Mathematical finance
(extended from lectures of Fall 2012
by Martin Schweizer
transcribed by Peter Gracar and Thomas Hille)

Josef Teichmann

Contents

CHAPTER 1

Arbitrage Theory

1. Stochastic Integration

It is one of the fascinating aspects in the history of sciences that deep discoveries developed in a “l’art pour árt” spirit often find unintentionally important and far-reaching applications. This happened to general stochastic integration theory being developed in the second half of the twentieth century in,. e.g, France, Soviet Union and USA, and being applied since the eighties in *all* its depth in mathematical finance. Therefore a proper understanding of mathematical finance needs all the main concepts of stochastic integration theory for a proper working with models.

Often being asked by students if this knowledge is necessary to work in financial industry I can provide two answers: first: of course not, since models in industry can be understood from a far more concrete point of view. This is the answer by business schools. However, second, if one wants to develop models, fully understand the pitfalls of modeling, then of course the proper knowledge all possible models is fundamental and stochastic integration is the key to this knowledge. Motivation is coming in later sections, therefore I go right away towards the first theorem: the Bichteler-Dellacherie Theorem telling that the set of *good integrators* for stochastic integrals coincides with the set of semi-martingales.

We consider here a time horizon $T = 1$ and a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying the usual conditions in order to guarantee that martingales, sub-martingales and super-martingales always have càdlàg (right continuous with left limits) versions (Doob’s regularity theorem). In Finance it is usually not restrictive to assume the stochastic processes modeling asset prices have càdlàg trajectories, since if we deviate from continuous processes with think of jumps not being announced from the left.

Basic definitions and properties are usually taken from Philip Protter’s excellent book on stochastic integration [**11**], or from Olav Kallenberg’s incredibly useful book on general probability theory [**9**]. We assume here acquaintance with martingale regularization and Doob’s optional sampling theorem and the predictable σ-algebra. We also use deliberately notions like S^* for the process of the running supremum $\sup_{s \leq t} S_s$.

Definition 1.1. *A property is said to hold* locally *for an adapted stochastic process with càdlàg trajectories* $(S_t)_{t\in[0,1]}$ *if the process* $S^{\tau_n} := (S_{t\wedge\tau})_{t\in[0,1]}$ *fulfills this property and if the sequence of* $[0, 1] \cup \{\infty\}$*-valued stopping times* τ_n *satisfies* $\mathbb{P}(\tau_n = \infty) \to 1$ *as* $n \to \infty$*. Notice that this is equivalent to the standard definition of locality on* $[0, \infty[$ *by extending all processes beyond* 1 *in a constant way.*

Definition 1.2. *A semi-martingale S is an adapted process with càdlàg trajectories such that there exists an adapted local martingale M with càdlàg trajectories and an adapted process of finite total variation A representing S as sum $S = M+A$.*

Definition 1.3. *We call*

$$H = \sum_{i=0}^{n} H_i 1_{]\tau_i,\tau_i+1]} \,,$$

with $n \in \mathbb{N}$, stopping times $0 = \tau_0 \le \tau_1 \le \dots \le \tau_n \le \tau_{n+1} = 1$ and H^i that is $\mathbb{R}^d$-valued, and $\mathcal{F}_{\tau_{i-1}}$-measurable, a simple predictable process. We write $H \in b\mathcal{E}$.

Remark 1.4. In the realm of mathematical finance we shall often speak of (simple) predictable strategies instead of processes underlining that H corresponds to trading strategies.

We write $\vartheta \in b\mathcal{E}_{det}$ if in addition the τ_i (but not the H_i) are deterministic. Furthermore we consider $b\mathcal{E}$ equipped with the topology of uniform convergence induced by the norm

$$\|H\|_\infty = \Big\| \sup_{0\le t\le 1} |H_t| \Big\|_\infty$$

for $H \in b\mathcal{E}$. On the space of simple predictable processes we define a linear operator $\mathcal{I}_S$, called stochastic integral with respect to a càdlàg, adapted process S, by

$$\mathcal{I}_S(H) := (H \bullet S)_1 := \sum_{i=1}^n h_i (S_{\tau_i} - S_{\tau_{i-1}})$$

mapping to the space of random variables $L^0(\mathbb{P})$ equipped with the topology of convergence in probability.

Definition 1.5. *We call an adapted process $(S_t)_{t\in[0,1]}$ with càdlàg trajectories a* good integrator *if the map $\mathcal{I}_S : b\mathcal{E} \to L^0(\mathbb{P})$ is continuous.*

Remark 1.6. If one interprets the stochastic integral $\mathcal{I}_S(H)$ as cumulative gains and losses process in a trading context, then good integrators are those models, where the outcome of a trading strategy is robust for small uniform changes of the portfolio strategy. This is an extremely reasonable requirement for models of asset prices. Hence it is of interest to understand which processes are good integrators. Notice that it might be reasonable to consider already here discounted values of price processes S to make prices at different times comparable.

Remark 1.7. Notice that a local good integrator is a good integrator.

In the sequel we characterize the set of good integrators. For a long time this characterization was technically quite involved, however, recently more elementary proofs have been obtained from re-inspecting classical proofs and replacing Dunford-Pettis type arguments by more elementary Komlos type arguments (see the works of Mathias Beiglböck, Walter Schachermayer, Pietro Siorpaes, Bezirgen Veliev, et al). We follow these elementary approaches in these lecture notes very closely in spirit and proofs, in particular the preprint [**3**].

First we show that bounded, adapted, càdlàg processes S, which are good integrators, have bounded mean variation.

Let $S = (S_t)_{0\le t\le 1}$ be an adapted process such that $S_t \in L^1(\mathbb{P})$ for all $t \in [0,1]$. Given a random partition $\pi = \{0 = \tau_0 \le \tau_1 \le \ldots \le \tau_n = 1\}$ of $[0,1]$, the *mean variation* of S *along* π is defined as

$$\mathrm{MV}(S,\pi) = \mathbb{E}\Big[\sum_{\tau_i\in\pi} \big|\mathbb{E}[S_{\tau_{i+1}} - S_{\tau_i} | \mathcal{F}_{\tau_i}]\big| \Big] .$$

The mean variation $\mathrm{MV}(S)$ is the supremum over all partitions π of $\mathrm{MV}(S,\pi)$, i.e.

$$\mathrm{MV}(S) := \sup_\pi \mathrm{MV}(S,\pi) .$$

Mean variation can be interpreted as the cumulative sum of absolute values of conditional expectations of returns. It is quite intuitive that adapted stochastic processes with càdlàg trajectories have bounded mean variation if the "hidden" drift is not overwhelmingly big. The drift "hidden" in a return is a colloquial notion for the conditional expectation of the return.

