Mathematical Finance HS24 - Prof. Beatrice Acciaio

Jonatan Wächter (jwaechter@ethz.ch)

February 2025

Contents

Chapter 1

First ideas in arbitrage theory

Basic idea: In reasonable models of a financial market, money pumps should not be possible. How is this formalized? How are good models 'characterised'?

Basic set up: $(\Omega, \mathcal{F}, P)$, $[0, T]$, $S^0 = 1$ and $S = (S^1, ..., S^d)$ is adapted and RCLL.

1.1 (Simple) strategies, admissibility, arbitrage

Definition 1.1. *A strategy is a-admissible for $a > 0$, if $V_t(\varphi) \geq -a$ P-a.s. $\forall t$.*

If we can choose a RCLL version of $V(\varphi)$, then equivalently we satisfy the stronger condition that $P(V_t(\varphi) \geq -a \quad \forall t \in [0, T]) = 1$.

Definition 1.2. *φ is admissible if it is a-admissible for some $a \geq 0$.*

Remarks:

- For $\varphi = (v_0, \theta)$ self-financing with $v_0 = 0$ then it is a-admissible if integrand θ is a-admissible in the sense $G(\theta) \geq$ const., resp $-a$.
- For $v_0 \neq 0$ things can be very different when looking for lower bounds for $V(\varphi)$ or $G(\theta)$
  - v_0 deterministic, then it is trivial
  - for $v_0 \in L(\mathcal{F}_0)$, non-constant, it is tricky

Definition 1.3. *A simple integrand, denoted by $\theta \in b\mathcal{E}$ satisfies $\theta = \sum_{i=1}^n h_i \mathbf{1}_{]\tau_{i-1},\tau_i]}$ for n a natural number, τ_i increasing stopping times bounded by T and h_i $\mathbb{R}^d$-valued, bounded $\mathcal{F}_{\tau_{i-1}}$-measurable (written $h_i \in L^\infty(\mathcal{F}_{\tau_{i-1}};\mathbb{R}^d)$) and so in particular predictable.*

The subset of $b\mathcal{E}$ with τ_i deterministic is called very simple and denoted $b\mathcal{E}_{\text{det}}$.

Note that for $\theta \in b\mathcal{E}$, $\int \theta dS$ is well-defined for any $\mathbb{R}^d$-valued process (i.e. we don't require that it is a semimartingale) and we have

$$G(\theta) = \int \theta dS = \sum_{i=1}^n h_i(S^{\tau_i} - S^{\tau_{i-1}})$$
$$G_T(\theta) = \sum_{i=1}^n h_i(S_{\tau_i} - S_{\tau_{i-1}})$$

since the stopping times are bounded by T. If S is adapted and RCLL, the same is true of $G(\theta)$.

Definition 1.4. *For a general S we define a simple arbitrage opportunity to be a $\theta \in b\mathcal{E}$, admissible, with $G_T(\theta) \in L^0_+(\mathcal{F}_T)\backslash\{0\}$, i.e.*

$$G_T(\theta) \geq 0$$
$$P(G_T(\theta) > 0) > 0$$

Interpretation: Starting with $v_0 = 0$, keeping bounded debts, following $\varphi = (0, \theta)$ self-financing we arrive in T with $V_T(\varphi) = G_T(\theta) \geq 0$ and with positive probability > 0.

Definition 1.5. *For S a semimartingale a general arbitrage opportunity is θ integrand for S, admissible with $G_T \in L^0_+\backslash\{0\}$.*

Being an integrand for S means $\mathbb{R}^d$-valued, predictable, S-integrable so that $\int \theta dS$ makes sense.

Notation

$\Theta_{\text{adm.}}$ is the set of all admissible integrands for S. We write $G_T(\Theta_{\text{adm.}}) = \{G_T(\theta) : \theta \in \Theta_{\text{adm.}}\}$.

We write (NA) to mean no arbitrage, a characteristic of a financial market model. We add super- and subscripts to specify the integrands used in the following conditions.

1.2 Absence of arbitrage conditions

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

Note that for the last condition we require that S is a semimartingale whereas the first three work for any S, since no notion of stochastic integral is needed.

Definition 1.6. *An equivalent (local) martingale measure E(L)MM for S is a positive measure $Q \sim P$ (equivalent to) s.t. S is a (local) martingale under Q*

Lemma 1.7. *If there exists an ELMM then both (NA) and $(NA^{adm.}_{elem.})$ hold.*

Note that the other two kinds of no arbitrage conditions might fail. Interpretation of lemma: "Market is good enough to avoid some kinds of arbitrage opportunities".

Proof. Failure of $(\mathrm{NA}_{\mathrm{det.}})$ left as an exercise. Since it is a subset of $(\mathrm{NA}_{\mathrm{elem.}})$ the claim follows for this condition as well. In a similar vein $(\mathrm{NA}^{\mathrm{adm.}}_{\mathrm{elem.}})$ follows from (NA) so we are only left to show the latter.

We show that if there exists Q ELMM, then for all $\theta \in \Theta_{\mathrm{adm.}}$ s.t. $G_T(\theta) \geq 0$ a.s. we have necessarily that $G_T(\theta) = 0$ a.s.

First note that since $S \in \mathcal{M}_{loc}(Q)$ and $Q \sim P$, by Girsanov, S is P-semimartingale so that talking about (NA) under P makes sense. Recall from stochastic calculus that S semimartingale and θ predictable and S-integrable then $G(\theta)$ is a semimartingale. However the fact that S is a local martingale does not guarantee that $G(\theta)$ is a local martingale. (Counterexample by Émery, S jumps). There are additional sufficient conditions to ensure that $G(\theta)$ is a local martingale as well. One example is the result by Ansel-Stricker: For S a local martingale and $G(\theta)$ bounded from below, the latter is a local martingale. We will use this result to show (NA).

$S \in \mathcal{M}_{loc}(Q)$, $\theta \in \Theta_{\mathrm{adm.}}$, then a.s. $G(\theta) \in \mathcal{M}_{loc.}$ and bounded from below. $G.(\theta)$ is Q-supermartingale, $G_0(\theta) = 0$ so that

$$E_Q[G_T(\theta)] \leq E_Q[G_0(\theta)] = 0 \tag{1.1}$$

So if $G_T(\theta) \geq 0$ P-a.s. then Q-a.s. the same is true, and $E_Q[G_T(\theta)] = 0$ and so $G_T(\theta) = 0$ Q-a.s., hence also P-a.s. which allows us to conclude that $G_T(\Theta_{\text{adm.}}) \cap L^0_+ = \{0\}$, which is the definition of (NA).

$\square$

This raises the following question: Does the reverse hold, i.e. does no arbitrage imply the existence of a ELMM? The answer is no in general, and we will see an example of this shortly. If we however restrict to discrete time, this is true, see theorem by Dalang-Morton-Willinger.

1.3 A counterexample in infinite discrete time

Let Y_n take values in $\{\pm 1\}$, $P(Y_n = +1) = \frac{1}{2}(1+\alpha_n)$. Define $S_0 = 1$ and $\Delta S_n = S_n - S_{n-1} = \beta_n Y_n$. S is a random walk with drift. Note that $\mathbb{F} = \mathbb{F}^S = \mathbb{F}^Y$. If S is to be a $(\mathbb{F}, Q)$-martingale, $Q(Y_{n+1} = 1 | \mathcal{F}_{n-1}) = \frac{1}{2}$ is the only candidate for Q. Using Kakutani's dichotomy theorem (Williams 4.17) we have that $Q \sim P$ iff $\sum_{n\geq 1} \alpha_n^2 < \infty$, otherwise $Q \perp P$. Note that $Q \sim P$ on $\mathcal{F}_n$, i.e. locally, so it's the infinitely many times that may cause problems. If we are able to choose β_n such that the sum does not converge, we have proved that there does not exist a Q which is EMM for S.

If $\sum \beta_n < \infty$, then S is bounded and so the set of EMMs for S coincides with that of ELMMs. If no EMM exists, the same is true for existence of ELMMs. We will show that for a suitable choice of β_n's S satisfies $(\text{NA}_{\text{elem.}})$

An exercise is the following: There does not exist an arbitrage opportunity in $b\mathcal{E}$ iff there does not exist an arbitrage opportunity of the form $\theta = h\mathbf{1}_{]\sigma,\tau]}$ where h is $\mathcal{F}_\sigma$-measurable and σ, τ are stopping times.

Taking this for granted, we pick $\beta_n = 3^{-n}$ so that $\sum_{k=m+1}^{\infty} \beta_k < \beta_m$. This ensures $\text{sgn}(S_n - S_m) = \text{sgn}(Y_{m+1})$ or even the same equality but under multiplication by a random variable g.

Take $\theta = \mathbf{1}_{[\sigma,\tau]}$ and let $A_m = \{\sigma = m, \tau > m\} \in \mathcal{F}_m$. Then $G_\infty(\theta) = h(S_\tau - S_\sigma)$. We get $\text{sgn}(G_\infty(\theta)) = \text{sgn}(hY_{m+1})$ on A_m so if $G_\infty(\theta) \geq 0$ we have $\forall m$ that $\mathbf{1}_{A_m}\text{sgn}(hY_{m+1}) \geq 0$. But $A_m \in \mathcal{F}_m$, h is $\mathcal{F}_\sigma$-measurable and $\sigma = m$ on A_m. We also know that Y_{m+1} is independent of $\mathcal{F}_m$ and with values $\{\pm 1\}$. So we must have $h = 0$ on A_m a.s. $\forall m$, i.e. $h = 0$ a.s. and so $G_\infty(\theta) = 0$ a.s.. This is exactly what we wanted to show.

Why is $h = 0$ on A_m? h is $\mathcal{F}_\sigma$-measurable, $\sigma = m$ on A_m and so $h\mathbf{1}_{A_m}$ is $\mathcal{F}_m$-measurable. We saw that $0 \leq \mathbf{1}_{A_m}\text{sgn}(h\mathbf{1}_{A_m}Y_{m+1})$ for all m. This implies that $h\mathbf{1}_{A_m}$ must have the same sign on $A_m \cap \{Y_{m+1} = 1\}$ and $A_m \cap \{Y_{m+1} = -1\}$. We get $h\mathbf{1}_{A_m} = 0$ for all m and so $h = 0$.

Remarks: Check Back-Pliska 91', Schachermayer 94', Delbaen-Schachermayer P 5.1.7

1.4 A counterexample in continuous time

We let $W = (W_t)_{t\in[0,T]}$ be a Brownian Motion, $\mathbb{G}$ the augmented natural filtration. $K(t) = \frac{1}{\sqrt{T-t}}$ and define

$$Z := \mathcal{E}(-\int KdW) \quad \text{on}[0,T] \tag{1.2}$$

$$Z_T := 0 \tag{1.3}$$

so that

$$Z_t = \exp(-\int KdW - \frac{1}{2}\int K^2 ds) \quad \text{for} \quad 0 \leq t \leq T \tag{1.4}$$

Define $\tau = \inf\{t \in [0,T] : Z_t \geq 2\} \wedge T$. We then have that Z^τ is a bounded martingale and $P(\tau < T) = \frac{1}{2}$. Define

$$S_t = \mathbf{1}(t \leq \tau)(W_t + \int kds) + \mathbf{1}(t \geq \tau)(S_\tau) \tag{1.5}$$

Take the filtration $\mathcal{F}_t = \mathcal{G}_{t\wedge\tau}$ on $0 \leq t \leq \tau$ so that $\mathbb{F} = \mathbb{F}^S = \mathbb{F}^{W^\tau}$ up to nullsets. From a (general form) of the martingale representation theorem for Brownian Motion, all $(\mathbb{F}, P)$-local martingales, null at zero, are stochastic integrals w.r.t. W^τ so by Girsanov theorem, only $Q \ll P$ which makes $S = S^\tau$ a local Q-martingale removes the drift, so that $\frac{dQ}{dP} = Z_\tau$. But $Z_\tau = 0$ on $\{\tau = T\}$ which has $P(\tau = T) = \frac{1}{2}$ so that $P(Z_\tau = 0) > 0$ implies $Q \not\sim P$. Hence there does not exist an ELMM for S.

We now want to show that S satisfies (NA).

Note that S a local Q-martingale means that S satsifies (NA) under Q. But we only have $Q \ll P$, $Q \not\sim P$. This means that L^0_+ and (NA) conditions are different under P and under Q. Note that $\frac{dQ}{dP}|_{\mathcal{F}_t} = Z^\tau_t > 0$ for $t < T$ since $Z > 0$ for $t < T$ which means $Q \sim P$ pn $\mathcal{F}_t$ for every $t < T$.

Note also that S is continuous and S_T well defined since $\int_0^T K(s)ds < \infty$. So the stochastic integral $G(\theta)$ is well defined P- and Q-a.s..

Now take $\theta \in \Theta_{adm}(P)$ and suppose that $G_T(\theta) \geq 0$ P-a.s. This holds also for Q. So S and by continuity also $G(\theta)$ are local Q-martingales and being bounded from below, $G(\theta)$ is a Q-supermartingale, null at zero. So $E_Q[G_T(\theta)] \leq 0$ implies $G_T(\theta) = 0$ Q-a.s.. (We need more, i.e. that this is true P-a.s.).

Fix $\epsilon > 0$ and set $\sigma = \inf\{t \in [0,T] : G_t(\theta) \geq \epsilon\} \wedge T$ and $\theta' = \mathbf{1}_{[0,\sigma]}\theta$. Then $G(\theta') = G(\theta)^\sigma$. In particular $\theta' \in \Theta_{\text{adm.}}$ and $G_T(\theta') = G_\sigma(\theta) = \epsilon$ on $\{\sigma < T\}$ and $= G_T(\theta)$ otherwise, and so by assumption it is ≥ 0 P-a.s..

Since S satisifes (NA) under Q we need to have $Q(\{\sigma < T\}) = 0$, so $G(\theta)$ can never go above the level ϵ, Q-a.s., i.e. $G_t(\theta) < \epsilon$ for all $t < T$ Q-a.s.. But this means that $G_t(\theta) < \epsilon$ also P-a.s., since $Q \sim P$ on $\mathcal{F}_t$ for $t < T$. Using continuity we get the (weaker, but sufficient) inequality $G_T(\theta) \leq \epsilon$ P-a.s.. Since $\epsilon > 0$ was arbitrary, we conclude that $G_T(\theta) \leq 0$ P-a.s. and so combined with our assumption we have $G_T(\theta) = 0$ P-a.s., i.e. S satisfies (NA) under P.

References for this example is the paper Delbaen-Schachermayer 1994 and the book by the same authors, example 9.7.7.

We used a general representation theorem above. To show that every $M \in \mathcal{M}_{0,loc.}(\mathbb{F})$ is a stochastic integral w.r.t. W^τ one could argue that $M = N^\tau$ for some $N \in \mathcal{M}_{0,loc}(\mathbb{F}^W)$ and then apply the classical martingale representation theorem $N = \int HdW$ and stop at τ. To get N one could represent M as a time change of W...

1.5 Recap of important results in infinite discrete time

The setup is a probability space, time steps $0, 1, ..., T$. $\Theta = \{\theta \quad \mathbb{R}^d-$valued, predictable$\}$, $G_\cdot(\theta) = \int_0^\cdot \theta ds$ and $\mathcal{G} = G_T(\Theta)$, the possible payoffs at time T. We similarly define $\Theta_{\text{adm.}}$ and $\mathcal{G}_{\text{adm.}}$ where the condition on admissile strategies is again that there is a s.t. $G(\theta) \geq -a$ a.s.

Let $\mathcal{C} := \mathcal{C}^0 := \mathcal{G} - L_0^+$. These are the payoffs that we can dominate (super-replicate) starting with capital 0 at time 0 and following a self-financing strategy. Define $\mathcal{C}_{adm.}$ in the obvious way.

In this notation, classical no-arbitrage boils down to the following:

$$(\text{NA}) : \mathcal{G}_{\text{adm.}} \cap L_+^0 = \{0\} \tag{1.6}$$

Let X be a set. Then $\bar{X}^{L^0}$ is the closure of this set in L^0, i.e. with respect to convergence in probability.

Theorem 1.8. *In a finite discrete time financial market, 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* $\mathcal{C}^0 = \bar{\mathcal{C}^0}^{L^0}$, *i.e. closure in* L^0
6. $\bar{\mathcal{C}^0}^{L^0} \cap L^0_+ = \{0\}$
7. $\exists$ *EMM* Q *for* S *s.t.* $\frac{dQ}{dP} \in L^\infty$
8. $\exists$ *EMM* Q *for* S
9. $\exists$ *ElMM* Q *for* S

Corollary 1.9. *(Dalang-Morton-Willinger)*
In finite discrete time no-arbitrage is equivalent to the existence of an equivalent martingale measure (EMM) Q

Some remarks: $1 \Leftrightarrow 2$ as well as $3 \Leftrightarrow 4$ are elementary equivalences. $1 \implies 3$ uses an argument that only works in discrete time. $4 \implies 5$ is a major argument that only works in finite discrete time. The main point is the closedness of $\mathcal{C}^0$. $5 \implies 6$, $7 \implies 8$ and $8 \implies 9$ are clear. $9 \implies 1$ is a standard supermartingale argument. $6 \implies 7$ is the second major argument which relies on separation.

The first step is to realize that we can without loss of generality assume that P is such that $S_t \in L^1(P)$ for all t. Indeed, we can change any P to $P' \sim P$ with this property and it doesn't change the set of random variables L^0.

Define $\mathcal{C}^1 = \mathcal{C}^0 \cap L^1 \subseteq L^1$. It is a convex cone, $\mathcal{C}^1 \supseteq -L^0_+$ and $\mathcal{C}^1$ closed in L^1. We can use the Kreps-Yan theorem to get the existence of $Q \sim P$ s.t. $\frac{dQ}{dP} \in L^\infty$ and $E_Q[Y] \leq 0$ for all $Y \in \mathcal{C}^1$.

We now use $\theta^{(\pm)} := \pm \mathbf{1}_{A \times \{t\}} e^i$ with $A \in \mathcal{F}_{t-1}$, $e^i \in \mathbb{R}^d$.
$Y^{(\pm)} := G_T(\theta^{(\pm)})$ both in $\mathcal{C}^1$. Deduce that each S^i, and therefore S, is a Q-martingale.

For general models in continuous time, we will use similar, but technically more advanced, arguments. We mention some of the changes and challenges inherent in the adaptation to the continuous time framework:

- We do not work in L^1 (which depends on P) but in L^∞ (which is independent of P as long as we are allowed to take an equivalent measure.) The topology and closure will be different.

- We need stronger conditions than (NA) or $\mathcal{C}^0_{adm} \cap L^0_+ = \{0\}$, namely that $\overline{\mathcal{C}^0_{adm} \cap L^\infty}^{L^\infty} \cap L^0_+ = \{0\}$. This can be rephrased as the need to exclude not only direct money pumps but also their limits.

- Again show that no-arbitrage condition implies closedness of $\mathcal{C}^0_{adm} \cap L^\infty$ but for a different topology. Since the terminal payoff comes from a stochastic integral this is technically more involved.

- We will again use Kreps-Yan to get $Q \sim P$ (this time $\frac{dQ}{dP} \in L^1(P)$) with $E_Q[Y] \leq 0$ for all $Y \in \mathcal{C}^0_{adm} \cap L^\infty$. We call this an equivalent separating measure.

- We need extra work to show that $\exists Q' \sim P$ s.t. S is Q' σ-martingale. This is also technically difficult since it uses a semimartingale characterisation.

Chapter 2

Stochastic integration and semimartingales

2.1 Setup and definitions

The setting is a probability space $(\Omega, \mathcal{F}, P)$, time horizon $T \in (0, \infty)$ and $\mathbb{F} = (\mathcal{F}_t)_t$ satisfying the usual conditions. We choose RCLL versions for sub-/supermartingales. In general we will take S to be an integrator or semimartingale, H an integrand and X a generic process.

Definition 2.1. *X satisifies property (A) locally if $\exists$ stopping times (τ_n) with $\tau_n \nearrow T$ stationarily and s.t. $\forall n$ X^{τ_n} satisifies property (A).*

Stationarily $\tau_n \nearrow T$ means $\tau_n \nearrow T$ and $n \mapsto \tau_n(\omega)$ constant $= T$ for $n > n_0(\omega)$, both properties P-a.s.. Equivalently $\tau_n \nearrow T$ P-a.s. and $P(\tau_n = T) \to 1$ as $n \to \infty$. Hence, something like $\tau_n = T - 1/n$ is not allowed.

Definition 2.2. *A semimartingale S is an adapted, RCLL process s.t. $S = S_0 + M + A$ where M is a RCLL local martingale, A adapted RCLL finite variation (FV) trajectories. We have by convention $M_0 = A_0 = 0$.*

Recall that $H \in b\mathcal{E}$ means $H = \sum_{i=0}^{n} h_i \mathbf{1}_{[\tau_i, \tau_{i+1}]}$ (see chapter 1). For any adapted RCLL $\mathbb{R}^d$-valued process S define the map $I_S : b\mathcal{E} \to L^0$ by

$$
\begin{aligned}
H \mapsto \sum_{i=0}^{n} h_i(S_{\tau_{i+1}} - S_{\tau_i}) && (2.1)\\
=: \int_0^T H_u dS_u && (2.2)\\
= (H \bullet S)_T && (2.3)\\
= I_S(H) && (2.4)
\end{aligned}
$$

which is clearly a linear operator. The goal as such will be to extend this operator from the rather small space $b\mathcal{E}$ to something bigger.

Definition 2.3. *A good integrator S is adapted, RCLL, $\mathbb{R}^d$-valued and such that $I_S : (b\mathcal{E}, ||\cdot||_\infty) \to L^0$ is continuous.*

In L^0 we think of the topology of convergence in probability. Said otherwise, this definition means that for $H^n \to H$ uniformly (in (ω, t)) with $H^n, H \in b\mathcal{E}$, then $I_S(H^n) \to I_S(H)$ in probability.

If S is locally a good integrator, then S is a good integrator. Indeed, $\forall H$, τ_m and $\epsilon > 0$ we have

$$
\begin{aligned}
\{|I_S(H) - I_{S^{\tau_m}}(H)| > \epsilon\} &\subseteq \{I_S(H) \neq I_{S^{\tau_m}}(H)\} && (2.5)\\
&\subseteq \{\tau_m \neq T\} && (2.6)
\end{aligned}
$$

and we can then use the fact that $\tau_n \to T$ stationarily.

Interpretation: Continuity of I_S means robustness of trading outcomes under small portfolio changes.

The first main goal will be to show that semimartingales are good integrators (or actually the same thing). This is the Bircheller-Dellacherie theorem.

Definition 2.4. *Let X be integrable (i.e. $X_t \in L^1\ \forall t$). For a deterministic partition $\pi = \{0 = t_0 < ... < t_{n+1} = T\}$ define*

$$
MV(X, \pi) := \sum_{i=0}^{n} E[|E[X_{t_{i+1}} - X_{t_i} | \mathcal{F}_{t_i}]|] \qquad (2.7)
$$

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

Remark: By the triangle inequality $\pi \subseteq \pi'$ implies $MV(X,\pi) \leq MV(X,\pi')$ so that $\pi \mapsto MV(X,\pi)$ is increasing. Examples of quasimartingales are: Martingales, sub-/supermartingales (conditional expectation all have same sign, yielding a telescoping sum), processes of integrable variation, and any linear combination of the above classes of processes.

2.2 Good integrators and local quasimartingales

Theorem 2.5. *If S is a good integrator and bounded. Then $\forall \epsilon > 0$ there exists a stopping time $\rho \leq T$ s.t. $P(\rho = T) \geq 1 - 3\epsilon$ and S^ρ is a quasimartingale (S locally quasimartingale)*

The proof relies on an L^2 version of Komlos lemma:

Lemma 2.6.

1. *If $(g_n)_n$ be a bounded sequence in L^2, then there exists $h_n \in C_n := conv(g_n, g_{n+1}, \dots)$ such that $h_n \to h$ P-a.s. and in L^2, for some $h \in L^2$.*
2. *If $(f_n)_n$ is UI, then $h_n \in conv(f_n, f_{n+1}, \dots)$ such that $h_n \to h$ in L^1 for some $h \in L^1$.*

Proof. 1. By convexity and Jensen

$$\|h_n\|^2 \leq \sup_{m \geq n} \|g_m\|^2 \leq \sup_{m \in \mathbb{N}} \|g_m\|^2 < \infty \forall n, \forall h_n \in C_n \tag{2.8}$$

Set $A_n := \inf_{h \in C_n} \|h\|^2$ increasing in n since $C_n \searrow$ 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 every $n \in \mathbb{N}$, choose $h_n \in C_n$ s.t. $\|h_n\|^2 \leq A_n + \frac{1}{n} \leq A + \frac{1}{n}$.

For every $\epsilon > 0$, take n large enough s.t. $A_n \geq A - \epsilon$. Then for $k, m \geq n$, $h_k \in C_k \subseteq C_n$ and $h_m \in C_m \subseteq C_n$. Define $\tilde{h} := \frac{1}{2}(h_k + h_m) \in C_n$ so from above $\|\tilde{h}\|^2 = \frac{1}{4}\|h_k + h_m\|^2 \geq A_n \geq A - \epsilon$. This implies, for all $k, m \geq n$

$$\|h_k - h_m\|^2 = 2\|h_k\|^2 + 2\|h_m\|^2 - \|h_k + h_m\|^2 \tag{2.9}$$

$$\leq 4(A + \frac{1}{n}) - 4(A - \epsilon) = 4(\frac{1}{n} + \epsilon) \tag{2.10}$$

So $(h_n)_n$ is a Cauchy sequence in L^2. Hence it is convergent in L^2 to $h \in L^2$, so also P-a.s. covergent passing to a subsequence.

2. See Beiglböck, Schachermayer and Velinger 2012, using part 1 and a diagonalization argument. ☐

We move on to the theorem:

Proof. For any X RCLL bounded process we have $\lim_{t\searrow s} E[X_t - X_s|\mathcal{F}_s] = 0$.

$D_n := [0,T] \cap \frac{T}{2^n}\mathbb{N}$, the n-dyadic partition of $[0,T]$. With this we can approximate any non-random partition π. Exercise: $MV(X) = \lim_{n\to\infty} MV(X, D_n)$.

1. Since S is a good integrator, $\forall \delta > 0, \epsilon > 0$ there exists $\eta_0 > 0$ s.t. $\|H\|_\infty \leq \eta_0$ implies $P(|I_S(H)| \geq \delta) \leq \epsilon$. Fix $\delta > 0$ and set $C := \frac{\delta}{\eta_0} + 2\|S\|_\infty$.

Then for all $H \in b\mathcal{E}$ with $\|H\|_\infty \leq 1$, using above we have

$$
\begin{aligned}
& P(I_S(H) \geq C - 2\|S\|_\infty) && (2.11)\\
& \quad \leq P(|I_S(\eta_0 H)| \geq \delta) \leq \epsilon && (2.12)
\end{aligned}
$$

For $n \in \mathbb{N}$, define $H^n := \sum_{t_i \in D_n} \operatorname{sgn}\left(E[S_{t_{i+1}} - S_{t_i}|\mathcal{F}_{t_i}]\right) \mathbf{1}_{(t_i,t_{i+1}]}$. And $\rho_n := \inf\{t_i \in D_n : (H \bullet S)_{t_i} \geq C - 2\|S\|_\infty\} \wedge T$.

Since $\|H^n\|_\infty \leq 1$ we can use the bounds above to get on $\{\rho_n < T\}$ that $I_S(H^n \mathbf{1}_{(0,\rho_n]}) = (H^n \bullet S)_{\rho_n} \geq C - 2\|S\|_\infty$. By the inequality before we have $P(\rho_n = T) \geq 1 - \epsilon$. Note that jumps of S, as well as jumps of $H^n \bullet S$ (since $\|H^n\|_\infty \leq 1$) are bounded by $2\|S\|_\infty$, so that by the definition of ρ_n, we have $(H^n \bullet S)_{\rho_n} \leq C$.

By definition of H^n,

$$
\begin{aligned}
C &\geq E[(H \bullet S)_{\rho_n}] && (2.13)\\
&= \sum_{t_i \in D_n} E[\mathbf{1}_{t_i < \rho_n} |E[S_{t_{i+1}} - S_{t_i}|\mathcal{F}_{t_i}]|] && (2.14)\\
&=: MV(S^{\rho_n^+}, D_n) && (2.15)\\
&\geq MV(S^{\rho_n}, D_n) - 2\|S\|_\infty && (2.16)
\end{aligned}
$$

where the last inequality follows by arguing the following strings of inequalities

$$
\begin{aligned}
& MV(S^\tau, D_n) - MV(S^{\tau^+}, D_n) = && (2.17)\\
& \quad = \sum_{t_i \in D_n} E[|E[S^\tau_{t_i+1} - S^\tau_{t_i}|\mathcal{F}_{t_i}]| - \mathbf{1}(t_i < \tau)|E[S_{t_i+1} - S_{t_i}|\mathcal{F}_{t_i}]|] && (2.18)
\end{aligned}
$$

and

$$E[S^\tau_{t_i+1} - S^\tau_{t_i}|\mathcal{F}_{t_i}] = \mathbf{1}(t_i < \tau)\left(E[S_{t_i+1} - S_{t_i}|\mathcal{F}_{t_i}] + E[S_{t_{i+1}\wedge\tau} - S_{t_i+1}|\mathcal{F}_{t_i}]\right) \tag{2.19}$$

so $MV(S^\tau, D_n) - MV(S^{\tau^+}, D_n) \leq 2\|S\|_\infty$ as at most one term in the sum is nonzero.

2. We want to get rid of dependence of n in ρ^n. Use the previous lemma, part 1 for $g_n := \mathbf{1}(\rho_n = T)$ to get convex weights μ^n_j and $h_n := \sum_{j=1}^{N_n} \mu^n_j g_{n_j}$ where $n_j \geq n$ for all j. We have $h_n \to h$ P-a.s. and in L^2.

Clearly $0 \leq h_n \leq 1$ and the same holds for h. Also $E[g_n] = P(\rho_n = T) \geq 1-\epsilon$ so that by Lebesgue we have $E[h] \geq 1 - \epsilon$.

We show that $P(h < 2/3) < 3\epsilon$.

$$\begin{aligned} 1 - \epsilon \leq E[h] &< 2/3P(h < 2/3) + 1P(h > 2/3) && (2.20)\\ &= 1 - 1/3P(h < 2/3) && (2.21) \end{aligned}$$

so $1 - 3\epsilon < P(h \geq 2/3)$ and recalling that $h = \lim h_n$ we can use Egorov's theorem and by uniform convergence on a large set A we have $h_n \geq 1/2$ for all $n \geq n_0$ with $P(A) \geq 1 - 3\epsilon$.

Define $B^n_t := \sum_{j=1}^{N_n} \mu^n_j \mathbf{1}_{[0,\rho_{n_j}]}(t)$. Each B^n is a decreasing adapted left-continuous process with $B^n_T = h_n$ because $\mathbf{1}_{[0,\rho_{n_j}]}(T) = \mathbf{1}(\rho_{n_j} = T) = g_{n_j}$.

Define the stopping times

$$\begin{aligned} \sigma_n &:= \inf\{t \in [0,T] : B^n_t < 1/2\} \wedge T && (2.22)\\ \rho &:= \inf_{n \geq n_0} \sigma_n && (2.23) \end{aligned}$$

Then $B^n \geq 1/2$ on $[0, \sigma_n]$ so that for all $n \geq n_0$

$$\mathbf{1}_{[0,\rho]} \leq \mathbf{1}_{[0,\sigma_n]} \leq 2B^n = 2\sum_{j=1}^{N_n} \mu^n_j \mathbf{1}_{[0,\rho_{n_j}]} \tag{2.24}$$

Moreover, $\rho < T$ implies that for some $n \geq n_0$, $t < T$ we have $B^n_t < 1/2$ so that we are in A^C. Then $A \subseteq \{\rho = T\}$ and therefore $P(\rho = T) \geq 1 - 3\epsilon$.

Finally, using the estimate from the first step and $n_j \geq n$ as well as the fact that $\pi \mapsto MV(X^{\tau^+}, \pi)$ is increasing, we get

$$
\begin{aligned}
MV(S^\rho, D_n) - 2\|S\|_\infty &\le MV(S^{\rho^+}, D_n) && (2.25)\\
&= \sum_{t_i \in D_n} E[\mathbf{1}(t_i < \rho)|E[S_{t_{i+1} - S_{t_i}}|\mathcal{F}_{t_i}]|] && (2.26)\\
&\le 2 \sum_{t_i \in D_n} \sum_{j=1}^{N_n} \mu_j^n (E[\mathbf{1}(t_i < \rho_{n_j}|E[S_{t_{i+1}} - S_{t_i}|\mathcal{F}_{t_i}]| + 2\|S\|_\infty)]) && (2.27)\\
&= 2\sum_{j=1}^{N_n} \mu_j^n MV(S^{\rho^+_{n_j}}, D_n) + 4\|S\|_\infty && (2.28)\\
&\le 2C + 4\|S\|_\infty && (2.29)
\end{aligned}
$$

so we get $MV(S^\rho) = \lim MV(S^\rho, D_n) \le 2C + 6\|S\|_\infty < \infty$.

We conclude that S^ρ is a quasimartingale.

□

Theorem 2.7. *(Rao) Any quasimartingale can be written as a difference of two RCLL non-negative supermartingales.*

Proof. Consider D_n dyadics, define

$$
Y_s^n := E[\sum_{\substack{t_i \in D_n \\ t_i \ge s}} (E[X_{t_i} - X_{t_{i+1}}|\mathcal{F}_{t_i}])^+|\mathcal{F}_s] \tag{2.30}
$$

$$
Z_s^n := E[\sum_{\substack{t_i \in D_n \\ t_i \ge s}} (E[X_{t_i} - X_{t_{i+1}}|\mathcal{F}_{t_i}])^-|\mathcal{F}_s] \tag{2.31}
$$

Note that $Y^n, Z^n \ge 0$, $Y_s^n - Z_s^n = X_s - E[X_T|\mathcal{F}_s]$ for all $s \in D_n$.

