\vartheta)\right) 1_{\{G_{\tau_{k^*-1}}(\vartheta)<0\}} \\ &\geq -G_{\tau_{k^*-1}}(\vartheta) 1_{\{G_{\tau_{k^*-1}}(\vartheta)<0\}} \in L^0_+ \setminus \{0\}. \end{aligned}$$

Hence in both cases we have an arbitrage opportunity. $\square$

**Exercise 3.2**: Theorem 3.4 tells us that every semimartingale is a good integrator, it is an immediate consequence of Lemma 3.2 which has a rather straightforward but technical proof. During the proof of Lemma 3.2 we made use of the following identity:

$$\mathbb{E}(M_T A_T^n) = \mathbb{E}\left(\int_0^T M_{s-} dA_s^n\right),$$

where $(M_t)_{0\le t\le T}$ is a non-negative bounded martingale and $A^n$ is increasing integrable RCLL, predictable, null at 0. Verify the above formula.

*Proof.* For any partition $\pi$ of the interval $[0,T]$, we can write

$$M_T A_T = \sum_{i=1}^{n} M_T(A_{t_i} - A_{t_{i-1}}).$$

Because $A$ is predictable, $A_{t_i}$ is $\mathcal{F}_{t_i-}$-measurable, and because $M$ is a martingale, we get

$$\mathbb{E}(M_T A_T) = \mathbb{E}\left[\sum_{i=1}^{n} M_{t_i-}(A_{t_i} - A_{t_{i-1}})\right].$$

As $|\pi| \to 0$, the sum inside the expectation converges to $\int_0^T M_{s-} dA_s$, and because $M$ is bounded and $A$ is increasing and integrable, a majorant for all sums is $\|M\|_\infty A_T \in L^1$. So dominated convergences establishes that

$$\begin{aligned}
\mathbb{E}(M_T A_T) &= \lim_{|\pi|\to 0} \mathbb{E}\left[\sum_{t_i\in\pi} M_{t_i-}(A_{t_i} - A_{t_{i-1}})\right] \\
&= \mathbb{E}\left[\lim_{|\pi|\to 0} \sum_{t_i\in\pi} M_{t_i-}(A_{t_i} - A_{t_{i-1}})\right] \\
&= \mathbb{E}\left[\int_0^T M_{s-} dA_s\right].
\end{aligned}$$

□

**Exercise 4.1**: a) Show that $D_0(Z^1, Z^2) := \mathbb{E}(1 \wedge |Z^1 - Z^2|)$ is a metric on $L^0$ and that $Z_n \to Z$ in probability iff $d_0(Z_n, Z) \to 0$. So $d_0$ metrizes the convergence in probability.

*Proof.* We start by showing that $d_0 = d_{L^0}$ is a metric on $L^0$. Let $Z^1, Z^2, Z^3 \in L^0$ denote arbitrary random variables (i.e. $L^0$ as the space of random variables).

- $\mathbb{E}(1 \wedge |Z^1 - Z^2|) \geq 0$ is clear.
- $\mathbb{E}(1 \wedge |Z^1 - Z^2|) = 0$ holds if and only if $Z^1 = Z^2$ $\mathbb{P}$-a.s.
- $\mathbb{E}(1 \wedge |Z^1 - Z^2|) = \mathbb{E}(1 \wedge |Z^2 - Z^1|)$ gives the symmetry.
- Finally we have $1 \wedge |Z^1 - Z^3| \leq 1 \wedge (|Z^1 - Z^2| + |Z^2 - Z^3|) \leq 1 \wedge |Z^1 - Z^2| + 1 \wedge |Z^1 - Z^3|$. Taking expectation gives the triangle inequality.

The above shows that $d_0$ is indeed a metric on $L^0$. Let us next show that it metrizes the convergence in probability. Recall the basic estimate for $0 < \delta \leq 1$ given by

$$\mathbb{P}(Z > \delta) \leq \frac{1}{\delta}\mathbb{E}(1 \wedge Z).$$

The above then establishes that convergence in $d_0$ implies convergence in probability. Next assume that $Z^n \to Z$ in probability, then $\mathbb{E}(1 \wedge |Z^n - Z|) \to 0$ by bounded convergence theorem, indeed $1 \wedge |Z^n - Z| \leq 1$. $\square$

b) Recall from the lecture $\mathbb{L}$ (resp. $\mathbb{D}$) of adapted LCRL (resp. RCLL) processes. Define

$$d(X^1, X^2) := \mathbb{E}(1 \wedge (X^1 - X^2)^*_T) = E\left[1 \wedge \sup_{0 \leq s \leq T} |X^1_s - X^2_s|\right]$$

on $\mathbb{L}$ (resp. on $\mathbb{D}$). Assume that the filtration $\mathbb{F}$ is complete (important here!). Show that both $(\mathbb{L}, d)$ and $(\mathbb{D}, d)$ are complete metric spaces.

*Proof.* We only show that $(\mathbb{L}, d)$ is a complete metric space. The proof for $(\mathbb{D}, d)$ is analogous. The fact that $d$ is a metric can be proved as in part a) and will be skipped here. Now let $(X^n)_{n\mathbb{N}} \subset \mathbb{L}$ be a Cauchy sequence in $d$. Set $n_1 = 1$. After $n_{k-1}$ is defined we choose $n_k$ such that

$$\mathbb{P}[(X^n - X^m)^*_T > 2^{-k}] < 2^{-k} \text{ for all } m, n \geq n_k$$

In particular, this gives

$$\sum_{k=1}^{\infty} \mathbb{P}[(X^n - X^m)^*_T > 2^{-k}] < \infty$$

and by the Borel-Cantelli lemma it follows that $(X^{n_k})_{k\in\mathbb{N}}$ is $\mathbb{P}$-a.s. Cauchy sequence under the **uniform** convergence on $[0,T]$. So for each $t\in[0,T]$ the there exists a limit $X_t=\sum_{k\to\infty}X_t^{n_k}$. Since each $X^{n_k}$ is adapted and $\mathbb{F}$ is complete, $X$ will also be adapted. Using the estimate

$$\sup_{s\in[0,T]}|X_{s+}^{n_k}-X_{s+}^{n_l}|\leq\sup_{s\in[0,T]}|X_s^{n_k}-X_s^{n_l}|$$

establishes that $(X_+^{n_k})_{k\in\mathbb{N}}$ is also Cauchy under uniform convergence on $[0,T]$ and has also a limit $X_+$. Using the uniform convergence on $[0,T]$, we can switch the following limits and obtain

$$\begin{aligned}\lim_{s\nearrow t}X_s&=\lim_{s\nearrow t}\lim_{k\to\infty}X_s^{n_k}=\lim_{k\to\infty}\lim_{s\nearrow t}X_s^{n_k}=X_t,\\ \lim_{s\searrow t}X_s&=\lim_{s\searrow t}\lim_{k\to\infty}X_s^{n_k}=\lim_{k\to\infty}\lim_{s\searrow t}X_s^{n_k}=X_{t+}.\end{aligned}$$

Since $\mathbb{L}$ is the space of LCRL adapted processes (on $[0,T]$) we get that $X\in\mathbb{L}$. Recall that for adaptedness we had to make use of the fact that $\mathbb{F}$ is complete.