The mean variation along π is an increasing function of π, i.e. we have $\mathrm{MV}(S,\pi) \leq \mathrm{MV}(S,\pi')$ whenever π' refines π. Let S be bounded with càdlàg trajectories: having a sequence of refining partitions $(\pi^n)_{n\geq 1}$ such that the mesh tends to zero in probability, then the limit along this sequence tends to $MV(S)$ due to $\lim_{t\searrow s} E(S_t - S_s|\mathcal{F}_s] = 0$.

LEMMA 1.8. *Let S be a bounded, adapted with càdlàg trajectories and assume that S is a good integrator, then for every $\epsilon > 0$ there is a stopping time ϱ taking values in $[0,1]\cup\{\infty\}$ such that $\mathbb{P}(\varrho = \infty) \geq 1-\epsilon$ and such that S^ϱ has bounded mean variation (we say that* locally *S has bounded mean variation).*

REMARK 1.9. See [**1**]: for the proof we need the (easy) L^2-version of Komlos' Lemma: let $(g_n)_{n\geq 1}$ be a bounded sequence in $L^2(\mathbb{P})$, then we can find elements $h_n \in C_n := \mathrm{conv}(g_n, g_{n+1}, \ldots)$ which converge almost surely and in $L^2(\mathbb{P})$ to some element h. For the proof of Komlos' Lemma we take

$$A = \sup_{n\geq 1} \inf_{g\in C_n} \|g\|^2 ,$$

then there are elements $h_n \in \mathrm{conv}(g_n, g_{n+1}, \ldots)$ such that $\|h_n\|^2 \leq A + \frac{1}{n}$. Fix $\epsilon > 0$, then there is n large enough such that for all $k, m \geq n$ the inequality $\|h_k + h_m\|^2 > 4(A-\epsilon)$ holds true, since the sup is along an non-decreasing sequence!. By the parallelogram-identity we then obtain

$$\|h_k - h_m\|^2 = 2\|h_k\|^2 + 2\|h_m\|^2 - \|h_k + h_m\|^2 < 4(A + \frac{1}{n}) - 4(A-\epsilon) = 4\epsilon + \frac{1}{n} ,$$

which yields the assertion of $L^2(\mathbb{P})$ convergence by completeness. By passing to a subsequence the almost sure convergence follows, too.

In an analog manner an L^1 version for uniformly integrable sequences $(f_n)_{n\geq 1}$ can be proven by truncation from the L^2 statement and an appropriate diagonalization argument: indeed let $g_n^m := (f_n 1_{\{|f_n|\leq i\}})_{1\leq i\leq m}$ be the truncated sequence, then we know that there are forward looking convex combinations in $\mathrm{conv}(g_n^m, g_{n+1}^m, \ldots)$ converging in $L^2(\mathbb{P})$ and almost surely to a limit h^m, for every $m \geq 1$. In other words we can choose weights which work for the first m truncated sequences simultaneously. Uniform integrability then yields that the diagonal sequence of convex combinations does the job since $\lim_{i\to\infty} \|(f_n 1_{\{|f_n|\leq i\}} - f_n\|$ uniformly in n.

PROOF. We follow closely [**3**]: the proof is in spirit typical for mathematical finance (we shall see this later!): whenever a functional depending on the future path is given, one tries to mimick it by a stochastic integral, i.e. the outcome of a trading strategy. In this very case this is particularly easy. S is a good integrator, hence for $\epsilon > 0$ there exists $C > 0$ such that for all simple processes H with $\|H\|_\infty \leq 1$ we have $\mathbb{P}((H\bullet S)_1 \geq C - 2\|S\|_\infty) \leq \epsilon$ (this is just a translation of what it means to be continuous from the uniform topology to convergence in probability).

For each $n \geq 1$ we can define a simple process H^n and a stopping time ϱ_n

$$\begin{aligned} H^n &:= \textstyle\sum_{t_i\in D_n} 1_{(t_i,t_{i+1}]}\, \mathrm{sign}\left(\mathbb{E}[S_{t_{i+1}} - S_{t_i}|\mathcal{F}_{t_i}]\right), \\ \varrho_n &:= \inf\{t \in D_n : (H^n \bullet S)_t \geq C - 2\|S\|_\infty\}. \end{aligned}$$

allowing to mimick (in expectation) the bounded variation functional. On the set $\{\varrho_n < \infty\}$,

$$(H^n 1_{(0,\varrho_n]} \bullet S) = (H^n \bullet S)_1^{\varrho_n} \geq C - 2\|S\|_\infty$$

holds true, and therefore $\mathbb{P}(\varrho_n = \infty) \geq 1 - \epsilon$. Furthermore S is bounded, so the jumps of S are bounded by $2\|S\|_\infty$, and whence

$$C \geq (H^n \bullet S)_1^{\varrho_n}$$

always holds true. Putting together these insights we arrive at

$$\begin{aligned} C \geq \mathbb{E}[(H^n \bullet S)_1^{\varrho_n}] &= \mathbb{E}\Big[\sum_{t_i \in D_n} \mathbb{1}_{\{t_i < \varrho_n\}} \operatorname{sign}\big(\mathbb{E}[S_{t_{i+1}} - S_{t_i} | \mathcal{F}_{t_i}]\big)(S_{t_{i+1}} - S_{t_i})\Big] = \\ &= \mathbb{E}\Big[\sum_{t_i \in D_n} \mathbb{1}_{\{t_i < \varrho_n\}} \big|\mathbb{E}[(S_{t_{i+1}} - S_{t_i}) | \mathcal{F}_{t_i}]\big|\Big] = \mathrm{MV}(S^{\varrho_n}, D_n) \ , \end{aligned}$$

which concludes the first estimate.