Both Y^n and Z^n are supermartingales on D_n. Moreover, for every $u \le r \le v$, by Jensen

$$
\begin{aligned}
(E[X_v - X_u|\mathcal{F}_u])^\pm &= 2E\left[\frac{1}{2}E[X_v - X_r|\mathcal{F}_r] + \frac{1}{2}E[X_r - X_u|\mathcal{F}_u]\middle|\mathcal{F}_u\right]^\pm && (2.32)\\
&\le E[(E[X_v - X_r|\mathcal{F}_r])^\pm + (E[X_r - X_u|\mathcal{F}_u])^\pm|\mathcal{F}_u] && (2.33)
\end{aligned}
$$

Using this for $v = t_{i+1} \in D_n$, $u = t_i \in D_n$, $r = \frac{1}{2}(u + v) \in D_{n+1}$, we have that summing up and conditioning on $\mathcal{F}_s$

$$n \mapsto Y_s^n \tag{2.34}$$
$$n \mapsto Z_s^n \tag{2.35}$$

are both increasing P-a.s., hence we can take limits

$$Y_s := \lim Y_s^n + E[X_T^+|\mathcal{F}_s] \tag{2.36}$$
$$Z_s := \lim Z_s^n + E[X_T^-|\mathcal{F}_s] \tag{2.37}$$

both existing P-a.s., and because

$$E[Y_s^n + Z_s^n] \leq MV(X, D_n) \tag{2.38}$$
$$\leq MV(X) \tag{2.39}$$
$$< \infty \tag{2.40}$$

monotone integration yields $Y_s, Z_s \in L^1$ for all $s \in D_n$. Call $D = \cup_{n\in\mathbb{N}} D_n$. By Fatou, both Y and Z are ≥ 0, supermartingales on D and $X_s = Y_s - Z_s$ for all $s \in D$.

By a standard argument using right-continuity of $\mathbb{F}$, we can extend Y, Z to RCLL supermartingales ≥ 0 on $[0, T]$, see DM5 1-4.

□

2.3 Doob Meyer for supermartingales

Definition 2.8. *X adapted RCLL is said to be of class (D) if the family $\{X_\tau : \tau$ stopping time $[0, T] -$ valued$\}$ is UI.*

Theorem 2.9. *(Doob-Meyer) Any supermartingale X of class (D) has a unique decomposition $X = X_0 + M - A$ where M is a martingale, $M_0 = 0$, RCLL. A is predictable, $A_0 = 0$, increasing, integrable.*

Proof. Uniqueness: a.) FV local martingales which are continuous must be constant, see BMSC.

b.) Any local martingale which is predictable must be continuous by predictable stopping thm.:

1.) If L is a local martingale, τ predictable stopping time ($[\tau, \infty)$ predictable), then $E[\Delta L_\tau|\mathcal{F}_{\tau-}] = 0$.

2.) If X predictable, τ any stopping time, then ΔX_τ is $\mathcal{F}_{\tau^-}$-measurable.

1.)+2.) $L \in \mathcal{M}_{0,loc}$ predictable, then jump times σ of L are predictable, so $\Delta L_\sigma = E[\Delta L_\sigma|\mathcal{F}_{\sigma-}] = 0$, this yields b.)

From a.) and b.) we have uniqueness, since if $X - X_0 = M - A = \tilde{M} - \tilde{A}$. Hence $M - \tilde{M}$ as well as $A - \tilde{A}$ are constant and therefore zero.

Existence: We first look at dyadics D_n: For every n, since X is discrete-time supermartingale on D_n, we write classical Doob decomposition $X - X_0 = M^n - A^n$ on D_n with $A^n_{t_{i+1}} - A^n_{t_i} := E[X_{t_i} - X_{t_{i+1}}|\mathcal{F}_{t_i}] \geq 0$ and $M^n_{t_i} = X_{t_i} - X_0 + A^n_{t_i}$ for all $t_i \in D_n$.

We want to show that $(M^n_T)_n$ is UI, in order to apply part b.) of the lemma. W.l.o.g., by replacing X_t by $X_t - E[X_T|\mathcal{F}_t]$ we can assume $X \geq 0$ and $X_T = 0$.

Then $M^n_T = A^n_T - X_0$, so for any stopping time τ on D_n by discrete time stopping theorem, we have

$$X_\tau = M^n_\tau + X_0 - A^n_\tau = E[A^n_T|\mathcal{F}_\tau] - A^n_\tau \tag{2.41}$$

For $c > 0$, define stopping times along D_n

$$\tau_n(c) := \inf\{t_i \in D_n : A^n_{t_{i+1}} > c\} \wedge T \tag{2.42}$$

so that $A^n_{\tau_n(c)} \leq c$, thus

$$X_{\tau_n(c)} \geq E[A^n_T|\mathcal{F}_{\tau_n(c)}] - c \tag{2.43}$$

Using that $\{A^n_T > c\} = \{\tau_n(c) < T\}$ we have

$$\begin{aligned} E[A^n_T \mathbf{1}(A^n_T > c)] &= E[E[A^n_T|\mathcal{F}_{\tau_n(c)}]\mathbf{1}(\tau_n(c) < T)] && (2.44)\\ &\leq E[X_{\tau_n(c)}\mathbf{1}(\tau_n(c) < T)] + cP(\tau_n(c) < T) && (2.45)\end{aligned}$$

Note that $\{\tau_n(c) < T\} \subseteq \{\tau_n(c/2) < T\}$, so using

$$\begin{aligned} A^n_{\tau_n(c/2)} &\leq c/2 && (2.46)\\ \{A^n_T > c\} &= \{\tau_n(c) < T\} && (2.47)\end{aligned}$$

we get

$$E[X_{\tau_n(c/2)}\mathbf{1}(\tau_n(c/2) < T)] = E[(A^n_T - A_{\tau_n(c/2)})\mathbf{1}(\tau_n(c/2) < T)] \tag{2.48}$$

$$\geq E[(A^n_T - A^n_{\tau_n(c/2)})\mathbf{1}(\tau_n(c) < T)] \tag{2.49}$$

$$\geq \frac{c}{2}P(\tau_n(c) < T) \tag{2.50}$$

This gives

$$E[A^n_T\mathbf{1}(A^n_T > c)] \leq E[X_{\tau_n(c)}\mathbf{1}(\tau_n(c) < T)]+ \tag{2.51}$$

$$+\,2E[X_{\tau_n(c/2)}\mathbf{1}(\tau_n(c/2) < T)] \tag{2.52}$$

But X is of class (D) and

$$P(\tau_n(c) < T) = P(A^n_T > c) \tag{2.53}$$

$$\leq E[A^n_T]/c \tag{2.54}$$

$$= E[M^n_T + X_0]/c \tag{2.55}$$

$$= E[X_0]/c \to 0 \quad \text{as } c \to \infty \tag{2.56}$$

We have just shown that $(A^n_T)_n$ is UI, and so is $(M^n_T)_n$.

Using the latter fact we can use part b.) of the lemma to find convex weights λ^n_j, $j = 1, ..., N_n$, $n_j \geq n$, s.t. $L^n_T := \sum_{j=1}^{N_n} \lambda^n_j M^{n_j}_T \to M_T$ in L^1 with the limit also in L^1. Consider the associated RCLL martingales $L^n_t := E[L^n_T|\mathcal{F}_t]$, $M_t := E[M_T|\mathcal{F}_t]$ for all $t \in [0, T]$. Then by Jensen, we also have $L^n_t \to M_t$ in L^1 for all $t \in [0, T]$. We extend each A^n to a process on $[0, T]$ by piecewise constant LCRL interpolation along D_n and get $B^n := \sum_{j=1}^{N_n} \lambda^n_j A^{n_j}$.

Then $A := M - X + X_0$ is RCLL and for all $t \in D$,

$$B^n_t = L^n_t - X_t + X_0 \to M_t - X_t + X_0 \quad \text{in } L^1 \tag{2.57}$$

Along a subsequence we have P-a.s. convergence simultaneously for all $t \in D$, then the process A is P-a.s. increasing on D (as are all A^n and B^n processes).

By right-continuity, A has also increasing trajectories on the whole $[0, T]$ P-a.s.

We are left to show that A is predictable. But all the A^n and B^n are predictable, since they are adapted and left-continuous, so it's enough to show that

$$A_t(\omega) = \limsup_{n\to\infty} B_t^n(\omega) \quad \forall t \in [0,T] \quad P\text{-a.s.} \tag{2.58}$$

If $f_n, f : [0,T] \to \mathbb{R}$ increasing with f right-continuous and $f_n(t) \to f(t)$ for all $t \in D$, then

$$\limsup_n f_n(t) \le f(t) \quad \forall t \in [0,T] \tag{2.59}$$

$$\lim_n f_n(t) = f(t) \quad \text{if } f \text{ is cont. at } t \tag{2.60}$$

So the previous P-a.s. $\limsup$ can only fail at discontinuity points of A, and these can be exhausted by a sequence of stopping times because A is adapted RCLL: $\{\Delta A \neq 0\} \subseteq \cup_{m=1}^{\infty}[\tau_m]$. Hence we are left to show

$$\limsup_n B_\tau^n = A_\tau \quad P\text{-a.s. for all stopping times } \tau \tag{2.61}$$

From above we know $\limsup_n B_\tau^n \le A_\tau$ P-a.s. and

$$0 \le B_\tau^n \le B_T^n \to A_T \quad \text{in } L^1 \tag{2.62}$$

We now use that B^n are convex combination of the A^n and Fatou to get

$$\liminf_n E[A_\tau^n] \le \limsup_n E[B_\tau^n] \tag{2.63}$$

$$\le E[\limsup_n B_\tau^n] \tag{2.64}$$

$$\le E[A_\tau] \tag{2.65}$$

so the $\limsup$ for stopping times will follow if we show that $\lim_n E[A_\tau^n] = E[A_\tau]$.

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 LCRL construction of A^n.

Now we use that X is RC and of class (D)

$$E[A_\tau^n] = E[A_{\sigma_n}^n] \tag{2.66}$$

$$= E[M_{\sigma_n}^n - X_{\sigma_n} + X_0] \tag{2.67}$$

$$= E[X_0 - X_{\sigma_n}] \tag{2.68}$$

$$\to E[X_0 - X_\tau] \tag{2.69}$$

$$= E[M_\tau - X_\tau + X_0] \tag{2.70}$$

$$= E[A_\tau] \tag{2.71}$$

□

2.4 Bichteller-Dellacherie theorem

Theorem 2.10. *Every good integrator is a semimartingale.*

Proof. For S RCLL adapted, we can define

$$J^1_\cdot := \sum_{0 \le s \le \cdot} \Delta S_s \mathbf{1}(|\Delta S_s| > 1) \tag{2.72}$$

Exercise: J^1 is adapted, RCLL, FV, hence it is a good integrator and a semimartingale.

We next consider $S - J^1$ and to complete the proof we need to show that it is a semimartingale.

$S - J^1$ adapted RCLL and it has bounded jumps so in particular is locally bounded. Since S and J^1 are good integrators so is their difference. By a previous theorem $S - J^1$ is locally a quasimartingale. Then by another theorem it is locally the difference of two supermartingales.

We next show the following: Every supermartingale X is locally of class (D). Indeed, take $\tau_n = \inf\{t \in [0,T] : |X_t| > n\} \wedge T$. Then for all stopping times $\sigma \le T$ we have $|X^{\tau_n}_\sigma| \le n + |X_{\tau_n}| \in L^1$ since $|X|$ is a submartingale then $|X_{\tau_n}| \le E[|X_T||\mathcal{F}_{\tau_n}]$. So the claim is true since for all n X^{τ_n} is of class (D). Then by a previous theorem $S - J^1$ is locally a sum of a martingale and a difference of two predictable increasing FV processes. Then $S = (S - J^1) + J^1$ is a sum of a local semimartingale and a semimartingale, hence a semimartingale. □

Remark: In the above proof we showed that any locally bounded good integrator can be written as $S = S_0 + M + A$ with the latter FV, predictable. As in a previous theorem, one can prove that this decomposition is unique, and this is called the canonical decomposition (for general semimartingale S, we don't have predictability of A nor uniqueness.)

Notation: For any RCLL process Y, we write $Y^*_t = \sup_{s\in[0,t]} |Y_s|$, where $t \in [0,T]$.

Lemma 2.11. *For any martingale L, $H \in b\mathcal{E}$ and $c > 0$*

$$cP(|I_L(H)| \ge c) \le cP((H \bullet L)^*_T \ge c) \tag{2.73}$$
$$\le 34\|H\|_\infty \|L_T\|_{L^1} \tag{2.74}$$

Proof. 1. We first suppose $L \geq 0$ and define $Z := L \wedge c$ so that Z is a bounded supermartingale. By the previous theorem we can write $Z = M - A$ where $M_0 = Z_0$ and A is integrable, increasing, predictable and RCLL.

Then $M = Z + A \geq 0$ and $M_T^2 \leq 2(Z_T^2 + A_T^2) \leq 2c^2 + 2A_T^2$. We want to show that M is square intergable, so we consider

$$A_T^2 = 2\int_0^T A_{s-}dA_s + \sum_{0\leq s\leq t} (\Delta A_s)^2 \tag{2.75}$$

$$= \int_0^T A_{s-}dA_s + \int_0^T A_s dA_s \tag{2.76}$$

$$\leq 2\int_0^T A_s dA_s \tag{2.77}$$

In the same way, setting $A_\cdot^n = A_\cdot \wedge n$ we get

$$A_T A_T^n = \int_0^T A_{s-}dA_T^n + \int_0^T A_s^n dA_s \tag{2.78}$$

Putting things together and using $A = M - Z$ we and the bound on Z we get

$$\int_0^T A_s^n dA_s = \int_0^T (A_T - A_{s-})dA_s^n \tag{2.79}$$

$$= \int_0^T (M_T - M_{s-})dA_s^n - \int_0^T (Z_T - Z_{s-})dA_s^n \tag{2.80}$$

$$\leq \int_0^T (M_T - M_{s-})dA_s^n + cA_T^n \tag{2.81}$$

Since M is a nonzero martingale and A^n is increasing and predictable, then (by DM 5.5.7)

$$E[M_T A_T^n] = E[\int_0^T M_{s-}dA_s^n] \tag{2.82}$$

Therefore

$$E[\int_0^T A_s^n dA_s] \leq cE[A_T^n] \leq cE[A_T] \tag{2.83}$$

Since $A^n \nearrow A$, $\int A^n dA \nearrow \int AdA$ and so monotone integration leads to

$$E[A_T^2] \leq 2 \lim_{n\to\infty} E[\int_0^T A_s^n dA_s] \tag{2.84}$$
$$\leq 2cE[A_T] \leq \infty \tag{2.85}$$

So A is square integrable and so is M.

We can compute the following:

$$E[A_T] = E[M_T - Z_T] \leq E[M_T] \tag{2.86}$$
$$= E[M_0] \tag{2.87}$$
$$= E[Z_0] \leq E[L_0] = E[L_T] \tag{2.88}$$

$$E[M_T^2] \leq 2(E[Z_T^2] + E[A_T^2]) \tag{2.89}$$
$$\leq 2cE[L_T] + 2 \cdot 2cE[A_T] \tag{2.90}$$
$$\leq 6cE[L_T] \tag{2.91}$$

2. Now take $H \in b\mathcal{E}$ and w.l.o.g. $\|H\|_\infty = 1$. We see that $L = Z = L \wedge c$ on $\{L_T^* \leq c\}$, to get

$$P((H \bullet L)_T^* \geq c) \leq P(L_T^* > c) + P((H \bullet Z)_T^* \geq c) \tag{2.92}$$

By Doob's maximal inequality,

$$P(L_T^* > c) \leq \frac{1}{c} E[L_T] \tag{2.93}$$

Next $Z = M - A$ and $\|H\|_\infty = 1$ imply that $H \bullet Z \leq H \bullet M + A$ which is a submartingale. Now use Doob's maximal inequality for submartingales

$$P((H \bullet Z)_T^* \leq c) \leq P(((H \bullet M + A)_T^*)^2 \geq c^2) \tag{2.94}$$
$$\leq \frac{1}{c^2} E[(H \bullet M + A)_T^2] \tag{2.95}$$
$$\leq \frac{2}{c^2} E[(H \bullet M)_T^2 + A_T^2] \tag{2.96}$$

By Itô isometry, since M is square integrable

$$E[(H \bullet M)_T^2] \leq \|H\|_\infty E[M_T^2] \tag{2.97}$$

We plug in the above estimates and get

$$cP((H \bullet L)^*_T \geq c) \leq cP(L^*_T \geq c) + 2/c(E[M_T^2] + E[A_T^2]) \tag{2.98}$$
$$\leq 17E[L_T] \tag{2.99}$$

This concludes the proof in the case $L \geq 0$.

3. Take any martingale L. It can be written as the difference of two non-negative martingales (DM 4.33). For each of those to we have the bound above, so the extra factor 2 yields a final constant 34, also replacing $E[L_T]$ by $\|L_T\|_{L^1}$ and 1 by $\|H\|_\infty$.

□

We are now in a position to prove a converse to theorem 2.10.

Theorem 2.12. *Every semimartingale is a good integrator.*

Proof. Sums of GI and local GI are GIs, so it is sufficient to show that both martingales and FV processes are GIs.

For a martingale L and $H \in b\mathcal{E}$ by the previous lemma we have

$$P(|I_L(H)| \geq \delta) \leq \frac{34}{\delta}\|H\|_\infty \|L_T\|_{L^1} \to 0 \quad \text{as} \quad \|H\|_\infty \to 0 \tag{2.100}$$

For a FV process A

$$|I_A(H)| = |\int_0^T H_s dA_s| \leq \|H\|_\infty \int_0^T |dA_s| \to 0 \tag{2.101}$$

P-a.s. as $\|H\|_\infty \to 0$.

□

Remarks: 1. The proof of the previous theorem relies only on the estimate for $I_L(H)$ from the lemma 2.11. This could be proved by arguing in discrete time becuase $H \in b\mathcal{E}$ is piecewise constant. But later we will need the estimate for the middle term, and since $H \bullet L$ is not piecewise constant, one cannot estimate its supremum directly but only using discrete times.

2. If we had a square integrable martingale L, we could use Itô isometry to get

$$E[(I_L(H))^2] = \|(H \bullet L)_T\|^2_{L^2} \tag{2.102}$$
$$\leq \|H\|^2_\infty E[\langle L\rangle_T] \tag{2.103}$$
$$= \|H\|^2_\infty E[L_T^2] \tag{2.104}$$

and argue that $I_L(H) \to 0$ in L^2 as $\|H\|_\infty \to 0$.

Chapter 3

General stochastic integration

The goal is to extend stochastic integration from $b\mathcal{E}$ to a larger class of integrands. The setup is again a probability space, filtration satisfying the usual condition and a finite time horizon.

3.1 Some metrics, and spaces of integrands

Remark: For every random variable $Z \geq 0$ and any constant $c > 0$, we have

$$\begin{aligned} 1 \wedge Z &= (1 \wedge Z)\mathbf{1}(Z > c) + (1 \wedge Z)\mathbf{1}(Z \leq c) && (3.1) \\ &\geq (1 \wedge c)\mathbf{1}(Z > c) && (3.2) \end{aligned}$$

For $c = \delta \leq 1$, this gives $P(Z > \delta) \leq \frac{1}{\delta}E[1 \wedge Z]$ and for $c = m \geq 1$ $P(Z > m) \leq E[1 \wedge Z]$. For every $c > 0$, $E[1 \wedge Z] \leq P(Z > c) + c$.

Definition 3.1. *We introduce $\mathbb{L}$ and $\mathbb{D}$, the spaces of $\mathbb{R}^d$-valued adapted LCRL and RCLL processes respectively, and define*

$$\begin{aligned} d(X^1, X^2) &:= E[1 \wedge (X^1 - X^2)^*_T] && (3.3) \\ &= E[1 \wedge \sup_{0 \leq s \leq t} |X^1_s - X^2_s|] && (3.4) \end{aligned}$$

and we indentify X^1 and X^2 if $d(X^1, X^2) = 0$. Then d is a metric making both of the introduced spaces complete.

The metric d describes uniform convergence in probability. Recall that for random variables in L^0 we consider $d_{L^0}(X, Y) = E[1 \wedge |X - Y|]$ metrizing convergence in L^0.

If we add to $b\mathcal{E}$ all H of the form

$$H = h_0 \mathbf{1}_{[0]} \tag{3.5}$$

with $h_0 \in L^\infty(\mathcal{F}_0; \mathbb{R}^d)$ then the resulting space is denoted by $b\mathcal{E}_0$ and for $H = h_0\mathbf{1}_{[0]}$ we set $H \bullet X = h_0 X_0$

Definition 3.2. *We define two metrics*

$$d_E(X^1, X^2) := \sup_{\substack{H \in b\mathcal{E}_0 \\ \|H\|_\infty \leq 1}} E[1 \wedge |(H \bullet (X^1 - X^2))_T|] \tag{3.6}$$

$$d'_E(X^1, X^2) := \sup_{\substack{H \in b\mathcal{E}_0 \\ \|H\|_\infty \leq 1}} E[1 \wedge |(H \bullet (X^1 - X^2))^*_T|] \tag{3.7}$$

Note that $d_E \leq d'_E$.

Lemma 3.3. *If $(X^n)_n \subset \mathbb{D}$ satisifes $d_E(X^n, X) \to 0$, then also $d(X^n, X) \to 0$.*

Proof. For $Y \in \mathbb{D}$, $0 < \delta < 1$, set stopping time $\tau := \inf\{t \in [0, T] : |Y_t| \geq \delta\} \wedge T$. Note that on $\{\tau < T\}$, then $|Y_\tau| \geq \delta$ by right-continuity and

$$\{Y^*_T \geq \delta\} \subseteq \{|Y_\tau| \geq \delta\} \tag{3.8}$$

Take $H = \mathbf{1}_{[0,\tau]} \in b\mathcal{E}_0$ with $\|H\|_\infty \leq 1$ and $Y_\tau = (H \bullet Y)_T$. So we get

$$\begin{aligned} P(Y^*_T \geq \delta) \leq P(|Y_\tau| \geq \delta) &\leq \frac{1}{\delta} E[1 \wedge |Y_\tau|] && (3.9) \\ &= \frac{1}{\delta} E[1 \wedge |(H \bullet Y)_T|] && (3.10) \\ &\leq \frac{1}{\delta} d_E(Y, 0) && (3.11) \end{aligned}$$

Applying this to $Y := X^n - X$ and using $d(Y, 0) \leq P(Y^*_T \geq \delta) + \delta$ gives the result.

□

Lemma 3.4. *For $(X^n)_n \subset \mathbb{D}$, $X \in \mathbb{D}$ we have*

$$d'_E(X^n, X) \to 0 \tag{3.12}$$

$$\iff \tag{3.13}$$

$$(H^n \bullet (X^n - X))_T \to 0 \tag{3.14}$$

where the latter limit is in L^0 for every sequence $(H^n) \subseteq b\mathcal{E}_0$ with $\|H^n\|_\infty \le 1$.

Proof. For the direction "$\Longrightarrow$" we need to show that $E[1 \wedge |(H^n \bullet (X^n - X))_T|] \to 0$. This follows from $E[1 \wedge Z] \le P(Z \ge \delta) + \delta$ and for $\delta \in (0,1)$ and $Y = X^n - X$,

$$P(|(H^n \bullet Y)_T| \ge \delta) \le \frac{1}{\delta} E[1 \wedge |(H^n \bullet Y)_T|] \tag{3.15}$$

$$\le \frac{1}{\delta} d_E(Y, 0) \tag{3.16}$$

$$\le \frac{1}{\delta} d'_E(Y, 0) \tag{3.17}$$

For the direction $\Longleftarrow$ if $d'_E(X^n, X) \not\to 0$ then there exists a $\delta_0 > 0$ and a sequence $(H^n)_n \subseteq b\mathcal{E}_0$ with $\|H\|_\infty \le 1$ s.t.

$$P((H^n \bullet (X^n - X))^*_T > \delta_0) > \delta_0 \tag{3.18}$$

for all n.

Define $\tau_n := \inf\{t \in [0,T] : |(H^n \bullet (X^n - X))_T| > \delta_0\} \wedge T$ and $\tilde{H}^n := H^n \mathbf{1}_{[0,\tau_n]}$.

$\tilde{H}^n \in b\mathcal{E}_0$, $\|\tilde{H}^n\|_\infty \le 1$ and by construction $(\tilde{H}^n \bullet (X^n - X))_T$ does not converge to 0 in L^0.

□

Corollary 3.5. *For $(X^n)_n \subseteq \mathbb{D}$, $X \in \mathbb{D}$, we have $d'_E(X^n, X) \to 0 \iff d_E(X^n, X) \to 0$*

Proof. "$\Longrightarrow$" is clear since $d'_E(X, Y) \ge d_E(X, Y)$ for all X, Y.

"$\Longleftarrow$" as seen in proof of previous lemma 3.4, $\delta \in (0,1)$ and for every $H^n \in b\mathcal{E}_0$ with $\|H^n\|_\infty \le 1$ we have

$$P(|H^n \bullet Y_T| \ge \delta) \le \frac{1}{\delta} d_E(Y, 0) \tag{3.19}$$

Using this for $Y := X^n - X$ shows that $d_E(X^n, X) \to 0$ implies that $(H^n \bullet (X^n - X))_T \to 0$ in L^0 for every $(H^n)_n \subseteq b\mathcal{E}_0$ with $\|H^n\|_\infty \leq 1$ and so by previous lemma 3.4 we have $d'_E(X^n, X) \to 0$.

$\square$

Definition 3.6. *Denote by $\mathcal{S} \subseteq \mathbb{D}$ the space of $\mathbb{R}^d$-valued semimartingales.*

Theorem 3.7. *With each of the metrics d_E and d'_E, $\mathcal{S}$ is a complete topological vector space.*

Proof. $\mathcal{S}$ with d_E or d'_E is a metric space and a TVS because addition and multiplication (by scalars) are continuous (exercise). We will show completeness, that is every Cauchy sequence converges to a limit in the space.

Let $(X^n)_n \subseteq \mathcal{S}$ be Cauchy for d_E or d'_E, then it is also Cauchy for d by lemma 3.3 and by completeness of $(\mathbb{D}, d)$ we have $d(X^n, X) \to 0$.

It remains to show that $X \in \mathcal{S}$ and that $X^n \to X$ w.r.t. d_E and d'_E respectively.

We first show $d_E(X^n, X) \to 0$. From $d(X^n, X) \to 0$ we get that $X^n_\tau \to X_\tau$ in L^0 for all stopping times τ. Hence

$$H \bullet X^n_T \to H \bullet X_T \quad \text{in } L^0 \ \forall H \in b\mathcal{E}_0 \tag{3.20}$$

Then

$$\begin{aligned} E[1 \wedge |H \bullet X^m_T - H \bullet X_T|] &= \lim_{n\to\infty} E[1 \wedge |H \bullet X^m_T - H \bullet X^n_T|] \qquad (3.21) \\ &\leq \limsup_n d_E(X^m, X^n) \qquad (3.22) \end{aligned}$$

Now we take the supremum over $H \in b\mathcal{E}_0$ and $\|H\|_\infty \leq 1$ and get $d_E(X^m, X) \leq \limsup d_E(X^m, X^n)$. This goes to zero as $m \to \infty$ since $(X^n)_n$ is Cauchy for d_E.

We next need to show that $X \in \mathcal{S}$ which is equivalent to showing that X is a good integrator, i.e. that $I_X : (b\mathcal{E}, \|\cdot\|_\infty) \to L^0$ is continuous.

Exercise: This is equivalent to showing that $\mathcal{X}_{(1)} := \{I_X(H) = H \bullet X_T : H \in b\mathcal{E}, \|H\|_\infty \leq 1\}$ is bounded in L^0, which means

$$\lim_{m\to\infty} \sup_{Y\in\mathcal{X}_1} P(|Y| \geq m) = 0. \tag{3.23}$$

To prove this we use $P(Z \geq m) \leq E[1 \wedge Z]$ for all $m \geq 1$ to get

$$P(|I_X(H)| \geq 2m) \leq P(|I_{X^n}(H)| \geq m) + P(|H \bullet (X^n - X)_T| \geq m) \tag{3.24}$$
$$\leq P(|I_{X^n}(H)| \geq m) + d_E(X, X^n) \tag{3.25}$$

with the latter term less than ϵ for n large enough. For fixed n, $X^n \in \mathcal{S}$ so that $\mathcal{X}^n_{(1)}$ is bounded in L^0 and then the first term is less than ϵ for large enough m, uniformly in $H \in b\mathcal{E}$ with $\|H\|_\infty \leq 1$. So we obtain

$$\sup_{\substack{H\in b\mathcal{E}\\ \|H\|_\infty\leq 1}} P(|I_X(H)| \geq 2m) \to 0 \tag{3.26}$$

as $m \to \infty$, which means that $\mathcal{X}_{(1)}$ is bounded in L^0 or equivalently, that $X \in \mathcal{S}$.

$\square$

3.2 Extension of the integral to $\mathbb{L}$

We already saw that $(\mathbb{L}, d)$ is complete. Any $H \in \mathbb{L}$ is adapted left-continuous, thus locally bounded (exercise), and bounded adapted left-continuous process can be approximated uniformly (in (ω, t)) by processes in $b\mathcal{E}_0$ (this uses a left-cont. discretization in time of H).

So $b\mathcal{E}_0$ is dense in $\mathbb{L}$ w.r.t. d. This will allow us to extend the notion of stochastic integral: For every $S \in \mathcal{S}$, we will extend $I_S : b\mathcal{E} \to L^0$ to a map $I_S : \mathbb{L} \to L^0$ which is continuous w.r.t. d, meaning $(H^n)_n \subset \mathbb{L}$, $H \in \mathbb{L}$ with $d(H^n, H) \to 0$ implies $I_S(H^n) = H^n \bullet S_T \to H \bullet S_T = I_S(H)$ in L^0.

By setting $H \bullet S_t := I_S(H\mathbf{1}_{[0,t]})$ we could also define a stochastic process indexed by t. However, this does not give any regularity for the process $H \bullet S$. The next theorem is therefore much stronger.

Theorem 3.8. *For every semimartingale $S \in \mathcal{S}$ the map $I_S : b\mathcal{E} \to L^0$ can be extended to a continuous map $J_S : (\mathbb{L}, d) \to (\mathcal{S}, d'_E)$*

Remark: The theorem contains two statements.

- For $H \in \mathbb{L}$, the stochastic integral process $\int H dS = H \bullet S = J_S(H)$ is well-defined and a semimartingale

- If $(H^n)_n \subseteq \mathbb{L}$, $H \in \mathbb{L}$ satisfy $d(H^n, H) \to 0$, then $d'_E(H^n \bullet S, H \bullet S) \to 0$.

Proof. As $b\mathcal{E}_0$ is dense in $\mathbb{L}$ for d, any $H \in \mathbb{L}$ admits $(H^n)_n \subseteq b\mathcal{E}_0$ s.t. $d(H^n, H) \to 0$. Then $(H^n)_n$ is Cauchy for d, and $X^n := H^n \bullet S = J_S(H^n)$ is in $\mathcal{S}$.

If we show that $(X^n)_n$ is Cauchy for d'_E, then from a previous theorem we have $d'_E(X^n, X) \to 0$ for some $X \in \mathcal{S}$ and we set $J_S(H) := X = H \bullet S = \int H dS$. Continuity of J_S for d and d'_E follow by how we defined J_S.

So we are left to show that $(X^n)_n$ is Cauchy for d'_E. If $S = A$ FV, this follows from standard results of Lebesgue-Stieltjes integration (we can argue for each ω). So suppose $S = M$ and by localization furthermore a martingale (exercise: prove for martingale, then argue using localizing sequence to extend).

It is enough to show if $(H^n)_n \subseteq b\mathcal{E}$ satisifes $d(H^n, 0) \to 0$ then we have $d'_E(H^n \bullet M, 0) \to 0$ (from this follows: $(H^n)_n$ Cauchy for d implies that $(H^n \bullet M)_n$ Cauchy for d'_E)

Take $K \in b\mathcal{E}_0$ with $\|K\|_\infty \leq 1$ and write $H^n = H^n \mathbf{1}((H^n)^* \leq b) + H^n \mathbf{1}((H^n)^* > b)$, name them $H^{n'}$ and $H^{n''}$ and note that both terms are in $b\mathcal{E}_0$.

Then $K \bullet (H^n \bullet M) = (KH^n) \bullet M$ and by lemma 2.11

$$P((K \bullet (H^n \bullet M))^*_T \geq c) \leq P((H^n)^*_T > b) + P(((K(H^n)') \bullet M)^*_T \geq c) \tag{3.27}$$

with the latter term $\leq \frac{34}{c} \|H^{n'}\|_\infty \|M_T\|_1$.

Then

$$d'_E(H^n \bullet M, 0) = \sup_{\substack{K \in b\mathcal{E}_0 \\ \|K\|_\infty \leq 1}} E[1 \wedge (K \bullet (H^n \bullet M))^*_T] \tag{3.28}$$

$$\leq c + P((H^n)^*_T > b) + \frac{34b}{c} \|M_T\|_{L^1} \tag{3.29}$$

Since $d(H^n, 0) \to 0$ we have that the second term goes to zero as n goes to infinity for any $b > 0$, and the first and third terms can be made small for c small and then b small. So $d'_E(H^n \bullet M, 0) \to 0$ as $n \to \infty$.

□

3.3 Extension of the integral to $b\mathcal{P}$ (w.r.t. martingale)

Remark: To extend the class of integrands beyond $\mathbb{L}$, we need a different approximation, thus a new concept. For a semimartingale S, the process $S_- \in \mathbb{L}$

so $\int S_- dS$ is well defined. Moreover, approximating S_- in $(\mathbb{L}, d)$ by

$$\sum S_{\tau_i} \mathbf{1}_{]\tau_i, \tau_{i+1}]} \tag{3.30}$$

for $(\tau_i)_{i=0,\dots,I^n}$ a sequence of increasing stopping times gives

$$\int S_- dS = \lim_{n\to\infty} \sum_{i=0}^{I^n-1} S_{\tau_i}(S^{\tau_{i+1}} - S^{\tau_i}) \tag{3.31}$$

where convergence is w.r.t. $d'_E \geq d$. hence uniformly on compact time intervals in probability (ucp).

Definition 3.9. *For every semimartingale S, we define the optional quadratic variation of S*

$$[S]_t = S_t^2 - S_0^2 - 2\int_0^t S_- dS \tag{3.32}$$

which is an adapted RCLL process, null at zero.

Lemma 3.10. $[S]$ *is P-a.s. increasing*

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

$$\int S_- dS = \lim_n \sum_{t_i \in D_n} S_{t_i}(S^{t_{i+1}} - S^{t_i}) \tag{3.33}$$

Compute

$$S^2 - S_0^2 - 2\sum_{t_i} S_{t_i}(S^{t_{i+1}} - S^{t_i}) = \sum_{t_i} \left((S^{t_{i+1}})^2 - (S^{t_i})^2 - 2S_{t_i}(S^{t_{i+1}} - S^{t_i})\right) \tag{3.34}$$

$$= \sum_{t_i} (S^{t_{i+1}} - S^{t_i})^2 =: V^n \tag{3.35}$$

From this we see that $[S] = \lim_{n\to\infty} V^n$ ucp and since each V^n is increasing on D_n, the limit $[S]$ is increasing on $D = \bigcup_n D_n$ $(D_n \subseteq D_{n+1})$, dyadics are dense in $[0, T]$ and $[S]$ is RCLL, so $[S]$ must also be increasing on $[0, T]$.