Finally, we need to show that $(X^n)_{n\in\mathbb{N}}$ converges to $X$ w.r.t. to our norm $d$ for $\mathbb{L}$. Let $\epsilon>0$ be arbitrary. We can find $k$ such that $d(X^{n_k},X)<\epsilon/2$ and $1/2^k<\epsilon/2$. Then for all $n\geq n_k$ we have by the triangle inequality

$$d(X^n,X)\overset{\Delta}{\leq}d(X^n,X^{n_k})+d(X^{n_k},X)<\epsilon.$$

This shows that $d(X^n,X)\to0$ as $n\to\infty$ and hence $(\mathbb{L},d)$ is indeed a complete metric space. □

c) Show that $\mathbf{b}\mathcal{E}_0$ is dense in $\mathbb{L}$ for $d$.

*Proof.* Let $X\in\mathbb{L}$, i.e. $X$ is an adapted LCRL process $X=(X_t)_{0\leq t\leq T}$. Any such process is locally bounded (see Exercise 4.3), hence by localization we may assume that $X$ is bounded. Set $Y=X_+$, observe that then $Y_-=X$ and consider for $\epsilon>0$ arbitrary $\tau_0^\epsilon:=0$ and iteratively

$$\tau_{n+1}^\epsilon:=\inf\{t>\tau_n^\epsilon:|Y_t-Y_{\tau_n^\epsilon}|>\epsilon\}\wedge T.$$

The right continuity of $Y$ implies that the $\tau_n^\epsilon$ are stopping times and $\tau_n^\epsilon\nearrow T$ stationarily. Define

$$X^\epsilon=X_0 1_{\{0\}}+\sum_{n=1}^{k}Y_{\tau_n^\epsilon}1_{(\!(\tau_n^\epsilon,\tau_{n+1}^\epsilon]\!]}\in\mathbf{b}\mathcal{E}_0.$$

By construction we have

$$\mathbb{P}[(X^\epsilon - X)^*_T > \epsilon] \leq \mathbb{P}(\tau_m \neq T) + \mathbb{P}[(X^\epsilon - X^{\tau_m})^*_T > \epsilon].$$

Since $\tau_m \nearrow T$ stationarily we have $\mathbb{P}(\tau_m \neq T) \to 0$ as $m \to \infty$, the second term converges to 0 as $m \to \infty$ by construction. $\square$

**Exercise 4.3**: Show that every $X \in \mathbb{L}$ is locally bounded.

*Proof.* By definition $\mathbb{L}$ is the space of all processes $X = (X_t)_{0 \leq t \leq T}$ which are adapted (i.e. stochastic processes) and LCRL (or caglad), i.e. left continuous with right limits. In order to verify that $X$ is locally bounded, we need by definition find the existence of a sequence of $[0,T]$-valued stopping times $(\tau_n)_{n\in\mathbb{N}}$ with $\tau_n \nearrow T$ stationarily, and such that **each** $X^{\tau_n}$ is bounded (in particular the bound $c = c_n$ may very well depend on $n$).

Since $X$ is LCRL, it satisfies for all $t \in [0,T]$:

$$\lim_{s \nearrow t} X_s = X_t, \text{ and } \lim_{s \searrow t} X_s = X_{t+} \text{ exists.}$$

The first property above can also be written as $X_{t-} = X_t$.
We define $Y = X_+$, that is

$$Y_t = X_{t+} = \lim_{s \searrow t} X_s,$$

which is well-defined for all $t \in [0,T]$ because RL exists for $X$. We will now show that with this definition $Y$ is a RCLL process. Indeed we need to verify that $Y_{t+} = Y_t$ and $Y_{t-}$ exists, in fact we even have $Y_- = X$.

$$\lim_{s \searrow t} Y_s = \lim_{s \searrow t} \lim_{h \searrow s} X_h = \lim_{s \searrow t} X_s = X_{t+} = Y_t. \tag{RC}$$

$$\lim_{s \nearrow t} Y_s = \lim_{s \nearrow t} \lim_{h \searrow s} X_h = \lim_{s \nearrow t} X_{s+} = X_t. \tag{LL}$$

Where we used in the second assertion that when $s \nearrow t$ and then particular $s+ \nearrow t$ and then the left continuity of $X$. Next we define the stopping times

$$\tau_n := \inf\{t > 0 : |Y_t| > n\} \wedge T = \inf\{t > 0 : |X_{t+}| > n\} \wedge T.$$

since $Y$ is right-continuous, these stopping times are well defined (since it is the first hitting time of an adapted RC process for an open set). We have $\tau_n \nearrow T$ stationarily and by definition

$$|Y^{\tau_n}| 1_{\{t<\tau_n\}} = |Y_t| 1_{\{t<\tau_n\}} \leq n.$$

**Careful**, since $Y$ is RCLL the above is only true for $t < \tau_n$ and not (necessarily) at the jump $t = \tau_n$ (cannot control it, in fact we have $|Y_{\tau_n}| > n$, draw picture!). But since $Y_- = X$ we can take $t \to t-$ in the above to obtain

$$|X^{\tau_n}|1_{\{t- \leq \tau_n\}}|X_t|1_{\{t- \leq \tau_n\}} \leq n,$$

and since $X$ is LCRL (in particular left continuous), the event $\{t- \leq \tau_n\}$ is the same as $\{t \leq \tau_n\}$. Hence we do have for all $t \in [0, T]$ that

$$|X^{\tau_n}| \leq n =: c_n$$

which asserts that $X$ is locally bounded. $\square$

**Exercise 4.2:** a) Show that a nonempty subset $C \subset L^0$ is bounded in $L^0$ iff for every sequence $\lambda_n \searrow 0$ in $(0, \infty)$, $\lambda_n c_n \to 0$ in $L^0$ for every sequence $(c_n)_{n \in \mathbb{N}} \subset C$.

b) Show that an adapted RCLL process $X$ is a good integrator iff the set $I_X(H) = \{H \cdot X_T : H \in \mathbf{b}\mathcal{E}, \|H\|_\infty \leq 1\}$ is bounded in $L^0$, i.e.

$$\lim_{m \to \infty} \sup_{Y \in I_X(H)} \mathbb{P}(|Y| \geq m) = 0.$$

*Proof of a).* Recall that to be bounded in $L^0$ for a subset $C \subset L^0$ means

$$\lim_{m \to \infty} \sup_{Y \in C} \mathbb{P}(|Y| \geq m) = 0.$$

Hence this part of the Exercise gives us a more "workable" version of this abstract definition.