Next we would like to replace the stopping time ϱ_n by a "sort" of accumulation point ϱ such that we can conclude the desired statement. For this very purpose we apply a Komlos-type Lemma to the random variables $X_n = 1_{\{\varrho_n = \infty\}} \in L^2(\mathbb{P}), n \geq 1$ to obtain for each n convex weigths $\mu_n^n, \ldots, \mu_{N_n}^n$ such that

$$Y_n := \mu_n^n X_n + \ldots + \mu_n^{N_n} X_{N_n}$$

converges to some random variable X in $L^2(\mathbb{P})$. By passing to a subsequence we can assume that convergence is almost surely.

From $0 \leq X \leq 1$ and $\mathbb{E}[X] \geq 1 - \epsilon$ we deduce that $\mathbb{P}(X < 2/3) < 3\epsilon$. Since $\mathbb{P}(\lim_m Y_m \geq 2/3) > 1 - 3\epsilon$, by Egoroff's theorem we deduce that there exists a set A with $\mathbb{P}(A) \geq 1 - 3\epsilon$ such that $Y_n \geq 1/2$ on the set A, for all n greater or equal than some $n_0 \in \mathbb{N}$, which we can assume to be equal to 1.

We now define the desired stopping time ϱ by

$$\varrho = \inf_{n \geq 1} \inf\{t : \mu_n^n \mathbb{1}_{[0,\varrho_n]}(t) + \ldots + \mu_n^{N_n} \mathbb{1}_{[0,\varrho_{N_n}]}(t) < 1/2\} \ .$$

Then clearly $A \subseteq \{\varrho = \infty\}$ and we arrive at $\mathbb{P}(\varrho = \infty) \geq 1 - 3\epsilon$. The stopping time ϱ apparently has the desired properties, since

(1.1)

$$\mathbb{E}\Big[\sum_{t_i \in D_n} \mathbb{1}_{\{t_i < \varrho\}} \big|\mathbb{E}[S_{t_{i+1}} - S_{t_i} | \mathcal{F}_{t_i}]\big|\Big] \leq 2\mathbb{E}\Big[\sum_{t_i \in D_n} \sum_{k=n}^{N_n} \mu_k^n \mathbb{1}_{\{t_i < \varrho_k\}} \big|\mathbb{E}[S_{t_{i+1}} - S_{t_i} | \mathcal{F}_{t_i}]\big|\Big] \, .$$

The left hand side differs from $MV(S^\varrho, D_n)$ at most by $2\|S\|_\infty$, whereas the right hand side is bounded by

$$2 \sum_{k=n}^{N_n} \mu_k^n (MV(S^{\varrho_n}, D_n) + 2\|S\|_\infty) \leq C + 4\|S\|_\infty,$$

which yields $MV(S^\varrho, D_n) \leq C + \|S\|_\infty$, i.e. bounded mean variation of S. □

The next step towards a characterization of good integrators is to understand that processes with bounded mean variation are nothing else than differences of non-negative super-martingales. This assertion is also called Rao's theorem:

Proposition 1.10. *Let S be an adapted, L^1-process with càdlàg trajectories and bounded mean variation, then S is the difference of two adapted, càdlàg super-martingales.*

Proof. This is a classical proof which can also be found in [**3**]: in the proof we directly construct the two super-martingales on the dyadic grid $D = \cup_{n \geq 1} D_n$

with $D_n = \{\frac{j}{2^n} \mid 0 \leq j \leq 2^n\}$ by

$$Y_s^n = \mathbb{E}\Big[\sum_{t_i \in D_n, t_i \geq s} \mathbb{E}[S_{t_i} - S_{t_{i+1}} | \mathcal{F}_{t_i}]^+ | \mathcal{F}_s\Big]$$

$$Z_s^n = \mathbb{E}\Big[\sum_{t_i \in D_n, t_i \geq s} \mathbb{E}[S_{t_i} - S_{t_{i+1}} | \mathcal{F}_{t_i}]^- | \mathcal{F}_s\Big]$$

for $s \in D_n$, $n \geq 1$. Apparently the discrete processes Y^n and Z^n are non-negative super-martingales, $Y_s^n - Z_s^n = S_1 - E[S_1|\mathcal{F}_s]$ if $s \in D_n$ and due to Jensen's inequality

$$Y_s^n \leq Y_s^{n+1}$$

for $s \in D_n$, $n \geq 1$. Addionally we have the bound $E[Y_s^n + Z_s^n] \leq MV(S)$, which yields L^1 convergence by monotone convergence. Therefore we can define limit processes

$$Y_s := \lim_{n\to\infty} Y_s^n + E[S_1^+|\mathcal{F}_s]$$

$$Z_s := \lim_{n\to\infty} Z_s^n + E[S_1^-|\mathcal{F}_s]$$

and hence $Y_s - Z_s = S_s$ for $s \in D$. By Doob's regularity theorem we can extend the super-martingales to $[0, 1]$ as càdlàg processes. Since the trajectories of S are càdlàg, too, the decomposition holds for all times. □

So far we have shown that locally a good integrator is locally a difference of two non-negative super-martingales. Finally we have to show that a super-martingale can be written as a difference of a local martingale and a predictable process of finite total variation. This statement, whose proof is again elementary, is called the Doob-Meyer decomposition theorem.

Definition 1.11. *An adapted process S with càdlàg trajectories is called of class (D) if the set $\{S_\tau \mid \tau$ stopping time $\}$ is uniformly integrable.*

Theorem 1.12. *Let S be a super-martingale of class (D), then there is a martingale M and a predictable process A, both with càdlàg trajectories such that*

$$S = M + A$$

holds true. The decomposition is unique.

Remark 1.13. If $(S_t)_{t\in[0,1]}$ is only a super-martingale without belonging to class (D), then it is still locally of class (D) and therefore locally the Doob-Meyer decomposition holds, which yields a *global* decomposition of S into a local martingale M and a predictable process A. Indeed considering the stopping time T_n, when $|S|$ is greater than n for the first time, $n \geq 1$, then for *any other* stopping time σ we have $|S_\sigma^{T_n}| \leq n + |S_{1\wedge T_n}|$ which is integrable by optional sampling, whence S^{T_n} is of class (D) and $\mathbb{P}(T^n = \infty) \to 1$ as $n \to \infty$.