□

Definition 3.11. $\mathcal{H}_0^1 := \{M \in \mathcal{M}_{0,loc} : M_T^* = \sup_{t\in[0,T]} |M_t| \in L^1\}$. *We identify M and M' if their difference has the same norm $\|M\|_{\mathcal{H}_0^1} := \|M_T^*\|_{L^1}$, and so $\mathcal{H}_0^1$ becomes a Banach space.*

Note: $M \in \mathcal{H}_0^1$ are UI martingales. Also $\mathcal{H}_{0,loc}^1 = \mathcal{M}_{0,loc}$, that is, every local martingale null at zero is locally in $\mathcal{H}_0^1$. Indeed, let $(\sigma_n)_n$ be a localizing sequence for $M \in M_{0,loc}$, then $\forall n$, $M^{\sigma_n \wedge n}$ is closed on the right so UI. If M^τ is UI, $M_\tau \in L^1$ by stopping theorem. Take $\tau_n := \inf\{t \in [0,T] : |M_t| > n\} \wedge T \wedge \sigma_n$. Note the following (left as an exercise):

$$(M^{\tau_n})_T^* \leq n + |\Delta M_{\tau_n}| \tag{3.36}$$

$$\leq 2n + |M_{\tau_n}| \in L^1 \tag{3.37}$$

Theorem 3.12. *(Davis' inequality) $\exists$ constants $0 < c < C < \infty$ s.t. $\forall M \in \mathcal{M}_{0,loc}$*

$$cE[([M]_T)^{\frac{1}{2}}] \leq E[M_T^*] \tag{3.38}$$

$$\leq CE[([M]_T)^{\frac{1}{2}}] \tag{3.39}$$

Proof. 1.) We first show that if we have the result for discrete time martingales, then we can show the general result by passing to the limit in the following way. Both M^* and $[M]$ are increasing, so consider the stopped processes $M^{\tau_n} \in \mathcal{H}_0^1$. Next note that $\Delta[M] = (\Delta M)^2$ which gives $[M]^{1/2}$ locally integrable being equivalent to M locally integrable. W.l.o.g. all terms in the statement of the theorem are finite by localization. So after localization, M is a martingale in $\mathcal{H}_0^1$ and $Y^n := \sum_{t_i}(M^{t_{i+1}} - M^{t_i})^2 \to [M]$ ucp along a subsequence uniformly on $[0,T]$, P-a.s.

So $Y := \sup_{n\in\mathbb{N}} Y^n$ RCLL with $\Delta Y_t \leq \sup_{n\in\mathbb{N}} \Delta Y_t^n \leq 4(M_T^*)^2 \in L^{1/2}$. Hence $\sup_{0\leq t\leq T}(\Delta Y_t)^{1/2} \in L^1$.

Along D_n (since we are assuming that the result holds in discrete time), we have

$$cE[(Y_T^n)^{1/2}] \leq E[\sup_{t_i\in D_n} |M_{t_i}|] \tag{3.40}$$

$$\leq CE[(Y_T^n)^{1/2}] \tag{3.41}$$

Taking the limit $n \to \infty$ and using dominated convergence on the first and third terms, monotone convergence on the second, we get to the statement of the theorem.

2.) For discrete time martingales there is an elementary proof in Beiglböck, Siorpaes (2015). They define

$$X_n^* := \max_{k=0,\dots,n} |X_k| \tag{3.42}$$

$$[X]_n := X_0^2 + \sum_{k=1}^{n} (X_n - X_{n-1})^2 \tag{3.43}$$

$$(h \bullet X)_n := \sum_{k=1}^{n} h_{k-1}(X_k - X_{k-1}) \tag{3.44}$$

Take $h_k = \frac{X_n}{([X]_k + (X_k^*)^2)^{1/2}}$ bounded and take it from there...

□

Corollary 3.13. *If $M \in \mathcal{M}_{0,loc}$, then $[M]^{1/2}$ is locally integrable.*

Definition 3.14. *For $M \in \mathcal{M}_{0,loc}$, $\mathcal{L}^1(M)$ is the family of predictable H s.t.*

$$\|H\|_{\mathcal{L}^1(M)} := E\left[\left(\int_0^T H_s^2 d[M]_s\right)^{1/2}\right] < \infty \tag{3.45}$$

Identifying H, H' if their difference has norm zero, we get the space $L^1(M) = \mathcal{L}^1(M)/N$.

Remarks:

1. Above we do this for $d = 1$. For larger d, $[M]$ is a process valued in positive semidefinite $d \times d$ matrices and

$$\|H\|_{L^1(M)} = E\left[\left(\int_0^T H_s^{tr} d[M]_s H_s\right)^{1/2}\right] \tag{3.46}$$

2. For $M \in \mathcal{H}_0^1$, $[M]^{1/2}$ is integrable and $L^1(M)$ contains family $b\mathcal{P}$ of bounded predictable processes.

3. For A FV, define

$$L_{var}(A) = \{H \quad \text{predictable}, \int_0^T |H_s||dA_s| < \infty \quad P-\text{a.s.}\} \tag{3.47}$$

Note that $b\mathcal{P}_{loc} \subseteq L_{var}(A)$

Lemma 3.15. *If $M \in \mathcal{M}_{0,loc}$, then $b\mathcal{E} \cap L^1(M)$ is dense in $L^1(M)$ for $\|\cdot\|_{L^1(M)}$*

Proof. By previous corollary 3.13 we can take $\tau_m \nearrow T$ stationarily s.t. $([M]^{\tau_m})^{1/2}$ is integrable.

If $H \in L^1(M)$, define $H^m := H\mathbf{1}_{[0,\tau_m]} \to H$ in $L^1(M)$ with each term of the sequence being in the corresponding space $L^1(M^{\tau_m})$.

Note also that if $\left(\int_0^T H^2 d[M]\right)^{1/2} \in L^1$, then

$$E\left[\left(\int_0^T H^2 d[M]\right)^{1/2} \mathbf{1}(\tau_m < T)\right] \to 0 \quad \text{as } m \to \infty \tag{3.48}$$

So w.l.o.g. we may assume $M \in \mathcal{H}_0^1$ and $H \in L^1(M)$ and show the lemma in this case. So define

$$H^k := H\mathbf{1}(|H| \leq k) \in b\mathcal{P} \tag{3.49}$$

$$H^k \to H \in L^1(M) \tag{3.50}$$

by dominated convergence. Therefore it is enough to show approximation of $H \in b\mathcal{P} \subseteq L^1(M)$. For such H we can find $H^j = \sum_k \lambda_k^j \mathbf{1}_{D_j^n}$ (with the lambdas in the real numbers and the indicator process predictable), s.t. $\|H^j - H\|_\infty \to 0$ as $j \to \infty$ (implies $H^j \to H$ in $L^1(M)$) and $\|H^j\|_\infty \leq \|H\|_\infty$. Note however that $H^j \notin b\mathcal{E}$ but any H of that form, can be approximated in $L^1(M)$ by a sequence $(H^n)_n \subseteq b\mathcal{E}$ with $\|H^n\|_\infty \leq \|H\|_\infty = |\lambda|$ by martingale convergence theorem (exercise).

□

By the previous lemma and completeness of $\mathcal{H}_0^1$, we can extend, for a local martingale M the stochastic integral to $H \in L^1_{loc}(M)$ and in particular to $b\mathcal{P}_{loc}$. Here we go for a stronger result, stronger continuity and less dependence on M.

Lemma 3.16. *Let $(M^n)_{n\in\mathbb{N}}$, Y a martingale s.t. $|\Delta M^n_\tau| \leq |\Delta Y_\tau|$ for all n and for all stopping times τ. If $E[1 \wedge [M^n]_T] \to 0$ (this is the same as L^0 convergence), then $d'_E(M^n, 0) \to 0$.*

Proof. By lemma 3.4, it is enough to show that for all $(H^n)_n \subseteq b\mathcal{E}_0$ with $\|H^n\|_\infty \leq 1$, setting $X^n := H^n \bullet M^n$, $X^n_T \to 0$ in L^0.

Note that $[X^n]_T \leq [M^n]_T \to 0$ in L^0, so along a subsequence $P([X^{n_k}]_T \geq 2^{-k}) \leq 2^{-k}$ for all k. By Borel-Cantelli, $A_\cdot := \sum_{k=1}^\infty [X^{n_k}]_\cdot$ is finite valued P-a.s. and $[X^{n_k}]_T \to 0$ P-a.s. as $k \to \infty$.

Define $\tau_m := \inf\{t \in [0,T] : A_t \geq m \text{ or } |Y_t| \geq m\} \wedge T$. So for all k

$$[X^{n_k}]_{\tau_m} \leq A_{\tau_m -} + (\Delta X^{n_k}_{\tau_m})^2 \tag{3.51}$$
$$\leq m + (m + |Y_{\tau_m}|)^2 \tag{3.52}$$

Hence, $\sup_{k\in\mathbb{N}}[X^{n_k}]^{1/2}_{\tau_m} \in L^1$. By theorem 3.12 and Lebesgue's theorem

$$E[(X^{n_k})^*_{\tau_m}] \leq CE[[X^{n_k}]^{1/2}_{\tau_m}] \to 0 \tag{3.53}$$
$$P(\tau_m = T) \to 1 \tag{3.54}$$

the first limit w.r.t k and the second w.r.t. m. Then $(X^{n_k})^*_T = (H^{n_k} \bullet M^{n_k})^*_T \to 0$ in L^0. Therefore every subsequence of $(H^n \bullet M^n)_n$ has a further subsequence which converges uniformly (in t) in L^0 to zero. Hence $(H^n \bullet M^n)_n$ itself is converging to zero uniformly in L^0. By lemma 3.4 $d'_E(M^n, 0) \to 0$.

$\square$

Corollary 3.17. *For M a martingale, $(H^n)_n \subseteq b\mathcal{E}$ s.t. $\|H^n\|_\infty \leq 1$, if $H^n \to 0$ pointwise, then $d'_E(H^n \bullet M, 0) \to 0$.*

Proof. Set $M^n := H^n \bullet M$, $Y := M$, then $|\Delta M^n_\tau| = |H^n_\tau \Delta M_\tau| \leq |\Delta M_\tau|$. Moreover, $H^n \to 0$ pointwise and $[M]_T < \infty$ P-a.s. So by Lebesgue $[M^n]_T = \int_0^T (H^n_s)^2 d[M]_s \to 0$ P-a.s. By lemma 3.16 we are done. $\square$

Theorem 3.18. *For any martingale $M \in \mathcal{H}^1_0$, the map $I_M : (b\mathcal{E}, \|\cdot\|_{L^\infty}) \to (L^0, d_{L^0})$ admits a unique linear extension to $J_M : b\mathcal{P} \to \mathcal{S}$ s.t. $J_M(H) =: H \bullet M$ satisfies $[H \bullet M] = \int H^2 d[M]$ and $\Delta(H \bullet M) = H\Delta M$ and having dominated convergence property that $(H^k)_k \subseteq b\mathcal{P}$ with $\|H^k\|_\infty \leq 1$, $H^k \to 0$ pointwise then $d'_E(H^k \bullet M, 0) \to 0$. Moreover, $J_M : (b\mathcal{P}, \|\cdot\|_{L^\infty}) \to (\mathcal{S}, d'_E)$ is continuous.*

Proof. We want to define $H \bullet M$ for $H \in b\mathcal{P} \subseteq L^1(M)$ so we use the proof of lemma 3.15 to get $(H^n)_n \subseteq b\mathcal{E}$ with $\|H^n\|_\infty \leq \|H\|_\infty$ and $H^n \to H$ in $L^1(M)$. Then $H^n - H^m \to 0$ as $n, m \to \infty$, thus $H^n - H^{m_n} \to 0$ as $n \to \infty$ for all subsequences $m_n \geq n$. So $d'_E((H^n - H^{m_n}) \bullet M, 0) \to 0$ by corollary 3.17 Then $d'_E((H^n - H^m) \bullet M, 0) \to 0$ as $n, m \to \infty$, that is $(H^n \bullet M)_n$ is Cauchy in $(\mathcal{S}, d'_E)$.

By completeness (thm 3.7), we can define the limit

$$H \bullet M := \lim_{n\to\infty}^{d'_E} H^n \bullet M \tag{3.55}$$

Uniqueness and linearity are clear. Also, we know that $[H^n \bullet M] = \int (H^n)^2 d[M]$ and $\Delta(H^n \bullet M) = H^n \Delta M$ for all n since $H^n \in b\mathcal{E}$ and then those properties follow for H by taking limits.

The dominated convergence property follows by corollary 3.17 and continuity for $(b\mathcal{P}, \|\cdot\|_\infty)$ is then clear. $\square$

3.4 Extension w.r.t. semimartingales, the Émery metric and a continuity result

Remark: Here we see that we can easily define $J_S : b\mathcal{P}_{loc} \to \mathcal{S}$ and that is is continuous w.r.t. the stated metrics. A clear extension of this theorem from $M \in \mathcal{H}_0^1$ to semimartingales $S \in \mathcal{S}$. Indeed, for a FV process A and $H \in b\mathcal{P}$, we can define $H \bullet A = \int H_s dA_s$ pathwise (classical Lebesgue-Stieltjes integrals satisfy properties in this theorem). Then we only need to argue for $M \in \mathcal{M}_{0,loc} = \mathcal{H}^1_{0,loc}$, so we use this theorem on each M^{τ_n}, where τ_n is a localizing sequence ($M^{\tau_n} \in \mathcal{H}_0^1$). With one more localizing sequence, we can define integrals for $H \in b\mathcal{P}_{loc}$.

$$H \bullet S = H \bullet M + H \bullet A \tag{3.56}$$

for $S = M + A$. Define the two following metrics on $\mathcal{S}$:

$$\tilde{d}_E(S^1, S^2) := \sup_{\substack{H \in b\mathcal{P}_0 \\ \|H\|_\infty \leq 1}} E[1 \wedge (H \bullet (S^1 - S^2))_T] \tag{3.57}$$

$$\tilde{d}'_E(S^1, S^2) := \sup_{\substack{H \in b\mathcal{P}_0 \\ \|H\|_\infty \leq 1}} E[1 \wedge (H \bullet (S^1 - S^2))^*_T] \tag{3.58}$$

$$\geq \max(\tilde{d}_E(S^1, S^2), d'_E(S^1, S^2)) \tag{3.59}$$

with the latter called the Émery metric (generating the Émery topology).

Then arguing exactly as in the case without tildes (corollary 3.5), $\tilde{d}_E$ can be shown to be equivalent to $\tilde{d}'_E$. Moreover, we can show that also d_E and $\tilde{d}_E$ are equivalent. Indeed, since $b\mathcal{E}_0 \subseteq b\mathcal{P}_0$, $d_E \leq \tilde{d}_E$. For the reverse, take any $H \in b\mathcal{P}_0$ with $\|H\|_\infty \leq 1$ and $(H^n)_n \subseteq b\mathcal{E}_0$ with $\|H^n\|_\infty \leq 1$ and $H^n \to H$ pointwise. This gives

$$
\begin{aligned}
E[1 \wedge (H \bullet (S^1 - S^2))_T] &= d_{L^0}(H \bullet S^1, H \bullet S^2) && (3.60)\\
&\leq d_{L^0}(H \bullet S^1_T, H^n \bullet S^1_T) + d_{L^0}(H^n \bullet S^1_T, H^n \bullet S^2_T) && (3.61)\\
&+ d_{L^0}(H^n \bullet S^2_T, H \bullet S^2_T) && (3.62)\\
&\leq d'_E((H - H^n) \bullet S^1, 0) + d'_E(S^1, S^2) && (3.63)\\
&+ d'_E((H - H^n) \bullet S^2, 0) && (3.64)
\end{aligned}
$$

By theorem 3.18 and its extension to semimartingales, the first and third terms of the RHS go to zero as $n \to \infty$ and so taking the supremum over all $H \in b\mathcal{P}_0$, $\|H\|_\infty \leq 1$ gives $\tilde{d}_E \leq d'_E$. This means $d_E, d'_E, \tilde{d}_E, \tilde{d}'_E$ are all equivalent.

Theorem 3.19. *For every semimartingale S and every $H \in b\mathcal{P}_{loc}$, the stochastic integral $H \bullet S$ is well-defined and a semimartingale. Moreover, the mapping*

$$
(b\mathcal{P} \times \mathcal{S}, \|\cdot\|_\infty \times \tilde{d}'_E) \to (\mathcal{S}, \tilde{d}'_E) \tag{3.65}
$$

is continuous.

Proof. Take $(H^n)_n \subseteq b\mathcal{P}$ with $\|H^n - H\|_\infty \to 0$ and take $(S^n)_n \subseteq \mathcal{S}$ with $\tilde{d}'_E(S^n, S) \to 0$. Then

$$
\begin{aligned}
\tilde{d}'_E(H^n \bullet S^n, H \bullet S) &\leq \tilde{d}'_E(H^n \bullet S^n, H^n \bullet S) + \tilde{d}'_E(H^n \bullet S, H \bullet S) && (3.66)\\
&= \sup_{\substack{k \in b\mathcal{P}_0 \\ \|k\|_\infty \leq 1}} E[1 \wedge ((kH^n) \bullet (S^n - S))^*_T] && (3.67)\\
&+ \tilde{d}'_E((H^n - H) \bullet S, 0) && (3.68)
\end{aligned}
$$

where the first term goes to zero due to $\tilde{d}'_E(S^n, S) \to 0$ and the fact that $\|H^n\|_\infty \leq b$ for all n. And the second term goes to zero by the dominated convergence property (theorem 3.18 and its extension). □

In theorem 3.18 we defined $H \bullet M$ for a martingale $M \in \mathcal{H}^1_0$ and $H \in b\mathcal{P}$ and showed that $H \bullet M \in \mathcal{S}$. There are examples where $H \bullet M$ is not a martingale (see e.g. Herdegen/Hermann), but they are always local martingales:

Theorem 3.20. *If $M \in \mathcal{H}^1_0$, $H \in b\mathcal{P}$, then $H \bullet M \in \mathcal{H}^1_0$.*

Proof. Using proof of lemma 3.15, we find $(H^n)_n \subseteq b\mathcal{E}_0$ s.t. $\|H^n\|_\infty \leq \|H\|_\infty$ and $H^n \to H$ in $L^1(M)$. Then $H^n \bullet M \to H \bullet M$ for d'_E (by proof of thm. 3.18). By theorem 3.12, $[M]^{1/2}_T \in L^1$ and so by Lebesgue theorem

$$E\left[\left(\int_0^T (H_s^n - H_s)^2 d[M]_s\right)^{1/2}\right] \to 0 \tag{3.69}$$

so again by thm. 3.12, $(H^n \bullet M)_n$ is Cauchy in $\mathcal{H}_0^1$ so its limit $H \bullet M$ is also in $\mathcal{H}_0^1$. $\square$

3.5 The space $\mathcal{L}(S)$

Definition 3.21. *For a semimatingale S, a predictable process H is S-integrable (denoted $H \in \mathcal{L}(S)$) if $(H^n \bullet S)_n$, $H^n = H\mathbf{1}(|H| \leq n) \in b\mathcal{P}$, is Cauchy in Émery topology (i.e. for $\tilde{d}'_E$). The limit in $\mathcal{S}$ is denoted by $H \bullet S$.*

Remark: If $H, H' \in \mathcal{L}(S)$ with $|H| \leq |H'|$ then $\tilde{d}'_E(H \bullet S, 0) \leq \tilde{d}'_E(H' \bullet S, 0)$. Indeed, $\{H \neq 0\} \subseteq \{H' \neq 0\}$ implies that for all $J \in b\mathcal{P}$ with $\|J\|_\infty \leq 1$:

$$\begin{aligned} J \bullet (H \bullet S) &= (JH) \bullet S && (3.70)\\ &= (JH\mathbf{1}(H \neq 0)) \bullet S && (3.71)\\ &= \left(J \bullet \frac{H}{H'}\mathbf{1}(H \neq 0)H'\right) \bullet S && (3.72)\\ &= K \bullet (H' \bullet S) && (3.73) \end{aligned}$$

where $K = J \bullet \frac{H}{H'}\mathbf{1}(H \neq 0) \in b\mathcal{P}$ with $\|K\|_\infty \leq 1$. Now the definition of $\tilde{d}'_E$ gives the result.

Theorem 3.22. *For $S \in \mathcal{S}$, $H \in \mathcal{P}$, then $H \in \mathcal{L}(S)$ iff $\tilde{d}'_E(K^n \bullet S, 0) \to 0$ for all sequences $(K^n)_n \subseteq b\mathcal{P}$ with $|K^n| \leq |H|$ and $K^n \to 0$ pointwise.*

Proof. " $\implies$ " $H^n := H\mathbf{1}(|H| \leq n)$, by definition $(H^n \bullet S)_n$ Cauchy for $\tilde{d}'_E$ and $\tilde{d}'_E((H\mathbf{1}(|H| > n)) \bullet S, 0) = \tilde{d}'_E(H^n \bullet S, H \bullet S) \to 0$. Then $\tilde{d}'_E((K^n\mathbf{1}(|H| > n)) \bullet S, 0) \to 0$ by the remark. For any $m \leq n$ we have

$$\begin{aligned} (K^n\mathbf{1}(|H| \leq n)) \bullet S &= (K^n\mathbf{1}(|H| \leq m)) \bullet S && (3.74)\\ &\quad + (K^n\mathbf{1}(m \leq |H| \leq n)) \bullet S && (3.75) \end{aligned}$$

For fixed m, the first term on the right hand side goes to zero for $\tilde{d}'_E$ by dominated convergence.

Since $|K^n| \leq |H|$, then

$$\tilde{d}'_E((K^n \mathbf{1}(m \le |H| \le n)) \bullet S, 0) \le \tilde{d}'_E((H \mathbf{1}(m \le |H| \le n)) \bullet S, 0) \tag{3.76}$$
$$= \tilde{d}'_E(H^m \bullet S, H^n \bullet S) \tag{3.77}$$
$$\to 0 \quad \text{as } n, m \to \infty \tag{3.78}$$

by the Cauchy property.

So we can make the second term small by taking n, m large, then also first term can be small by taking $n \ge m$ so $\tilde{d}'_E(K^n \bullet S, 0) \to 0$.

"$\Longleftarrow$": For any $m_n \ge n$, define $K^n := H^{m_n} - H^n = H\mathbf{1}(n < |H| \le m_n)$ so that $|K^n| \le |H|$ and $K^n \to 0$ pointwise. Hence $\tilde{d}'_E(K^n \bullet S, 0) \to 0$. But $m_n \ge n$ was arbitrary, so $(H^n \bullet S)_n$ Cauchy in $\tilde{d}_E$, i.e. $H \in \mathcal{L}(S)$.

□

Remark: Some of the above arguments are done for $d = 1$. If S and H are $\mathbb{R}^d$-valued, then $H \bullet S$ is $\mathbb{R}$-valued but $[S]$ is $\mathbb{R}^{d \times d}$-valued with entries $[S^i, S^j] = \frac{1}{4}([S^i + S^j] - [S^i - S^j])$ quadratic covariation and $[H \bullet S] = \int H^{tr} d[S] H$.

Moreover $H \bullet S$ can be well-defined even if the individual $H^i \bullet S^i$ are not, so $H \bullet S \ne \sum H^i \bullet S^i$ in general. Vector stochastic integration is needed (see Cherny/Shiryaev 2002).

3.6 Closedness of stochastic integrals

We later need closedness results for spaces of stochastic integrals. For this, fix a semimartingale S and define $d_S(H, H') := \tilde{d}'_E(H \bullet S, H' \bullet S)$. We identify H, H' if $d_S(H, H') = 0$ and write $L(S) = \mathcal{L}(S)/\sim_{d_S}$.

Lemma 3.23. *Suppose $(\gamma_k)_k \subseteq b\mathcal{P}$ satisfying $\sum d_S(\gamma_k, 0) < \infty$ and define $[0, \infty]$-valued process $G := \sum_{k=1}^{\infty} |\gamma_k|$. Then*

1. *$\forall t \in [0, T]$, $\{\int_0^t H_u dS_u, H \in b\mathcal{P}, |H| \le G\}$ is bounded in L^0.*
2. *If $(K^n)_n \subseteq b\mathcal{P}$ satisfies $|K^n| \le G$ and $K^n \to 0$ pointwise, then $\tilde{d}'_E(K^n \bullet S, 0) \to 0$*
3. *$\int H\mathbf{1}(G = \infty) dS = 0$ for all $H \in b\mathcal{P}$.*

Proof. Take $K \in b\mathcal{P}$ with $|K| \le G$, and set

$$G^m := \sum_{k=1}^{m} |\gamma_k| \nearrow G \tag{3.79}$$

$$K^m := (K \wedge G^m) \vee (-G^m) \tag{3.80}$$

Then $|K^m| \leq |K| \in b\mathcal{P}$ and $K^m \to K$ pointwise so $d_S(\lambda K^m, 0) \to d_S(\lambda K, 0)$ for all $\lambda \geq 0$ by theorem 3.19.

Moreover,

$$|K^m| \leq \sum_{k=1}^{m} (|K| \wedge |\gamma^k|) \tag{3.81}$$

so for all $\lambda \geq 0$ fixed

$$d_S(\lambda K, 0) = \lim_{m\to\infty} d_S(\lambda K^m, 0) \tag{3.82}$$

$$\leq \lim_{m\to\infty} \sum_{k=1}^{m} d_S(\lambda(|K| \wedge |\gamma_k|), 0) \tag{3.83}$$

$$= \sum_{k=1}^{\infty} d_S(\lambda(|K| \wedge |\gamma_k|), 0) \tag{3.84}$$

Now take $(\lambda_n)_n \subseteq [0,1]$ and $(K^n)_n \in b\mathcal{P}$ and get analogous estimates for $d_S(\lambda_n K^n, 0)$. If $\lambda_n \to 0$ or $K^n \to 0$ pointwise, then by theorem 3.19 we have for all k that

$$\lim_{n\to\infty} d_S(\lambda_n(|K^n| \wedge |\gamma_k|), 0) = 0 \tag{3.85}$$

Moreover,

$$d_S(\lambda_n(|K^n| \wedge |\gamma_k|), 0) \leq d_S(\gamma_k, 0) \tag{3.86}$$

which is summable over k by assumption. So we can use Lebesgue for sums to get

$$\lim_{n\to\infty} d_S(\lambda_n K^n, 0) \leq \lim_{n\to\infty} \sum_{k=1}^{\infty} d_S(\lambda_n(|K^n| \wedge |\gamma_k|), 0) \tag{3.87}$$

$$= \sum_{k=1}^{\infty} \lim_{n\to\infty} d_S(\lambda_n(|K^n| \wedge |\gamma_k|)) = 0 \tag{3.88}$$

whenever $\lambda_n \to 0$ or $K^n \to 0$ pointwise.

1.) Take any $(K^n)_n \subseteq b\mathcal{P}$ with $|K^n| \le G$ and $\lambda_n \to 0$. Then the above implies

$$\lambda_n \int_0^t K_u^n dS_u \to 0 \tag{3.89}$$

in L^0 for all t and boundedness in L^0 easily follows (exercise).

2.) Since $d_S(H,0) = \tilde{d}'_E(H \bullet S, 0)$, then this follows from above for $\lambda_n = 1$.

3.) Since G predictable, for all $H \in b\mathcal{P}$, $H\mathbf{1}(G = \infty) \in b\mathcal{P}$.

For $c_n \to \infty$, define $H^n := c_n H \mathbf{1}(G = \infty)$. Then $|H^n| \le G$. This implies that $\left(c_n \int_0^t \mathbf{1}(G_u = \infty) H_u dS_u\right)_{n \in \mathbb{N}}$ is bounded in L^0 by point 1.).

But as $c_n \to \infty$, this can only be true if $\int_0^t \mathbf{1}(G_u = \infty) H_u dS_u = 0$, P-a.s.. And since t is arbitrary, we are done.

□

Theorem 3.24. *(Memin)*
If S $\mathbb{R}^d$-valued semimartingale, then $(L(S), d_S)$ is a complete metric space. Equivalently, the space

$$\{X = \int H dS : H \in L(S)\} \tag{3.90}$$

is closed in the Émery topology (that is for $\tilde{d}'_E$). This is the space of all stochastic integral processes of S.

Proof. The equivalence follows by definition of d_S and by completeness of $(\mathcal{S}, \tilde{d}'_E)$.

Take $(\tilde{H}^n)_n \subseteq L(S)$ which is Cauchy for d_S and approximate each $\tilde{H}^n$ in d_S by $H^n \in b\mathcal{P}$ (existence by definition of d_S and $L(S)$), and choose a subsequence, still called $(H^n)_n$ with $d_S(H^n, H^{n-1}) \le 2^{-n}$ (using Cauchy property).

This will be enough to show that $(H^n \bullet S)$ converges for $\tilde{d}'_E$ to some $H \bullet S$ with $H \in L(S)$. Define $\gamma_n := H^{n+1} - H^n$ and $G := \sum_{n=1}^\infty |\gamma_n|$. For any $|K^n| \le G\mathbf{1}(G < \infty)$ and $K^n \to 0$, use lemma 3.23(2) to get $\tilde{d}'_E(K^n \bullet S, 0) \to 0$. Hence $G\mathbf{1}(G < \infty) \in L(S)$ by thm. 3.22.

On $\{G < \infty\}$, define

$$\bar{H} := \sum_{n=1}^{\infty} \gamma_n = \sum_{n=1}^{\infty} (H^{n+1} - H^n) \tag{3.91}$$

$\bar{H}\mathbf{1}(G < \infty)$ is predictable and $|\bar{H}\mathbf{1}(G < \infty)| \leq G$.

By using lemma 3.23(2) and then thm. 3.22 we obtain $\bar{H}\mathbf{1}(G < \infty) \in L(S)$ and so

$$H := (H^1 + \bar{H})\mathbf{1}_{G<\infty} \in L(S) \tag{3.92}$$

(here $H^1 \in b\mathcal{P} \subseteq L(S)$)

Note that

$$H^1 + \bar{H} - H^n = \sum_{k=n}^{\infty} \gamma_k \tag{3.93}$$

This implies

$$\begin{aligned} d_S(H, H^n\mathbf{1}(G < \infty)) &= d_S((H^1 + \bar{H})\mathbf{1}(G < \infty), H^n\mathbf{1}(G < \infty)) && (3.94) \\ &= d_S((H^1 + \bar{H} - H^n)\mathbf{1}(G < \infty), 0) && (3.95) \\ &\leq \sum_{k=n}^{\infty} d_S(\gamma_k, 0) \to 0 && (3.96) \end{aligned}$$

As $H^n \in b\mathcal{P}$, by lemma 3.23(3) $\int H^n\mathbf{1}(G < \infty) dS = H^n \bullet S$ so

$$\tilde{d}'_E(H \bullet S, H^n \bullet S) = d_S(H, H^n\mathbf{1}(G < \infty)) \to 0 \tag{3.97}$$

□

Remark: We have seen in Introduction to Mathematical Finance (ETH Spring semester course, finite discrete time) that

$$G_T(\Theta) = \left\{ \int_0^T \theta_u dS_u : \theta \ \mathbb{R}^d - \text{valued, predictable} \right\} \tag{3.98}$$

is always closed in L^0. One can ask if $\{(H \bullet S)_T : H \in L(S)\}$ is also closed in L^0. This is not true in general. The reason is that if we have that the final values $(H^n \bullet S)_T$ are Cauchy in L^0 (so converging to some bounded random variable), we cannot get any control over the whole process $H^n \bullet S$.

Chapter 4

The Fundamental Theorem of Asset Pricing (FTAP)

Goal: Precise formulation of the equivalence between "absence of arbitrage" and "existence of an equivalent martingale measure (EMM)" in general continuous time models.

Setup: $(\Omega, \mathcal{F}, \mathbb{F}, P)$ with $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$, $S^0 = B = 1$, S adapted RCLL $\mathbb{R}^d$-valued. We assume that S is a semimartingale (equivalently, it is a good integrator). Recall $L(S)$ the space of predictable S-integrable processes θ and $G(\theta) = \int \theta dS = \theta \bullet S$.

We call θ a-admissible, $a \geq 0$ if $G(\theta) \geq -a$.

$$\Theta^a = \text{space of a-admissible } \theta \in L(S) \tag{4.1}$$
$$\Theta_{adm} := \bigcup_{a\geq 0} \Theta^a \tag{4.2}$$
$$\mathcal{G}^a := G_T(\Theta^a) = \left\{ \int_0^T \theta_u dS_u : \theta \in \Theta^a \right\} \tag{4.3}$$
$$\mathcal{G}_{adm} := G_T(\Theta_{adm}) = \bigcup_{a\geq 0} \mathcal{G}^a \tag{4.4}$$

Definition 4.1. *No arbitrage (NA):*

$$\mathcal{G}_{adm} \cap L^0_+ = \{0\} \tag{4.5}$$

Recall: $Q \sim P$ is ELMM for S if $S \in \mathcal{M}_{loc}(Q)$.

We have already seen lemma 1.7, here restated and given the new number 4.2:

Lemma 4.2. *If S satisfies that the space of ELMMs for S is nonempty, then S satsifies no arbitrage.*

4.1 Ansel-Stricker

As previously mentioned, we will need some preliminary results.