" $\implies$ " Let us assume that $C \subset L^0$ is bounded in $L^0$. Let $(\lambda_n)_{n \in \mathbb{N}} \subset (0, \infty)$ be a sequence with $\lambda_n \searrow 0$ and $(c_n)_{n \in \mathbb{N}} \subset C$. Let $\epsilon > 0$ be arbitrary. Since $C$ is bounded in $L^0$ we find by definition a $K > 0$ with $\sup_n \mathbb{P}(|c_n| > K) \leq \epsilon$. Then

$$\begin{aligned}
\mathbb{P}(|\lambda_n c_n| > \epsilon) &= \mathbb{P}(|\lambda_n c_n| > \epsilon, |c_n| > K) + \mathbb{P}(|\lambda_n c_n| > \epsilon, |c_n| \leq K) \\
&\leq \sup_n \mathbb{P}(|\lambda_n c_n| > \epsilon, |c_n| > K) + \mathbb{P}(|\lambda_n c_n| > \epsilon, |c_n| \leq K) \\
&\leq \sup_n \mathbb{P}(|c_n| > K) + \mathbb{P}(|\lambda_n c_n| > \epsilon, |c_n| \leq K) \\
&\leq \epsilon + \mathbb{P}(|\lambda_n c_n| > \epsilon, |c_n| \leq K).
\end{aligned}$$

Where we used that $\{|\lambda_n c_n| > \epsilon, |c_n| > K\} \subset \{|c_n| > K\}$. The second term goes to $0$ as $n \to \infty$ because $\lambda_n \searrow 0$ for $n \to \infty$. This shows that $\lambda_n c_n \to 0$ in probability.

"$\impliedby$" Assume that for every sequence $\lambda_n \searrow 0$ in $(0, \infty)$, we have $\lambda_n c_n \to 0$ in $L^0$ for every sequence $(c_n)_{n \in \mathbb{N}} \subset C$. Assume for contradiction that $C$ is not bounded in $L^0$, then there exists $\delta > 0$ and a sequence $(c_n)_{n \in \mathbb{N}} \subset C$ such that $\mathbb{P}(|c_n| > n) > \delta$ for all $n \in \mathbb{N}$. It follows that $\mathbb{P}(|c_n/n| > 1) > \delta$ for all $n \in \mathbb{N}$. But if we now specialize on $\lambda_n = 1/n \in (0, \infty)$, then $\lambda_n$ decreases to $0$ as $n \to \infty$ and hence we must (by assumption) have $\lambda_n c_n = c_n/n \to 0$ in $L^0$, but this is not the case since $\mathbb{P}(|c_n/n| > 1) > \delta$ for all $n \in \mathbb{N}$ as argued above, hence a contradiction. Thus $C \subset L^0$ is bounded in $L^0$. $\square$

*Proof of b).* Recall that we say that an adapted RCLL process $X = (X_t)_{0\le t\le T}$ ($\mathbb{R}^d$-valued) is called a good integrator if the map $I_X : (\mathbf{b}\mathcal{E}, \|\cdot\|_\infty) \to L^0$ is continuous. That translates to that if $(H^n)_{n\in\mathbb{N}}, H \in \mathbf{b}\mathcal{E}$ with $H^n \to H$ uniformly (in $(\omega, t)$), then $I_X(H^n) \to I_X(H)$ in probability.

" $\implies$ " Assume for contradiction that the set $I_X(H)$ is not bounded in $L^0$, then there exists $\delta > 0$ and a sequence $(H^n) \subset \mathbf{b}\mathcal{E}$ with $\|H^n\|_\infty \le 1$ such that $\mathbb{P}(|H^n\cdot X| > n) > \delta$ for all $n \in \mathbb{N}$. Let us set $\tilde{H}^n := H^n/n$, then we have $(\tilde{H}^n) \subset \mathbf{b}\mathcal{E}$ and $\tilde{H}^n \to 0$ uniformly in $(\omega, t)$ as $n \to \infty$. Hence, since we assume that $X$ is a good integrator, we have that $I_X(\tilde{H}^n) = \tilde{H}^n \cdot X_T \to I_X(0) = 0$ in probability, but

$$\mathbb{P}(|\tilde{H}^n \cdot X| > 1) = \mathbb{P}(|H^n \cdot X| > n) > \delta,$$

and hence $I_X(\tilde{H}^n)$ does not converge in probability to 0 which is a contradiction.

"$\Longleftarrow$" Let $(H^n) \subset \mathbf{b}\mathcal{E}$ with $\|H^n\|_\infty \to 0$. W.L.O.G, we can assume that $\|H^n\|_\infty > 0$ for all $n \in \mathbb{N}$. Let us set $\tilde{H}^n = H^n/\|H^n\|_\infty$, such that $(\tilde{H}^n) \subset \mathbf{b}\mathcal{E}$ with $\|\tilde{H}^n\| \le 1$. Hence, by assumption, the set $\{\tilde{H}^n \cdot X_T : n \in \mathbb{N}\}$ is bounded in $L^0$.

Therefore by part a), with $\lambda_n = \|H^n\|_\infty \in (0, \infty)$ with $\lambda_n \searrow 0$ as $n \to \infty$ we have

$$H^n \cdot X_T = \lambda_n \tilde{H}^n \cdot X_T \to 0 \text{ in } L^0.$$

Which shows us by part a), that $X$ is indeed a good integrator. $\square$

**Exercise** 4.5: Recall that the space $\mathcal{H}_0^1$ (on the time interval $[0,T]$) is defined by

$$\mathcal{H}_0^1 := \left\{ M \in \mathcal{M}_{\text{loc}} : M_0 = 0, M_T^* := \sup_{0 \le t \le T} |M_t| \in L^1 \right\}.$$

a) Show that every local martingale $M$ in $\mathcal{H}_0^1$ is of class $D$; that is, the family $\{M_\tau : \tau \le T \text{ stopping time}\}$ is uniformly integrable (UI).

*Proof.* Obviously, if $M \in \mathcal{H}_0^1$, then for any stopping time $\tau \le T$, it holds that $|M_\tau| \le M_T^*$. Then, since $M_T^* \in L^1$, we must have by dominated convergence

$$\lim_{K\to\infty} \left( \sup_{\tau \le T} \mathbb{E}(|M_\tau| 1_{\{|M_\tau| \ge K\}}) \right) \le \lim_{K\to\infty} \mathbb{E}(|M_T^*| 1_{\{|M_T^*| \ge K\}}) = 0,$$

whence the uniform integrability of the family $\{M_\tau : \tau \le T \text{ stopping time}\}$ (by its definition). □

b) Show that every local martingale in $\mathcal{H}_0^1$ is a martingale.