Proof. We follow closely [**1**], where again the proof relies in a discrete insight combined with a Komlos type limiting procedure (see also [**11**]): the discretely sampled process $(S_t)_{t\in D_n}$ for $n \geq 1$ has a Doob-Meyer decomposition, namely the processes A^n and M^n with $A_0^n = 0$,

$$A_{t_{i+1}} - A_{t_i} = E[S_{t_{i+1}} - S_{t_i} \mid \mathcal{F}_{t_i}]$$

for $t_i \in D_n$ and $M_t^n = S_t - A_t^n$, for $t \in D_n$, $n \geq 1$. The increments of the martingale M^n are the conditional drift corrected increments of S.

We now try to prove uniform integrability of the sequence $(A_1^n)_{n\geq 1}$ in order to apply the Komlos argument for convergence: By subtracting $E[S_1|\mathcal{F}_t]$ from S_t, for

$t \in [0,1]$ we can assume that $S_t \geq 0$ and $S_1 = 0$. For the discrete martingales we have – by optional sampling – that

$$S_\tau = A^n_\tau - E(A^n_1 \,|\, \mathcal{F}_\tau]$$

since $M^n_1 = -A^n_1$ for $n \geq 1$, and any stopping time τ with respect to the sampled filtration. We define for $c > 0$ and $n \geq 1$ a stopping time

$$\tau_n(c) = \inf\big\{(j-1)/2^n : A^n_{j/2^n} < -c\big\} \wedge 1\,.$$

From $A^n_{\tau_n(c)} \geq -c$ we obtain $S_{\tau_n(c)} \geq -E[A^n_1|\mathcal{F}_{\tau_n(c)}] - c$. Thus,

$$-\int_{\{-A^n_1 > c\}} A^n_1\, d\mathbb{P} = -\int_{\{\tau_n(c)<1\}} \mathbb{E}[A^n_1|\mathcal{F}_{\tau_n(c)}]\, d\mathbb{P} \leq c\,\mathbb{P}\big[\tau_n(c) < 1\big] + \int_{\{\tau_n(c)<1\}} S_{\tau_n(c)}\, d\mathbb{P}\,.$$

Since $\{\tau_n(c) < 1\} \subseteq \{\tau_n(\frac{c}{2}) < 1\}$, we have again

$$\begin{aligned}\int_{\{\tau_n(\frac{c}{2})<1\}} (S_{\tau_n(\frac{c}{2})})\, d\mathbb{P} &= \int_{\{\tau_n(\frac{c}{2})<1\}} (-A^n_1 + A^n_{\tau_n(\frac{c}{2})})\, d\mathbb{P} \\ &\geq \int_{\{\tau_n(c)<1\}} (-A^n_1 + A^n_{\tau_n(\frac{c}{2})})\, d\mathbb{P} \geq \frac{c}{2}\,\mathbb{P}[\tau_n(c) < 1]\,.\end{aligned}$$

Combining the above two inequalities we obtain

$$-\int_{\{-A^n_1 > c\}} A^n_1\, d\mathbb{P} \leq 2\int_{\{\tau_n(\frac{c}{2})<1\}} S_{\tau_n(\frac{c}{2})}\, d\mathbb{P} + \int_{\{\tau_n(c)<1\}} S_{\tau_n(c)}\, d\mathbb{P}\,. \tag{1.2}$$

On the other hand

$$\mathbb{P}[\tau_n(c) < 1] = \mathbb{P}[-A^n_1 > c] \leq -\mathbb{E}[A^n_1]/c = \mathbb{E}[M^n_1]/c = \mathbb{E}[S_0]/c,$$

hence, as $c \to \infty$, $\mathbb{P}[\tau_n(c) < 1]$ goes to 0, uniformly in n. As the process S is of class (D), (1.2) implies that $(A^n_1)_{n\geq 1}$ is uniformly integrable and hence also $(M^n_1)_{n\geq 1} = (S_1 - A^n_1)_{n\geq 1}$.

We can extend M^n to a càdlàg martingale on $[0,1]$ by setting $M^n_t := \mathbb{E}[M^n_1|\mathcal{F}_t]$. By Komlos' Lemma there exist $M_1 \in L^1(\mathbb{P})$ and for each n convex weights $\lambda^n_n, \ldots, \lambda^n_{N_n}$ such that the discrete martingales

$$\mathcal{M}^n := \lambda^n_n M^n + \ldots + \lambda^n_{N_n} M^{N_n} \tag{1.3}$$

converge at the end point $t = 1$, i.e. $\mathcal{M}^n_1 \to M_1$ in $L^1(\mathbb{P})$. By Jensen's inequality $\mathcal{M}^n_t \to M_t := \mathbb{E}[M_1|\mathcal{F}_t]$ for all $t \in [0,1]$. For each $n \geq 1$ we extend A^n to $[0,1]$ by

$$A^n := \textstyle\sum_{t_i \in \mathcal{D}_n} A^n_t \mathbb{1}_{(t_i,t_{i+1}]} \tag{1.4}$$

$$\text{and set}\quad \mathfrak{A}^n := \lambda^n_n A^n + \ldots + \lambda^n_{N_n} A^{N_n}\,, \tag{1.5}$$

where we use the same weights as in (1.3). Then the càdlàg process $A := S - M$ satisfies for every $t \in \mathcal{D}$

$$\mathfrak{A}^n_t = (S_t - \mathcal{M}^n_t) \;\to\; (S_t - M_t) = A_t \quad \text{in } L^1(\mathbb{P}).$$

Passing to a subsequence we obtain that convergence holds true almost surely. Consequently, A is almost surely increasing on $\mathcal{D}$ and, by right continuity, also on $[0,1]$.