Theorem 4.3. *(Ansel-Stricker) If M is a local martingale and $\theta \in L(M)$ s.t. $\int \theta dM \geq -b$ for some $b \geq 0$, then $G(\theta) = \int \theta dM$ is a local martingale (and a supermatingale).*

Lemma 4.4. *(de Donno-Pratelli) Let X be an RCLL process, $(\gamma_n)_n \subseteq L^1$, localizing sequence $(\tau_n)_n$ s.t. $X^{\tau_n} \geq \gamma_n$ for all n and let $(M^n)_n$ be a sequence of martingales s.t. $d(M^n, X) \to 0$ and $(\Delta M^n_\sigma)^\pm \leq (\Delta X_\sigma)^\pm$ for all n and all stopping times σ. Then X is a local martingale.*

Proof. Assume $M^n_0 = X_0 = 0$. Define the stopping times

$$\rho_n := \inf\{t : X_t > n \text{ or } M^n_t > X_t + 1 \text{ or } M^n_t < X_t - 1\} \wedge T \tag{4.6}$$

and compute

$$\begin{aligned} P(\inf_{n\geq m} \rho_n = T) &= 1 - P(\bigcup_{n\geq m} \{\rho_n < T\}) && (4.7)\\ &\geq 1 - \sum_{n\geq m} P(\rho_n < T) \to_{m\to\infty} 1 && (4.8) \end{aligned}$$

as $m \to \infty$. Here we used the Borel-Cantelli lemma, as indeed, by pairing to a subsequence, we may assume that $\sum_{n=1}^\infty P(\rho_n < T) < \infty$. (Recall that by assumption, $d(M^n, X) \to 0$ so that $M^n \to X$ uniformly on $[0, T]$ in L^0). Then

$$\sigma_m := \tau_m \wedge \inf_{n\geq m} \rho_n \nearrow T \text{ stationarily} \tag{4.9}$$

We are going to show that X^{σ_m} is a martingale. Note that

$$\begin{aligned} (\Delta M^n_{t\wedge\sigma_m})^- &\leq (\Delta X_{t\wedge\sigma_m})^- && (4.10)\\ &\leq (m - \gamma_m)\mathbf{1}_{\Delta X_{t\wedge\sigma_m} \leq 0} && (4.11) \end{aligned}$$

Moreover for $n \geq m$ and $t < \sigma_m$ we have $M^n_{t\wedge\sigma_m} = M^n_t \geq X_t - 1 \geq \gamma_m - 1$. Therefore, putting the two above inequalities together $M^n_{t\wedge\sigma_m} \geq \gamma_m - 1 - (m - \gamma_m)\mathbf{1}_\Delta \in L^1$.

This allows us to use Fatou and implies $X_{t\wedge\sigma_m} \in L^1$, since $M^n_{t\wedge\sigma_m} \to X_{t\ \wedge\sigma_m}$ in L^0. Also $\Delta M_{t\wedge\sigma_m} \in L^1$ because $X_{\sigma_m -} \leq m$ and $\geq \gamma_m \in L^1$.

Analogously,

$$M^n_{t\wedge\sigma_m} \leq m + 1 + (\Delta M^n_{\sigma_m})^+ \tag{4.12}$$

$$\leq m + 1 + |\Delta X_{t\wedge\sigma_m}| \in L^1 \tag{4.13}$$

so we can use dominated convergence for $n \to \infty$:

$$M^n_{t\wedge\sigma_m} \to X_{t\wedge\sigma_m} \text{ in } L^0 \tag{4.14}$$

also converges in L^1. Since M^n are all martingales then so is X^{σ^m}.

□

We now prove the prop./theorem:

Proof. We know that $\mathcal{M}_{loc} = \mathcal{H}^1_{loc}$, so by stopping, assume that $M \in \mathcal{H}^1$.

Take $\theta \in L(M)$ and define $\theta^n := \theta\mathbf{1}(|\theta| \leq n)$ and $M^n := \int \theta^n dM$. Then $M^n \to \int \theta dM =: X$ for the Émery metric $\tilde{d}'_E$ (by thm. 3.19) and each $M^n \in \mathcal{H}^1$ by thm 3.20. Since $X \geq -b$, all assumptions of lemma 4.4 are satisfied and hence $\int \theta dM = X \in \mathcal{M}_{loc}$. Then it is also a supermartingale by Fatou. □

4.2 Existence of ELMM gives no arbitrage

We are now ready for the proof of the very first lemma, sketched already in the very first chapter:

Proof. By theorems 2.10 and 2.12, a semimartingale is the same as a good integrator, and so a good integrator under P is a good integrator under any Q equivalent to P, $Q \sim P$. So if $S \in \mathcal{M}_{loc}(Q)$, then S is a P-semimartingale and $L(S)$ makes sense.

If $\theta \in \Theta_{adm}$, then $\int \theta dS \geq$ const. is a local Q-martingale and a Q-supermartingale by thm 4.3. Then $E_Q[G_T(\theta)] \leq 0$. So if $G_T(\theta) \in L^0_+$, it would need to be the zero random variable, and we are done.

□

4.3 NFLVR and NUBPR

We saw in the first lemma 1.7 (or 4.2) that the existence of an ELMM implies (NA). In order to prove some form of converse, we need a stronger version of (NA).

Recall:

- $\mathcal{C}^\infty := \big(G_T(\Theta_{adm}) - L^0_+\big) \cap L^\infty =: \mathcal{C}$ is the set of bounded payoffs at time T that can be superreplicated starting with 0 wealth at time zero and following a self-financing strategy.
- (NA): $G_T(\Theta_{adm}) \cap L^0_+ = \{0\}$ or equivalently (exercise) $\mathcal{C}^\infty \cap L^\infty_+ = \{0\}$.

Definition 4.5. *(NFLVR) A semimartingale S satisfies no free lunch with vanishing risk if*

$$\overline{\mathcal{C}^\infty}^{L^\infty} \cap L^\infty_+ = \{0\} \tag{4.15}$$

With this condition we strenghten the NA concept. The next proposition shows how much stronger this concept is. This gap will be covered by the following:

Definition 4.6. *No unbounded profit with unbounded risk (NUPBR): $\mathcal{G}^1 = G_T(\Theta^1)$ is bounded in L^0 (i.e. $\sup_{g\in\mathcal{G}^1} P(|g| \geq n) \to 0$).*

Theorem 4.7. *For a semimartingale S, the following are equivalent*

1. *S satisfies (NFLVR)*
2. *for any sequence $g_n = G_T(\theta^n)$ in $\mathcal{G}_{adm}$ with $G_T^-(\theta^n) \to 0$ in L^∞, then $g_n \to 0$ in L^0 (that is, if losses go to zero uniformly, then gains go to zero in probability)*
3. *S satisfies (NA) + (NUPBR)*

In order to show the proposition we will use the following result in the spirit om Komlós.

Lemma 4.8. *For all sequences $(X_n)_n \subseteq L^0_+$, there exist convex combinations $\tilde{X}_n \in conv(X_n, X_{n+1}, ...)$ such that $\tilde{X}_n \to \tilde{X}_\infty$ P-a.s. random variable with values in $[0,\infty]$. Moreover, if $P(X_n \geq \alpha) \geq \delta > 0$ for some*

$\alpha > 0$, then $P(\tilde{X}_\infty > 0) > 0$. If $conv(X_1, X_2, ...)$ is bounded in L^0, then $\tilde{X}_\infty < \infty$ P-a.s.

Proof. See appendix B. □

For thm. 4.7:

Proof. "1.) $\implies$ 2.)": Take (g_n) as in 2.) and suppose by contradiction that $g_n \not\to 0$ in L^0. Then there is a subsequence which we still call $(g_n)_n$ s.t. $P(g_n \geq \alpha) \geq \delta > 0$ for some $\alpha > 0$.

Then define

$$f_n := g_n \wedge 1 = g_n - (g_n - g_n \wedge 1) \in (\mathcal{G}_{adm} - L^0_+) \cap L^\infty = \mathcal{C}^\infty \tag{4.16}$$

Since $f_n^- = g_n^- \to 0$ in L^∞ by assumption, we can assume that $f_n \geq -a$ for some $a > 0$. Moreover, for α above (which w.l.o.g. can be taken < 1) $P(f_n \geq \alpha) = P(g_n \geq \alpha) \geq \delta > 0$. By lemma 4.8, there are

$$\tilde{f}_n \in \text{conv}(f_n, f_{n+1}, ...) \text{ s.t.} \tag{4.17}$$

$$\tilde{f}_n \to \tilde{f}_\infty \quad P\text{-a.s.} \tag{4.18}$$

$$\tilde{f}_\infty \geq -a \tag{4.19}$$

$$P(\tilde{f}_\infty > 0) =: \beta > 0 \tag{4.20}$$

Since $f_n^- \to 0$ in L^∞ we can take $a > 0$ as small as wanted, hence $\tilde{f}_\infty \geq 0$. This gives $\tilde{f}_\infty \in L^\infty_+ \backslash \{0\}$. Since $\mathcal{C}^\infty$ is convex, $(\tilde{f}_n)_n$ is still a sequence in $\mathcal{C}^\infty$.

by Egorov's theorem, P-a.s. converging implies converging in $\mathcal{C}^\infty$ on subset of measure as close as wanted to 1.

So there is $B \in \mathcal{F}$ with $P(B) \geq 1 - \frac{\beta}{2}$ and $\tilde{f}_n \mathbf{1}_B$ in L^∞. Note

$$\tilde{f}_n \mathbf{1}_B - \tilde{f}_n^- \mathbf{1}_{B^C} = \tilde{f}_n - \tilde{f}_n^+ \mathbf{1}_{B^C} \tag{4.21}$$

is a sequence in $\mathcal{C}^\infty$ which converges in L^∞ to $\tilde{f}_\infty \mathbf{1}_B$. So $\tilde{f}_\infty \mathbf{1}_B \in \overline{\mathcal{C}}^{L^\infty} \cap L^\infty_+$, but

$$P(\tilde{f}_\infty \mathbf{1}_B > 0) \geq P(\tilde{f}_\infty > 0) - P(B^C) \tag{4.22}$$

$$\geq \beta - \beta/2 = \beta/2 > 0 \tag{4.23}$$

which contradicts (NFLVR), so 1.) $\implies$ 2.).

"2.) $\implies$ 3.)": Any $f \in \mathcal{C}^\infty \cap L^\infty_+$ has the form $0 \leq f = g - Y$ with $Y \geq 0$, hence $g \geq 0$.

Now use 2.) with $g_n = g$. This gives $g = 0$ which means (NA) holds.

If (NUPBR) fails, this means that $\mathcal{G}^1$ is not bounded in L^0. Then there are $\gamma_n \nearrow \infty$ with $P(g_n \geq \gamma_n) \geq \delta > 0$ for a sequence $(g_n)_n \subset \mathcal{G}^1$. But $g_n = G_T(\theta^n)$ with θ^n 1-admissible, which means $\tilde{g}_n := \frac{1}{\gamma_n} g_n \in \mathcal{G}_{adm}$ and $\|\tilde{g}_n^-\|_\infty \leq \frac{1}{\gamma_n} \to 0$.

But we also have $P(\tilde{g}_n \geq 1) \geq \delta > 0$, so that $(\tilde{g}_n)_n$ does not go to zero in L^0 and therefore 2.) is violated.

"3.) $\implies$ 1.)": Suppose (NA) holds but (NFLVR) does not. We then show that (NUPBR) fails.

Since (NFLVR) fails, there exists $(f_n)_n \subseteq \mathcal{C}^\infty$ and $f \in L^\infty_+ \backslash \{0\}$ s.t. $\|f_n - f\|_\infty \leq \frac{1}{n}$ for all n. But $f_n \leq g_n$ for some $g_n \in \mathcal{G}_{adm}$, so $\|g_n^-\|_\infty \leq \|f_n^-\|_\infty \leq 1/n$, and due to (NA), $g_n = G_T(\theta^n)$ with $g_n \geq -1/n$ implies that $G_\cdot(\theta^n) \geq -1/n$ (exercise), so that $n\theta^n$ is 1-admissible.

Using lemma 4.8 gives $\tilde{g}_n \in \operatorname{conv}(g_n, g_{n+1}, ...)$ s.t. $\tilde{g}_n \to \tilde{g}_\infty$ P-a.s. and $g_n \geq f_n$ yields $\tilde{g}_n \geq f$ so that $P(\tilde{g}_\infty > 0) > 0$.

However, $(n\tilde{g}_n)_n \subseteq \mathcal{G}^1$ is not bounded in L^0, i.e. (NUPBR) does not hold.

□

4.4 σ-martingales, EσMMs and ESMs

Definition 4.9. *An adapted RCLL $\mathbb{R}^d$-valued process X is a σ-martingale if*

$$X = X_0 + \int \psi dM \tag{4.24}$$

for M an $\mathbb{R}^d$-valued local P-martingale and $\psi \in L(M)$ a one-dimensional integrand (meaning $\psi \in L(M^i)$ for all $i = 1, ..., d$), $\psi > 0$.

Definition 4.10. *(EσMM) An Equivalent σ-martingale measure for S is a probability measure $Q \sim P$ on $\mathcal{F}_T$ s.t. S is a Q-σ-martingale.*

Definition 4.11. *(ESM) An equivalent separating measure for S is a probability measure $Q \sim P$ on $\mathcal{F}_T$ with $E_Q[G_T(\theta)] \leq 0$ for all $\theta \in \Theta_{adm}$.*

Remarks:

1. A martingale is a local martingale which is a σ-martingale, by taking $\psi = 1$, but the converses are not true in general (more on this below).

2. X a σ-martingale and $X - X_0 \geq$ const. implies that X is a local martingale by thm. 4.3 (Ansel-Stricker). This is useful for us for $S \geq 0$ (and S_0 nonrandom). More generally: A σ-martingale which is locally bounded below (e.g. continuous) is a local martingale.

3. If S is a Q-σ-martingale and $\theta \in \Theta_{adm}$, then $G(\theta) = \int \theta dS = \int \psi\theta dM^{(Q)} \geq -a$ is a local Q-martingale and Q-supermartingale by thm. 4.3. Then :

$$EMM \implies ELMM \implies E\sigma MM \implies ESM \tag{4.25}$$

Conversely, if S is (locally) bounded, then $ESM \implies E(L)MM$ (exercise).

4.5 Example: A σ-martingale that is not a local martingale

(σ-martingale $\not\Rightarrow$ local martingale)

$\tau \sim \exp(1)$ and $Z \perp \tau$ with values ± 1 with probability 1/2 each. Define $M_t := Z\mathbf{1}_{t \geq \tau}$ for $t \geq 0$, $\mathbb{F} = \mathbb{F}^M$. Then $M_t - M_s$ for $s < t$ is not zero iff $s < \tau \leq t$. Then since $\tau \perp Z$:

$$E[(M_t - M_s)h(M_u; u \leq s)] = \text{const.} E[Z]P(s < \tau \leq t) = 0 \tag{4.26}$$

Noting that for $\tau > s \geq u$, $M_u = Z\mathbf{1}(u \geq \tau) = 0$ so that $\mathcal{F}_s^M \cap \{\tau > s\} = \{A \cap \{\tau > s\} : A \in \mathcal{F}_s^\tau\}$ is trivial, so $h(M_u; u \leq s)$ constant on $\{\tau > s\}$.

So M is a martingale, and of course FV. Define $\psi_t := \frac{1}{t}$, clearly predictable and > 0 and in $L(M)$ (M constant up to a single jump at time τ).

So for $t > 0$, $\int_0^t \psi_u dM_u = \psi_\tau \Delta M_\tau \mathbf{1}(\tau \leq t)$.

For any stopping time σ w.r.t. $\mathbb{F}^M$, σ must be constant on $\{\sigma < \tau\}$ because $M = 0$ before τ and if $\sigma \not\equiv 0$ then $\sigma \geq \tau$ on $\{\tau \leq \epsilon\}$ for some $\epsilon > 0$. So

$$\left|\int_0^\sigma \psi_u dM_u\right| = \frac{|Z|}{\tau}\mathbf{1}(\tau \le \sigma) \tag{4.27}$$

$$= \frac{1}{\tau}\mathbf{1}(\tau \le \sigma) \tag{4.28}$$

$$\ge \frac{1}{\tau}\mathbf{1}(\tau \le \epsilon) \notin L^1 \tag{4.29}$$

because

$$E\left[\frac{1}{\tau}\mathbf{1}(\tau \le \epsilon)\right] = \int_0^\epsilon \frac{1}{u}e^{-u}du = \infty \tag{4.30}$$

This shows that $(\psi \bullet M)_\sigma \notin L^1$ for all $\sigma \not\equiv 0$, so $\psi \bullet M$ cannot be locally integrable, and therefore not a local martingale, but of course it is a σ-martingale by construction.

4.6 Fundamental Theorem of Asset Pricing

We saw in lemma 4.2 that if S admits an ELMM, then it satisfies (NA) and the proof shows that it is enough if S admits EσMM. The next result gives a converse (with stronger condition (NFLVR)).

Theorem 4.12. *Fundamental Theorem of Asset Pricing (FTAP) (Delbaen-Schachermayer) Let S be $\mathbb{R}^d$-valued semimartingale. TFAE:*

1. *S satisfies (NFLVR)*
2. *S admits ESM*
3. *S admits EσMM*

Proof. Outline of the proof: "3.) $\implies$ 1.)": Is easy and essentially goes as the proof of lemma 4.2 (exercise).

"1.) $\implies$ 2.)": Conceptually similar to proof of Thm. 1.9 (DMW)

- Show that (NFLVR) $\implies$ $\mathcal{C}^\infty$ is $w^* = \sigma(L^\infty, L^1)$-closed in L^∞. (we'll see this in Thm 4.13).
- Use Kreps-Yan (Thm A.1) for $\mathcal{C}^\infty$ and $p = \infty$, to get $Q \sim P$ on $\mathcal{F}_T$ s.t. $E_Q[Y] \le 0$ for all $Y \in \mathcal{C}^\infty$.
- For any $g \in \mathcal{G}_{adm}$, $n \in \mathbb{N}$, $g \wedge n = g - (g - g\wedge n) \in \mathcal{C}^\infty$. Hence $E_Q[g\wedge n] \le 0$ for all n. Also, there exists a s.t. $g \ge -a$, so by Fatou $E_Q[g] \le 0$ for all $g \in \mathcal{G}_{adm}$, hence Q is an ESM.

"2.) $\implies$ 3.)": Under extra assumptions of S locally bounded, then every ESM is an ELMM. But in general this is not true. If S is not bounded then it can happen that $\mathcal{G}_{adm} = \{0\}$ (exercise), then $E_Q[g] \leq 0$ for all $g \in \mathcal{G}_{adm}$ does not give much information. See Delbaen-Schachermayer section 8.3 and section 14.3 and 14.4 for a proof. One can show that the set of EσMM is dense in the set of ESMs (if the latter $\neq \emptyset$).

$\square$

Theorem 4.13. *S satisfies (NFLVR) then $\mathcal{C}^\infty = (\mathcal{G}_{adm} - L^0_+) \cap L^\infty$ is w^*-closed in L^∞.*

Note that in this case:

$$\mathcal{C}^\infty \subseteq \overline{\mathcal{C}^\infty}^{L^\infty} \subseteq \overline{\mathcal{C}^\infty}^{w^*} = \mathcal{C}^\infty \tag{4.31}$$

where the last equality is furnished by NFLVR.

Definition 4.14. *$A \subseteq L^0$ is Fatou-closed if any sequence in A which is uniformly bounded from below and P-a.s. convergent has a limit in A. If A is a cone, then it is enough to show this for the lower bound -1.*

Proof. 1.) A result from functional analysis guarantees: A convex set $C \subset L^\infty$ is w^*-closed in L^∞ iff for any uniformly bounded sequence in C which converges P-a.s. then the limit is in C. It suffices to show "Fatou-closedness".

2.) We show that the convex cone $\mathcal{C}^0_{adm}$ is Fatou-closed. For this, take $(f_n)_n \subseteq \mathcal{C}^0_{adm}$ s.t. $f_n \geq -1$ for all n and $f_n \to f$ P-a.s..

Then $-1 \leq f_n \leq g_n$ for some $g_n = G_T(\theta^n) \in \mathcal{G}_{adm}$. By thm 4.7, (NFLVR) implies (NA), and since $G_T(\theta^n) \geq -1$ implies $G_\cdot(\theta^n) \geq -1$ so $\theta^n \in \Theta^1$ and $g_n \in \mathcal{G}^1$.

By lemma 4.8, we get $\tilde{g}_n \in \text{conv}(g_n, g_{n+1}, ...) \subseteq \mathcal{G}^1$ and $\tilde{g}_n \to \tilde{g}_\infty$ P-a.s. and clearly $\tilde{g}_\infty \geq -1$.

Moreover, $g_n \geq f_n$, $f_n \to f$ P-a.s. so $\tilde{g}_\infty \geq f$. Therefore,

$$\tilde{g}_\infty \in \mathcal{D}_f := \{g \in L^0 : g \geq f\} \cap \overline{\mathcal{G}^1}^{L^0} \tag{4.32}$$

If $\mathcal{G}^1$ were closed in L^0 or if we could show that $\tilde{g}_\infty = G_T(\theta)$ for some $\theta \in \Theta_{adm}$, then we would have the desired conclusion

$$f = G_T(\theta) - (\tilde{g}_\infty - f) \in \mathcal{C}^0_{adm} \tag{4.33}$$

but the above two properties are not true in general (convergence only of $\tilde{g}_n = G_T(\tilde{\theta}^n)$ gives no information about $G.(\tilde{\theta}^n)$ in general.).

Definition 4.15. *$a \in A \subseteq L^0$ is called maximal in A if $h \in A$ and $h \geq a$ P-a.s. implies $h = a$ P-a.s.*

3.) $\mathcal{D}_f = \{g \in L^0 : g \geq f\} \cap \overline{\mathcal{G}^1}^{L^0} \neq \emptyset$ is bounded in L^0 since NFLVR implies NUPBR (thm 4.7) and closed in L^0, since it is the intersection of two closed sets.

Since every closed bounded nonempty $A \subseteq L^0$ has a maximal element by Zorn's lemma, we call h_0 the maximal element in $\mathcal{D}_f$. Then $h_0 \geq f$ and $h_0 = \lim_n G_T(\theta^n)$ in L^0, with θ^n 1-admissible.

Then $f = h_0 - (h_0 - f) \in \mathcal{C}^0_{adm}$ if we show $h_0 \in \mathcal{G}_{adm}$.

4.): h_0 maximal in $\mathcal{D}_f$ and $h_0 = \lim_n G_T(\theta^n)$ in L^0. Claim: $G^*_T(\theta^n - \theta^m) = \sup_{0 \leq t \leq T} |G_t(\theta^n - \theta^m)| \to 0$ in L^0 as $n, m \to \infty$. Hence $(G(\theta^n))_n \subseteq \mathbb{D}$ Cauchy and therefore convergent for the metric d. We show the claim by contradiction: So we suppose that there exist $i_k, j_k \to \infty$ as $k \to \infty$ s.t.

$$P(G^*_T(\theta^{i_k} - \theta^{j_k}) > \alpha) \geq \delta > 0 \tag{4.34}$$

Define

$$\tau_k := \inf\{t : G_t(\theta^{i_k} - \theta^{j_k}) > \alpha\} \wedge T \tag{4.35}$$

this satisfies $P(\tau_k < T) \geq \delta$. Also define

$$\tilde{\theta}^k := \theta^{i_k} \mathbf{1}_{[0,\tau_k]} + \theta^{j_k} \mathbf{1}_{(\tau_k, T]} \in \Theta^1 \tag{4.36}$$

On $[0, \tau_k]$, $G.(\tilde{\theta}^k) = G.(\theta^{i_k}) \geq -1$. On $(\tau_k, T]$, that is take $t > \tau_k$:

$$G_t(\tilde{\theta}^k) = G_{\tau_k}(\theta^{i_k}) + G_t(\theta^{j_k}) - G_{\tau_k}(\theta^{j_k}) \geq -1 + \alpha \geq -1 \tag{4.37}$$

Moreover at time T

$$G_T(\tilde{\theta}^k) = G_T(\theta^{i_k})\mathbf{1}(\tau_k = T) + G_T(\theta^{j_k})\mathbf{1}(\tau_k < T) + \xi_k \tag{4.38}$$

where $\xi_k = (G_{\tau_k}(\theta^{i_k}) - G_{\tau_k}(\theta^{j_k}))\mathbf{1}(\tau_k < T) \geq 0$ and s.t.

$$P(\xi_k \geq \alpha) \geq \delta > 0 \tag{4.39}$$

Then by lemma 4.8 and using convexity of $\mathcal{G}^1$, we get in $\mathcal{D}_f$ an element $h_0 + \eta \in \mathcal{D}_f$, where η is the limit of $(\tilde{\xi}_n)_n$ in $L^0_+ \backslash \{0\}$. This is a contradiction to the maximality of h_0 in $\mathcal{D}_f$.

$\square$

To conclude the proof of theorem 4.13 we use a result from Cuchiero/Teichmann (2015):

> **Theorem 4.16.** *S (NUPBR), suppose $(\theta^n)_n \subseteq \Theta^1$ s.t. $(G(\theta^n))_n$ converges for d to a process X s.t. X_T is maximal in $\overline{\mathcal{G}^1}^{L^0}$. Then we have $G(\theta^n) \to X$ for Émery metric $\tilde{d}'_E$. As a consequence, $X = G(\theta)$ for some $\theta \in \Theta^1$ ($X_T \in \mathcal{G}^1$).*

Let's first see how, by using the above theorem, we can conclude our proof. From step 4, $(G(\theta^n))_n \subseteq \mathbb{D}$ with $\theta^n \in \Theta^1$ Cauchy for d, hence convergent for d to some $X \in \mathbb{D}$. Moreover, $X_T = \lim_n G_T(\theta^n) = h_0$ maximal in $\overline{\mathcal{G}^1}^{L^0}$, so by the previous theorem 4.16 $G(\theta^n) \to X$ for $\tilde{d}'_E$. Now we use that the space of stochastic integrals with respect to S is closed in $\tilde{d}'_E$, hence $X = G(\theta)$ for some $\theta \in L(S)$. Since θ^n is 1-admissible and $G(\theta^n) \to G(\theta) = X$ we have that θ is 1-admissible and so $h_0 = X_T \in \mathcal{G}^1$. This is what we wanted for Fatou-closedness of $\mathcal{C}^0_{adm}$.

Main ideas for the proof of previous theorem 4.16:

1. Show that (NUPBR) implies that $(G(\theta^n))_n$ satisfies (P-UT): Boundedness in Émery topology can be used to control convergence in decomposition of $G(\theta^n)$ except FV part.

2. Exploit maximality to argue that FV part converges as well.

Chapter 5

No-Arbitrage properties in some model classes

5.1 NUPBR, $E\sigma$MDs/ELMDs and numéraire portfolios

Setup: S $\mathbb{R}^d$-valued semimartingale, probability space, filtration etc. Recall: $\Theta^1 = \{\theta \in L(S) : G.(\theta) = \theta \bullet S \geq -1\}$. Notation:

$$\mathcal{X}^1 := 1 + G(\Theta^1) \tag{5.1}$$
$$\mathcal{X}^1_T := \{X_T : X \in \mathcal{X}^1\} \tag{5.2}$$
$$\mathcal{X}^1_{++} := \{X \in \mathcal{X}^1 : X > 0, X_- > 0\} \tag{5.3}$$

Recall: (NUPBR) means that $\mathcal{G}^1 = G_T(\Theta^1)$ is bounded in L^0.

Definition 5.1. *An equivalent σ-martingale density (EσMD) for S is a local martingale $Z > 0$, $Z_0 = 1$ s.t. ZS is a P-σ-martingale. If ZS is even a local P-martingale, then Z is called an equivalent local martingale density (ELMD) for S.*

Remark: If $Q \sim P$ is EσMM or ELMM for S and $Z = Z^{Q,P}$ its density process w.r.t. P, then ZS is a P-σ-martingale or P-local martingale (by Bayes), and $Z > 0$ is a true P-martingale, $Z_0 = 1$ iff $Q = P$ on $\mathcal{F}_0$. This explains the difference between EσMD and ELMD.

Recall: Every ELMD is clearly EσMD. If S is continuous, then the converse holds as well (exercise). Use lemma 4.4 and $\Delta(ZS) = S\Delta Z$ by continuity, S locally bounded and ΔZ like Z locally in L^1, hence the first term is locally in L^1.

Definition 5.2. *A numéraire portfolio is a process $X^{np} \in \mathcal{X}^1_{++}$ s.t. X/X^{np} is a P-supermartingale for all $X \in \mathcal{X}^1_{++}$.*

Intuition: X/X^{np} is decreasing on average and X^{np} is the unique element with better performance than all others in the set.

Remark: If Z is EσMD or ELMD for S, then using product rule we study the process $ZG(\theta)$:

$$d(ZG(\theta)) = Z_- dG(\theta) + G(\theta)_- dZ + d[Z, G(\theta)] \tag{5.4}$$

$$= (G(\theta)_- - \theta S_-)dZ + \theta d(ZS) \tag{5.5}$$

so $ZG(\theta)$ is a stochastic integral w.r.t a local P-martingale for all $\theta \in L(S)$. Indeed, both Z and ZS are local P-martingales, or the latter is a stochastic integral w.r.t. one.

For every $X \in \mathcal{X}^1$, $ZX = Z + ZG(\theta)$ for some $\theta \in \Theta^1$, so ZX is a local P-martingale and since $ZX \geq 0$ a P-supermartingale. We call those Z with the P-supermartingale property (for ZX for all $X \in \mathcal{X}^1$) a supermartingale deflator for S.

If in addition, $\frac{1}{Z}$ is in $\mathcal{X}^1_{++}$, then $X^{np} := \frac{1}{Z}$ is a numéraire portfolio.

Theorem 5.3. *For $\mathbb{R}^d$-valued semimartingales S, TFAE*

1. *S satisfies (NUPBR)*
2. *There exists EσMD for S*
3. *There exists a numéraire portfolio X^{np} for S*

Proof. "1.) $\iff$ 2.)" Takaoka Theorem 2.6.

"1.) $\iff$ 3.)" Karatzas/Kardaras Theorem 4.12. □

5.2 The structure condition (SC)

We consider S $\mathbb{R}^d$-valued, continuous and adapted. Recall: Q is EσMM is equivalent to being ELMM in that case. And the analogous statement holds for the density process.

Theorem 5.4. *S continuous s.t. there exists Z EσMD, then the following structure condition (SC) holds: S is a continuous semimartingale with the following decomposition $S = S_0 + M + A$ with $A = \int \lambda d\langle M\rangle$ where $M \in \mathcal{M}^c_{0,loc}$, $A \in FV^c_0$ and $\lambda \in L^2_{loc}(M)$ is $\mathbb{R}^d$-valued predictable.*

Proof. $Z > 0$, local martingale, hence a supermartingale and also $Z_- > 0$. Since S continuous, Z EσMD is also ELMD, hence SZ is a local martingale and therefore a semimartingale. Then write

$$S = (ZS)\frac{1}{Z} \tag{5.6}$$

so by Itô, using $Z_- > 0$, we have that $S = S_0 + M + A$ is a semimartingale.

Here Z is a local martingale $Z_0 = 1$, $Z, Z_- > 0$. Therefore we can write $Z = \mathcal{E}(N)$, with $N = \frac{1}{Z_-} \bullet Z$ a local martingale, because with this choice $dZ = Z_- dN$.

Now we use a generalized version of the Kunita-Watanabe decomposition:

$$N = \int \vartheta dM + L \tag{5.7}$$

with $\vartheta \in L^2_{loc}(M)$ and L strongly orthogonal to M, i.e. $[L, M] = \langle L, M\rangle \equiv 0$.

We use that $ZS - Z_0S_0$ is local martingale. By the product rule

$$\begin{aligned} ZS - Z_0S_0 &= Z_- \bullet S + S_- \bullet Z + [Z, S] && (5.8)\\ &= (Z_- \bullet M + S_- \bullet Z) + (Z_- \bullet A + \langle Z, S\rangle) && (5.9)\end{aligned}$$

where we have used the continuity of S. The second term is of finite variation, and from the equality it is also a local martingale. All predictable local martingales are continuous, and if they are also of FV, then they are constant. Therefore

$$\begin{aligned} 0 &\equiv Z_- \bullet A + \langle Z, S\rangle && (5.10)\\ &= Z_- \bullet A + \langle Z_- \bullet N, S\rangle && (5.11)\\ &= Z_- \bullet (A + \langle N, M\rangle) && (5.12)\end{aligned}$$

So that $A = -\langle N, M\rangle = -\int \vartheta d\langle M\rangle$. Hence $\lambda = -\vartheta$ gives the statement. $\square$

We have a parametrization of all EσMD /ELMD for S continuous.

Corollary 5.5. *Let S continuous, satisfying (SC), then*

1. *A local martingale $Z > 0$, $Z_0 = 1$ is EσMD/ELMD for S iff it has the 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 EσMD and ELMD.*

2. *$Q \sim P$ on $\mathcal{F}_T$ is an EσMM and ELMM iff its density process $Z := Z^{Q,P}$ has the form $Z = Z_0 \hat{Z}\mathcal{E}(L)$ with L as in 1.) and such that it is a true martingale > 0 on $[0, T]$.*

Proof. 2.) follows easily from 1.), so we need to show 1.).

By the previous theorem, any EσMD Z has the form

$$Z = \mathcal{E}(-\int \lambda dM + L) \tag{5.13}$$

(We use Yor's formula: $\mathcal{E}(X + Y)e^{\langle X,Y\rangle} = \mathcal{E}(X)\mathcal{E}(Y)$)

This gives $Z = \mathcal{E}(-\int \lambda dM)\mathcal{E}(L)$ since $\langle L, M\rangle \equiv 0$.

$$\begin{aligned}
d\mathcal{E}(L) &= \mathcal{E}(L)_- dL && (5.14)\\
&\implies \Delta\mathcal{E}(L) = \mathcal{E}(L)_-\Delta L && (5.15)\\
&\implies \mathcal{E}(L) = \mathcal{E}(L)_-(1 + \Delta L) > 0 && (5.16)\\
&\implies \Delta L > -1 && (5.17)
\end{aligned}$$

Conversely, if $Z = \hat{Z}\mathcal{E}(L)$, then $Z = \mathcal{E}(-\int \lambda dM + L)$ and the computation as in the proof of the previous thm. 5.4 gives $ZS - Z_0S_0 = Z_- \bullet M + S_- \bullet Z$ local martingale.

□

The process $\hat{Z} := \mathcal{E}(-\int \lambda dM)$ is called the minimal ELMD for S. If it is a true martingale, then the corresponding probability measure $\hat{P} \sim P$ s.t. $\frac{d\hat{P}}{dP} := \hat{Z}_T$ is called the minimal martingale measure for S.

The process $\hat{K} := \langle \int \lambda dM\rangle = \int \lambda^{tr} d\langle M\rangle\lambda$ is called the mean-variance trade-off process.

Remark: From the above arguments, a continuous adapted process admits an EσMD/ELMD iff it satisfies (SC).