*Proof.* Suppose that $M$ is a local martingale in $\mathcal{H}_0^1$, and let $(\tau_n)_{n\ge1}$ be a localising sequence, i.e. $(\tau_n)_{n\ge1}$ is a sequence of increasing stopping times in $[0,T]$ such that $\mathbb{P}(\tau_n = T)$ tends to 1 and **every** stopped process $M^{\tau_n}$ is a martingale. The latter means that for any $0 \le s < t \le T$ and $A \in \mathcal{F}_s$, it holds that

$$\mathbb{E}(M_t^{\tau_n} 1_A) = \mathbb{E}(M_s^{\tau_n} 1_A).$$

Note that obviously $|M_r^{\tau_n}| \le M_T^*$ holds for all $n \ge 1$ and for all $r \in [0,T]$, and because $M \in \mathcal{H}_0^1$ implies (by definition of the space) that $M_T^* \in L^1$, we can apply dominated convergence theorem to both sides of the equation above and use the fact that $\lim_{n\to\infty} M_r^{\tau_n} = M_r$ for all $r \in [0,T]$ to obtain

$$\mathbb{E}(M_t 1_a) = \mathbb{E}(M_s 1_A),$$

which implies that $M$ is in fact a (true) martingale. □

c) Shows that every local martingale is locally in $\mathcal{H}_0^1$, i.e. $\mathcal{M}_{0,\text{loc}} = \mathcal{H}_{0,\text{loc}}^1$.

*Proof.* We first notice that by definition $\mathcal{H}_0^1 \subset \mathcal{M}_{0,\text{loc}}$ and thus also $\mathcal{H}_{0,\text{loc}}^1 \subset \mathcal{M}_{0,\text{loc}}$. We also recall that if $C$ denotes a class of stochastic processes, then the processes which are locally in $C$ will be denoted by $C_{\text{loc}}$. Letting $\tau_n$ be the stopping times taking the constant value $T$ shows that every process in $C$ is also locally in $C$, i.e. $C \subset C_{\text{loc}}$. Hence we have in particular $\mathcal{H}_0^1 \subset \mathcal{H}_{0,\text{loc}}^1$. If we show that $\mathcal{M}_{0,\text{loc}} \subset \mathcal{H}_{0,\text{loc}}^1$ then we have

$$\mathcal{H}_{0,\text{loc}}^1 \subset \mathcal{M}_{0,\text{loc}}^1 \subset \mathcal{H}_{0,\text{loc}}^1,$$

and hence equality throughout.

Let $X = (X_t)_{t\in[0,T]}$ be a local martingale, null at 0. Then there exists an increasing sequence of $[0,T]$-valued stopping times $(\tau_n)_{n\geq 1}$ and such that $\mathbb{P}(\tau_n = T)$ tends to 1 as $n \to \infty$ and for each $n$, the stopped process $X^{\tau_n}$ is a martingale. Since $\tau_n$ is bounded by a finite $T$, by the optional stopping theorem, $X^{\tau_n}$ is even a uniformly integrable martingale and therefore $X^{\tau_n}_\sigma = X_{\tau_n\wedge\sigma}$ is integrable for any stopping time $\sigma$.

Let us now define another sequence of stopping times $(\sigma_n)_{n\geq 1}$ by

$$\sigma_n := \inf\{t \geq 0 : |X_t| > n\} \wedge T.$$

Clearly, $(\sigma_n)_{n\geq 1}$ is an increasing sequence of $[0,T]$-valued stopping times ($X$ is a local martingale and hence in particular RCLL, especially right-continuous) and satisfies $\lim_{n\to\infty} \mathbb{P}(\sigma_n = T) = 1$. Moreover, by definition we have for each $n$ that $|X_t| \leq n$ for all $t < \sigma_n$ (careful not necessarily for $t \leq \sigma_n$) and therefore

$$|X_{\sigma_n -}| = \left|\lim_{t\nearrow\sigma_n} X_t\right| \leq n.$$

As a result, for each $n$ we obtain that

$$\begin{aligned}(X^{\tau_n\wedge\sigma_n})^*_T = \sup_{t\in[0,T]} |X^{\tau_n\wedge\sigma_n}_t| &\leq \sup_{t<\tau_n\wedge\sigma_n} |X_t| + |X_{\tau_n\wedge\sigma_n}| \\ &\leq n + |X_{\tau_n\wedge\sigma_n}|.\end{aligned}$$

We have already argued (above) that $X_{\tau_n\wedge\sigma_n}$ is in $L^1$, hence the inequality above shows that $(X^{\tau_n\wedge\sigma_n})^*_T$ is in $L^1$ as well. On the other hand, by the optional stopping theorem it holds that $X^{\tau_n\wedge\sigma_n}$ is a local martingale null at time 0 (standard argument). Hence, we conclude that $X^{\tau_n\wedge\sigma_n} \in \mathcal{H}^1_0$ for all $n \geq 1$. Finally, since $(\tau_n \wedge \sigma_n)_{n\geq 1}$ is an increasing sequence of $[0,T]$-valued stopping times which satisfies that $\lim_{n\to\infty} \mathbb{P}(\tau_n \wedge \sigma_n = T) = 1$, it follows that $X \in \mathcal{H}^1_{0,\mathrm{loc}}$ and the claim follows. □

**Remark 16.1.** This important Exercise shows that when we work with a local martingale $M$, null at 0, then it is locally in $\mathcal{H}^1_0$, and we know that local martingales in $\mathcal{H}^1_0$ are in fact true martingales in fact they are even martingales of class $D$.

The previous exercise, together with the Remark above has immediate consequences which we will just briefly list:

**Exercise 4.4:** By localisation, we can assume w.l.o.g. that $M \in \mathcal{M}_{0,\mathrm{loc}}$ is a (true) martingale.

**Exercise 4.7:** Let $M \in \mathcal{M}_{0,\mathrm{loc}}$, then $M^{\tau_m} \in \mathcal{H}^1_0$ for $\tau_m \nearrow T$ stationarily.

**Exercise 4.8**: In order to prove the Theorem of Mémin, we first required an auxiliary result, i.e. Lemma 4.6. During the prove of said Lemma we've established that for any $(K^n)_{n\in\mathbb{N}} \subset \mathbf{b}\mathcal{P}$ with $|K^n| \leq G$ and $\lambda_n \searrow 0$ we have $\lambda_n \int_0^t K_u^n dS_u \to 0$ in $L^0$, for each $t \in \mathbb{R}_+$. Show that this ensures that the set

$$\left\{ \int_0^t K_u dS_u : K \in \mathbf{b}\mathcal{P}, |K| \leq G \right\} \subset L^0$$

is bounded in $L^0$.