The processes A^n and $\mathfrak{A}^n$ are left-continuous and adapted, hence predictable. To obtain that A is predictable, we show that for almost every ω and every $t \in [0,1]$

$$\limsup\nolimits_{n\to\infty} \mathfrak{A}^n_t(\omega) = A_t(\omega). \tag{1.6}$$

Let $f_n, f : [0,1] \to \mathbb{R}$ are increasing functions such that f is right continuous and $\lim_{n\to\infty} f_n(t) = f(t)$ for $t \in \mathcal{D}$, then

$$\limsup_{n\to\infty} f_n(t) \le f(t) \text{for all } t \in [0,1] \text{ and}$$
$$\lim_{n\to\infty} f_n(t) = f(t) \text{ if } f \text{ is continuous at } t.$$

Consequently, (1.6) can only be violated at discontinuity points of A. As A is càdlàg, every path of A can have only finitely many jumps larger than $1/k$ for $k \in \mathbb{N}$. It follows that the points of discontinuity of A can be exhausted by a countable sequence of stopping times, and therefore it is sufficient to prove $\limsup_{n\to\infty} \mathfrak{A}^n_\tau = A_\tau$ for every stopping time τ.

By the previous inequalities $\limsup_{n\to\infty} \mathfrak{A}^n_\tau \le A_\tau$ and as $\mathfrak{A}^n_\tau \le \mathfrak{A}^n_1 \to A_1$ in $L^1(\mathbb{P})$ we deduce from Fatou's Lemma that

$$\liminf_{n\to\infty} \mathbb{E}\big[A^n_\tau\big] \le \limsup_{n\to\infty} \mathbb{E}\big[\mathfrak{A}^n_\tau\big] \le \mathbb{E}\big[\limsup_{n\to\infty} \mathfrak{A}^n_\tau\big] \le \mathbb{E}\big[A_\tau\big]\,.$$

Therefore it suffices to prove $\lim_{n\to\infty} \mathbb{E}[A^n_\tau] = \mathbb{E}[A_\tau]$. For $n \ge 1$ set

$$\sigma_n := \inf\{t \in \mathcal{D}_n : t \ge \tau\}\,.$$

Then $A^n_\tau = A^n_{\sigma_n}$ and $\sigma_n \downarrow \tau$. Using that S is of class D, we obtain

$$\mathbb{E}[A^n_\tau] = \mathbb{E}[A^n_{\sigma_n}] = \mathbb{E}[S_{\sigma_n}] - \mathbb{E}[M_0] \to \mathbb{E}[S_\tau] - \mathbb{E}[M_0] = \mathbb{E}[A_\tau]\,,$$

which ends the proof. □

THEOREM 1.14. *Let S be a good integrator, then S is the sum of a local martingale and a process of finite total variation A.*

PROOF. We follow [**3**]: the process S can written, as any adapted, càdlàg process, as the sum of a locally bounded process and an adapted càdlàg process of finite total variation, namely

$$S_t = (S_t - J_t) + J_t$$

with $J_t := \sum_{s\le t}(S_s - S_{s-})1_{|S_s - S_{s-}|\ge 1}$. Hence we can assume with loss of generality that S is locally bounded. Since S is a locally bounded good integrator, S has locally bounded mean variation and is therefore locally the difference of two supermartingales. By the Doob-Meyer decomposition, any supermartingale is locally the sum of a local martingale and a process of finite total variation. □

For later purposes we also prove the following easy consequence from Doob's optional sampling theorem for super-martingales

THEOREM 1.15. *Let S be a non-negative super-martingale and $\tau = \inf\{0 \le t \le 1 \mid S_{t-} \wedge S_t = 0\}$ (first hitting time of 0), then S vanishes on the stochastic interval $[\tau, 1]$.*

PROOF. It is sufficient to show the result when S is of class (D), the rest follows by localization. Consider the stopping times $\tau_n := \inf\{0 \le t \le 1 \mid S_t \le 1/n\}$, then $\tau_n \le \tau$ and $\tau_n \nearrow \tau$. In particular we have that $\tau_n = \tau$ for all $n \ge 0$ if and only if $X_\tau = 0$. Hence the stochastic integral along the simple strategy $1_{]\tau_n,\tau]}$ has expectation

$$E[S_\tau - S_{\tau_n}] \le 0$$

due to optional sampling. By dominated convergence we obtain $E[\Delta S_\tau 1_{\{S_{\tau-}=0\}}] \le 0$. This means that upwards jumps at τ cannot happen, so either one jumps to 0 at $\tau-$ from above or we have a point of continuity there. Hence $S_\tau = 0$. Since $E[S_t - S_\tau 1_{\{\tau\le t\}}] \le 0$, we obtain that $S_t = 0$ on the stochastic interval $[\tau, 1]$. □

After having established that good integrators are indeed semi-martingales we have to understand the other direction, i.e. whether all semi-martingales are good integrators, which is based on an interesting inequality valid for super-martingales in the L^1-case or on Ito's fundamental insight in the local L^2-case, respectively:

THEOREM 1.16. *Every semi-martingale is a good integrator.*

PROOF. Since sums of good integrators are good integrators and locally good integrators are good integrators it is sufficient to provide arguments for finite variation processes and for martingales. Since a martingale is not necessarily locally square integrable (jumps!) we have to provide a proof for martingales, which is the tricky bit.

- Any process of finite variation A is a good integrator, since for every simple (not even predictable) process it holds that
$$|\int_0^1 H_s dA_s| \leq \|H\|_\infty \int_0^1 d|A|_s$$
almost surely.
- By Ito's fundamental insight square integrable martingales S are good integrators, since
$$E\big[(H \bullet S)^2\big] \leq \|H\|_\infty E[S_1^2]$$
holds true for simple, bounded and predictable processes $H \in b\mathcal{E}$.
- By the following elementary inequality due to Burkholder we can conclude that martingales are good integrators: for every martingale S and every simple, bounded and predictable process H it holds that
$$c\mathbb{P}(|(H \bullet S)|_1^* \geq c) \leq 18\|H\|_\infty \|S_1\|_1$$
for all $c \geq 0$. For an easy proof of this inequality see [**4**]. Since the inequality has some importance for our treatment, we shall give it here too. Notice that we are just dealing with elementary integrals, so all use of stochastic integration is in fact easily justified: let S be a non-negative martingale first and H bounded predictable with $\|H\|_\infty \leq 1$, then $Z := S \wedge c$ is a supermartingale and we have
$$c\mathbb{P}(|(H \bullet S)|_1^* \geq c) \leq c\mathbb{P}(|S|_1^* \geq c) + c\mathbb{P}(|(H \bullet Z)|_1^* \geq c)\,.$$
Since Z is a supermartingale we have obtain by Doob-Meyer $Z = M - A$ and $(H \bullet Z) \leq (H \bullet M) + A$, which is a submartingale. Hence we can conclude by Doob's maximal inequalities for $p = 2$ in case of the second term and $p = 1$ (weak version) in case of the first term that
$$c\mathbb{P}(|(H \bullet S)|_1^* \geq c) \leq cE[S_1] + 2\frac{1}{c}E[(H \bullet M)_1^2 + A_1^2]\,.$$
Ito's insight allows to estimate the variance of the stochastic integral at time 1 by $E[M_1^2]$. Both quantities M and A of the Doob-Meyer decomposition may however be estimated through $E[A_1^2] \leq E[M_1^2] \leq 2cE[Z_0]$, since Z is non-negative (so $A \leq M$ holds true) and $Z \leq c$. For an easy proof see [**4**, Lemma 3.6]. This leads to an upper bound
$$c\mathbb{P}(|(H \bullet S)|_1^* \geq c) \leq 9E[S_0]\,.$$
Writing a martingale as difference of two non-negative martingales leads to the desired result. Apparently the result translates directly to the fact that S is a good integrator.