5.3 Model class: Itô processes

An important class of models are the so-called Itô processes:

- W is an $\mathbb{R}^n$-valued Brownian Motion on a general probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$
- Undiscounted prices $\tilde{B}$ and $\tilde{S} = (\tilde{S}^i)_{i=1,\dots,d}$ given by

$$d\tilde{B}_t = \tilde{B}_t r_t dt \quad \tilde{B}_0 = 1 \tag{5.18}$$

$$d\tilde{S}^i_t = \tilde{S}^i_t \mu^i_t dt + \tilde{S}^i_t \sum_{j=1}^n \sigma^{ij}_t dW^j_t \quad \tilde{S}^i_0 = s^i > 0 \tag{5.19}$$

- In discounted terms (discounted prices):

$$S^i = \frac{\tilde{S}^i}{\tilde{B}} \tag{5.20}$$

$$dS^i_t = S^i_t (b^i_t dt + \sum_{j=1}^n \sigma^{ij}_t dW^j_t) \tag{5.21}$$

  with $b^i_t = \mu^i_t - r_t$. That is

$$dS_t = \mathrm{diag}(S_t)(b_t dt + \sigma_t dW_t) \tag{5.22}$$

- $S = S_0 + M + A$ with

$$dA_t = \mathrm{diag}(S_t) b_t dt \tag{5.23}$$

$$dM_t = \mathrm{diag}(S_t)\sigma_t dW_t \tag{5.24}$$

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

- Assume $d < n$ and that $\mathrm{rank}(\sigma_t) \equiv d$ so that $\sigma_t \sigma^{tr}_t$ is invertible for all t, we can then define

$$\bar{\lambda}_t := \sigma^{tr}(\sigma_t \sigma^{tr}_t)^{-1} b_t \tag{5.25}$$

  so we have

$$dS_t = \mathrm{diag}(S_t)\sigma_t(\bar{\lambda}_t dt + dW_t) \tag{5.26}$$

  and

$$\begin{aligned} dA_t &= \mathrm{diag}(S_t)\sigma_t \bar{\lambda}_t dt && (5.27)\\ &= \mathrm{diag}(S_t)\sigma_t \sigma^{tr}_t \mathrm{diag}(S_t)\mathrm{diag}(S_t)^{-1}(\sigma_t \sigma^{tr}_t)^{-1} b_t dt && (5.28)\\ &= d\langle M \rangle_t \lambda_t && (5.29) \end{aligned}$$

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

$$\lambda_t := \text{diag}(S_t)^{-1}(\sigma_t\sigma_t^{tr})^{-1}b_t \tag{5.30}$$
$$= \text{diag}(S_t)^{-1}(\sigma_t\sigma_t^{tr})^{-1}(\mu_t - r_t\mathbf{1}_d) \tag{5.31}$$

Note that

$$\int \lambda dM = \int b^{tr}(\sigma_t\sigma_t^{tr})^{-1}\sigma_t dW_t \tag{5.32}$$
$$= \int \bar{\lambda} dW \tag{5.33}$$

and

$$\hat{K} = \int \lambda_t^{tr} d\langle M\rangle_t \lambda_t \tag{5.34}$$
$$= \int b_t^{tr}(\sigma_t\sigma_t^{tr})^{-1}b_t dt \tag{5.35}$$
$$= \int |\bar{\lambda}_t|^2 dt \tag{5.36}$$

so S satisfies the structure condition (SC) iff this process is finite valued. In this case EσMDs/ELMDs for S are parametrized as

$$Z = \hat{Z}\mathcal{E}(L) = \mathcal{E}(-\int \bar{\lambda} dW)\mathcal{E}(L) \tag{5.37}$$

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

- We have analyzed this in the context of a general filtration $\mathbb{F}$, but can say more if we take $\mathbb{F} = \mathbb{F}^W$ for W the Brownian Motion.

5.4 With the Brownian filtration

Lemma 5.6. *Suppose $\mathbb{F} = \mathbb{F}^W$, $M = \int diag(S)\sigma dW$, $L \in \mathcal{M}_{0,loc}$. Then we have*

1. *$\Delta L > -1$ (actually, L continuous, and $\Delta L \equiv 0$)*
2. *L strongly orthogonal to M iff $L = \int \vartheta dW$ with $\vartheta \in L^2_{loc}(W)$ and $\sigma\vartheta \equiv 0$*

As a consequence, EσMDs/ELMDs are parametrized by

$$Z = \hat{Z}\mathcal{E}(\int \vartheta dW) = \mathcal{E}(-\int (\bar{\lambda} - \vartheta) dW) \tag{5.38}$$

with $\bar{\lambda} = \sigma^{tr}(\sigma\sigma^{tr})^{-1}b$, $\sigma\vartheta \equiv 0$.

Proof. 1.) Is immediate from Itô's representation theorem: $L \in \mathcal{M}_{0,loc}$ in $\mathbb{F}^W$ has representation: $L = \int \vartheta dW$, $\vartheta \in L^2_{loc}(W)$, hence L is continuous and $\Delta L \equiv 0$.

2.) Note that $L \perp M$ iff $\langle L, M \rangle \equiv 0$, and

$$\langle L, M \rangle = \int \mathrm{diag}(S)\sigma\vartheta dt \tag{5.39}$$

□

Let's consider the specific case of $n = d$, in particular σ is a $d \times d$ matrix. σ has full rank iff its kernel is the zero vector: $\sigma v \equiv 0$ iff $v \equiv 0$.

Then the only EσMD/ELMD is $\hat{Z} = \mathcal{E}(-\int \bar{\lambda} dW)$ with $\bar{\lambda} = \sigma^{-1}b$, hence

$$\hat{Z} = \mathcal{E}(-\int \sigma^{-1}bdW) \tag{5.40}$$

This implies that in this market we have only one candidate for an EσMM/ELMM, that is $\hat{\mathbb{P}}$ with density $\hat{Z}$ with respect to P. This will be an EσMM/ELMM iff $\hat{Z}$ is a true martingale.

5.5 Example: Black-Scholes model

Itô process $n = d = 1$, all constant coefficients r, μ, σ:

$$dS_t = S_t(\mu - r)dt + S_t\sigma dW_t \tag{5.41}$$

with minimal EσMD/ELMD

$$\hat{Z} = \mathcal{E}(-\frac{\mu - r}{\sigma} W) \tag{5.42}$$

which is a true martingale (e.g. by Novikov's condition). So $\hat{P}$ with density $\hat{Z}$ is an ELMM.

Actually S is not only local but also a true martingale under $\hat{P}$, hence $\hat{P}$ EMM:

$$S = S_0\mathcal{E}(\sigma\hat{W}) \tag{5.43}$$

where $\hat{W}$ is a $\hat{P}$-Brownian Motion given by

$$\hat{W}_t = W_t + \frac{\mu - r}{\sigma} t \quad \text{(Girsanov's theorem)} \tag{5.44}$$

Note that here $\mathbb{F} = \mathbb{F}^W = \mathbb{F}^S$.

5.6 Numéraire portfolio under continuous S with (SC)

We go back to general $(\Omega, \mathcal{F}, \mathbb{F}, P)$ and $\mathbb{R}^d$-valued continuous adapted S. We already know that S admits EσMD/ELMD iff S satisfies (SC), and in this case

$$\hat{Z} = \mathcal{E}(-\int \lambda dM) \tag{5.45}$$

is an EσMD/ELMD.

We now look at $\frac{1}{\hat{Z}}$ and use continuity of S and specific form of FV part $A = \int d\langle M \rangle \lambda$ due to (SC):

$$\begin{aligned}
\frac{1}{\hat{Z}} &= \frac{1}{\mathcal{E}(-\int \lambda dM)} && (5.46)\\
&= \exp\left(\int \lambda dM + \frac{1}{2}\left\langle \int \lambda dM \right\rangle\right) && (5.47)\\
&= \exp\left(\int \lambda dM + \int \lambda^{tr} d\langle M \rangle \lambda - \frac{1}{2}\int \lambda^{tr} d\langle M \rangle \lambda\right) && (5.48)\\
&= \exp\left(\int \lambda dS - \frac{1}{2}\left\langle \int \lambda dS \right\rangle\right) && (5.49)\\
&= \mathcal{E}(\int \lambda dS) && (5.50)\\
&= 1 + \int \lambda \mathcal{E}(\int \lambda dS) dS && (5.51)
\end{aligned}$$

Hence $\frac{1}{\hat{Z}} \in \mathcal{X}^1$, and both $\frac{1}{\hat{Z}}$and $\frac{1}{\hat{Z}_-} > 0$ by continuity. Hence $\frac{1}{\hat{Z}} \in \mathcal{X}^1_{++}$. We then get:

Corollary 5.7. *If S is continuous, satisfying (SC), then numéraire portfolio X^{np} exists and it's equal to $\frac{1}{\hat{Z}}$.*

Proof. From the remark before Thm.5.3 because $\hat{Z}$ is EσMD with $\frac{1}{\hat{Z}} \in \mathcal{X}^1_{++}$. □

The converse of this corollary:

Theorem 5.8. *Let S continuous semimartingale and X^{np} exists. Then S satisfies (SC).*

Proof. By thm. 5.3, the existence of X^{np} implies existence of an EσMD, which gives (SC) by Thm. 5.4.

However, as we haven't proved thm 5.3, let's give a direct proof of this theorem. To simplify notation, take $d = 1$. Any $X \in \mathcal{X}^1_{++}$ can be rewritten as follows:

$$X = 1 + \theta \bullet S \tag{5.52}$$

$$= 1 + (X_- \frac{\theta}{X_-}) \bullet S \tag{5.53}$$

$$= 1 + X_- \bullet (\pi \bullet S) \tag{5.54}$$

where $\pi := \frac{\theta}{X_-} \in L(S)$. Therefore $X = \mathcal{E}(\pi \bullet S)$.

We now use continuity of $S = S_0 + M + A$ to rewrite $\frac{X}{\bar{X}}$, where $\bar{X}$ is any other element in $\mathcal{X}^1_{++}$ (in particular $\bar{X} = \mathcal{E}(\bar{\pi} \bullet S)$):

$$\frac{X}{\bar{X}} = \exp\left((\pi - \bar{\pi}) \bullet S - \frac{1}{2}\pi^2 \bullet \langle S \rangle + \frac{1}{2}\bar{\pi}^2 \bullet \langle S \rangle\right) \tag{5.55}$$

$$= \exp\left((\pi - \bar{\pi}) \bullet M - \frac{1}{2}(\pi - \bar{\pi})^2 \bullet \langle M \rangle\right) \cdot \tag{5.56}$$

$$\cdot \exp\left((\pi - \bar{\pi}) \bullet A + \frac{1}{2}(\pi - \bar{\pi})^2 \bullet \langle M \rangle - \frac{1}{2}(\pi^2 - \bar{\pi}^2) \bullet \langle M \rangle\right) \tag{5.57}$$

The first exponential factor is $\mathcal{E}((\pi - \bar{\pi}) \bullet M)$, which is a local martingale and the latter exponential factor is predictable and FV.

In particular, if we take $\bar{X} = X^{np}$, then by definition $\frac{X}{X^{np}}$ is a supermartingale.

But a supermartingale > 0 has a unique multiplicative decomposition as a product of a local martingale and a predictable decreasing RCLL process (exercise). Hence the second exponential factor in the decomposition above has to be decreasing, that is

$$(\pi - \bar{\pi}) \bullet A - \pi\bar{\pi} \bullet \langle M \rangle + \bar{\pi}^2 \bullet \langle M \rangle = (\pi - \bar{\pi}) \bullet (A - \bar{\pi} \bullet \langle M \rangle) \tag{5.58}$$

has to be decreasing for all π. Therefore $A - \bar{\pi} \bullet \langle M \rangle$ must be $\equiv 0$. (Note that here $\bar{\pi} \in L^2_{loc}(M)$). Therefore $A = \bar{\pi} \bullet \langle M \rangle$ which implies that the structure condition holds.

□

Corollary 5.9. *For S an $\mathbb{R}^d$-valued continuous semimartingale, TFAE:*

1. *S admits EσMD/ELMD*
2. *S satisfies the structure condition (SC)*
3. *There exists a numéraire portfolio X^{np} for S and in this case $X^{np} = \frac{1}{\hat{Z}}$*

(cf. with Thm. 5.3, without continuity.)

Therefore we show:

Lemma 5.10. *Let S adapted $\mathbb{R}^d$-valued RCLL process. If S admits EσMD, then S is a semimartingale and satisfies (NUPBR).*

Proof. The semimartingale property of $S = \frac{1}{Z}(ZS)$ follows as in Thm. 5.4 (here ZS is a σ-martingale, thus a semimartingale, for all Z EσMD).

Moreover, as seen in the discussion before Thm. 5.3, for every $X \in \mathcal{X}^1$, we have ZX supermartingale, starting at 1 and is nonnegative. So

$$E[Z_T X_T] \leq E[Z_0 X_0] = 1 \tag{5.59}$$

for all $X \in \mathcal{X}^1$. This shows that the set

$$Z_T \mathcal{X}_T^1 = \{Z_T g : g \in \mathcal{X}_T^1\} \tag{5.60}$$

is bounded in L^1, thus also in L^0, and from this the following set is bounded in L^0, since $0 < Z_T < \infty$:

$$\mathcal{G}^1 = \mathcal{X}_T^1 - 1 = \frac{1}{Z_T}(Z_T \mathcal{X}_T^1 - Z_T) \tag{5.61}$$

(exercise).

□

Chapter 6

Pricing and hedging by replication

Consider time T payoff $H \in L^0_+(\mathcal{F}_T)$ in an arbitrage-free market. There are two main questions:

1. What is a reasonable time-t price of H, for $t < T$?
2. How can we manage risk resulting from selling H?

6.1 Basic ideas and results

Given $H \in L^0_+(\mathcal{F}_T)$, a replicating strategy for H is a self-financing admissible strategy $\varphi = (v_0, \theta)$ with $V_T(\varphi) = H$ P-a.s. In this case we say that H is attainable or replicable or that it can be hedged (by φ).

The fundamental idea of valuation/pricing by replication is: The value in $t < T$ for $H \in L^0_+(\mathcal{F}_T)$ replicable by φ is $V_t(\varphi)$, otherwise we can easily construct an arbitrage opportunity: Consider the time interval $[t, T]$ and the two strategies

1. Buy H at its time-t price in market, say $\pi_t(H)$ and wait until time T
2. Use self-financing admissible strategy $\varphi = (V_t(\varphi), \theta)$

Both strategies have zero cash flow on (t, T) and the same value at time T, namely $H = V_T(\varphi)$. So if in $t < T$ we would have $\pi_t(H) \neq V_t(\varphi)$ then we can buy cheaper and sell the more expensive strategy, thus producing an arbitrage opportunity in time t.

For easy computation of $V_t(\varphi)$, note that if H is attainable by $\varphi = (v_0, \theta)$, then $H = V_T(\varphi) = v_0 + \int_0^T \theta_u dS_u$ P-a.s. and at any time $t < T$, $V_t(\varphi) =$

$v_0 + \int_0^t \theta_u dS_u$. So if Q is an EσMM for S and Q is nice enough so that $\int \theta dS$ is a Q-martingale, then

$$V_t(\varphi) = E_Q[V_T(\varphi)|\mathcal{F}_t] = E_Q[H|\mathcal{F}_t] \tag{6.1}$$

for $0 \leq t \leq T$. This is called "valuation/pricing by risk-neutral expectation".

Comments:

1. The key argument is to transform $V_t(\varphi)$ by riskless dynamic trading on $[t, T]$ into H. In particular
   - Hedging by dynamic trading (then by (NA) the valuation is a byproduct)
   - Q is an auxiliary tool, and Q-probabilties are artificial, not corresponding to probabiltites of events in the market in an obvious way.
   - The LHS of the above pricing rule does not depend on φ used to replicate H. If other strategies also replicate H, say φ', then $V_t(\varphi) = V_t(\varphi')$ and the RHS does not depend on which martingale measure Q or Q' is used.
2. The basic structure of the pricing rule is simple: Discounted time-t value of a payoff is conditional expectation of the discounted payoff under EσMM Q.

In finite discrete time things are easier to scale. For the rest of this subsection, consider $S = (S_k)_{k=0,\ldots,T}$, $\mathbb{R}^d$-valued, adapted and $S^0 \equiv 1$. Denote by $\mathcal{P}_{e(loc)}(S)$ the set of all E(L)MMs for S and recall by Cor. 1.9 that S satisfies (NA) iff $\mathcal{P}_e(S) \neq \emptyset$. Call a market $(S, \mathbb{F})$ complete if every $H \in L^0_+(\mathcal{F}_T)$ is attainable.

Lemma 6.1. *If $\mathcal{F}_0$ is trivial, then TFAE:*

1. *$(S, \mathbb{F})$ is complete*
2. *Every $H \in L^0_+(\mathcal{F}_T)$ admits a representation $H = H_0 + \int_0^T \theta_u dS_u$ P-a.s.*

Proof. The proof is easy and is left as an exercise. □

So completeness means that, up to constants, S spans all $\mathcal{F}_T$-measurable random variables via stochastic integrals.

Recall the following results from IMF:

Theorem 6.2. *Let $\mathcal{F}_0$ trivial, S satisfying (NA). Then for any $H \in L^0_+(\mathcal{F}_T)$, TFAE:*

1. *H is attainable*
2. $\sup_{Q \in \mathcal{P}_{e,loc}(S)} E_Q[H] < \infty$ *is attained in some $Q^* \in \mathcal{P}_{e,loc}(S)$.*
3. *The mapping $\mathcal{P}_{e,loc}(S) \to \mathbb{R}$ given by $Q \mapsto E_Q[H]$, is constant (and equal to v_0, initial capital of replicating strategy).*

From this it follows:

Theorem 6.3. *Let $\mathcal{F}_0$ trivial, S satisfying (NA), $\mathcal{F}_T = \mathcal{F}$, then TFAE:*

1. *$(S, \mathbb{F})$ is complete*
2. *The number of elements in $\mathcal{P}_{e,loc}(S) = 1$, that is there exists a unique ELMM for S.*

6.2 Example: The Black-Scholes formula

Bank account: $\tilde{B}_t = \exp(rt)$. One stock $\tilde{S}_t = s_0 \exp(\sigma W_t + (\mu - \sigma^2/2)t)$.

Discounted terms: $B = \frac{\tilde{B}}{B} \equiv 1$, $S = \frac{\tilde{S}}{\tilde{B}}$ satisfies the SDE

$$dS_t = S_t((\mu - r)dt + \sigma dW_t) \tag{6.2}$$

Consider a call option $\tilde{H} = (\tilde{S}_T - \tilde{K})^+$ and the discounted payoff

$$H = \frac{\tilde{H}}{\tilde{B}_T} = (S_T - K)^+ \tag{6.3}$$

$$K = \tilde{K} \exp(-rT) \tag{6.4}$$

Question: What is the value of $t < T$?

Assume $\mathbb{F} = \mathbb{F}^W$ the filtration generated by the Brownian Motion and completed. By lemma 5.6 and the reasoning right after, the only candidate for an EσMD is

$$\hat{Z}_t = \mathcal{E}(-\frac{\mu - r}{\sigma} W)_t = \exp(-\frac{\mu - r}{\sigma} W_t - \frac{1}{2}(\frac{\mu - r}{\sigma})^2 t) \tag{6.5}$$

$\hat{Z} > 0$ is a true martingale. Hence $\hat{P}$ defined by

$$\frac{d\hat{P}}{dP} = \hat{Z}_T \tag{6.6}$$

is ELMM. Actually $\hat{\mathbb{P}}$ is EMM since S is not only a local but a true martingale under $\hat{P}$:

$$dS_t = S_t \sigma d\hat{W}_t \tag{6.7}$$

where $\hat{W}_t := W_t + \frac{\mu - r}{\sigma} t$ is a Brownian Motion under $\hat{P}$ (use Girsanov's theorem).

Comparing with the discrete time case, we guess that this market is complete (i.e. existence of a unique EMM). So let's compute

$$\begin{aligned}
\hat{V}_t &:= E_{\hat{P}}[H|\mathcal{F}_t] && (6.8)\\
&= \hat{E}[(S_T - K)^+|\mathcal{F}_t] && (6.9)\\
&= \hat{E}[(a\exp(bZ - c) - d)^+]|_{a,b,c,d} && (6.10)\\
&=: \hat{v}(t, S_t) && (6.11)
\end{aligned}$$

where $a = S_t$, $b = \sigma\sqrt{T-t}$, $c = b^2/2$ and $d = K$, where $Z \sim \mathcal{N}(0,1)$. We used that $S_T = \frac{S_T}{S_t} S_t$. This is computable using some integral transformations and properties of the lognormal distribution.

Undiscounted value $\tilde{V}_t = \tilde{B}_t \hat{V}_t =: \tilde{v}(t, \tilde{S}_t)$. Computing gives

$$\tilde{v}(t, \tilde{S}_t) = \tilde{S}_t \Phi(d_1) - \tilde{K}\exp(-r(T-t))\Phi(d_2) \tag{6.12}$$

with Φ the cdf of $\mathcal{N}(0,1)$ and where

$$d_{1,2} = \frac{\log(\frac{\tilde{S}_t}{\tilde{K}\exp(-r(T-t))}) \pm \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}} \tag{6.13}$$

The last two equations are known as the Black-Scholes formula.

Comment: The fact that price in t of call option is a function only of t and S_t comes from $H = f(S_T)$.

To justify that $\tilde{V}_t = \tilde{v}(t, \tilde{S}_t)$ is a reasonable valuation at time t we need to check whether H is attainable. One way is to use Itô's theorem: Since $\mathbb{F} = \mathbb{F}^W = \mathbb{F}^{\hat{W}}$ any $H \in L^1(\mathcal{F}_T, \hat{P})$ has a unique representation:

$$H = \hat{E}[H] + \int_0^T \psi d\hat{W} = \hat{E}[H] + \int_0^T \theta dS \tag{6.14}$$

$$\text{taking } \theta_t = \frac{\psi_t}{\sigma S_t} \tag{6.15}$$

with the latter terms $\hat{P}$-martingales. Moreover, if $H \geq 0$ (as for the example of a call option), then $\int \theta dS \geq -\hat{E}[H]$ implies $\varphi = (\hat{E}[H], \theta)$ is admissible. Therefore H is the terminal value of a self-financing admissible strategy whose wealth $V(\varphi) = \hat{E}[H] + \int \theta dS$ is a $\hat{P}$-martingale. And so H is attainable.

Alternative argument to see that call option H is attainable for any $\mathbb{F}$: Consider the function $\tilde{v}(t,x)$ from the Black-Scholes formula. One can show that $\tilde{v}$ satisfies the following PDE:

$$\frac{\partial \tilde{v}}{\partial t} + rx\frac{\partial \tilde{v}}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 \tilde{v}}{\partial x^2} - r\tilde{v} = 0 \tag{6.16}$$

$$\tilde{v}(T,x) = (x - \tilde{K})^+ \tag{6.17}$$

Now we have that $\tilde{S}_t = \tilde{B}_t S_t = \exp(rt)S_t$ satisfies

$$d\tilde{S}_t = \tilde{S}_t(rdt + \sigma d\hat{W}_t) \tag{6.18}$$

so applying Itô's formula and using the above PDE gives

$$\begin{aligned} d\tilde{V}_t &= d\tilde{v}(t, \tilde{S}_t) && (6.19)\\ &= r\tilde{V}_t dt + \tilde{v}_x(t,\tilde{S}_t)\sigma \tilde{S}_t d\hat{W}_t && (6.20)\\ &= \tilde{v}_x(t,\tilde{S}_t)d\tilde{S}_t + (r\tilde{V}_t - \tilde{v}_x(t,\tilde{S}_t)r\tilde{S}_t)dt && (6.21)\\ &=: \theta_t d\tilde{S}_t + \varphi_t^0 r\tilde{B}_t dt && (6.22) \end{aligned}$$

hence

$$d\tilde{V}_t = \theta_t d\tilde{S}_t + \varphi_t^0 d\tilde{B}_t \tag{6.23}$$

and $\tilde{V}_t = \tilde{v}_t(t,\tilde{S}_t) = \theta_t \tilde{S}_t + \varphi_t^0 \tilde{B}_t$.

Now integrate from t to T and use that $\tilde{v}(T,\tilde{S}_T) = \tilde{H} = (\tilde{S}_T - \tilde{K})^+$ to get

$$(\tilde{S}_T - \tilde{K})^+ = \tilde{V}_T = \tilde{V}_t + \int_t^T \theta_u d\tilde{S}_u + \varphi_u^0 d\tilde{B}_u \tag{6.24}$$

so we have that $\varphi = (\varphi^0, \theta)$ is a self-financing strategy for the undiscounted process $(\tilde{B}, \tilde{S})$, and admissible since $\tilde{v} \geq 0$. Therefore $\tilde{H}$ is attainable.

Chapter 7

Super-replication and optional decomposition

Basic question: How to "hedge" and price general (non-replicable) payoffs (in an incomplete market)?

Setup: $(\Omega, \mathcal{F}, \mathbb{F}, P)$, $S^0 \equiv 1$, S risky assets defined on $[0,T]$. S is $\mathbb{R}^d$-valued semimartingale satisfying (NFLVR), so that $\mathcal{P}_{e,\sigma}$ (the set of all EσMM) for S is non-empty. Fix $H \in L^0_+(\mathcal{F}_T)$. If H is attainable, this means that it is replicable and that we can easily find a fair price. If it is not attainable, then there does not exist a self financing strategy with $V_T(\varphi) = H$ P-a.s. and this raises the question as to how we are to price H.

7.1 The super-replication price $\pi^S(H)$

The idea behind super-replication: Look for strategies that at time T produce at least H ($H \leq V_T(\varphi)$ a.s.) and find cheapest among them.

Definition 7.1. *Super-replication price of $H \in L^0_+(\mathcal{F}_T)$ is*

$$\pi^S(H) := \inf \Big\{ v_0 \in \mathbb{R} : \exists \theta \text{ admissible with} \tag{7.1}$$

$$v_0 + \int_0^T \theta_t dS_t \geq H \ P\text{-a.s.} \Big\} \tag{7.2}$$

$$= \inf\{ v_0 \in \mathbb{R} : H - v_0 \in \mathcal{G}_{adm} - L^0_+ = \mathcal{C}^0_{adm} \} \tag{7.3}$$

Intuition: Sell H in $t = 0$ for $\pi^S(H)$ so that no risks in T if I follow strategy $\varphi = (\pi^S(H), \theta)$. This seems like a reasonable price from the seller's perspective.

Note: We do not know if infimum is attained, i.e. if there is such θ such that $v_0 := \pi^S(H)$ can be taken to superreplicate H.

Lemma 7.2. *Assume (NFLVR), i.e $\mathcal{P}_{e,\sigma} \neq \emptyset$. Then for any $H \in L^0_+(\mathcal{F}_T)$, we have $\pi^S(H) \geq \sup_{Q \in \mathcal{P}_{e,\sigma}} E_Q[H]$.*

Proof. Since $\inf \emptyset = \infty$ by convention, w.l.o.g. we assume that there exists $v_0 \in \mathbb{R}$, $\theta \in \Theta_{adm}$ s.t. $v_0 + G_T(\theta) \geq H$ P-a.s.. For any $Q \in \mathcal{P}_{e,\sigma}$, $G(\theta) = \int \theta dS$ is a stochastic integral, and greater than some constant. It is also a Q-local martingale and hence by thm 4.3 a Q-supermartingale.

Therefore $E_Q[H] \leq v_0 + E_Q[G_T(\theta)] \leq v_0$, so by taking the supremum over Q in $\mathcal{P}_{e,\sigma}$ and infimum over v_0 satisfying the appropriate conditions laid out in the definition above, we get the claim. $\square$

The goal is to show that equality holds in Lemma 7.2 and that the infimum for $\pi^S(H)$ is attained.

Fix $H \in L^0_+(\mathcal{F}_T)$ and define the following "process" (it is $[0, \infty]$-valued)

$$U_t := \operatorname*{ess\,sup}_{Q \in \mathcal{P}_{e,\sigma}} E_Q[H|\mathcal{F}_t] \quad 0 \leq t \leq T \tag{7.4}$$

If $\mathcal{F}_0$ is trivial, then $U_0 = \sup_{Q \in \mathcal{P}_{e,\sigma}} E_Q[H]$.

Theorem 7.3. *Assume (NFLVR), i.e. $\mathcal{P}_{e,\sigma} \neq \emptyset$. Take $H \in L^0_+(\mathcal{F}_T)$. If $\sup_{Q \in \mathcal{P}_{e,\sigma}} E_Q[H] < \infty$, then U is a Q-supermartingale for all $Q \in \mathcal{P}_{e,\sigma}$.*

Proof. U adapted and nonnegative. We fix Q in $\mathcal{P}_{e,\sigma}$ and show Q-supermartingale property: $E_Q[U_t|\mathcal{F}_s] \leq U_s$ for all $s < t$. Then Q-integrability follows in the same way via $E_Q[U_t] \leq \sup_{Q \in \mathcal{P}_{e,\sigma}} E_Q[H]$.

Fix $Q \in \mathcal{P}_{e,\sigma}$. For every $t \in [0, T]$, consider the set $\mathcal{Z}_t := \{Z : Z$ density process of some $R \in \mathcal{P}_{e,\sigma}$ w.r.t. Q s.t. $R = Q$ on $\mathcal{F}_t\}$ where the latter condition means that $Z_s = 1$ for $s \leq t$. Now, $\mathcal{Z}_t \neq \emptyset$, since $Z \equiv 1 \in \mathcal{Z}_t$ (for $R = Q$). For all $s < t$, $\mathcal{Z}_t \subseteq \mathcal{Z}_s$.

Claim: $\mathcal{Z}_t = \{\frac{Z^R_{t\vee\cdot}}{Z^R_t} : Z^R$ density of $R \in \mathcal{P}_{e,\sigma}$ w.r.t. $Q\}$.

proof of claim: "$\subseteq$" for $Z = Z^R \in \mathcal{Z}_t$ with $Z_s = 1$ for $s \in [0, t]$ so

$$Z_\cdot = \mathbf{1}(\cdot \leq t) + Z_\cdot \mathbf{1}(\cdot > t) = \frac{Z^R_{t\vee\cdot}}{Z^R_t} \tag{7.5}$$

"$\supseteq$" for $R \in \mathcal{P}_{e,\sigma}$ with density process Z^R w.r.t. Q, we define $Z_\cdot = \frac{Z^R_{t\vee\cdot}}{Z^R_t}$, hence $Z > 0$, $Z_s = 1$ for all $s \leq t$. Since Z^R is a Q-martingale, so is Z, hence Z is a density process of some $R' \sim Q \sim P$ w.r.t. Q.

Since $Q \in \mathcal{P}_{e,\sigma}$ and $R \in \mathcal{P}_{e,\sigma}$, both S and $Z^R S$ (by Bayes) are Q-σ-martingales and then so is

$$Z_\cdot S_\cdot = S_\cdot \mathbf{1}(\cdot \leq t) + \frac{Z^R_\cdot S_\cdot}{Z^R_t} \mathbf{1}(\cdot > t) \tag{7.6}$$

and then S is a R'-σ-martingale by Bayes and $R' \in \mathcal{P}_{e,\sigma}$ which proves the claim.

By Bayes rule

$$U_t = \operatorname*{ess\,sup}_{R \in \mathcal{P}_{e,\sigma}} E_R[H|\mathcal{F}_t] \tag{7.7}$$
$$= \operatorname*{ess\,sup}_{R \in \mathcal{P}_{e,\sigma}} E_Q[H \frac{Z^R_T}{Z^R_t}|\mathcal{F}_t] \tag{7.8}$$
$$= \operatorname*{ess\,sup}_{Z \in \mathcal{Z}_t} E_Q[H Z_T|\mathcal{F}_t] =: \operatorname*{ess\,sup}_{Z \in \mathcal{Z}_t} \Gamma_t(Z) \tag{7.9}$$

We are going to show that the family $\{\Gamma_t(Z) : Z \in \mathcal{Z}_t\}$ is directed upward (if I take two elements in it, then their supremum is also an element of the set):

- if $Z, Z' \in \mathcal{Z}_t$, $A \in \mathcal{F}_t$, then $Z\mathbf{1}_A + Z'\mathbf{1}_{A^C} \in \mathcal{Z}_t$ (exercise)
- and for this element we have

$$\Gamma_t(Z'') = E_Q[H(Z\mathbf{1}_A + Z'\mathbf{1}_{A^C})|\mathcal{F}_t] \tag{7.10}$$
$$= \mathbf{1}_A E_Q[HZ_T|\mathcal{F}_t] + \mathbf{1}_{A^C} E_Q[HZ'_T|\mathcal{F}_t] \tag{7.11}$$
$$= \max\{\Gamma_t(Z), \Gamma_t(Z')\} \tag{7.12}$$

for the specific choice of $A := \{\Gamma_t(Z) \geq \Gamma_t(Z')\} \in \mathcal{F}_t$

which shows upward-directedness.

Thanks to this property, there exists $(Z^n_t)_n \subseteq \mathcal{Z}_t$ s.t.

$$U_t = \nearrow \lim_{n\to\infty} \Gamma_t(Z^n) \tag{7.13}$$
$$= \nearrow \lim_{n\to\infty} E_Q[HZ^n_T|\mathcal{F}_t] \tag{7.14}$$

for any $s \leq t$, take conditional expectation and use monotone integration

$$E_Q[U_t|\mathcal{F}_s] = \nearrow \lim_{n\to\infty} E_Q[HZ_T^n|\mathcal{F}_s] \tag{7.15}$$
$$\leq \operatorname*{ess\,sup}_{Z\in\mathcal{Z}_s} E_Q[HZ_T|\mathcal{F}_s] = U_s \tag{7.16}$$

implying that U satisfies the Q-supermartingale property, for all $Q \in \mathcal{P}_{e,\sigma}$. Furthermore U is clearly adapted.

□

One can show that U admits RCLL version (see e.g. Kramkov, Dellacherie, Meyer). We still call this version U.

7.2 Optional decomposition

Example of a process which is a Q-supermartingale for all $Q \in \mathcal{P}_{e,\sigma}$: Start with $x \in \mathbb{R}$, $\theta \in L(S)$, C adapted, increasing RCLL and $C_0 = 0$. Define

$$V^{x,\theta,C} := x + \int \theta dS - C \tag{7.17}$$

where x is initial capital, θ is holdings on risky assets and C is the cumulative consumption process (C_t is consumption from 0 to t). (x, θ, C) is a generalized strategy.

Note that $C \geq 0$, $V^{x,\theta,C} + C = x + \int \theta dS$, so

- if $V^{x,\theta,C}$ is bounded from below, then the strategy $\varphi = (x, \theta)$ is admissible
- wherever φ admissible, then by Ansel/Stricker (thm 4.3) we know that $x + \int \theta dS$ is a Q-supermartingale for all $Q \in \mathcal{P}_{e,\sigma}$
- if $V^{x,\theta,C}$ bounded below, then $0 \leq C \leq$ const. $+ \int \theta dS$ Q-integrable. Therefore $V^{x,\theta,C} = x + \int \theta dS - C$ also a Q-supermartingale for all $Q \in \mathcal{P}_{e,\sigma}$

Next result shows that example above is the only example of processes with that property.