*Proof.* In Exercise 4.2 we have established a more general statement, namely a (nonempty) subset $C \subset L^0$ is bounded in $L^0$ iff for every sequence $\lambda_n \searrow 0$ in $(0,\infty)$, $\lambda_n c_n \to 0$ in $L^0$ for every sequence $(c_n)_{n\in\mathbb{N}} \subset C$. The claim then follows by this more general exercise. □

**Exercise 5.1:** For a general setup, where $S$ denotes a semimartingale, we recall the following important notions: By definition NA $:\iff G_T(\Theta_{\text{adm}}) \cap L^0_+ = \{0\}$, where $G_T(\Theta_{\text{adm}}) = \left\{ G_T(\vartheta) = \int_0^T \vartheta_u dS_u : \vartheta \in \Theta_{\text{adm}} \right\}$. This set interprets as the final wealth of all admissible, self-financing trading strategies $\varphi \stackrel{\wedge}{=} (0, \vartheta)$ starting from 0 initial wealth. The NA condition then ensures that if $G_T(\vartheta) \geq 0$ $\mathbb{P}$-a.s. then in fact $G_T(\vartheta) = 0$ $\mathbb{P}$-a.s., which again translates to that we cannot make money (in a self-financing way) out of nothing. We then further introduce

$$\mathcal{C}^0_{\text{adm}} := \mathcal{C}_{\text{adm}} = G_T(\Theta_{\text{adm}}) - L^0_+ = \{G_T(\vartheta) - Y : \vartheta \in \Theta_{\text{adm}}, Y \geq 0\},$$

that is all payoffs one can dominate/superreplicate from zero initial wealth via self-financing and admissible trading. Finally we define $\mathcal{C}^\infty = \mathcal{C}_{\text{adm}} \cap L^\infty = (G_T(\Theta_{\text{adm}}) - L^0_+) \cap L^\infty$, is set of all **bounded** payoffs which can be dominated/superreplicated by final wealth of some admissible self-financing strategy with 0 initial wealth.
With the above recap in mind, show that:

a) NA $\iff \mathcal{C}^0_{\text{adm}} \cap L^0_+ = \{0\}$.

b) NA $\iff (\mathcal{C}^0_{\text{adm}} \cap L^\infty) \cap L^0_+ = \{0\}$.

*Proof.* a) "$\implies$" Let us assume that NA holds. Let $H = G_T(\vartheta) - Y \geq 0$ for some $\vartheta \in \Theta_{\text{adm}}$ and $Y \geq 0$ ($Y \in L^0_+$). We then have $G_T(\vartheta) = H + Y \geq 0$, and thus $G_T(\vartheta) \in G_T(\Theta_{\text{adm}}) \cap L^0_+$. Since NA (by definition $G_T(\Theta_{\text{adm}}) \cap L^0_+ = \{0\}$) holds we have $G_T(\vartheta) = 0$ $\mathbb{P}$-a.s. whence $0 = H + Y \geq 0$, i.e. $H = Y = 0$ and thus $\mathcal{C}^0_{\text{adm}} \cap L^0_+ = \{0\}$.

"$\impliedby$" Let us assume that $\mathcal{C}^0_{\text{adm}} \cap L^+_0 = \{0\}$. That is we have for all $\vartheta \in \Theta_{\text{adm}}$ $H = G_T(\vartheta) - 0 \geq 0$ $\mathbb{P}$-a.s. which implies (by setting $Y = 0$ that $G_T(\vartheta) = 0$ $\mathbb{P}$-a.s., hence NA.

b) "$\implies$" From a) we know that NA implies that $\mathcal{C}^0_{\text{adm}} \cap L^0_+ = \{0\}$, since $(\mathcal{C}^0_{\text{adm}} \cap L^\infty) \cap L^0_+$ is clearly a subset, we must also have that $(\mathcal{C}^0_{\text{adm}} \cap L^\infty) \cap L^0_+ = \{0\}$.

"$\impliedby$" Let us show that $\mathcal{C}^0_{\text{adm}} \cap L^0_+ = \{0\}$, which by a) is equivalent to NA. So let $H \in \mathcal{C}^0_{\text{adm}} \cap L^0_+$ be arbitrary. Set $H^n := H \wedge n$, then $H^n \in (\mathcal{C}^0_{\text{adm}} \cap L^\infty) \cap L^0_+$. Indeed, we have $H = G_T(\vartheta) - Y$ and then $H \wedge n = H - (H-n)^+ = G_T(\vartheta) - (Y + (H-n)^+) = G_T(\vartheta) - Y'$ where $Y' \geq 0$. Hence, by assumption, $H^n = 0$ for all $n \in \mathbb{N}$ and since $H^n \to H$ $\mathbb{P}$-a.s., we conclude that $H = 0$ $\mathbb{P}$-a.s. as desired. $\square$

**Remark 16.2.** From b) it follows immediately that NA $\iff \mathcal{C}^\infty \cap L^\infty_+ = \{0\}$. Indeed, just intersect by $L^\infty_+$ and use $\{0\} \subset L^\infty_+ \subset L^0_+$. Note that the more general NFLVR is defined by $\overline{\mathcal{C}^\infty}^{L^\infty} \cap L_+\infty = \{0\}$.

**Exercise 5.2:** In a general financial market. Let $S$ be a semimartingale and assume that $S$ satisfies NA. Prove that if $\vartheta \in \Theta_{\text{adm}}$ satisfies $G_T(\vartheta) \geq -c$ $\mathbb{P}$-a.s. for some $c \geq 0$, then $G_.(\vartheta) \geq -c$ $\mathbb{P}$-a.s.

**Hint**: Use that if $\vartheta$ is $S$-integrable and $C \in \mathcal{P}$ is a predictable set, then $1_C\vartheta$ is $S$-integrable too.

*Proof.* Since $\vartheta \in \Theta_{\text{adm}}$ we have $a \geq 0$ such that $G_.(\vartheta) \geq -a$ $\mathbb{P}$-a.s. (recall definition is $G_t(\varphi) \geq -a$ $\mathbb{P}$-a.s. for **all** $t \in [0,T]$). By the right-continuity of the paths of $G_.(\vartheta)$, it suffices to show $G_t(\vartheta) \geq -c$ $\mathbb{P}$-a.s. for any $t \in [0,T)$. Seeking a contradiction, assume there is $t \in [0,T)$ such that $\mathbb{P}(G_t(\vartheta) < -c) > 0$. But then

$$\vartheta^* := \vartheta 1_{\{G_t(\vartheta)<-c\}\times(t,T]}$$

is predictable, $S$-integrable, because $C = \{G_t(\vartheta) < -c\}\times(t,T] \in \mathcal{P}$ is a predictable set (see given Hint). Moreover $\vartheta^*$ satisfies

$$G_.(\vartheta^*) = (G_.(\vartheta) - G_t(\vartheta))1_{\{G_t(\vartheta)<-c\}\times(t,T]}) \geq -a + c$$

which shows that $\vartheta^*$ is admissible. Moreover

$$G_T(\vartheta^*) = (G_T(\vartheta) - G_t(\vartheta))1_{\{G_t(\vartheta)<-c\}} \geq (-c - G_t(\vartheta))1_{\{G_t(\vartheta)<-c\}} > 0.$$

Which shows that $S$ fails NA (for the strategy $\varphi \triangleq (0,\vartheta^*)$), in contradiction to the assumption. $\square$

**Remark 16.3.** Interesting about this exercise is, that under the assumption of (NA), the final wealth gain $G_T(\vartheta)$ of an admissible strategy provides some control for the gain process (i.e. $G_t(\vartheta)$ for all $t \in [0,T]$).