- Any pre-local good integrator is a good integrator, hence all semimartingales are good integrators since sums of good integrators are good integrators. A process S is said to satisfy a property *pre-locally* if there is an increasing sequence of stopping times τ_n taking values in $[0,1] \cup \{\infty\}$ such that $P(\tau_n = \infty) \to 1$ as $n \to \infty$ and
$$S^{\tau_n -} := S_t 1_{\{0 \leq \tau_n < t\}} + S_{\tau_n -} 1_{\{t \geq \tau_n\}}$$
satisfies the property for all $n \geq 0$. This provides another proof for integrability of martingales since every local martingale is pre-locally square integrable by stopping when the process leaves a bound.

□

Having established by minimal requirements the class of good integrators it is our goal to extend for a given semi-martingale S the class of integrands H. There is a first direct step, which extends the set of integrands towards càglàd processes (left continuous with right limits existing), which is just the closure of bounded, simple predictable processes with respect to the metric "uniform convergence along paths in probability" on càdlàg or càglàd processes given, e.g., by
$$d(S_1, S_2) := E\big[|(S_1 - S_2)|_1^* \wedge 1\big] .$$

THEOREM 1.17. *For every semi-martingale S the map I_S defined on $b\mathcal{E}$ extends to a continuous map J_S from the space $\mathbb{L}$ of càglàd processes to $\mathbb{D}$ of càdlàg processes. Notice that the spaces $\mathbb{L}$ and $\mathbb{D}$ are complete with respect to the metric d.*

PROOF. See the next result, which is proved for an even stronger topology on the image space. □

The Emery topology on the set of semimartingales $\mathbb{S}$ is defined by the metric
$$d_E(S_1, S_2) := \sup_{K \in b\mathcal{E},\, \|K\|_\infty \leq 1} E\big[|(K \bullet (S_1 - S_2))|_1^* \wedge 1\big] .$$
We can by means of the Bichteler-Dellacherie theorem easily prove the following important theorem, which goes back to Michel Emery, see [**7**] (notice that Michel Emery defines the metric by a supremum over all bounded predictable processes, we cannot do that at the moment since this integral is not defined yet).

THEOREM 1.18. *The set of semi-martingales $\mathbb{S}$ is a topological vector space and complete with respect to the Emery topology.*

PROOF. Obviously d_E defines a metric and a Cauchy sequence $(S_n)_{n \geq 1}$ in d_E is a Cauchy sequence in d, so there is a càdlàg process S which is the pathwise uniform limit of the semi-martingales S_n. We have to show that S is a semi-martingale. We show that by proving that I_S is continuous on $b\mathcal{E}$ with respect to the uniform topology, which is equivalent to the fact that the set $\{(K \bullet S)_1 \mid K \in b\mathcal{E}, \|K\|_\infty \leq 1\}$ is bounded in probability. Fix $1 > \epsilon > 0$, then for $c > 0$
$$\mathbb{P}\big((K \bullet S)_1 \geq c\big) \leq \mathbb{P}\big((K \bullet S - S_n)_1 \geq c\big) + \mathbb{P}\big((K \bullet S_n)_1 \geq c\big) \tag{1.7}$$
$$\leq \frac{d_E(S, S_n)}{c} + \mathbb{P}\big((K \bullet S_n)_1 \geq c\big) . \tag{1.8}$$
Now we choose n large enough to make the first term smaller than $\frac{\epsilon}{c}$. Since S_n is a semi-martingale I_{S_n} is continuous, hence the set $\{(K \bullet S_n)_1 \mid K \in b\mathcal{E}, \|K\|_\infty \leq 1\}$ is bounded in probability, which in turn means that we can choose c large enough such that the second term is smaller than ϵ. Hence both terms are small and therefore S is a good integrator. □

THEOREM 1.19. *For every semi-martingale S the map I_S defined on $b\mathcal{E}$ extends to a continuous map J_S from the space $\mathbb{L}$ of càglàd processes to $\mathbb{S}$ of semi-martingales.*

PROOF. It is sufficient to show the result for martingales S, since the rest follows by localization and the respective theorem for finite variation processes. Let S be a martingale. Take a sequence H_n which converges pathwise uniformly in probability to 0, i.e. $\mathbb{P}(|H_n|_1^* \geq b) \to 0$ as $n \to \infty$. Fix furthermore $K \in b\mathcal{E}$ with $\|K\|_\infty \leq 1$. We can decompose $H_n = H_n' + H_n''$ where $H_n' := H_n 1_{\{|H_n|^* \geq b\}}$ for some $b \geq 0$. This decomposition is of course done in $b\mathcal{E}$. Observe that $H_n' H_n'' = 0$ for all $n \geq 1$. Now we can estimate through

$$\{|(KH_n \bullet S)|_1^* \geq c\} \subset \{|H_n|_1^* \geq b\} \cup \{|(KH_n'' \bullet S)|_1^* \geq c\}$$