Theorem 7.4. *(Optional decomposition theorem)*
Let S satisfy (NFLVR) ($\mathcal{P}_{e,\sigma} \neq \emptyset$). If a process $U \geq 0$, is a Q-supermartingale for all $Q \in \mathcal{P}_{e,\sigma}$, U_0 bounded, then there exists $\theta \in \Theta_{adm}$, and C adapted increasing RCLL with $C_0 = 0$ s.t.

$$U = U_0 + \int \theta dS - C \tag{7.18}$$

(if $U_0 \in \mathbb{R}$, e.g. if $\mathcal{F}_0$ is trivial, then $U = V^{U_0,\theta,C}$)

Proof. For proof in full generality, see Kramkov and Föllmer-Kabanov (quite involved). We will show the statement in the case where $\mathbb{F}$ is continuous, i.e. all local martingales (under P, equivalently, under any $Q \sim P$ on $\mathcal{F}_T$) are continuous. (classical example $\mathbb{F} = \mathbb{F}^W$)

Step 1: Parametrize EσMMs Q via density processes as in chapter 5.

Since S satisifes (NFLVR), S is a σ-martingale under some $Q \sim P$ on $\mathcal{F}_T$, so it's a stochastic integral for some local Q-martingale, therefore S is continuous. So we write $S = S_0 + M + A$ where $M \in \mathcal{M}^c_{0,loc}(P)$ and $A = \int \lambda d\langle M\rangle$ with λ predictable in $L^2_{loc}(M)$. Minimal EσMD: $\hat{Z} := \mathcal{E}(-\int \lambda dM)$ hence $\hat{Z}S \in \mathcal{M}_{loc}(P)$ by continuity of S.

For simplicity (but not needed) assume that $\hat{Z}$ is a time P-martingale. Then minimal EσMM/ELMM $\hat{P}$ for S exists with density $\hat{Z}$ w.r.t. P, so $S \in \mathcal{M}_{loc}(\hat{P})$.

By Cor. 5.5, any $Q \in \mathcal{P}_{e,\sigma}$ has density process w.r.t. P of the form $Z = Z_0\hat{Z}\mathcal{E}(N)$, with $N \in \mathcal{M}_{0,loc}(P)$ and $N \perp M$ (since $\mathbb{F}$ continuous we actually have $N \in \mathcal{M}^c_{0,loc}(P)$).

Moreover, $Z_0\hat{Z}\mathcal{E}(N) = Z$ is a P-martingale (since it is the density process of Q w.r.t. P) so $Z_0\mathcal{E}(N)$ is a $\hat{P}$-martingale, > 0 and so a density process of Q w.r.t. $\hat{P}$.

Step 2: Decompose U under $\hat{P}$ via Doob-Meyer and Kunita-Watanabe.

Since U is a $\hat{P}$-supermartingale , it has a D-M decomposition $U = U_0 + \hat{L} - \hat{B}$ where $\hat{L} \in \mathcal{M}_{0,loc}(\hat{P})$ and $\hat{B}$ is adapted, increasing, $\hat{B}_0 = 0$ RCLL, predictable. We now use K-W decomposition $\hat{L} = \int \hat{\theta} dS + \hat{N}$ where $\hat{N} \in \mathcal{M}_{0,loc}(\hat{P})$ and $\hat{N} \perp S$. Note: $\langle N, M\rangle = \langle N, S\rangle \equiv 0$.

We have $\hat{Z} = \mathcal{E}(-\int \lambda dM)$ and so $\langle \hat{N}, \hat{Z}\rangle \equiv 0$. Now using Itô's formula and continuity of M and $\hat{Z}$ to find $\langle \hat{N}, \frac{1}{\hat{Z}}\rangle \equiv 0$. Hence also the two $\hat{P}$-local martingales $\hat{N}$ and $\frac{1}{\hat{Z}}$ are strongly orthogonal, therefore $\hat{N}\frac{1}{\hat{Z}}$ is also a $\hat{P}$-local martingale (by the product rule.) Note that $\frac{1}{\hat{Z}} = \frac{dP}{d\hat{P}}$. We conclude that $\hat{N}$ is a P-local martingale. So we have:

$$U - U_0 = \hat{L} - \hat{B} = \int \hat{\theta} dS - \hat{B} + \hat{N} \tag{7.19}$$

Step 3: Use Girsanov to express $\hat{N}$ under $Q \in \mathcal{P}_{e,\sigma}$ and use Q-supermartingale property of U to get information about $\hat{N}$.

Take any $Q \in \mathcal{P}_{e,\sigma}$ and write its density w.r.t. P as $Z = \hat{Z} Z_0 \mathcal{E}(N)$. Recall that U is a Q-supermartingale and has the decomposition from the end of step 2. We use Girsanov from $\hat{P}$ to Q: $\hat{N}$ is $\hat{P}$-local martingale, $Z_0 \mathcal{E}(N) = Z^{\hat{P},Q}$ hence

$$L^Q := \hat{N} - \frac{1}{Z^{\hat{P},Q}} \bullet \langle Z^{\hat{P},Q}, \hat{N} \rangle \tag{7.20}$$

is a Q-local martingale, with the latter term equal to $\langle N, \hat{N} \rangle$. So under Q, from the decomposition of step 2 we get

$$U - U_0 = (\int \hat{\theta} dS + L^Q) - (\hat{B} - \langle \hat{N}, N \rangle) \tag{7.21}$$

with the first term a Q-local martingale and the second predictable FV. Since this is a unique decomposition it must coincide with the Doob-Meyer decomposition of U Q-supermartingale, so the second term is monotone, that is

$$\hat{B} - \langle \hat{N}, N \rangle \tag{7.22}$$

is increasing.

Step 4: We will show that $\hat{N} \equiv 0$. The last line from step 3 holds for any $N \in \mathcal{M}^c_{0,loc}(P)$, $N \perp M$. In particular, we can take $N = \alpha \hat{N}$ for any $\alpha > 0$, meaning that

$$\hat{B} - \alpha \langle \hat{N} \rangle \tag{7.23}$$

is increasing for all $\alpha > 0$. Thus $\langle \hat{N} \rangle \equiv 0$, so $\hat{N} \equiv 0$ and $L^Q \equiv 0$.

So looking at the decomposition from step 3 we get

$$U - U_0 = \int \hat{\theta} dS - \hat{B} \tag{7.24}$$

so that

$$\int \hat{\theta} dS = U - U_0 + \hat{B} \geq \text{const.} \tag{7.25}$$

and we can conclude that $\hat{\theta}$ is admissible.

□

7.3 Hedging duality

A key consequence of the optional decomposition theorem is that in lemma 7.2 we actually have equality (we get a hedging duality result).

Theorem 7.5. *Assume (NFLVR) ($\mathcal{P}_{e,\sigma} \neq \emptyset$), $\mathcal{F}_0$ trivial. Then for any $H \in L^0_+(\mathcal{F}_T)$,*

$$\pi^S(H) := \inf\{v_0 \in \mathbb{R} : H - v_0 \in \mathcal{G}_{adm} - L^0_+\} \tag{7.26}$$

$$= \sup_{Q \in \mathcal{P}_{e,\sigma}} E_Q[H] \tag{7.27}$$

Moreover, infimum is a minimum if $\sup_{Q\in\mathcal{P}_{e,\sigma}} E_Q[H] < \infty$.

Proof. "$\geq$" was proved in lemma 7.2. To show "$\leq$", let U be RCLL version of $U_t := \operatorname{ess\,sup}_{Q\in\mathcal{P}_{e,\sigma}} E_Q[H|\mathcal{F}_t]$, $0 \leq t \leq T$. Note that

- $U_T = H$
- $U_0 = \sup_{Q\in\mathcal{P}_{e,\sigma}} E_Q[H] < \infty$

By thm 7.3., U is a Q-supermartingale for all $Q \in \mathcal{P}_{e,\sigma}$. Therefore, by thm 7.4, it has a decomposition $U = U_0 + \int \theta dS - C$ so we have

$$H - U_0 = U_T - U_0 = G_T(\theta) - C_T \in \mathcal{G}_{adm} - L^0_+ \tag{7.28}$$

Therefore $\pi^S(H) \leq U_0$ and U_0 is the minimizer (attains infimum).

□

Comments:

- $\pi^S(H)$ super-replication price is good from seller's perspective, since we can set up a portfolio with this initial capital that at time T gives a value always greater than H.
- We may find no buyer willing to buy H at this price. For example, it can happen that H is bounded and $\pi^S(H) = \|H\|_{L^\infty}$ or that $S \geq 0$ and $\pi^S((S_T - K)_+) = S_0$.
- So far we took $H \geq 0$ (equivalently we can take H bounded from below) and considered $\pi^S(H)$. Similarly we can take H bounded from above and define the buyer's price $\pi^B(H) =$ supremum of initial price to set up a sub-replicating strategy for H. We can show that $\pi^B(H) = \inf_{Q\in\mathcal{P}_{e,\sigma}} E_Q[H]$. For H bounded (from above and below): $\pi^B(H) = -\pi^S(-H)$.

- One drawback of the above considerations is that they are for H bounded. Moreover, for $H \geq 0$ we don't get much information if $\mathcal{G}_{adm} = \{0\}$ since in this case

$$\{v_0 \in \mathbb{R} : H - v_0 \in \mathcal{G}_{adm} - L^0_+\} = \{v_0 \in \mathbb{R} : H \leq v_0\} \neq \emptyset \qquad (7.29)$$

  is equivalent to H bounded from above. We need analysis for more general (non-bounded) payoffs.

Remark: As in finite discrete time (see chapter 6), we can use optional decomposition or hedging duality to characterize payoffs "attainable" in a suitable sense. More precisely we can call $H \in L^0_+(\mathcal{F}_T)$ attainable if

- $H = H_0 + G_T(\theta)$ for some $(H_0, \theta) \in \mathbb{R} \times \Theta_{adm}$ with the extra property that
- $G(\theta)$ is a true Q^*-martingale, for some $Q^* \in \mathcal{P}_{e,\sigma}$

In this case one can show that H is attainable in the above sense iff $\sup_{Q \in \mathcal{P}_{e,\sigma}} E_Q[H] < \infty$ and attained at some $Q^* \in \mathcal{P}_{e,\sigma}$.

Chapter 8

Duality in the unbounded case

Goal: Extend hedging duality (Thm 7.5) to possibly unbounded payoffs. Idea: Random lower bound for $G.(\theta)$

8.1 Feasible weight functions and (a, W)-admissibility

Recall that under (NFLVR), there exists an EσMM Q for S s.t. $S - S_0 = \int \psi dM$ where $M \in \mathcal{M}_{0,loc}(Q)$, $\psi > 0$ one-dimensional and $\psi \in L(M^i)$. We have $dS = \psi dM$.

Even more, we can get $M \in \mathcal{H}^1_0(Q)$, i.e. $M^*_T := \sup_{t\in[0,T]} |M_t| \in L^1(Q)$. Let $\xi^\psi := \frac{1}{\psi}$, so $dM = \xi^\psi dS$, $\xi^\psi > 0$

$$M^*_T = \sup_{t\in[0,T]} \left| \int_0^t \xi^\psi_u dS_u \right| = (\xi^\psi \bullet S)^*_T \in L^1(Q) \tag{8.1}$$

Moreover, if S is locally bounded and Q is an ELMM for S, we can even choose $\xi^\psi = \sum_{n=1}^\infty \gamma_n \mathbf{1}_{]\tau_{n-1},\tau_n]}$ and $(\xi^\psi \bullet S)^*_T \leq$ const..

Definition 8.1. *Feasible weight function: random variable W s.t. $W \geq 1$ satisfying*

1. $\exists 0 < \xi \in L(S)$ *s.t.* $(\xi \bullet S)^*_T \leq W$
2. $\exists Q \in \mathcal{P}_{e,\sigma}$ *s.t.* $E_Q[W] < \infty$, *that is* $W \in \bigcup_{Q\in\mathcal{P}_{e,\sigma}} L^1(Q)$

Remarks:

1. If S satisfies (NFLVR), as seen above there exists a feasible weight function.

2. w.l.o.g. we can take W $\mathcal{F}_T$-measurable.

3. If S satisfying (NFLVR) is also locally bounded every constant greater or equal than one is f.w.f.

4. Good (random) lower bound for $G_\cdot(\theta)$ should be large enough so we can use many strategies, so handle many payoffs, but not too large to avoid doubling strategies.

Notation: For any random variable $W \geq 1$, define $\mathcal{P}^W_{e,\sigma} := \{Q \in \mathcal{P}_{e,\sigma} : E_Q[W] < \infty\}$ (for $W \equiv 1$, $\mathcal{P}^W_{e,\sigma} = \mathcal{P}_{e,\sigma}$). If S satisfies (NFLVR) and W f.w.f., then $\mathcal{P}^W_{e,\sigma} \neq \emptyset$.

Definition 8.2. *Let $W \geq 1$ be a random variable with $E_Q[W] < \infty$ for some $Q \in \mathcal{P}_{e,\sigma}$. For a constant a, we call an integrand θ $(a, W)-$admissible if $G_t(\theta) \geq -aE_Q[W|\mathcal{F}_t]$ P-a.s. for all $t \in [0, T]$ and for all $Q \in \mathcal{P}_{e,\sigma}$.*

Then we write $\theta \in \Theta^a_W$, and set

$$\Theta_W := \bigcup_{a \geq 0} \Theta^a_W \tag{8.2}$$

with elements in this set called W-admissible.

$$\mathcal{G}^a_W := \{g = G_T(\theta) : \theta \text{ is } (a, W) - \text{admissible}\} \tag{8.3}$$

$$\mathcal{G}_W = \bigcup_{a \geq 0} \mathcal{G}^a_W \tag{8.4}$$

$$\mathcal{C}^\infty_W := \frac{1}{W}(\mathcal{G}_W - L^0_+) \cap L^\infty =: \mathcal{C}^0_W \cap L^\infty \tag{8.5}$$

The above notation denotes the W-analogues of $\mathcal{G}^a_{adm} = \mathcal{G}^a_1$ etc.

We can think of $\mathcal{C}^\infty_W$ as the set of all bounded W-discounted payoffs that can be superreplicated with zero initial wealth by a W-admissible self-financing strategy. For $W \equiv 1$, W-admissible reduces to admissible. For clarity, use $(a, 1)$-admissible instead of admissible.

Lemma 8.3. *Assume (NFLVR), let $W \geq 1$ be a random variable s.t. $E_Q[W] < \infty$ for some $Q \in \mathcal{P}_{e,\sigma}$. For any W-admissible θ, the stochastic integral process $G(\theta)$ is a Q-supermartingale. This implies*

1. *if $g \in \mathcal{G}_W$ and $Q \in \mathcal{P}^W_{e,\sigma}$ we have $E_Q[g] \leq 0$*

2. *if* $g \in \mathcal{G}_{adm}$ *and* $Q \in \mathcal{P}_{e,\sigma}$ *we have* $E_Q[g] \leq 0$

3. *if* $f \in \mathcal{G}_W - L^0_+$ *and* $Q \in \mathcal{P}^W_{e,\sigma}$ *we have* $E_Q[f] \leq 0$.

4. *Suppose* $g \in \mathcal{G}_W$ *so that* $g = G_T(\theta)$, θ W*-admissible and fix* $Q \in \mathcal{P}_{e,\sigma}$. *If* $\tilde{W} \geq W$ *s.t.* $E_Q[\tilde{W}] < \infty$ *and if* $g \geq -\tilde{W}$ P*-a.s. then* θ *is* $(1, \tilde{W})$*-admissible and in particular* $g \in \mathcal{G}^1_{\tilde{W}} \subseteq \mathcal{G}_{\tilde{W}}$

Proof. Since $Q \in \mathcal{P}_{e,\sigma}$, we have $G(\theta) = \int \theta\psi dM$ for some $M \in \mathcal{M}_{0,loc}(Q)$. Moreover, since θ is W-admissible and $E_Q[W] < \infty$ then $G(\theta) \geq -aE_Q[W|\mathbb{F}]$. Hence $G(\theta)$ is bounded below by a Q-martingale.

So $G(\theta)$ is a local Q-martingale and a Q-supermartingale by Thm 4.3, lemma 4.4 and subsequent remark. Then the rest follows easily:

1. $E_Q[g] = E_Q[G_T(\theta)] \leq E_Q[G_0(\theta)] = 0$ by the supermartingale property

2. follows from 1.) for $W \equiv 1$

3. Follows from 1.) since $f \leq g$ for some $g \in \mathcal{G}_W$ (and $g = G_T(\theta) \in L^1(Q)$ by 1.), hence $f^+ \in L^1(Q)$)

4. Take any $Q \in \mathcal{P}_{e,\sigma}$. If $E_Q[\tilde{W}] < \infty$, then $E_Q[W] < \infty$. Use Q-supermartingale property of $G(\theta)$ to get

$$G_t(\theta) \geq E_Q[G_T(\theta)|\mathcal{F}_t] \tag{8.6}$$
$$\geq -E_Q[\tilde{W}|\mathcal{F}_t] \tag{8.7}$$

   P-a.s. for all t.

□

Remark: Part 4.) of Lemma 8.3 is an extension of a result that under (NA), admissible θ with $G_T(\theta) \geq -b$ implies automatically that θ is b-admissible. (Same bound also for all $t \in [0, T]$)

8.2 A closedness result

We still consider setup with $\mathbb{R}^d$-valued semimartingale S on $(\Omega, \mathcal{F}, \mathbb{F}, P)$ over $[0, T]$ for $\tilde{S}^0$ discounted prices ($S^0 \equiv 1$). We still assume (NFLVR) for S and we want to extend hedging duality Theorem 7.5 to possibly unbounded payoffs H.

From chapter 4, a crucial consequence of (NFLVR) is Theorem 4.12 which says

$$\mathcal{C}^\infty = \mathcal{C}^\infty_{adm} = (\mathcal{G}_{adm} - L^0_+) \cap L^\infty \tag{8.8}$$

is w^*-closed in L^∞, $\mathcal{G}^1_{adm} - L^0_+$ is Fatou-closed where $\mathcal{G}^1_{adm} = \{g = G_T(\theta),\ \theta\ (1,1)-$adm.$\}$.

But for general unbounded S, we already pointed out that $\mathcal{G}_{adm}$ may collapse to $\{0\}$. So we need to look for larger sets (else super-replication with admissible strategies is impossible for unbounded payoffs).

We want to generalize the Fatou-closedness result to the current case.

Theorem 8.4. *Assume (NFLVR) and let $W \geq 1$ random variable s.t. there exists $Q \in \mathcal{P}_{e,\sigma}$ .t. $E_Q[W] < \infty$. Then the convex set $\mathcal{G}^1_W - L^0_+$ is closed in L^0.*

The proof is very technical (compactness results for bounded sequences of martingales). For proof, see DS Chapter 15 (C. 15.4.11).

Once we admit this, we have important consequences. First one is an analogue of theorem 4.12.

Corollary 8.5. *Suppose S satisfies (NFLVR) and let $W \geq 1$ a random variable s.t. $E_Q[W] < \infty$ for some $Q \in \mathcal{P}_{e,\sigma}$. Then the convex cone*

$$\mathcal{C}^\infty_W = \frac{1}{W}(\mathcal{G}_W - L^0_+) \cap L^\infty \tag{8.9}$$

is weak-closed in L^∞.*

Proof. Since $\mathcal{C}^\infty_W \subseteq L^\infty$ convex cone, we can use some criterion for weak*-closedness as in the first step of Thm. 4.12. Take a sequence $(f_n) \subseteq \mathcal{C}^\infty_W$ uniformly bounded by 1. and converging to some f P-a.s., then we show $f \in \mathcal{C}^\infty_W$:

Since $|f_n| \leq 1$ for all n we have $|f| \leq 1$ and so $f \in L^\infty$. Since $f_n \in \mathcal{C}^0_W = \frac{1}{W}(\mathcal{G}_W - L^0_+)$ for all n, then $\exists g_n \in \mathcal{G}_W$ s.t. $G_T(\theta^n) = g_n \geq W f_n \geq -W$. (Here θ^n is W-admissible)

Then, lemma 8.3(4), with $\tilde{W} = W$ given that θ^n is even $(1, W)$-admissible, so that $g_n \in \mathcal{G}^1_W$ so that $W f_n \in \mathcal{G}^1_W - L^0_+$.

But now we know that $W f_n \to W f$ P-a.s. and we use Thm. 8.4 for which $\mathcal{G}^1_W - L^0_+$ is L^0-closed, hence $Wf \in \mathcal{G}^1_W - L^0_+ \subseteq G_W - L^0_+$, or equivalently $f \in \frac{1}{W}(\mathcal{G}_W - L^0_+) = \mathcal{C}^\infty_W$.

□

8.3 Duality result using the bipolar theorem

We want to obtain a duality result by applying the bipolar theorem to the convex cone $\mathcal{C}_w^\infty \subseteq L^\infty$.

With a little abuse of notation, identify any probability $R \ll P$ with its density $\frac{dR}{dP} \in L^1(P)$ so we see $\mathcal{P}_{e,\sigma}$ as a subset of $L^1(P)$.

Notation:

$$W\mathcal{P}_{e,\sigma}^W \hat{=} \{WY : Y = \frac{dQ}{dP}, \quad Q \in \mathcal{P}_{e,\sigma}^W\} \tag{8.10}$$

Note that $E_P[WY] = E_Q[W] < \infty$ which gives that $W\mathcal{P}_{e,\sigma}^W \subseteq L^1(P)$.

More notation:

$$\mathcal{P}_{a,\sigma} := \{Q \ll P \text{ on } \mathcal{F}_T \text{ s.t. } S \text{ is a } Q\text{-}\sigma\text{-martingale}\} \tag{8.11}$$

and note that if $\mathcal{P}_{e,\sigma}^W \neq \emptyset$ then (exercise) it is $L^1(P)$-dense in $\mathcal{P}_{a,\sigma}^W = \{Q \in \mathcal{P}_{a,\sigma}$ s.t. $E_Q[W] < \infty\}$. In the same way, $W\mathcal{P}_{e,\sigma}^W$ is $L^1(P)$-dense in $W\mathcal{P}_{a,\sigma}^W$.

Theorem 8.6. *If S satisfies (NFLVR) and W is a feasible weight function, then*

1. *The polar of $\mathcal{C}_W^\infty$ in $L^1(P)$ is equal to cone$(W\mathcal{P}_{a,\sigma}^W)$. From this we obtain for all $f \in L^\infty$*
2. *$f \in \mathcal{C}_W^\infty$ iff $E_Q[Wf] \leq 0$ for all $Q \in \mathcal{P}_{e,\sigma}^W$.*

Recall: For a set $A \subset L^\infty$, the polar of A in L^1 is the set $A^\circ := \{Y \in L^1 : E[YX] \leq 1$ for all $X \in A\}$. If A is a convex cone with a vertex at zero, the 1 is replaced by a 0 in the inequality.

Proof. Since $W\mathcal{P}_{e,\sigma}^W$ is $L^1(P)$-dense in $W\mathcal{P}_{a,\sigma}^W$, RHS of the second item in the theorem says that f is in the polar of cone$(W\mathcal{P}_{a,\sigma}^W)$. On LHS we have that $\mathcal{C}_W^\infty$ is a convex cone with vertex at 0 and weak*-closed by Cor. 8.5, so $\mathcal{C}_W^\infty = (\mathcal{C}_W^\infty)^{\circ\circ}$ by the bipolar theorem (see Thm. D.1.). So the LHS of the second item says that f is in the polar of (polar of $\mathcal{C}_W^\infty$) and so the second item would follow from the first.

To see the first item: We begin with "$\supseteq$" in the first item, take $Q \in \mathcal{P}_{e,\sigma}^W$ and $f \in \mathcal{C}_W^\infty$ so that $Wf \in \mathcal{G}_W - L_+^0$. By Lemma 8.3(3), we have

$$E_P[fW\frac{dQ}{dP}] = E_Q[fW] \leq 0 \tag{8.12}$$

for all $f \in \mathcal{C}_W^\infty$. This extends to all $Q \in \mathcal{P}_{a,\sigma}^W$ by $L^1(P)$-denseness so the inclusion follows because $(\mathcal{C}_W^\infty)^\circ$ is a cone.

For the converse "$\subseteq$", by definition of polar set, every $Y \in (\mathcal{C}_W^\infty)^\circ$ is in $L^1(P)$ and satisfies $E_P[Yf] \leq 0$ for all $f \in \mathcal{C}_W^\infty$. Since $W \geq 1$, $\frac{Y}{W} \in L^1(P)$ and for any $h \in \mathcal{G}_W - L_+^0$ s.t. $\frac{h}{W}$ is bounded, i.e. $\frac{h}{W} \in \mathcal{C}_W^\infty$, we have

$$E_P[\frac{Y}{W}h] = E_P[Y\frac{h}{W}] \leq 0 \tag{8.13}$$

Take $h := -\mathbf{1}(Y < 0)$ and note that $\frac{h}{W}$ is bounded since $W \geq 1$ hence the previous eq. holds for this h which gives $Y \geq 0$ P-a.s., so we can define $\frac{dQ}{dP} :=$ const.$\frac{Y}{W}$, with the constant chosen to yield P-expectation 1, which gives a measure $Q \ll P$. With $E_Q[W] =$ const.$E_P[Y] < \infty$ and from the previous equation again $E_Q[h] \leq 0$ for all $h \in \mathcal{G}_W - L_+^0$ with $\frac{h}{W}$ bounded.

Below we will show that this implies that $Q \in \mathcal{P}_{a,\sigma}$ and so

$$Y = W\frac{Y}{W} = \text{const.}W\frac{dQ}{dP} \in \text{cone}(W\mathcal{P}_{a,\sigma}^W) \tag{8.14}$$

which shows "$\subseteq$" in the first item.

It then remains to show that S is a Q-σ-martingale. Since W is a feasible weight function, there exists $\xi \in L(S)$, $\xi > 0$ and s.t. $(\xi \bullet S)_T^* \leq W$. So if we take $s \leq t$ and $A \in \mathcal{F}_s$, $h_\pm^i := \pm I_A \int_s^t \xi_u dS_u^i \in \mathcal{G}_W$ and $h_\pm^i$ is bounded so that

$$E_Q[I_A \int_s^t \xi_u dS_u^i] = 0 \tag{8.15}$$

this is the Q-martingale property. Since also $E_Q[W] < \infty$, this shows that $M := \int \xi dS$ is a Q-martingale, and so $S - S_0 = \int \frac{1}{\xi} dM$ and therefore S is a Q-σ-martingale.

□

The duality in the second item of the previous theorem only holds for bounded payoffs $f \in L^\infty$. However, we can relax this condition:

Corollary 8.7. *Assume (NFLVR) and W f.w.f.. Then for any $f \geq -W$ we have that $f \in \mathcal{G}_W - L_+^0$ iff $E_Q[f] \leq 0$ for all $Q \in \mathcal{P}_{e,\sigma}^W$.*

Proof. " $\Longrightarrow$ " this is Lemma 8.3(3)

" $\Longleftarrow$ ": Since $f \geq -W$, then $\frac{f \wedge n}{W}$ is bounded from above and also from below as $W \geq 1$. So

$$E_Q[W \frac{f \wedge n}{W}] \leq E_Q[f] \leq 0 \tag{8.16}$$

for all $Q \in \mathcal{P}_{e,\sigma}^W$. Then Thm. 8.6 implies that $f_n \in \mathcal{C}_W^\infty \subseteq \frac{1}{W}(\mathcal{G}_W - L_+^0)$. So for each n, there exists $g_n \in \mathcal{G}_W$ with $g_n \geq W f_n \geq -W$.

Due to Lemma 8.3(4) with $\tilde{W} = W$, each g_n is in $\mathcal{G}_W^1$. Then $(W f_n)_n$ is a sequence in $\mathcal{G}_W^1 - L_+^0$ converging to f P-a.s. Since $\mathcal{G}_W^1 - L_+^0$ is closed in L^0 by Thm. 8.4 we get $f \in \mathcal{G}_W^1 - L_+^0 \subseteq \mathcal{G}_W - L_+^0$.

$\square$

The previous corollary 8.7 extends the hedging duality of theorem 7.5 to unbounded payoffs.

Remark: If S satisfies (NFLVR), is locally bounded, any constant ≥ 1 is a f.w.f.. Then Corollary 8.7 characterizes all f bounded from below that can be super-replicated via admissible strategies, and it says that this is possible iff one can start from initial capital at least $\sup_{Q \in \mathcal{P}_{e,\sigma}} E_Q[f]$.

Note: Take f_0 initial capital, then $f - f_0 \in \mathcal{G}_W - L_+^0$ iff $\sup_{Q \in \mathcal{P}_{e,\sigma}^W} E_Q[f] \leq f_0$.

Chapter 9

Superreplication, pricing and hedging

Goal: Study hedging and pricing for non-necessarily bounded payoffs. Study the seller's and buyer's price and see when they agree.

9.1 (W)-super- and sub-replication

Setting: $(\Omega, \mathcal{F}, \mathbb{F}, P)$ over $[0, T]$, $S^0 \equiv 1$, S an $\mathbb{R}^d$-valued semimartingale satisfying (NFLVR) (so $\mathcal{P}_{e,\sigma} \neq \emptyset$). Fix W a feasible weight function and note that most of the results we will see do depend on this weight function, even though it will not necessarily be made explicit in the notation.

Definition 9.1. *Fix $f \in L^0(\mathcal{F}_T)$. If $f \geq -W$, call f (W)-super-replicable at price $a \in \mathbb{R}$ if $f - a \in \mathcal{G}_W - L^0_+$.*

Set

$$\Gamma_+ := \{a \in \mathbb{R} : f - a \in \mathcal{G}_W - L^0_+\} \tag{9.1}$$

and call

$$\alpha := \alpha(f) := \inf \Gamma_+ \tag{9.2}$$

the superreplication price or seller price or ask price of f.

From the buyer's perspective: If $f \leq W$, call f (W)-subreplicable at price $b \in \mathbb{R}$ if $-f + b \in \mathcal{G}_W - L^0_+$. Set

$$\Gamma_- := \{b \in \mathbb{R} : -f + b \in \mathcal{G}_W - L^0_+\} \tag{9.3}$$

and call

$$\beta := \beta(f) = \sup \Gamma_- \tag{9.4}$$

the subreplication price, buyer's price or bid price of f.

Intuition: If we get a and use W-admissible strategy θ to generate $g = G_T(\theta)$ from 0 initial wealth, then $f - a \leq g$ means that the total result $a + g$ is enough to cover the payoff f. So we can sell for a without incurring risk. On the other hand: If we spend b and use some W-admissible strategy θ to generate g, then $-f + b \leq g$ means that total expense $b - g$ is still less than buying f outright at T, so should buy f at time 0 for b. Optimizing via infimum and supremum then gives competitive prices.

Remark: By definition, Γ_+ and Γ_- are both intervals and $\beta(f) = -\alpha(-f)$. So results for α directly translate to β.

The next theorem shows that α and β are attained if finite (analogue of thm. 7.5)

Theorem 9.2. *S satisfying (NFLVR) and W f.w.f.. If $f \geq -W$ and $\alpha(f) = \inf \Gamma_+$ is in $\mathbb{R}$, then $\alpha \in \Gamma_+$, i.e. $\Gamma_+ = [\alpha, \infty)$ closed interval and and there exists $g \in \mathcal{G}_W$ s.t. $a + g \geq f$ P-a.s. (f superreplicable at price α). Viceversa: if $f \leq W$ and $\beta(f) = \sup \Gamma_- \in \mathbb{R}$ then $\beta \in \Gamma_- = (\infty, \beta]$.*

Proof. It is enough to argue for α: Show that Γ_+ is closed. So we take $(a_n)_n \subseteq \Gamma_+$, with $a_n \searrow \alpha$. For each n there exists $g_n \in \mathcal{G}_W$ s.t. $f \leq a_n + g_n$. Since $a_n \searrow$, the sequence is bounded above by some constant $A > 0$, so $f \geq -W$ and $W \geq 1$ give $g_n \geq f - a_n \geq -W(1 + A) := -W'$. It is clear that $W' \geq W$, $\mathcal{G}_{W'} = \mathcal{G}_W$,

$$E_Q[W'] = (1 + A)E_Q[W] < \infty \tag{9.5}$$

for all $Q \in \mathcal{P}^W_{e,\sigma}$. By lemma 8.3(4) we have that for each $g_n \in \mathcal{G}^1_{W'}$

$$f - a_n \in \mathcal{G}^1_{W'} - L^0_+ \tag{9.6}$$

with the set on the RHS closed in L^0 by theorem 8.4 And furthermore that the LHS goes to $f - \alpha$ P-a.s., i.e. in L^0.

Hence the limit $f-\alpha$ is still in $\mathcal{G}^1_{W'} - L^0_+ \subseteq \mathcal{G}_{W'} - L^0_+ = \mathcal{G}_W - L^0_+$. Therefore $\alpha \in \Gamma_+$ and so $\Gamma_+ = [\alpha, \infty)$ is closed.

□

9.2 Duality

If $|f| \leq W$ and both α and β are finite, we expect that $\alpha \geq \beta$ (so that $\Gamma_+ \cap \Gamma_-$ contains at most one point) since we are happy to sell for α and buy for β. Let's show that indeed $\alpha \geq \beta$: For any $c \in \mathbb{R}$ with $f - c \in \mathcal{G}_W - L^0_+$, lemma 8.3(3) gives $E_Q[f-c] \leq 0$ for all $Q \in \mathcal{P}^W_{e,\sigma}$. If $-W \leq f \leq W$ and $\alpha, \beta \in \mathbb{R}$, then Thm. 9.2 gives that $f - \alpha$ and $-f + \beta$ are both in $\mathcal{G}_W - L^0_+$. Then $\alpha \geq \sup_{Q \in \mathcal{P}^W_{e,\sigma}} E_Q[f] \geq \inf_{Q \in \mathcal{P}^W_{e,\sigma}} E_Q[f] \geq \beta$. The next result shows that the outer inequalities in the string above are in fact equalities (cf. thm 7.5, hedging duality).

Theorem 9.3. *Let S satisfy (NFLVR) and W is f.w.f.. If $f \geq -W$, then*

$$\alpha = \inf\{a \in \mathbb{R} : f - a \in \mathcal{G}_W - L^0_+\} \tag{9.7}$$
$$= \sup_{Q \in \mathcal{P}^W_{e,\sigma}} E_Q[f] \tag{9.8}$$

and if these expressions are finite, then the infimum is attained. If $f \leq W$, then

$$\beta = \sup\{b \in \mathbb{R} : -f + b + \mathcal{G}_W - L^0_+\} \tag{9.9}$$
$$= \inf_{Q \in \mathcal{P}^W_{e,\sigma}} E_Q[f] \tag{9.10}$$

and if these are finite, the supremum is attained.