**Exercise 5.3:** Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space, $S = (S_t^1, \dots, S_t^d)_{t\in[0,T]}$ a $d$-dimensional semimartingale and $Q \approx \mathbb{P}$ on $\mathcal{F}_T$ an equivalent probability measure.

a) Assume that $\mathcal{F}_0$ is trivial and $Q$ is a separating measure for $S$. Show that if $S$ is (locally) bounded, then $Q$ is an equivalent (local) martingale measure for $S$.

b) Assume that $Q$ is an equivalent $\sigma$-martingale measure for $S$. Show that it is also an equivalent separating measure.

*Proof.* a) First, let us assume that $S$ is bounded. Note, that then every simple strategy is admissible. Moreover, $S$ is a uniformly integrable $Q$-martingale if and only if $E_Q(S_\tau - S_0) = 0$ for all stopping times $\tau$ with values in $[0,T]$. So let $\tau$ be an arbitrary $[0,T]$-valued stopping time, and consider the simple strategies $\vartheta^\pm := \pm 1_{(\!(0,\tau]\!]}$. Since by assumption $Q$ is a separating measure for $S$ we have $\mathbb{E}_Q(G_T(\vartheta)) \leq 0$ for all $\vartheta \in \Theta_{\text{adm}}$. Applying this to our strategies $\vartheta^\pm$ we obtain

$$0 \geq \mathbb{E}_Q(G_T(\vartheta^\pm)) = \mathbb{E}_Q(\vartheta^\pm \cdot S_T) = \pm\mathbb{E}_Q(S_\tau - S_0). \tag{$\star$}$$

The above yields that that $S$ is a $Q$-martingale.

Let us now assume that $S$ is (only) locally bounded. Then there exists an increasing sequence of stopping times $(\sigma_n)_{n\in\mathbb{N}}$ taking values in $[0,T]$ with $\lim_{n\to\infty} \mathbb{P}(\sigma_n = T) = 1$ such that $S^{\sigma_n}$ is bounded for every $n \in \mathbb{N}$. It suffices to show that for each $n \in \mathbb{N}$, $S^{\sigma_n}$ is a uniformly integrable $Q$-martingale. To this end, fix $n \in \mathbb{N}$ and then it suffices to show that for each stopping time $\tau$ with $\tau \leq \sigma_n$ $\mathbb{P}$-a.s., $\mathbb{E}_Q(S_\tau - S_0) = 0$.

So let $\tau$ be such a stopping time, and consider as above the simple strategies $\vartheta^\pm := \pm 1_{(\!(0,\tau]\!]}$. Then both strategies are admissible since $S$ is bounded on $[\![0,\sigma_n]\!]$ and $\tau \leq \sigma_n$ $\mathbb{P}$-a.s., and the same argument as in the first step (see ($\star$)) gives $E_Q(S_\tau - S_0) = 0$.

b) By assumption $Q$ is an equivalent $\sigma$-martingale measure for $S$, hence there exist a strictly positive predictable process $\Psi = (\Psi_t)_{t\in[0,T]}$ and an $\mathbb{R}^d$-valued local $Q$-martingale $M$ and an $\mathbb{R}^d$-valued $\mathcal{F}_0$-measurable random vector $S_0$ such that

$$S - S_0 = \int \Psi dM = \Psi \cdot M.$$

Let now $\vartheta \in \Theta_{\text{adm}}$, then we obtain by the associativity of the stochastic integral that

$$G(\vartheta) = \vartheta \cdot S = \vartheta \cdot (\Psi \cdot M) = (\vartheta\psi) \cdot M \geq -a.$$

**Remark 16.4.** We can also use the heuristic argument (which isn't very rigoros of course), name $dS = \Psi dM$ and hence

$$G(\vartheta) = \int \vartheta dS = \int \vartheta \Psi dM \overset{\vartheta \in \Theta_{\text{adm}}}{\geq} -a$$

As our calculations have shown, $(\vartheta\Psi) \cdot M$ is uniformly bounded from below by admissibility, hence by Ansel-Stricker theorem it is a local $Q$-martingale and then, by Fatou, it is also a $Q$-supermartingale and hence

$$\mathbb{E}_Q(G_T(\vartheta)) \leq \mathbb{E}_Q(G_0(\vartheta)) = 0.$$

Hence $Q$ is an equivalent separating measure. $\square$

**Remark 16.5.** The previous exercise (part b) establishes that

- EMM $\Longrightarrow$ ELMM $\overset{\Psi=1}{\Longrightarrow}$ E$\sigma$MM $\overset{b)}{\Longrightarrow}$ ESM.

If we further assume that $S$ is bounded, we established in part $a$) that ESM $\Longrightarrow$ EMM, i.e. we have

- EMM $\Longrightarrow$ ELMM $\Longrightarrow$ E$\sigma$MM $\Longrightarrow$ ESM$\overset{S \text{ bounded}}{\Longrightarrow}$EMM. Hence all these notions become one and the same.

In the case of $S$ being locally bounded, we still have ESM $\Longrightarrow$ ELMM and hence we get ELMM=E$\sigma$MM=ESM.

**Exericse 5.4:** The FTAP asserts that for a (general) $\mathbb{R}^d$-valued semimartingale $S$, the following statements are equivalent:

1. $S$ satisfies (NFLVR) i.e $\overline{\mathcal{C}^\infty}^{L^\infty} \cap L^\infty_+ = \{0\}$.

2. $S$ admits an equivalent separating measure.

3. $S$ admits an equivalent $\sigma$-martingale measure.

Show that 3) $\implies$ 1).

*Proof.* We know already that when $S$ admits an equivalent $\sigma$-martingale measure, then $S$ admits an equivalent separating measure (Ansel-Stricker and Fatou argument, see previous Exercise). That is there exists a probability measure $Q \approx \mathbb{P}$ such that

$$\mathbb{E}_Q(G_T(\vartheta)) \leq 0 \text{ for all } \vartheta \in \Theta_{\text{adm}}. \tag{$\star$}$$

**Remark 16.6** (Optional)**.** From ($\star$) one can actually deduce (NA). Indeed, if $G_T(\vartheta) \in L^0_+$, i.e. nonnegative $\mathbb{P}$-a.s., then also $Q$-a.s. (since $Q \approx \mathbb{P}$). Hence we must have $G_T(\vartheta) = 0$ $Q$-a.s. by ($\star$) and then also $\mathbb{P}$-a.s since $Q \approx \mathbb{P}$. We have thus verified that (NA) must hold.

However, we don't want to show (NA) but the stronger (NFLVR). To this extend, we will show that $\mathbb{E}_Q(G_T(\vartheta)) \leq 0$ for all $h \in \mathcal{C}^\infty$ and then show that this statement extends even to the norm closure of $\mathcal{C}^\infty$.