the probabilities directly

$$c\mathbb{P}(|(KH_n \bullet S)|_1^* \geq c) \leq c\mathbb{P}(|H_n|_1^* \geq b) + 18\|KH_n''\|_\infty \|S_1\|_1 ,$$

where we notice that $\|H_n''\|_\infty \leq b$ and $\mathbb{P}(|H_n|^* \geq b) \to 0$ as $n \to \infty$. This, however, yields that

$$\sup_{K \in b\mathcal{E},\, \|K\|_\infty \leq 1} E\big[|(KH_n \bullet S)|_1^* \wedge 1\big] \leq \mathbb{P}(|H_n|^* \geq b) + \frac{18}{c}\|H_n''\|_\infty \|S_1\|_1 + c$$

for each $c > 0$. For every chosen $b > 0$ and $c > 0$ we see that as n tends to ∞ the right hand side converges to $\frac{18b}{c}\|S_1\|_1 + c$ which is small for appropriate choices of the constants b, c. Consequently Cauchy sequences in $\mathbb{L}$ are mapped to Cauchy sequences in the Emery topology, which – due to completeness – converge to a semi-martingale. □

At this point we can formulate the single most important notion of semi-martingale theory: optional quadratic variation.

DEFINITION 1.20. *Let S be a semi-martingale, then the semi-martingale*

$$[S, S] = S^2 - 2(S_- \bullet S)$$

is called optional quadratic variation. $[S, S]$ is a non-negative increasing process. For properties see [**12**] *and* [**11**].

By means of quadratic variation we can introduce the Banach space in the theory of semi-martingales:

$$\mathcal{H}^1 := \left\{ M \in \mathcal{M}_{loc} \mid E[|M|_1^*] < \infty \right\} .$$

REMARK 1.21. Notice that every $\mathcal{H}^1$-martingale is uniformly integrable and that every local martingale is in fact locally $\mathcal{H}^1$: it is sufficient to see this for martingales, but there localization by crossing a finite limit is enough since the last jump is integrable.

On the $\mathcal{H}^1$ we have the single most important inequality of stochastic integration theory: Davis inequality.

THEOREM 1.22. *There are constants $0 < c < C$ such that for every local martingale M with $M_0 = 0$*

$$cE\big[\sqrt{[M, M]_1}\big] \leq E\big[|M|_1^*\big] \leq CE\big[\sqrt{[M, M]_1}\big] .$$

PROOF. For the full proof, in particular of the elementary deterministic inequalities, see [**2**]. It is sufficient to proof the inequality for discrete time martingales, since the result follows from a standard limiting procedure. For discrete martingales we can consider deterministic inequalities of the type

$$\sqrt{[x, x]_N} \leq 3|x|_N^* - (h \bullet x)_N$$

and

$$|x|_N^* \le 6\sqrt{[x,x]_N} + 2(h \bullet x)_N$$

for a "predictable" strategy

$$h_i := \frac{x_i}{\sqrt{[x,x]_i + (|x|_i^*)^2}}$$

for $i = 1, \ldots, n$. Here we consider sequences $0 = x_0, x_1, \ldots, x_N$, wherefrom Davis inequality follows immediately. □

The next extension towards bounded predictable processes is more delicate and cannot be achieved by simple continuous extensions. We do need an additional convergence property with respect to some weaker notion of convergence, which will be given by dominated convergence. In the literature there are several ways to approach this problem: the first one is to pass via L^2-integration theory (Ito's insight) applying the fact that every semi-martingale can be written as a sum of a locally square integrable martingale and a finite variation process, which needs from time to time elements of the general theory (Doob-Meyer decomposition, fundamental theorem of local martingales). The second way is to work directly with L^1-integration theory which essentially needs the Davis inequality (see, e.g., the excellent introduction to (vector-valued) stochastic integration [**12**]). The third one is to apply a continuity result for the Emery topology on the set of semi-martingales. We shall follow a somehow modified third approach here. We believe that this way provides us with a quick and direct way towards stochastic integration:

For this purpose we need the following important Lemma, which L^∞-version is due to Kostas Kardaras in [**10**]:

Lemma 1.23. *Let S^n be a sequence of martingales such that $|\Delta S^n_\tau| \le |\Delta Y_\tau|$, for all $n \ge 1$ and all stopping times τ, for some martingale Y, and let $E[[S^n, S^n] \wedge 1] \to 0$ as $n \to \infty$, then $S^n \to 0$ in the Emery topology.*

Proof. Consider an arbitrary sequence $K^n \in b\mathcal{E}$ of simple, predictable processes bounded by 1. We show first that $E[|(K^n \bullet S^n)_1| \wedge 1] \to 0$. We first observe that also

$$E[[(K^n \bullet S^n), (K^n \bullet S^n)]_1 \wedge 1] \to 0,$$

since K^n is uniformly bounded by 1. We use $M^n := (K^n \bullet S^n)$ as abbreviation and select a subsequence n_k such that $\mathbb{P}(|M_{n_k}| \ge 2^{-k}) \le 2^{-k}$ for all $k \ge 1$, then $A := \sum_k [M^{n_k}, M^{n_k}]$ is almost surely finite by Borel-Cantelli and we can consider the stopping time

$$\tau_m := \inf\{t \mid A_t \ge m \text{ or } |Y_t| \ge m\} \wedge 1\,,$$

which apparently leads to

$$[M^{n_k}_{\tau_m}, M^{n_k}_{\tau_m}] \le A_{\tau_m} + (\Delta M^n_{\tau_m})^2 \le m + (\Delta Y_{\tau_m})^2 \le m + (m + |Y_{\tau_m}|)^2$$

for each n_k, which leads – after taking square-roots – to $E[|M^{n_k}_{\tau_m}|_1^*] \to 0$ for $k \to \infty$ by Davis inequality. Since $\mathbb{P}(\tau_m = \infty) \to 1$ as $m \to \infty$ we do also have that $(K^{n_k} \bullet S^{n_k}) \to 0$ in probability. However this already yields the result, since we have proved that every sequence $((K^n \bullet S^n))_{n \ge 0}$ has a subsequence which converges. Applying this result twice, we see that $(K^n \bullet S^n) \to 0$ in probability for all bounded by 1 simple predictable strategies, which characterizes convergence in the Emery topology. □