Proof. We only argue for α, and we have already shown "$\geq$" above. To show the reverse direction "$\leq$", assume w.l.o.g. that $\sup_{Q \in \mathcal{P}^W_{e,\sigma}} E_Q[f] < \infty$. For any $a < \alpha = \inf \Gamma_+$, then by definition of Γ_+ $f - a \notin \mathcal{G}_W - L^0_+$. But since $f \geq -W$ and $W \geq 1$, then $f - a \geq -(1+|a|)W =: -W'$ and W' is f.w.f. and $\mathcal{G}_W = \mathcal{G}_{W'}$. So $f - a \geq -W'$ and not in $\mathcal{G}_{W'} - L^0_+$, so Cor. 8.7 implies that $E_{Q^*}[f-a] > 0$ for some $Q^* \in \mathcal{P}^{W'}_{e,\sigma} = \mathcal{P}^W_{e,\sigma}$. Then

$$a < E_{Q^*}[f] \leq \sup_{Q \in \mathcal{P}^W_{e,\sigma}} E_Q[f] < \infty \tag{9.11}$$

Since this is true for all arbitrary $a < \alpha$, we get "$\leq$" in the first equation of the theorem and $\alpha < \infty$. So α is attained due to Thm.9.2. □

9.3 Strict super- and sub-replication

By the above results, any payoff f with $|f| \leq W$ should have a price in the interval $[\beta, \alpha] = [\inf_{Q \in \mathcal{P}_{e,\sigma}^W} E_Q[f], \sup_{Q \in \mathcal{P}_{e,\sigma}^W} E_Q[f]]$.

If $\beta = \alpha$, then this is the unique price (we will later see what this means for f). If $\beta < \alpha$, we have the following lemma:

Lemma 9.4. *Let S satisfy (NFLVR) and W be a f.w.f.. Let f be a bounded payoff, with $|f| \leq W$. Then if $-\infty < \beta(f) < \alpha(f) < \infty$, we have*

1. *f is strictly superreplicable at price α, i.e. there exists $g \in \mathcal{G}_W$ with $f \leq \alpha + g$ P-a.s. and $P(f < \alpha + g) > 0$.*

2. *f is strictly subreplicable at price β.*

Proof. Again, we content ourselves with the argument for α. Since $\alpha < \infty$, $\alpha \in \Gamma_+$ by Thm.9.2, then there exists $g_0 \in \mathcal{G}_W$ s.t. $f \leq \alpha + g_0$, with $g_0 = G_T(\theta^0)$, θ^0 W-admissible. Set

$$\tau := \inf\{t \in [0, T] : G_t(\theta^0) \geq 1 + |\alpha| + \|f\|_\infty\} \tag{9.12}$$

Then define

$$\theta = \theta^0 \mathbf{1}_{]0,\tau]} \in L(S) \tag{9.13}$$

and

$$G_t(\theta) = G_t(\theta^0)\mathbf{1}(t < \tau) + G_\tau(\theta^0)\mathbf{1}(t \geq \tau) \tag{9.14}$$

By right-continuity, $G_\tau(\theta^0) \geq 1 + |\alpha| + \|f\|_\infty \geq \max(0, 1 + f - \alpha)$ valid on $\{\tau \leq T\}$. Note that θ is also W-admissible

$$\begin{aligned}
\alpha + G_T(\theta) &= (\alpha + g_0)\mathbf{1}(T < \tau) + (\alpha + G_\tau(\theta^0))\mathbf{1}(T \geq \tau) && (9.15)\\
&\geq f\mathbf{1}(T < \tau) + (1 + f)\mathbf{1}(T \geq \tau) && (9.16)\\
&= f + \mathbf{1}(T \geq \tau) && (9.17)\\
&\geq f && (9.18)
\end{aligned}$$

P-a.s.. So $g = G_T(\theta) \in \mathcal{G}_W$ superreplicates f at price α. Moreover, if $P(\tau \leq T) > 0$, it does so strictly. And then in that case, the statement of the theorem is true.

To complete the proof, we show that if $\tau > T$ P-a.s., then g_0 itself is strictly superreplicating f at price α. Indeed, in this case then $G_\cdot(\theta^0)$ bounded above (by $1+|\alpha|+\|f\|_\infty$). For each $Q \in \mathcal{P}^W_{e,\sigma}$, $G_\cdot(\theta^0)$ is by W-admissiblity also bounded from below by a constant times the Q-martingale $E_Q[W|\mathbb{F}]$. So by Ansel-Stricker (thm. 4.3) $G_\cdot(\theta^0)$ is a local Q-martingale and by the above bounds is also of class (D) under Q, thus it's a time Q-martingale. So $E_Q[g_0] = E_Q[G_T(\theta^0)] = 0$ and $f \leq \alpha + g_0$ a.s. implies that $P(f < \alpha + g_0) > 0$, since otherwise $f = \alpha + g_0$ P-a.s.. And this would imply $E_Q[f] = \alpha$ for all $Q \in \mathcal{P}^W_{e,\sigma}$ and so $\alpha = \beta$ which is a contradiction.

$\square$

Remarks:

1. The above stopping argument uses that f is bounded. If only $|f| \leq W$, then it is not clear what happens.
2. The assumption $|f| \leq W$ is harmless if f is bounded. Take any f.w.f. W_0 and set $W := (\|f\|_\infty \vee 1)W_0$ to get a f.w.f. with $|f| \leq W$ and $\mathcal{G}_{W_0} = \mathcal{G}_W$.

9.4 W-hedgeability

We want to study payoffs f for which buyer and seller prices agree. Let's first discuss attainability of a payoff. Recall: f superreplicable at price c if $f \leq c+g$ for some $g \in \mathcal{G}_W$. So tempting to call f replicable/attainable at price c if $f = c + g$ for some $g \in \mathcal{G}_W$, and one might want to define a new seller price for f as smallest such c, to avoid waste from condition $f \leq c + g$. However, there are two problems:

1. It may happen that $f = c + g$ P-a.s. is impossible whereas $f \leq c + g$ is achieved, so asking for equality is too much.
2. It can also happen that $f = c + g$ P-a.s. for some g but $f \leq c + g'$ for some other g' with $P(c+g' > f) > 0$. So θ from $g = G_T(\theta)$ would exactly reproduce f at cost c, but the strategy θ' from $g' = G_T(\theta')$ is more clever since it achieves more for the same price c. Translating back to f, we can afford $f = c + g$ at price c at time 0, but with the same initial price, f is not a good choice since better can be achieved ($f' = c + g'$).

We want to avoid the second listed problem and this motivates recalling the following definition.

Definition 9.5. *For a subset $A \subset L^0$, we call an element $a \in A$ maximal in A if the following holds: If $a' \in A$ s.t. $a' \geq a$ P-a.s., then $a' = a$ P-a.s.*

Definition 9.6. *Let W be a f.w.f.. Then $f \in L^0$ is W-hedgeable if $f = c+g$ where $c \in \mathbb{R}$ and $g \in \mathcal{G}_W$ maximal. (g maximal means we avoid "stupid"* *trading).*

Remark: Having non-maximal g does not imply that the market admits arbitrage: if $g, g' \in \mathcal{G}_W$ with $g' \geq g$ and $P(g' > g) > 0$, one would like to buy g' and sell g, to obtain some gain without risk ($g' - g \in L^0_+ \backslash \{0\}$). But this may not be W-admissible.

Theorem 9.7. *Suppose S satisfies (NFLVR), W is a f.w.f.. Let f be a payoff s.t. $|f| \leq W$. If $\alpha = \beta \in \mathbb{R}$, then f is W-hedgeable.*

Proof. By Thm. 9.2, $\alpha \in \Gamma_+$ and $\beta \in \Gamma_-$, so that $f - \alpha$ and $-f + \beta (= -(f - \alpha))$ are both in $\mathcal{G}_W - L^0_+$. Therefore there exists $g_1, g_2 \in \mathcal{G}_W$ with $f \leq \alpha + g_1$ and $-f \leq -\alpha + g_2$, so summing these up we get $g_1 + g_2 \geq 0$.

By lemma 8.3(1), we have $E_Q[g_1 + g_2] \leq 0$ for all $Q \in \mathcal{P}^W_{e,\sigma} \neq \emptyset$, hence $g_1 + g_2 = 0$ Q-a.s. or equivalently P-a.s..

Then $\alpha + g_1 = \alpha - g_2 \leq f \leq \alpha + g_1$ which implies that $f = \alpha + g_1$ P-a.s.. We are left to show that g_1 is maximal in $\mathcal{G}_W$.

By Thm 9.3, the assumption that $\alpha = \beta \in \mathbb{R}$ gives $\sup_{Q \in \mathcal{P}_{e,\sigma}} E_Q[f] = \inf_{Q \in \mathcal{P}_{e,\sigma}} E_Q[f]$. That is, the map

$$\mathcal{P}^W_{e,\sigma} \to [-\infty, \infty] \tag{9.19}$$

$$Q \mapsto E_Q[f] \tag{9.20}$$

is finite-valued and constant ($= \alpha = \beta$). Since $f = \alpha + g_1$, this gives $E_Q[g_1] = 0$ for all $Q \in \mathcal{P}^W_{e,\sigma}$. In particular (*) $E_Q[g_1] = 0$ for some $Q \in \mathcal{P}^W_{e,\sigma}$. We show that this implies g_1 maximal in $\mathcal{G}_W$. Indeed, let $g' \in \mathcal{G}_W$ s.t. $g' \geq g_1$, then by the previous and lemma 8.3(1) give $0 = E_Q[g_1] \leq E_Q[g'] \leq 0$ so all the inequalities are in fact equalities and $g_1 = g'$ a.s.. Hence g_1 is maximal in $\mathcal{G}_W$. $\square$

From the previous proof we have: f is W-hedgeable if it has same finite expectation under all $Q \in \mathcal{P}^W_{e,\sigma}$. This is analogous to discrete time result thm 6.2. Intuition: If f has unique price, then it is W-hedgeable. But in finite discrete time we also have the converse, which does not hold in general in the continuous-time setting. That is, we would expect that f is W-hedgeable ($\implies f = c + g$, for g maximal) implies c, the unique price of f to be computed as $c = E_Q[f]$ for all $Q \in \mathcal{P}^W_{e,\sigma}$. But in general this is wrong: We will see an example where $E_Q[g] < 0$ for some maximal $g \in \mathcal{G}_W$ and $Q \in \mathcal{P}^W_{e,\sigma}$ when g is unbounded. On the other hand, if g is bounded, then things work as in the discrete-time setting, that is f W-hedgeable implies a unique price c given by the constant value $E_Q[f]$ for all Q. In the next theorem we get an extension of thm 6.2 for characterisation of W-hedgeable payoffs.

Theorem 9.8. *Suppose S satisfies (NFLVR), W is a f.w.f.. For any payoff $f \geq -W$, TFAE:*

1. *f is W-hedgeable*
2. *$f = c+g$ for some $c \in \mathbb{R}$, $g \in \mathcal{G}_W$ s.t. $E_{Q^*}[g] = 0$ for some $Q^* \in \mathcal{P}^W_{e,\sigma}$*
3. *$f = c + G_T(\theta)$ for some $c \in \mathbb{R}$, $\theta \in \Theta_W$ s.t. $G(\theta)$ in $[0,T]$ is a Q^*-martingale for some $Q^* \in \mathcal{P}^W_{e,\sigma}$.*
4. *$\sup_{Q\in\mathcal{P}^W_{e,\sigma}} E_Q[f] < \infty$ and the supremum is attained.*

Proof. "2.) $\iff$ 3.)": For any $g \in \mathcal{G}_W$ with $g = G_T(\theta)$ where $\theta \in \Theta_W$, $G(\theta)$ is a Q-supermartingale for all $Q \in \mathcal{P}^W_{e,\sigma}$ by lemma 8.3(1), and $G_0(\theta) = 0$ so if for some $Q^* \in \mathcal{P}^W_{e,\sigma}$, we have

$$E_{Q^*}[g] = E_{Q^*}[G_T(\theta)] = 0 = G_0(\theta) \tag{9.21}$$

this is equivalent to $G(\theta)$ being a Q^*-martingale in $[0,T]$.

"2.) $\implies$ 4.)": For any $Q \in \mathcal{P}^W_{e,\sigma}$, from lemma 8.3(1) we have

$$E_Q[f] = c + E_Q[g] \leq c \tag{9.22}$$

so if for some Q^* we have equality, then such a Q^* achieves the supremum and is equal to c.

"4.) $\implies$ 1.)": Define $c := \sup_{Q\in\mathcal{P}^W_{e,\sigma}} E_Q[f] = E_{Q^*}[f]$ since by assumption $c < \infty$ (actually $c = \alpha$), which gives $c \in \Gamma_+$ by Thm. 9.3 and Thm. 9.2.

Then $f \leq c + g$ for some $g \in \mathcal{G}_W$. Agains by lemma 8.3(1):

$$c = E_{Q^*}[f] \leq c + E_{Q^*}[g] \leq c \tag{9.23}$$

so both inequalities are actually equalities. Therefore $f = c + g$ Q^*-a.s. or equivalently P-a.s. and $E_{Q^*}[g] = 0$.

Then we have that g is maximal in $\mathcal{G}_W$. So f is W-hedgeable.

"1.) $\implies$ 2.)": We want to prove that there exists Q^* as in 2.) by using a variant of the Kreps-Yan theorem (Prop. A.3) and since $f \geq -W$ has no upper bound, we first need a weight function. By assumption we have $f = c + g$ with $g \in \mathcal{G}_W$ maximal.

For any $g \in \mathcal{G}_W$, we will show that (**) g maximal in $\mathcal{G}_W$ implies $E_{Q^*}[g] = 0$ for some $Q^* \in \mathcal{P}^W_{e,\sigma}$.

Again, by lemma 8.3(1), $E_Q[g] \leq 0$ for all $Q \in \mathcal{P}^W_{e,\sigma}$ and thus $g^+ \in L^1(Q)$ for all $Q \in \mathcal{P}^W_{e,\sigma}$. Define $W' = W + g^+$ which is also a f.w.f. s.t. $\mathcal{P}^{W'}_{e,\sigma} = \mathcal{P}^W_{e,\sigma}$.

However, $\mathcal{G}_{W'} \supseteq \mathcal{G}_W$, so it is not immediate that g is still maximal in $\mathcal{G}_{W'}$. But this is true, indeed: For any $\tilde{g} \in \mathcal{G}_{W'}$ s.t. $\tilde{g} \geq g$ then $\tilde{g} \geq g \geq -aW$ for some $a \in \mathbb{R}_+$.

By lemma 8.3(4) also $\tilde{g} \in \mathcal{G}^1_{aW} \subseteq \mathcal{G}_W$. But then by maximality of g in $\mathcal{G}_W$, then $\tilde{g} = g$ P-a.s., hence g maximal in $\mathcal{G}_{W'}$ as well. Now, $g \geq -aW$ for some $a \geq 0$ and $g \leq g^+$, hence $\frac{g}{W'} = \frac{g}{W+g^+}$ is bounded.

Moreover, by maximality of g in $\mathcal{G}_{W'}$:

$$\left(\mathcal{C}^\infty_{W'} - \frac{g}{W'}\right) \cap L^\infty_+ = \{0\} \tag{9.24}$$

Indeed: $h \in \mathcal{C}^\infty_{W'} - \frac{g}{W'}$, then $h = \frac{1}{W'}(\tilde{g} - \tilde{Y}) - \frac{g}{W'}$, for some $\tilde{g} \in \mathcal{G}_{W'}$ and $\tilde{Y} \geq 0$.

If also $h \geq 0$, then $\tilde{g} = g + hW' + Y' \geq g$. Then by maximality of g in $\mathcal{G}_{W'}$ we have $hW' + \tilde{Y} = 0$ P-a.s., hence $h = 0$.

By Cor. 8.5, $\mathcal{C}^\infty_{W'}$ is weak*-closed in L^∞. So we can apply Kreps-Yan theorem in the variant of Prop. A.3 to get $Q^*_0 \sim P$ s.t. $E_{Q^*_0}[Y] \leq 0$ for all $Y \in \mathcal{C}^\infty_{W'}$ and $E_{Q^*_0}[\frac{g}{W'}] = 0$.

Define another probability $Q^* \sim P$ via $\frac{dQ^*}{dQ^*_0} = \text{const.} \frac{1}{W'}$. Clearly $E_{Q^*}[g] = 0$ and $E_{Q^*}[W'Y] \leq 0$ for all $Y \in \mathcal{C}^\infty_{W'}$.

For any $\tilde{Y} \in \mathcal{G}_{W'} - L^0_+$ s.t. $\frac{\tilde{Y}}{W'}$ is bounded then $\frac{\tilde{Y}}{W'} \in \mathcal{C}^\infty_{W'}$. Therefore $E_{Q^*}[\tilde{Y}] \leq 0$.

As in the last step of the proof of thm 8.6, this implies $Q^* \in \mathcal{P}_{e,\sigma}$. Then $Q^* \in \mathcal{P}^{W'}_{e,\sigma} = \mathcal{P}^W_{e,\sigma}$. That is, Q^* is the measure in 2.).

$\square$

Remark: Putting together (*) and (**), we have that for $g \in \mathcal{G}_W$ g maximal if and only if $E_{Q^*}[g] = 0$ for some $Q^* \in \mathcal{P}^W_{e,\sigma}$.

Remark: If f is W-hedgeable and bounded, then in part 3.) of Thm. 9.8, $G_T(\theta) = f - c$ is bounded, hence $G(\theta)$ is bounded. So $G(\theta)$ is a true Q-martingale for all $Q \in \mathcal{P}^W_{e,\sigma}$ and the map $Q \mapsto E_Q[f]$ is finite-valued and constant $(= c)$ on $\mathcal{P}^W_{e,\sigma}$. So for bounded payoffs the analogy to the discrete time result in thm. 6.2 is perfect.

9.5 Example

Example: We construct a model and a payoff $g \in \mathcal{G}_W$ s.t. $E_{Q'}[g] = 0$ for some $Q' \in \mathcal{P}^W_{e,\sigma}$ but $E_Q[g] < 0$ for another $Q \in \mathcal{P}^W_{e,\sigma}$. ($f \equiv g$, $c = 0$, g maximal, hence $f = g$ is W-hedgeable but no unique price $E_Q[f]$ under all $Q \in \mathcal{P}^W_{e,\sigma}$).

This example shows that a full analogue of the discrete-time case in Thm. 6.2 does not hold. For simplicity we work on $[0, \infty)$, we consider W and W' independent Brownian Motions under P, and define

$$X := \mathcal{E}(W) = \exp(W - \frac{1}{2}t) \tag{9.25}$$

$$Y := \mathcal{E}(W') \tag{9.26}$$

and consider the stoppping times:

$$\sigma := \inf\{t \geq 0 : X_t = \frac{1}{2}\} \tag{9.27}$$

$$\tau := \inf\{t \geq 0 : Y_t = 2\} = \inf\{t \geq 0 : W'_t := \ln 2 + \frac{t}{2}\} \tag{9.28}$$

We use the law of large numbers for Brownian Motion, for which $\lim_{t\to\infty} \frac{W_t}{t} = 0$ a.s.. This implies $X_t \to 0$ as $t \to \infty$, so that $\sigma < \infty$ P-a.s.. In the same way we have $Y_t \to 0$, but $\tau = \infty$ with probability $\frac{1}{2}$. We are going to show this and we use that $Y_\tau = 2$ on $\{\tau < \infty\}$ and $= Y_\infty$ on $\{\tau = \infty\}$.

Indeed, Y^τ martingale, non-negative and bounded by 2. Therefore,

$$1 = \lim_{t\to\infty} E[Y^\tau_t] \tag{9.29}$$

$$= E[\lim_{t\to\infty} Y^\tau_t] \tag{9.30}$$

$$= E[Y_\tau \mathbf{1}(\tau < \infty) + Y_\infty \mathbf{1}(\tau = \infty)] \tag{9.31}$$

$$= 2P(\tau < \infty) \tag{9.32}$$

So $P(\tau < \infty) = \frac{1}{2}$.

Consider $S := X^{\tau\wedge\sigma}$, so $S \in \mathcal{M}_{loc}(P)$, with $S_0 = 1$ and $S_\infty = X_{\tau\wedge\sigma} = X_\sigma \mathbf{1}(\tau = \infty) + X_{\tau\wedge\sigma}\mathbf{1}(\tau < \infty)$.

By the independence of W and W', we have that X and σ are independent of τ. We want to use this together with $E[X_{\sigma\wedge\tau}] = E[X_0] = 1$ by the stopping theorem. From those two facts, we have:

$$E[S_\infty] = \frac{1}{2}P(\tau = \infty) + E[E[X_{\tau\wedge\sigma}|\mathcal{F}_t]\mathbf{1}(\tau < \infty)] \tag{9.33}$$
$$= \frac{3}{4} < 1 \tag{9.34}$$

So the payoff $g = S_\infty - S_0 = \int_0^\infty 1dS \in \mathcal{G}_1$ since $S \geq 0$. Taking $Q := P \in \mathcal{P}_{e,\sigma}$, $E_Q[g] < 0$. We are left to find Q'. In order to do this, let $Z := Y^{\tau\wedge\sigma}$, which is a P-local martingale which is bounded by 2, hence a martingale of class (D). So we set $\frac{dQ'}{dP} := Z_\infty > 0$ a.s.. This defines a probability $Q' \sim P$, using that $\sigma < \infty$ a.s. and $Y > 0$ on $[0, \infty)$.

We are going to show that S is a Q'-martingale on $[0, \infty]$, which will then imply $E_{Q'}[g] = 0$. By Bayes rule, it is enough to show that SZ is a P-martingale on $[0, \infty]$.

$$SZ = (XY)^{\tau\wedge\sigma} \tag{9.35}$$
$$= (\mathcal{E}(W)\mathcal{E}(W'))^{\tau\wedge\sigma} \tag{9.36}$$
$$= (\mathcal{E}(W + W'))^{\tau\wedge\sigma} \tag{9.37}$$

local P-martingale, nonnegative and therefore a P-supermartingale. So it is enough to show that $E[S_\infty Z_\infty] = S_0 Z_0 = 1$. We know that Y^τ is a bounded P-martingale, so we can write $Y_{\tau\wedge\sigma} = E[Y_\tau|\mathcal{F}_{\tau\wedge\sigma}]$ with Y_τ equal to zero on $\tau = \infty$ and equal to 2 on $\tau < \infty$. We know $E[X_{\tau\wedge\sigma}\mathbf{1}(\tau < \infty)] = \frac{1}{2}$. Putting these things together:

$$E[S_\infty Z_\infty] = E[X_{\tau\wedge\sigma} Y_{\tau\wedge\sigma}] \tag{9.38}$$
$$= E[X_{\tau\wedge\sigma} Y_\tau] \tag{9.39}$$
$$= 2E[X_{\tau\wedge 2}\mathbf{1}(\tau < \infty)] \tag{9.40}$$
$$= 1 \tag{9.41}$$

Chapter 10

Utility maximization: primal problem

Goal: Study optimal portfolio choice problem with preferences given by expected utility from terminal wealth.

Setup: Horizon $T < \infty$, $(\Omega, \mathcal{F}, \mathbb{F}, P)$, bank account $S^0 \equiv 1$, discounted asset price of d risky assets given by a $\mathbb{R}^d$-valued semimartingale S. Assume $\mathcal{F}_0$ trivial and absence of arbitrage in the form $\mathcal{P}_{e,\sigma} \neq \emptyset$.

- Initial wealth $x > 0$
- We consider self-financing strategies $\varphi \hat{=} (x, \theta)$, with $\theta \in L(S)$ predictable and $\mathbb{R}^d$-valued. Therefore we have the wealth process $V(\varphi) = V(x, \theta) = x + \theta \bullet S$ and we want $V(\varphi) \geq 0$, so the strategy $\varphi = (x, \theta)$ is 0-admissible and integrand is x-admissible ($\theta \in \Theta^x_{adm}$)
- Utility maximization problem: Find 0-admissible (x, θ) to maximize expected utility $E[U(V_T(x, \theta))]$ from terminal wealth, where U is a utility function on $(0, \infty)$.

Reference: Biaghi-Friltelli (2005, 2008) for U more general than on $(0, \infty)$.

For $x > 0$, define

$$
\begin{aligned}
\mathcal{V}(x) &= \{V = V(x, \theta) = x + \int \theta dS : V \geq 0\} && (10.1)\\
&= \{x + \int \theta dS : \theta \in \Theta^x_{adm}\} && (10.2)
\end{aligned}
$$

Definition 10.1. *Utility function: A map $U : (0, \infty) \to \mathbb{R}$ strictly increasing, strictly concave, in C^1 and satisfying Inada conditions:*

$$U'(0) := \lim_{x \searrow 0} U'(x) = \infty \tag{10.3}$$

$$U'(\infty) := \lim_{x \to \infty} U'(x) = 0 \tag{10.4}$$

10.1 Primal optimization problem

$$u(x) := \sup_{V \in \mathcal{V}(x)} E[U(V_T)] \quad x > 0 \tag{10.5}$$

Interpretation:

- U quantifies subjective preferences by assigning to monetary amount z the "level of happiness" $U(z)$.
- U increasing means "more is better".
- U concave captures the idea that an extra dollar increases happiness more for a poor person than someone who is already rich.
- The indirect utility $u(x)$ is the maximal expected level of happiness one could achieve by time T starting with initial wealth x and being allowed to invest in the market (with 0-admissible φ).

Natural assumption: Problem is well-posed, so that there exists $x_0 > 0$ s.t. $u(x_0) < \infty$.

Remarks:

1. A priori U only defined on $(0, \infty)$, but we can define by monotonicity: $U(0) := \lim_{x \searrow 0} U(x)$ which exists in $[-\infty, \infty]$, so that $U(V_T)$ is well defined and in $[-\infty, \infty]$.
2. For any random variable $f \geq 0$, we set $E[U(f)] := -\infty$ if $(U(f))^- \notin L^1(P)$. This is harmless as we are taking the supremum and $\theta \equiv 0$ (which is feasible) gives $E[U(V_T(x, 0))] = U(x)$.
3. If S is not locally bounded, it may happen that $\Theta^x_{adm} = \{0\}$. Then the problem is trivial with $u(x) = U(x)$.
4. If S allows arbitrage and U unbounded then $u \equiv \infty$. So the utility maximization problem only makes sense in a market without arbitrage.

Questions:

1. Does there exist an optimal strategy $\theta^* \in \Theta^x_{adm}$ i.e.

$$E[U(V_T(x,\theta^*))] = u(x) = \sup_{V \in \mathcal{V}(x)} E[U(V_T)]? \tag{10.6}$$

   If so, how do we find it?

2. How does the optimal expected utility $u(x)$ behave as a function of x?

Exercise: If U is increasing and concave (not necessarily strictly) and if $x_0 > 0$ s.t. $u(x_0) < \infty$, then $x \mapsto u(x)$ is increasing and concave, and $u(x) < \infty$ for all $x > 0$.

10.2 The set $\mathcal{C}(x)$

In order to tackle the above questions, we reformulate the primal optimization problem over a bigger set. Define

$$\begin{aligned} \mathcal{C}(x) &:= \{f \in L^0_+(\mathcal{F}_T) : f \leq V_T \text{ for some } V \in \mathcal{V}(x)\} && (10.7) \\ &= (x + G_T(\Theta^x_{adm}) - L^0_+) \cap L^0_+(\mathcal{F}_T) && (10.8) \end{aligned}$$

Note $G_T(\Theta^x_{adm}) = \mathcal{G}^x_{adm}$. This is the space of all non-negative time T payoffs one can superreplicate with 0-admissible self-financing strategy $\varphi = (x, \theta)$ starting with initial wealth x.

Clearly $\mathcal{C}(x) \supseteq \{V_T : V \in \mathcal{V}(x)\}$.

We have:

$$\begin{aligned} u(x) &= \sup_{V \in \mathcal{V}(x)} E[U(V_T)] && (10.9) \\ &= \sup_{f \in \mathcal{C}(x)} E[U(f)] && (10.10) \end{aligned}$$

The inequality "$\leq$" is clear. For "$\geq$": Any $f \in \mathcal{C}(x)$ has $f \leq V_T$ for some $V \in \mathcal{V}(x)$ so $U(f) \leq U(V_T)$ because U is increasing and then $E[U(f)] \leq \sup_{V \in \mathcal{V}(x)} E[U(V_T)]$.

Remark: If $f^* \in \mathcal{C}(x)$ is optimal, that is $E[U(f^*)] = u(x)$, then $f^* \leq V_T(x, \theta^*)$ for some $\theta^* \in \Theta^x_{adm}$. but then θ^* is optimal because

$$u(x) = E[U(f^*)] \leq E[U(V_T(x,\theta^*))] \leq u(x) \tag{10.11}$$

so all inequalities are in fact equalities, so $f^* = V_T(x, \theta^*)$ and both of the above suprema are maxima.

The set $\mathcal{C}(x)$ is convex and solid, i.e. $f \in \mathcal{C}(x)$ and $0 \leq f' \leq f$ and so $f' \in \mathcal{C}(x)$.

Lemma 10.2. *Assume (NFLVR) and $\mathcal{F}_0$ trivial. Then*

$$\mathcal{C}(x) = \{f \in L^0_+(\mathcal{F}_T) : E_Q[f] \leq x \text{ for all } Q \in \mathcal{P}_{e,\sigma}\} \tag{10.12}$$

In particular, $\mathcal{C}(x)$ is bounded in $L^1(Q)$ for any $Q \in \mathcal{P}_{e,\sigma}$.

Proof. This is a variation of the hedging duality in thm. 7.5. For "$\subseteq$", note that $G(\theta) = \int \theta dS$ is a Q-supermartingale for all $Q \in \mathcal{P}_{e,\sigma}$ and for all $\theta \in \Theta^x_{adm}$ so $f \leq V_T(x,\theta) = x + G_T(\theta)$ which implies $E_Q[f] \leq x$ for all $Q \in \mathcal{P}_{e,\sigma}$.

"$\supseteq$": Define U as an RCLL version of $U_t := \operatorname{ess\,sup}_{Q \in \mathcal{P}_{e,\sigma}} E_Q[f|\mathcal{F}_t]$, $0 \leq t \leq T$. Note that $U_T = f$ because $f \in L^0_+(\mathcal{F}_T)$ and $U_0 = \sup_{Q \in \mathcal{P}_{e,\sigma}} E_Q[f] \leq x$ since $\mathcal{F}_0$ trivial.

By thm 7.3 and thm. 7.4, U has representation $U = U_0 + \int \theta dS - C$ where $\theta \in \Theta_{adm}$ and C is increasing and $C_0 = 0$.

So $\int \theta dS = U - U_0 + C \geq -U_0 \geq x$, hence $\theta \in \Theta^x_{adm}$. Finally, $f = U_T = U_0 + G_T(\theta) - C_T \leq x + G_T(\theta)$. Therefore $f \in \mathcal{C}(x)$. $\square$

In view of this previous lemma and of representation of $u(x)$ as supremum over $\mathcal{C}(x)$, we can rewrite our primal optimal problem as a supremum of $E[U(f)]$ over all $f \in L^0_+(\mathcal{F}_T)$ satisfying $E_Q[f] \leq x$ for all $Q \in \mathcal{P}_{e,\sigma}$.

This is an optimization problem with constraints, but writing a Lagrange function is not easy since we have infinitely many such constraints.

10.3 A first look at the dual problem

We want to find a suitable dual problem and for this it will turn out to be useful to generalize the set of EσMMs. Start with $Q \in \mathcal{P}_{e,\sigma}$, denote by $Z = Z^{Q;P}$ its density process, $V = V(x,\theta) = x + \int \theta dS \in \mathcal{V}(x)$. Then $Z > 0$, P-martingale, $Z_0 = 1$ since $\mathcal{F}_0$ is trivial, and ZV is a P-supermartingale by Bayes theorem since V is a Q-supermartingale by Ansel-Stricker. This is a motivation for the following definition.

Definition 10.3. *For $z > 0$, call $\mathcal{Z}(z)$ the family of all non-negative $\mathbb{F}$-adapted RCLL processes $Z = (Z_t)_{t\in[0,T]}$ with $Z_0 = z$, ZV is a P-*

supermartingale for all $V \in \mathcal{V}(1)$. Equivalently, for all $V \in \mathcal{V}(x)$ for all $x > 0$ because $\mathcal{V}(x) = x\mathcal{V}(1)$.

Note that $\theta \equiv 0$ gives $V(1,0) \equiv 1 \in \mathcal{V}(1)$, so $Z \in \mathcal{Z}(z)$ is a P-supermartingale, non-negative, $Z_0 = z$, and we also have $\mathcal{Z}(z) = z\mathcal{Z}(1)$, for all $z > 0$.

Similarly to the argument from $\mathcal{V}(x)$ to $\mathcal{C}(x)$, now we go from $\mathcal{Z}(z)$ to a set $\mathcal{D}(z)$: For any $z > 0$, define

$$\mathcal{D}(z) := \{h \in L^0_+(\mathcal{F}_T) : h \leq Z_T \text{ for some } Z \in \mathcal{Z}(z)\} \tag{10.13}$$

Now take $V \in \mathcal{V}(x)$, $Z \in \mathcal{Z}(z)$, then ZV is a P-supermartingale, starting at xz, hence $E[Z_T V_T] \leq zx$.

Define $J : (0,\infty) \to \mathbb{R}$ by $J(y) := \sup_{x>0}(U(x) - xy)$, $y > 0$. Then J is decreasing, convex, since pointwise supremum of affine functions $l_x = U(x) - xy$. Moreover, $U(V_T) \leq J(Z_T) + Z_T V_T$, so by taking expectations $E[U(V_T)] \leq E[J(Z_T)] + zx$. Now take supremum over $V \in \mathcal{V}(x)$ and infimum over $Z \in \mathcal{Z}(z)$, and so we obtain $u(x) \leq j(z) + zx$ for all $x, z > 0$ where we set $j(z) := \inf_{Z \in \mathcal{Z}(z)} E[J(Z_T)]$, $z > 0$. This is the dual problem that will be studied below.

Note that while $u(x)$ is maximization of a concave function, $j(z)$ is minimization of a convex function.