Let us consider $h \in \mathcal{C}^\infty$ arbitrary, we then have $h = G_T(\vartheta) - Y$ where $\vartheta \in \Theta_{\text{adm}}$ and $Y \in L^0_+$. Moreover

$$h = G_T(\vartheta) \underbrace{-Y}_{\leq 0} \leq G_T(\vartheta),$$

and hence $\mathbb{E}_Q(h) \leq \mathbb{E}_Q(G_T(\vartheta)) \leq 0$. This shows that $\mathbb{E}_Q(h) \leq 0$ for all $h \in \mathcal{C}^\infty$. Let now $h \in \overline{\mathcal{C}^\infty}^{L^\infty}$, then there exists a sequence $(h_n)_{n\in\mathbb{N}} \subset \mathcal{C}^\infty$ such that $\|h_n - h\|_{L^\infty} \to 0$ as $n \to \infty$. In particular we have (by dominated convergence)

$$\mathbb{E}_Q(h) = \mathbb{E}_Q(\lim_{n\to\infty} h_n) = \lim_{n\to\infty} \mathbb{E}_Q(h_n) \leq 0.$$

If we now also assume that $h \in L^\infty_+$, i.e. $h \geq 0$ $\mathbb{P}$-a.s., then also $Q$-a.s. since $Q \approx \mathbb{P}$, and then $h \equiv 0$ $Q$-a.s., hence (again since $Q \approx \mathbb{P}$) $\mathbb{P}$-a.s. and we conclude (NFLVR). $\square$

**Exercise 6.1:** Let $Z$ be a strictly positive local martingale with $Z_0 = 1$ and suppose that $S$ is an adapted **continuous** process with $S_0 = 0$ such that $ZS$ is a $\sigma$-martingale. Prove that $ZS$ is a local martingale.

*Proof.* We will verify this statement by using Lemma 5.2, i.e. De Donno/Pratelli. By definition $ZS$ is a $\sigma$-martingale, i.e. $ZS = \int \Psi dM$ for a local martingale $M$ and an integrand $\Psi \in L(M)$ with $\Psi > 0$. We also recall that $\mathcal{M}_{0,\text{loc}} = \mathcal{H}^1_{0,\text{loc}}$.

Let now $(\rho_n)_{n\geq 0}$ be a localizing sequence for $M$ such that $M^{\rho_n} \in \mathcal{H}^1_0$ for each $n$. As in the proof of the Ansel-Stricker theorem (in order to apply De Donne/Pratelli), for each $k$, we define $\Psi^k := \psi 1_{\{|\Psi|\leq k\}} \in \mathbf{b}\mathcal{P}$ and for each $n$ and $k$ define

$$M^{n,k} := \int \Psi^k dM^{\rho_n} \in \mathcal{H}^1_0.$$

Note that $M^{n,k} \in \mathcal{H}^1_0$ thanks to Proposition 4.1 since $\Psi^k \in \mathbf{b}\mathcal{P}$ and $M^{\rho_n} \in \mathcal{H}^1_0$. Now, by the very definition of $\Psi$ being $M$-integrable (i.e. in $L(M)$), we have that $M^{n,k} = \Psi^k \cdot M^{\rho_n} \to \Psi \cdot M^{\rho_n}$ as $k \to \infty$ with respect to the Emery topology, i.e. with respect to $d_S$, hence also for $d$ for each fixed $n$. We also have

$$(\Delta M^{n,k})^{\pm} \leq \left(\Delta \int \Psi dM^{\rho_n}\right)^{\pm} \quad \text{for all } k \in \mathbb{N}.$$

With the above it is easy to see that the inequality also holds for all stopping times $\sigma$, i.e. $(\Delta M^{n,k}_\sigma)^{\pm} \leq (\Delta \int_0^\sigma \Psi dM^{\rho_n})^{\pm}$. Assume for the moment that $\int \Psi dM$ has locally a lower bound in $L^1$, which means that there exists a localizing sequence (increasing to $T$ stationarily) and a sequence of random variables $(\gamma_m) \subset L^1$ such that $(\int \Psi dM)^{\tau_m} \geq \gamma_m$ for each $m$. In particular, we have $(\int \Psi dM^{\rho_n})^{\tau_m} \geq \gamma_m$ for each $n, m$. Thus, all the assumptions of De Donno/Pratelli are satisfied for each $n$, which implies that

$$\int \Psi dM^{\rho_n} = \left(\int \Psi dM\right)^{\rho_n} \in \mathcal{M}_{0,\text{loc}} \text{ for all } n \in \mathbb{N}.$$

Since $\rho_n \nearrow T$ stationarily, the above implies that $ZS = \int \Psi dM \in \mathcal{M}_{0,\text{loc}}$, i.e. $ZS$ is a local martingale. Therefore, it remains to show that $ZS$ has locally a lower bound in $L^1$.

**Remark 16.7.** So far we haven't used the fact that $S$ is **continuous**. We still need to argue the local lower bound in $L^1$ and we'll see that this is where the continuity argument of $S$ comes into play.

Let us now work on the final (missing) statement. Let $(T_n)_{n\in\mathbb{N}}$ be a localizing sequence for the local martingale $Z$ such that each $Z^{T_n}$ is an $\mathcal{H}^1$-martingale and define for each $n \in \mathbb{N}$ the stopping times $\sigma_n := \{t \geq 0 : |S_t| \geq n\}$ and $\hat{T}_n := \{t \geq 0 : |Z_t| \geq n\}$. Consider the sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ defined by $\tau_n := T_n \wedge \hat{T}_n \wedge \sigma_n$. By construction, $\tau_n$ converges to $\infty$ a.s. Moreover,

$$\sup_{t\geq 0} |Z_t^{\tau_n}| \leq n + |\Delta Z_{\tau_n}| \in L^1,$$

Also, since $S$ is continuous, we have $|S_t^{\tau_n}| \leq n$ i.e. $S_t^{\tau_n} \geq -n$ for all $t$ and thus we obtain for each $n$ that

$$(ZS)^{\tau_n} \geq -n(n + |\Delta Z_{\tau_n}|) := \gamma_n \in L^1.$$

which shows the local bound from below in $L^1$ (we used that for $f \geq -n$ and $g \leq \|g\|_\infty$, then $fg \geq -n\|g\|_\infty$). □

**Remark 16.8.** In the general setup where $S$ is a RCLL semimartingale we have, by definition, that every ELMD is a E$\sigma$MD. This Exercise shows that when $S$ is continuous, the converse holds as well, in particular in the continuous case the notions of ELMD and E$\sigma$MD are one and the same.

**Exercise 6.2:** Show that there is at most one numéraire portfolio.

*Proof.* Suppose that $X, Y \in \mathcal{X}^1_{++}$ are both numéraire portfolios. Hence, by definition of a numéraire portfolio, both $X/Y$ and $Y/X$ are supermartingales starting from 1 at time 0. Set $R := X/Y$ and $S := Y/X$. Then, since $R$ is a supermartingale, we have

$$\mathbb{E}(R_t) \leq \mathbb{E}(R_0) = 1 \text{ for all } t > 0.$$

On the other hand, since $S$ is a supermartingale, we have

$$\mathbb{E}(1/R_t) = \mathbb{E}(S_t) \leq \mathbb{E}(S_0) = 1 \text{ for all } t > 0. \tag{$\star$}$$

But, by Jensen's inequality, since $x \mapsto 1/x$ is (strictly) convex on $(0, \infty)$ we have together with $(\star)$

$$1 \overset{(\star)}{\geq} \mathbb{E}(1/R_t) \overset{\text{Jensen}}{\geq} 1/\mathbb{E}(R_t) \geq 1 \text{ for all } t > 0.$$

The above entails that $\mathbb{E}(1/R_t) = 1$ and $\mathbb{E}(R_t) = 1$ for all $t > 0$, but since $x \mapsto 1/x$ is in fact strictly convex on $(0, 1)$ this can happen if and only if $R = 1$ i.e. $X_t = Y_t$ $\mathbb{P}$-a.s for all $t \geq 0$. □

**Exercise 11.1 and 11.2**: We summarize the two exercises here as follows: Let $U : (0, \infty) \to \mathbb{R}$ be concave and increasing. For $x > 0$, define

$$u(x) := \sup_{V \in \mathcal{V}(x)} \mathbb{E}[U(V_T)],$$

where $\mathcal{V}(x) := \{x + \vartheta \cdot S : \vartheta \in \Theta^x_{\text{adm}}\}$ (see primal problem).

a) Show that $u$ is concave and increasing.

b) If additionally $u(x_0) < \infty$ for some $x_0 >$, show that $u(x) < \infty$ for all $x > 0$.

c) If $U$ is **strictly** increasing with $U(\infty) < \infty$ and $S$ satisfies (NFLVR), then $u(x) < U(\infty)$ for all $x > 0$. (Compare with Exercise 11.1, shows that primal problem only makes sense in arbitrage-free market!)

*Proof.* a) We first show that $u$ is increasing. Let $x, y \in (0, \infty)$ with $x \leq y$ and $\vartheta \in \Theta^x_{\text{adm}}$. Clearly we have

$$x + \vartheta \cdot S_T \leq y + \vartheta \cdot S_t.$$

Since $U$ is increasing, and $\Theta^x_{\text{adm}} \subset \Theta_t extadm^y$ (which implies that $\mathcal{V}(x) \subset \mathcal{V}(y)$) we have

$$\mathbb{E}(U(V_T(\vartheta)) = \mathbb{E}[U(x + \vartheta \cdot S_T)] \leq \mathbb{E}[U(y + \vartheta \cdot S_T)] \leq u(y).$$

Taking the supremum over $\mathcal{V}(x)$ on the LHS yields $u(x) \leq u(y)$, hence $u$ is increasing.

Now we prove the concavity of $u$. Let $\lambda \in [0, 1]$ and $x, y \in (0, \infty)$ with $x \leq y$. If $\vartheta^x \in \Theta^x_{\text{adm}}$ and $\vartheta^y \in \Theta^y_{\text{adm}}$, we clearly have $\lambda\vartheta^x + (1-\lambda)\vartheta^y \in \Theta^{\lambda x+(1-\lambda)y}_{\text{adm}}$. Since $U$ is concave, we obtain

$$\begin{aligned} u(\lambda x + (1-\lambda)y) &\geq \mathbb{E}[U(\lambda(x + \vartheta^x \cdot S_T) + (1-\lambda)(y + \vartheta^y \cdot S_T))] \\ &\geq \lambda\mathbb{E}[U(x + \vartheta^x \cdot S_T)] + (1-\lambda)\mathbb{E}[U(y + \vartheta^y \cdot S_T)]. \end{aligned}$$

Taking the supremum over $\mathcal{V}(x)$ and $\mathcal{V}(y)$ on the RHS yields $u(\lambda x + (1-\lambda)y) \geq \lambda u(x) + (1-\lambda)u(y)$. Hence $u$ is concave.

b) By part a), we only need to prove $u(x) < \infty$ for all $x \in (x_0, \infty)$, since $u$ is increasing. We have $u(x_0) < \infty$ by assumption, clearly we can find $\lambda \in (0, 1)$ and $(x_0 <)\ x < y$ such that $x_0 = \lambda x + (1-\lambda)y$. By the concavity of $u$, we then have $\lambda u(x) + (1-\lambda)u(y) \leq u(x_0)$, which can be rewritten as

$$u(x) \leq \frac{u(x_0) - (1-\lambda)u(y)}{\lambda} < \infty, \text{ for all } x \in (x_0, \infty).$$

Hence the claim follows.

c) Suppose to the contrary that we have $u(x) \geq U(\infty)$ for some $x \in (0, \infty)$. Clearly we have $U(V_T) \leq U(\infty)$ for all $V \in \mathcal{V}(x)$ and hence

$$u(x) = \sup_{V \in \mathcal{V}(x)} \mathbb{E}[U(V_T)] \leq U(\infty).$$

This entails that $u(x) = U(\infty)$, hence we can find $(V^n) \subset \mathcal{V}(x)$ be such that $\mathbb{E}(U(V_T^n)] \nearrow U(\infty)$. By Komlós Lemma (see Appendix), for each $n \in \mathbb{N}$, there exists $\tilde{V}_T^n \in \operatorname{conv}(V_T^n, V_T^{n+1}, \dots)$ such that $\tilde{V}_T^n \to \tilde{V}^\infty$ $\mathbb{P}$-a.s. We know that (NFLVR) = (NA) + (NUPBR), hence by assumption of (NFLVR) and in particular (NUPBR) implies that $\operatorname{conv}(V_T^1, V_T^2, \dots)$ is bounded in $L^0$, therefore (again by Komlós Lemma), we know that $\tilde{V}^\infty < \infty$ $\mathbb{P}$-a.s. The concavity of $U$ (together with $U$ being increasing) implies that

$$\mathbb{E}[U(\tilde{V}_T^n)] \overset{U \text{ concave}}{\geq} \inf_{k \geq n} \mathbb{E}[U(V_T^k)] \overset{U \text{ incr.}}{=} \mathbb{E}[U(V_T^n)].$$

Since $U(\tilde{V}_T^n) \leq U(\infty)$ for all $n \in \mathbb{N}$, applying the reverse Fatou lemma gives

$$\mathbb{E}[U(\tilde{V}^\infty)] \geq \limsup_{n \to \infty} \mathbb{E}[U(\tilde{V}_T^n)] \geq \liminf_{n \to \infty} \mathbb{E}[U(\tilde{V}_T^n)] = U(\infty).$$

So clearly $\mathbb{E}[U(\infty) - U(\tilde{V}^\infty)] = 0$. But $U$ is strictly increasing and we established that $\tilde{V}^\infty < \infty$ $\mathbb{P}$-a.s., so $U(\infty) - U(\tilde{V}_T^\infty) > 0$ $\mathbb{P}$-a.s. which gives a contradiction, hence $u(x) < U(\infty)$ for all $x > 0$ and we're done. $\square$