Since J is decreasing, similarly to the case of the primal problem, we can prove (exercise)

$$j(z) := \inf_{Z \in \mathcal{Z}(z)} E[J(Z_T)] = \inf_{h \in \mathcal{D}(z)} E[J(h)] \tag{10.14}$$

Finally, we also have

$$j(z) \geq \sup_{x>0}(u(x) - xz) \quad \forall z > 0 \tag{10.15}$$

$$u(x) \leq \inf_{z>0}(j(z) + zx) \quad \forall x > 0 \tag{10.16}$$

We shall see later that under an extra condition on U, we actually have equalities, and this will give a lot of results.

The next result slightly extends the first lemma.

Lemma 10.4. *Assume (NFLVR), $\mathcal{F}_0$ trivial. Then for any $f \in L^0_+(\mathcal{F}_T)$,* *we have $f \in \mathcal{C}(x)$ if and only if $\sup_{h \in \mathcal{D}(1)} E[fh] \leq x$. As a consequence, $\mathcal{C}(x)$ is closed in L^0 (while $\{V_T : V \in \mathcal{V}(x)\} =: \mathcal{V}_T(x)$ is not).*

Proof. " $\Longleftarrow$ ": For all $Q \in \mathcal{P}_{e,\sigma}$, $h = \frac{dQ}{dP} \in \mathcal{D}(1)$. Note that $h = Z_T$ for the process $Z_t := E[h|\mathcal{F}_t]$ defining a process $Z \in \mathcal{Z}(z)$. Then we can use lemma 10.2 to get that $f \in \mathcal{C}(x)$.

" $\Longrightarrow$ ": If $f \leq V_T$ for some $V \in \mathcal{V}(x)$ and $h \leq Z_T$ for some $Z \in \mathcal{Z}(1)$, we get $E[fh] \leq E[V_T Z_T] \leq x$ where the last inequality is argued by noting that VZ is a P-supermartingale starting at x.

Finally, if $(f_n)_n \subseteq \mathcal{C}(x)$ converging in L^0 to some f, then $f \in L^0_+(\mathcal{F}_T)$. We can assume, up to passing to a subsequence, that $f_n \to f$ P-a.s.. Then by Fatou

$$E[fh] \leq \liminf_{n \to \infty} E[f_n h] \leq x \tag{10.17}$$

for all $h \in \mathcal{D}(1)$. Therefore $f \in \mathcal{C}(x)$. $\square$

Chapter 11

Utility maximization: the dual problem

Goal: Prove existence of solution to the dual problem and show how this helps solve the primal problem.

Recall the utility function U defined on $(0, \infty)$, for which we defined

$$J(y) := \sup_{x>0}(U(x) - xy) \quad y > 0 \tag{11.1}$$

This is the Legendre transform or convex conjugate of the function $-U(-\cdot)$ in the sense of convex analysis, see Rockafeller, chapter 12, with the convention $U(x) := -\infty$ for $x < 0$.

We first collect properties of J.

Lemma 11.1. *If U is a utility function, then $J : (0, \infty) \to \mathbb{R}$ is strictly decreasing, strictly convex, in C^1, with $J'(0) = -\infty$ and $J'(\infty) = 0$, as well as $J(0) = U(\infty)$, $J(\infty) = U(0)$. We also have the following conjugacy relation:*

$$U(x) = \inf_{y>0}(J(y) + xy) \quad x > 0 \tag{11.2}$$

In addition

$$J' = -(U')^{-1} =: -I \tag{11.3}$$
$$J(y) = U(I(y)) - yI(y) \tag{11.4}$$

Proof. If $U \in C^2$, this is immediate to prove. For the general case, see Rockafeller (Thm 26.5). $\square$

Examples of classical utility functions on $(0, \infty)$:

1. $U(x) = \log(x)$, $J(y) = -\log(y) - 1$
2. $U(x) = \frac{x^\gamma}{\gamma}$ for $\gamma < 1$ and not zero, $J(y) = \frac{1-\gamma}{\gamma} y^{\frac{\gamma}{\gamma-1}}$. U is bounded above by 0 for $\gamma < 0$ and U is unbounded and non-negative for $\gamma > 0$. In the limit $\gamma \to 0$, this gives $U(x) = \log(x)$.

11.1 Dual problem

$$j(z) := \inf_{h \in \mathcal{D}(z)} E[J(h)] \quad z > 0 \tag{11.5}$$

Goal: Show that this problem has a unique solution $h_z^* \in \mathcal{D}(z)$ if $j(z) < \infty$, i.e. $h \mapsto E[J(h)] =: F(h)$ attains infimum over the set $\mathcal{D}(z)$ in a unique h_z^*.

Classical approach to show that continuous function has a minimizer on a compact set:

1. Approximate the infimum along a subsequence.
2. Use compactness to get a convergent subsequence.
3. The limit of such a subsequence is a candidate for the minimizer.
4. Use continuity to compute function value in the limit.

In our setting: We use insead a Komlós-type result (lemma 4.8) to produce a candidate; then show that $\mathcal{D}(z)$ is convex and closed in L^0 so that it also contains the candidate; then show that F is convex and lower semicontinuous to deduce that candidate is in fact a minimizer.

Theorem 11.2. *For every $z > 0$, the set $\mathcal{D}(z)$ is convex, solid, closed in L^0.*

Proof. $\mathcal{D}(z)$ solid is immediate from its definition, $\mathcal{D}(z)$ convex follows from convexity of $\mathcal{Z}(z)$. Proof of closedness is slightly more complicated than IMF (P. IV 3.2), due to continuous-time. Take $(h_n)_n \subseteq \mathcal{D}(z)$ converging in L^0 to some h, so $h \geq 0$, and $(Z^n)_n \subseteq \mathcal{Z}(z)$ s.t. $h_n \leq Z_T^n$ for all n. Up to passing to a subsequence (that we still denote by $(h_n)_n$) we can assume $h_n \to h$ P-a.s. and $h < \infty$ P-a.s.. $Z^n \geq 0$, so we can use lemma 4.8 (Komlós) and a diagonal argument to find a convex combination $\tilde{h}_n \in \text{conv}(h_n, h_{n+1}, ...)$ and $\tilde{Z}^n_r \in \text{conv}(Z^n_r, Z^{n+1}_r, ...)$ for all rationals $r \in [0,T]$ s.t. they are converging simultaneously P-a.s. to h_∞ and Z^∞_r.

We now have $h_n \to h$ P-a.s., $\tilde{h}_n \to h_\infty$ P-a.s.. This implies $h = h_\infty$ P-a.s.. Since $h_n \leq Z^n_T$, and because we use (thanks to diagonalization) the same convex combination for h_n and all the $\tilde{Z}^n_r$ for all rationals $r \in [0,T]$, we also get $h = h_\infty \leq Z^\infty_T$ P-a.s.. It only remains to show that $Z^\infty_T = Z_T$ for some $Z \in \mathcal{Z}(z)$. Each $\tilde{Z}^n$ is in $\mathcal{Z}(z)$ because $\mathcal{Z}(z)$ is convex, also $Z^\infty_0 = z$ and each $\tilde{Z}^n V$ is a P-supermartingale for all $V \in \mathcal{V}(1)$, non-negative.

Take any s, any rational $r \leq s$, then

$$E[Z^\infty_s V_s | \mathcal{F}_r] \leq \liminf_{n\to\infty} E[\tilde{Z}^n_s V_s | \mathcal{F}_r] \tag{11.6}$$

$$\leq \liminf_{n\to\infty} \tilde{Z}^n_r V_r = Z^\infty_r V_r \quad P-\text{a.s.} \tag{11.7}$$

with the first inequality due to Fatou. So $Z^\infty V$ is a P-supermartingale on $\mathbb{Q} \cap [0,T]$ (We can take $V \equiv 1$, hence Z^∞ itself is a P-supermartingale on rationals). From a standard construction from martingale theory (see DM, T VI.2), we have existence of a RCLL P-supermartingale $Z = (Z_t)_{t\in[0,T]}$ on $[0,T]$ with $Z_r \leq Z^\infty_r$ for all rationals r. We can take $Z_t = \lim_{r \searrow t} Z^\infty_r$ so $Z_T = Z^\infty_T$ and $Z_0 = Z^\infty_0 = z$ (uses that $\mathbb{F}$ right continuous). As above, using Fatou and construction of Z from Z^∞ yields that ZV is a P-supermartingale on $[0,T]$ for all $V \in \mathcal{V}(1)$. Then $Z \in \mathcal{Z}(z)$ and $h \leq Z^\infty_T = Z_T$ and therefore $h \in \mathcal{D}(z)$. □

Theorem 11.3. *For each $z > 0$, we have*

1. *$\mathcal{D}(z)$ is bounded in $L^1(P)$ and the family $\{(J(h))^- : h \in \mathcal{D}(z)\}$ is P-uniformly integrable.*
2. *$F : L^0_+(P) \to [-\infty, \infty]$ is $h \mapsto F(h) = E[J(h)]$ is lower semi-continuous 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 (P IV 3.3), or KS Lemma 3.2. □

Remark: In thm 11.2 (as we will see in thm 11.4) one should have argued that $h_\infty < \infty$ P-a.s..

Theorem 11.4. *Suppose $\mathcal{P}_{e,\sigma} \neq \emptyset$ (so in particular $\mathcal{D}(z) \neq \emptyset$ for all $z > 0$). For every $z > 0$ with $j(z) < \infty$, there exists a unique solution $h^*_z \in \mathcal{D}(z)$ to the dual problem:*

$$\inf_{h\in\mathcal{D}(z)} E[J(h)] =: j(z) = E[J(h_z^*)] \tag{11.8}$$

Proof. Uniqueness immediately follows from strict convexity of J. For existence, take $(h_n)_n \subseteq \mathcal{D}(z)$ s.t. $E[J(h_n)] = F(h_n)$ is decreasing to $j(z) < \infty$. As all $h_n \geq 0$, lemma 4.8 (Komlós) gives existence of $\tilde{h}_n \in \mathrm{conv}(h_n, h_{n+1}, ...)$ converging to h P-a.s., thus also in L^0. (To be more precise, a priori we could have that h takes values in $[0,\infty]$. But by thm. 11.3, $\mathcal{D}(z)$ is bounded in $L^1(P)$, hence also in L^0, and so lemma 4.8 implies that $h < \infty$ P-a.s., so that $h \in L^0$).

By Prop. 11.2, $\mathcal{D}(z)$ is convex, so $(\tilde{h}_n)_n \subseteq \mathcal{D}(z)$, and closed in L^0, so $h \in \mathcal{D}(z)$. By lemma 11.1, J is convex so that

$$\begin{aligned}
F(\tilde{h}_n) &= E[J(\tilde{h}_n)] && (11.9)\\
&\leq \sup_{k\geq n} E[J(h_k)] && (11.10)\\
&= \sup_{k\geq n} F(h_k) && (11.11)\\
&= F(h_n) && (11.12)
\end{aligned}$$

since $n \mapsto F(h_n)$ is decreasing. Since $\tilde{h}_n \in \mathcal{D}(z)$, we have

$$j(z) \leq F(\tilde{h}_n) \leq F(h_n) \searrow j(z) \tag{11.13}$$

Therefore by thm 11.3:

$$E[J(h)] = F(h) \leq \liminf_{n\to\infty} F(\tilde{h}_n) = j(z) \tag{11.14}$$

which implies that $h_z^* := h$ is optimal. □

Corollary 11.5. *The function j is decreasing, strictly convex on $\{z > 0 : j(z) < \infty\}$ and it's continuous on the interior of $\{j < \infty\}$ (This is some interval from some a to ∞).*

Proof. See IMF Cor. IV 3.5, or KS Lemma 3.3. □

11.2 From the dual to the primal

Now we want to see how to use the dual problem to solve the primal problem. Fix $x > 0$ and take $f \in \mathcal{C}(x)$ so that $f \leq V_T$ with $V \in \mathcal{V}(x)$. Take $z > 0$ and $h \in \mathcal{D}(z)$ so that $h \leq Z_T$ with $Z \in \mathcal{Z}(z)$. Then we have

$$E[fh] \leq E[V_T Z_T] \leq xz \tag{11.15}$$

since ZV is a P-supermartingale. From $U(f) \leq J(h) + fh$ (by definition of the function J) we then get

$$E[U(f)] \leq E[J(h)] + xz \tag{11.16}$$

maximizing the LHS over $f \in \mathcal{C}(x)$ gives us $u(x)$ and minimizing the RHS over $h \in \mathcal{D}(z)$ gives $j(z) + xz$. Each side provides a bound for the other side, and so we should get two optima (for primal and dual problems respectively) by making the bounds sharp via equality. So we aim for equalities everywhere.

Maximizer for $J(y) = \sup_{x>0}(U(x) - xy)$ is obtained for $U'(x) = y$, i.e. $x = (U')^{-1}(y) =: I(y)$. So as in lemma 11.1, $J(y) = U(I(y)) - yI(y)$.

We choose, for any $h \in \mathcal{D}(z)$, $f := I(h)$, so we get $U(f) = J(h) + fh$. In order to get $E[fh] = xz$, we then want $E[I(h)h] = xz$.

If we have that, then

$$\begin{aligned}
E[U(I(h))] &= E[J(h)] + E[hI(h)] && (11.17)\\
&\geq j(z) + xz && (11.18)\\
&\geq \inf_{z'>0}(j(z') + xz') && (11.19)\\
&\geq u(x) && (11.20)
\end{aligned}$$

where the last inequality is obtained towards the end of chapter 10. By theorem 11.4, the first inequality becomes an equality if we take $h = h_z^*$. The second inequality becomes an equality if z is minimizer for $z' \mapsto j(z') + xz'$, that is solving for z the equation $j'(z) = -x$.

Reverse engineering then suggests the following steps to construct a solution to the primal problem:

1. Start with $x > 0$ and define $z = z_x$ via $j'(z) = -x$.
2. Solve the dual problem for this z to get $h_z^* \in \mathcal{D}(z)$ as the unique optimizer. Define $f_x^* := I(h_z^*) = I(h_{z_x}^*)$
3. Show that $E[h_{z_x}^* I(h_{z_x}^*)] = xz_x$
4. Show that $f_x^* \in \mathcal{C}(x)$.

If all this can be done, then the above reasoning gives

$$E[U(f_x^*)] \geq u(x) = \sup_{f \in \mathcal{C}(x)} E[U(f)] \tag{11.21}$$

so f_x^* is the solution to the primal problem. Moreover, we also get $u(x) = \inf_{z>0}(j(z) + xz)$ for any $x > 0$. Now if this can be done for all $x > 0$ and if the resulting z_x span all $(0, \infty)$, then we have

$$\begin{aligned} j(z_x) &= E[J(h_{z_x}^*)] && (11.22)\\ &= E[U(I(h_{z_x}^*)) - h_{z_x}^* I(h_{z_x}^*)] && (11.23)\\ &= E[U(f_x^*)] - xz_x && (11.24)\\ &= u(x) - xz_x && (11.25) \end{aligned}$$

and since we always have $j(z) \geq u(x) - xz$ (as seen at the end of chapter 10), then

$$j(z) = \sup_{x>0}(u(x) - xz) \tag{11.26}$$

for all $z > 0$. This means that the conjugacy relation between U and J extends to the value functions u and j.

Chapter 12

Utility maximization: auxiliary results

In this chapter we collect results needed for the implementation of the 4 step procedure to solve the primal problem via the dual problem.

To start with, we need to make sure that step 1 is doable (define $z = z_x$ from x via $-j'(z) = x$), which needs information about the function j.

Standing assumption: $\mathcal{P}_{e,\sigma} \neq \emptyset$ and $u(x_0) < \infty$ for some $x_0 > 0$.

Lemma 12.1. *The map*

$$(0, \infty) \to L^0_+ \tag{12.1}$$

$$z \mapsto h^*_z \tag{12.2}$$

is continuous on the interior of $\{j < \infty\}$*: If* $z_n, z > 0$ *with* $z_n \to z$ *and* $j(z_n), j(z) > 0$*, and if* $z \in \mathit{int}(\{j < \infty\})$*, then* $h^*_{z_n} \to h^*_z$ *in* L^0*.*

Proof. See IMF Lemma IV.5.1, or KS L3.6. □

We need an extra condition on the utility function U for the next result.

Definition 12.2. *U has reasonable asymptotic elasticity (RAE) at $+\infty$ (we denote it $RAE_\infty(U)$) if*

$$AE_\infty(U) = \lim_{x\to\infty} \frac{xU'(x)}{U(x)} < 1 \tag{12.3}$$

Intuition of RAE: Marginal utility

$$U'(x) = U'(x)(x + 1 - x) = U'(x)(x + 1 - x) \approx U(x + 1) - U(x) \tag{12.4}$$

measures increase of utility per unit increase of wealth. On the other hand

$$\frac{U(x)}{x} = \frac{1}{x}\sum_{k=1}^{x}(U(k) - U(k-1)) + \frac{1}{x}U(0) \tag{12.5}$$

measures average increase of utility when wealth increases successively from 0 to x. Since U is strictly concave, we always have that $\mathrm{AE}_\infty(U) \leq 1$. So $\mathrm{RAE}_\infty(U)$ is about having a strict inequality. Having equality would mean that for large wealth, marginal and average utility behave in the same way, so that U would be almost linear for large x (which economically does not seem reasonable).

Examples:

1. $U(x) = \log x$. In this case $\frac{xU'(x)}{U(x)} = \frac{1}{\log x}$, so $\mathrm{AE}_\infty(U) = 0$.
2. $U(x) = \frac{x^\gamma}{\gamma}$, $\gamma < 1$. In this case $\frac{xU'(x)}{U(x)} = \gamma$, so $\mathrm{AE}_\infty(U) = \gamma < 1$.
3. $U(x) = \frac{x}{\log x}$. For large x, $\frac{xU'(x)}{U(x)} = 1 - \frac{1}{\log x} \to 1$, so $\mathrm{AE}_\infty(U) = 1$. In fact $U'(x) = 1 - \frac{1}{\log x} \approx 1$.

The condition $RAE_\infty(U)$ gives an estimate for U of the form $U'(x) \leq \beta\frac{U(x)}{x}$ for large x, where $\beta < 1$.

This condition translates to an analogous estimate for the conjugate function.

Lemma 12.3. *Let U be a utility function, with conjugate function J. If $RAE_\infty(U)$, then there are $y_0 > 0$ and $C > 0$ s.t.*

$$-J'(y) \leq C\frac{J(y)}{y} \quad \textit{for } y \in [0, y_0] \tag{12.6}$$

Proof. IMF Lemma IV.5.2, KS L6.3. ∎

Lemma 12.4. *If $RAE_\infty(U)$, then the map*

$$H : (0, \infty) \to \mathbb{R} \tag{12.7}$$

$$z \mapsto H(z) := E[h_z^* I(h_z^*)] \tag{12.8}$$

is continuous on the interior of the set $\{j < \infty\}$.

Proof. Because of continuity of $z \mapsto h_z^*$ in L^0 and continuity of I, the statement follows once we show uniform integrability. This uses thm 11.3, lemma 12.3, Cor. 11.5 as well as lemma 12.1. For details, see IMF L IV 5.3, KS L 3.7. $\square$

Theorem 12.5. *Suppose* $\mathcal{P}_{e,\sigma} \neq \emptyset$, U *a utility function (satisfying Inada conditions) and for some* $x_0 > 0$ *we have* $u(x_0) < \infty$. *Then we have the conjugacy*

$$j(z) = \sup_{x>0}(u(x) - xz) \quad \forall z > 0 \tag{12.9}$$

and therefore $j(z) < \infty$ *for* $z \geq z_0$ *for some* $z_0 > 0$. *If* $RAE_\infty(U)$, *then* $j(z) < \infty$ *for all* $z > 0$.

Proof. We give only the main ideas. Recall that

$$J(y) = \sup_{x>0}(U(x) - xy) \tag{12.10}$$

so it seems reasonable that

$$E[J(h)] = \sup_{f \in L_+^\infty} E[U(f) - fh] \tag{12.11}$$

From this we have

$$j(z) = \inf_{h \in \mathcal{D}(z)} E[J(h)] \tag{12.12}$$

$$= \inf_{h \in \mathcal{D}(z)} \sup_{f \in L_+^\infty} E[U(f) - fh] \tag{12.13}$$

We would like to exchange the order of inf and sup, so we need what is called a min-max theorem (we comment on this below).

If we can exchange the two, then we get

$$j(z) = \sup_{f \in L_+^\infty} \inf_{h \in \mathcal{D}(z)} E[U(f) - fh] \tag{12.14}$$

Looking at the inner problem,

$$\inf_{h \in \mathcal{D}(z)} E[-fh] = -\sup_{h \in \mathcal{D}(z)} E[fh] > -\infty \tag{12.15}$$

with the latter inequality true if and only if $f \in \mathcal{C}(x)$ for some $x > 0$, using that $f \in \mathcal{C}(x)$ iff $\sup_{h \in \mathcal{D}(z)} E[hf] \leq zx$.

Therefore, we should get

$$j(z) = \sup_{x>0} \sup_{f \in \mathcal{C}(x)} (E[U(f)] - xz) \tag{12.16}$$

because $\sup_{h \in \mathcal{D}(z)} E[fh] \leq xz$ for all $f \in \mathcal{C}(x)$. Then $j(z) = \sup_{x>0}(u(x) - xz)$.

This shows the idea of why the conjugacy relation should hold.

Let's comment on the exchange of inf and sup: Almost all min-max theorems need compactness for one of the two sets over which we are optimizing. We do not have this, so we approximate L^∞_+ by compact balls and then use a min-max argument there, subsequently passing to the limit. We view L^∞ as the dual of L^1 and equip it with the weak*-topology $\sigma(L^\infty, L^1)$. For each $n \in \mathbb{N}$, we consider the ball $B_n := \{f \in L^\infty_+ : f \leq n\}$. It is weak*-compact as a subset of L^∞ (Alaoglu's theorem). On the other hand, each $\mathcal{D}(z)$ is a convex subset of L^1 by thm 11.3.

Finally, the map

$$B_n \times \mathcal{D}(z) \to \mathbb{R} \tag{12.17}$$
$$(f, h) \mapsto E[U(f) - fh] \tag{12.18}$$

is concave in f and convex (actually linear) in h. Therefore, in this setting we can use the classical min-max results (e.g. Aubin (1979) T2.7.1), which gives

$$\sup_{f \in B_n} \inf_{h \in \mathcal{D}(z)} E[U(f) - fh] = \inf_{h \in \mathcal{D}(z)} \sup_{f \in B_n} E[U(f) - fh] \tag{12.19}$$

We omit the details.

Letting $n \to \infty$ we can arrive at the conjugacy relation of the theorem. For details, see IMF T IV 5.4., KS L3.4, T3.2, L3.8.

□

Lemma 12.6. *Suppose $RAE_\infty(U)$. Then j is in C^1 on $(0,\infty)$ with j' strictly increasing and*

$$-zj'(z) = E[h_z^* I(h_z^*)] \quad \forall z > 0 \tag{12.20}$$

Proof. See IMF L IV.5.5, KS L3.8. □

This lemma helps us with the recipe from chapter 11. Find for every $x > 0$ a $z > 0$ s.t. $j'(z) = -x$. As j' is continuous and strictly monotone, we have uniqueness of the solution z to the above equation, if it exists. For existence, we need to study the range of values of j'.

Lemma 12.7. *We always have $\lim_{z\to\infty} j'(z) = 0$. Moreover, if $RAE_\infty(U)$, then we also have $\lim_{z\searrow 0} j'(z) = -\infty$*

Proof. See IMF L IV.5.6., KS T.3.2. □

Lemma 12.8. *For all $z > 0$ and $h \in \mathcal{D}(z)$, we have $E[hI(h_z^*)] \le E[h_z^* I(h_z^*)]$*

Proof. See IMF L IV.6.1., KS L3.9. □

Chapter 13

Utility maximization: Solving the primal problem

Goal: Solve the utility maximization problem for terminal wealth

$$u(x) = \sup_{f \in \mathcal{C}(x)} E[U(f)] \quad \text{for } x > 0 \tag{13.1}$$

In chapter 11, we suggested the following steps:

1. For $x > 0$, define $z = z_x$ via $j'(z) = -x$. Then by lemma 12.6 and lemma 12.7: $j' : (0, \infty) \to (0, \infty)$ is continuous, strictly decreasing and surjective, so we have indeed existence and uniqueness of such a z_x.

2. Solve the dual problem for such a z_x to get unique dual optimizer $h_z^* = h_{z_x}^* \in \mathcal{D}(z_x)$. Set $f_x^* := I(h_{z_x}^*)$. By thm. 12.5 (using RAE), $j(z) < \infty$ for all $z > 0$. So we can use thm. 11.4 to ensure existence and uniqueness of $h_{z_x}^*$.

3. Show that $E[h_{z_x}^* I(h_{z_x}^*)] = xz_x$. By lemma 12.6 $E[h_z^* I(h_{z_x}^*)] = -zj'(z)$ for all $z > 0$. Using this for z_x gives $-z_x j'(z_x) = xz_x$ by the first step.

4. Show that $f_x^* \in \mathcal{C}(x)$. Recall that $\mathcal{D}(z) = z\mathcal{D}(1)$ for all $z > 0$. Take any $h \in \mathcal{D}(1)$, set $\tilde{h} := z_x h \in \mathcal{D}(z_x)$. Compute

$$E[hf_x^*] = \frac{1}{z_x} E[\tilde{h} f_x^*] \tag{13.2}$$

$$\leq \frac{1}{z_x} E[h_{z_x}^* I(h_{z_x}^*)] \tag{13.3}$$

$$= \frac{1}{z_x} xz_x \tag{13.4}$$

by using 12.8 for the inequality and step 3 for the last equality. So we showed that $E[hf_x^*] \leq x$ which implies by lemma 10.4 that $f_x^* \in \mathcal{C}(x)$.

Theorem 13.1. *Suppose $\mathcal{F}_0$ is trivial, $\mathcal{P}_{e,\sigma} \neq \emptyset$, U a utility function (satisfying Inada consditions), $u(x_0) < \infty$ for some $x_0 > 0$ and $RAE_\infty(U)$. Then for any $x > 0$, the primal problem of maximal expected utility from terminal 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 the dual problem for $z_x > 0$ defined by $-j'(z_x) = x$.*

Proof. From step 4. above, we know that f_x^* is a feasible element. Uniqueness of solution is clear from strict concavity of U. So we only need to show that $E[U(f_x^*)] = u(x)$.

Indeed, by using $U(I(y)) = J(y) + yI(y)$, we have

$$\begin{aligned}
u(x) &\geq E[U(f_x^*)] && (13.5)\\
&= E[U(I(h_{z_x}^*))] && (13.6)\\
&= E[J(h_{z_x}^*) + h_{z_x} I(h_{z_x}^*)] && (13.7)\\
&= j(z_x) + xz_x && (13.8)\\
&\geq \inf_{z>0}(j(z) + xz) && (13.9)\\
&\geq u(x) && (13.10)\\
&= \sup_{f \in \mathcal{C}(x)} E[U(f)] && (13.11)
\end{aligned}$$

with the latter inequality from the end of chapter 10. So all inequalities are equalities and f_x^* is the unique primal optimizer. □

Remark: From the proof of theorem 13.1 we also have that $u(x) = \inf_{z>0}(j(z) + zx)$ for all $x > 0$ which together with the already shown equality $j(z) = \sup_{z>0}(u(x) - xz)$ for all $z > 0$ from thm 12.5 shows that under the assumptions of thm 13.1 the conjugacy relation between U and J extends to the value functions u and j.

Remark: In KS, they provide an example to show that one cannot drop the assumption of RAE$_\infty(U)$ from the above theorem in the following sense: If not satisfied, one can construct an example for a market S s.t. there is no solution to the primal problem.

We can ask if there is a direct way to solve the primal problem without passing to the dual problem. On the one hand, by passing through the dual, we obtained a lot of properties and a characterisation of the primal problem. On the other hand, if we are only interested in showing existence of a solution to the primal problem and we are willing to add some assumptions (or even show that the market we are studying satisfies them) then we can entirely avoid the dual problem.

To this end, let's introduce the following condition. We say that U satisfies $(U+)$ if $(U(x))^+ \leq k(1+x^\beta)$ for $x>0$ for some constant $k>0$, $\beta \in (0,1)$.

This is for instance satisfied by the power utility $U(x) = \frac{x^\gamma}{\gamma}$, for $\gamma \in (0,1)$ (for $\gamma < 0$ we have the upper bound 0).

Lemma 13.2. *Suppose either $U(\infty) < \infty$ (i.e. U is bounded above), or U satisfies $(U+)$. Let's also suppose that there is $\tilde{Q} \in \mathcal{P}_{e,\sigma}$ s.t. $(\frac{d\tilde{Q}}{dP})^{-1}$ has moments of all orders. (Note from the proof that a weaker condition on the moments would actually suffice). Then $U^+(\mathcal{C}(x)) = \{(U(f))^+ : f \in \mathcal{C}(x)\}$ is P-UI for every $x>0$.*

Proof. If $U(\infty) < \infty$ then the statement is clear. So let's now consider the case where U satisfies $(U+)$, and take $p>1$ s.t. $\beta p < 1$, yielding

$$E[(U(f)^+)^p] \leq k^p E[(1+f^\beta)^p] \tag{13.12}$$
$$\leq A_p + B_p E[f^{\beta p}] \tag{13.13}$$

so if we show that $\sup_{f \in \mathcal{C}(x)} E[f^{\beta p}] < \infty$, then we get $U^+(\mathcal{C}(x))$ is bounded in $L^p(P)$ and then it is P-UI.

Let's call $\tilde{Z} = \frac{d\tilde{Q}}{dP}$, and use Hölder's inequality with $r = \frac{1}{\beta p} > 1$ and s conjugate to r.

$$E[f^{\beta p}] = E[(f\tilde{Z})^{\beta p} \tilde{Z}^{-\beta p}] \tag{13.14}$$
$$\leq E[f\tilde{Z}]^{1/r} E[\tilde{Z}^{-s\beta p}]^{1/s} \tag{13.15}$$
$$= E_{\tilde{Q}}[f]^{\beta p} E[\tilde{Z}^{-s\beta p}]^{1/s} \tag{13.16}$$
$$\leq x^{\beta p} \cdot \text{const.}(\tilde{Z}, s, \beta, p) \tag{13.17}$$

for all $f \in \mathcal{C}(x)$. □

Remark: If S is continuous and satisfies (SC) so that $S = S_0 + M + \int d\langle M \rangle \lambda$ and also $K = \int \lambda^{tr} d\langle M \rangle \lambda = \langle \int \lambda dM \rangle$ is bounded uniformly in (ω, t), then the minimal ELMM $\hat{P}$ exists and its density $\frac{d\hat{P}}{dP} = \mathcal{E}(-\int \lambda dM)_T$ is s.t. $(\frac{d\hat{P}}{dP})^{-1}$ has moments of all orders. (exercise)

Theorem 13.3. *Suppose $\mathcal{F}_0$ is trivial, $\mathcal{P}_{e,\sigma} \neq \emptyset$, $u(x_0) < \infty$ for some $x_0 > 0$. Suppose also that either $u(\infty) < \infty$ or that U satisfies $(U+)$ and that there exists a $\tilde{Q}$ ELMM s.t. $(\frac{d\tilde{Q}}{dP})^{-1}$ has moments of all orders. Then for any $x > 0$, the primal problem of expected utility maximization from terminal wealth admits a unique solution $f_x^* \in \mathcal{C}(x)$.*

Proof. Uniqueness is clear from strict concavity of U. For existence, take $(f_n)_n \subseteq \mathcal{C}(x)$ s.t. $E[U(f_n)] \nearrow u(x) = \sup_{f \in \mathcal{C}(x)} E[U(f)] < \infty$.

All f_n are ≥ 0 so we can use lemma 4.8 (Komlos) to obtain $\tilde{f}_n \in \text{conv}(f_n, f_{n+1}, ...)$ s.t. $\tilde{f}_n \to f_\infty$ a.s. for some f_∞ with values in $[0, \infty]$.

By lemma 10.2 $\mathcal{C}(x)$ is bounded in $L^1(Q)$ for all $Q \in \mathcal{P}_{e,\sigma}$, therefore also bounded in L^0 (which is the same for all $Q \sim P$). So again by lemma 4.8 we have that $f_\infty < \infty$ a.s., so $f_\infty \in L^0_+$, and since $\tilde{f}_n \to f_\infty$ P-a.s., then also in L^0.

Now we are using that $\mathcal{C}(x)$ is convex to say that $\tilde{f}_n \in \mathcal{C}(x)$, and using that $\mathcal{C}(x)$ is closed in L^0 to obtain that also $f_\infty \in \mathcal{C}(x)$.

We also have that $U(\tilde{f}_n) \to U(f_\infty)$ P-a.s. and we claim that $f_x^* := f_\infty$ is the primal optimizer.

Since U concave and $n \mapsto E[U(f_n)]$ is increasing to $u(x)$, then

$$E[U(\tilde{f}_n)] \geq \inf_{m \geq n} E[U(f_m)] = E[U(f_n)] \tag{13.18}$$

If $U(\infty) < \infty$, then all $U(\tilde{f}_n)$ are bounded above by a constant, so Fatou gives

$$\begin{aligned} E[U(f_\infty)] &\geq \limsup_{n \to \infty} E[U(\tilde{f}_n)] && (13.19) \\ &\geq u(x) && (13.20) \end{aligned}$$

and so f_∞ is optimal.

If $(U+)$ holds, lemma 13.2 ensures P-UI so that

$$\lim_{n \to \infty} E[U^+(\tilde{f}_n)] = E[U^+(f_\infty)] \tag{13.21}$$

On the other hand, since $U^- \geq 0$, we can use Fatou to get

$$E[U^-(f_\infty)] \leq \liminf_{n\to\infty} E[U^-(\tilde{f}_n)] \tag{13.22}$$

so by subtraction we get

$$E[U(f_\infty)] \geq \limsup_{n\to\infty} E[U(\tilde{f}_n)] \geq u(x) \tag{13.23}$$

hence f_∞ is optimal.

□

Theorem 13.4. *Suppose $\mathcal{F}_0$ is trivial, $u(x_0) < \infty$ for some $x_0 > 0$, $\mathcal{P}_{e,\sigma} \neq \emptyset$ and either $U(\infty) < \infty$ or $U \geq 0$ and $RAE_\infty(U)$. Then for any $x > 0$, the primal problem of expected utility maximization from terminal wealth admits a unique solution $f_x^* \in \mathcal{C}(x)$.*

Proof. See IMF T IV 7.2. and the subsequent remarks. □

Remark: The previous theorem does not cover log-utility $U(x) = \log x$, but thm 13.3 does at the extra cost of assuming the existence of $\tilde{Q}$ with the moments property.