Corollary 1.24. *Let S be a martingale and $H^n \to 0$ a sequence of simple, predictable bounded by 1 strategies, then $S^n = (H^n \bullet S)$ converges to 0 in the Emery topology.*

PROOF. Apparently the conditions on jumps and martingality are satisfied. It remains to show that $E[[S^n, S^n] \wedge 1] \to 0$, which is true since $[S^n, S^n] = ((H^n)^2 \bullet [S, S]) \to 0$ almost surely by the existence of $[S, S]$ and the fact that it is pathwise of finite variation. But from almost sure convergence we conclude convergence in probability. □

These considerations lead us to the main extension result, compare [**4**]:

THEOREM 1.25. *Let S be a martingale, then there is a unique continuous linear map $(.\bullet S)$ from bounded, predictable processes with respect to the uniform topology to semi-martingales $\mathbb{S}$ in the Emery topology extending J_S such that dominated convergence holds true, i.e. if $H_n - H_m \to 0$ pointwise, as $n, m \to \infty$ for a sequence of simple, predictable strategies with $\|H_n\|_\infty \leq 1$, for $n \geq 1$, then $((H_n - H_m)\bullet S) \to 0$ in the Emery topology as $n, m \to \infty$.*

PROOF. This can be seen by the following proof: for every subsequence $m_n \geq n$ we have that $((H_n - H_{m_n})\bullet S) \to 0$ in the Emery topology, hence $(H_n - H_m \bullet S) \to 0$ as $n, m \to \infty$.

The extension of J_S is defined by considering almost surely converging sequences $H_n \to H$ in $b\mathcal{E}$ being uniformly bounded, which yield – by the previous dominated convergence result – Cauchy sequences $((H_n \bullet S))_{n\geq 1}$ in the Emery topology. This, however, means that the limits are semi-martingales, uniquely defined and linearly depending on H. Finally the resulting map is continuous. □

THEOREM 1.26. *Let S be a $\mathcal{H}^1$ martingale and H be a bounded, predictable strategy, then $(H \bullet S) \in \mathcal{H}^1$.*

PROOF. See, e.g., [**12**]. Since every bounded, predictable strategies can be approximated by $H_n \in b\mathcal{E}$ with $|H_n| \leq \|H\|_\infty$ (consider for instance that $b\mathcal{E}$ is dense in the space of all predictable processes with $\|H\| = E[\sqrt{\int_0^1 H_s^2 d[S,S]_s}] < \infty$ with respect to the norm). Since

$$[(H_n \bullet S), (H_n \bullet S)] = \int_0^1 H_s^2 d[S, S]_s$$

we can conclude from Davis inequality that

$$E[|(H_n - H_m \bullet S)|_1^*] \leq EC[\sqrt{\int_0^1 (H_s^n - H_s^m)^2 d[S,S]_s}],$$

which yields that $((H_n \bullet S))_{n\geq 0}$ is a Cauchy sequence and hence converging to a $\mathcal{H}^1$-martingale. □

This result immediately generalizes to semi-martingales: let S be a semi-martingale, then there is a unique continuous linear map $(.\bullet S)$ from bounded, predictable processes with respect to the uniform norm to semi-martingales $\mathbb{S}$ in the Emery topology extending J_S such that dominated convergence holds true: if $H_n \to 0$ pointwise for $\|H_n\| \leq 1$, for $n \geq 1$, then $(H_n \bullet S) \to 0$ in the Emery topology.

It is sufficient to see the statement for local martingales: we have to show that bounded convergence holds for local martingales. Let H^n be a sequence of simple, predictable bounded by 1 strategies converging almost surely to 0 and let τ_m be a localizing sequence for S, then $(K^n H^n \bullet S^{\tau_m}) \to 0$ as $n \to \infty$ in probability for all simple, predictable and bounded by 1 sequences K^n. This means that $(K^n H^n \bullet S) \to 0$ in probability, which in turn yields the statement.

Finally this leads us to the following structure: let S be a semi-martingale, then $H \mapsto (H \bullet S)$ is a continuous map, where we consider pathwise uniform convergence in probability on the set of bounded predictable strategies and the Emery topology on the set of semi-martingales. We can re-define the defining metric for the Emery topology by taking the supremum over all bounded predictable strategies. By the previous continuous extension result both metrics coincide, since every value $(K \bullet S)$ can be approximated by values $(K_n \bullet S)$ where K_n is bounded, simple and predictable.

Additionally we have the property that $L^\infty_{pred}(\Omega \times [0,1]) \times \mathbb{S} \to \mathbb{S}$, $(H,S) \mapsto (H \bullet S)$ is continuous by definition of the Emery topology.

It is our final goal, after having achieved a characterization of good integrators and a stochastic integral for bounded predictable strategies to create the somehow largest set of integrands for a given semi-martingale.

Definition 1.27. *Let H be a predictable process: consider $H_n := H 1_{\{\|H\| \leq n\}}$, for $n \geq 1$. If $(H_n \bullet S)$ is a Cauchy sequence in the Emery topology, then we call H integrable with respect to S, in signs $H \in L(S)$ and we write $(H \bullet S) = \lim_{n\to\infty}(H_n \bullet S)$.*

Remark 1.28. By the very definition of the Emery topology the following lemma is clear: let $(H_n \bullet S) \to 0$ in the Emery topology and $|K_n| \leq |H_n|$, for $n \geq 0$, then also $(K_n \bullet S) \to 0$. Notice that we use here that the supremum goes over all predictable strategies, so changing signs works.

Theorem 1.29. *Let S be a semi-martingale. Then $H \in L(S)$ if and only if H is predictable and for all sequences $(K_n)_{n\geq 0}$ of bounded, predictable processes with $|K_n| \leq H$ and $K_n \to 0$ pointwise, it holds that $(K_n \bullet S) \to 0$.*

Proof. Let $H \in L(S)$ be fixed, then we know that $H_n := H 1_{\|H\| \leq n}$, for $n \geq 1$ leads to a converging sequence $(H_n \bullet S) \to (H \bullet S)$ in the Emery topology. Take a sequence $(K_n)_{n\geq 0}$ of bounded, predictable processes with $|K_n| \leq H$ and $K_n \to 0$ pointwise, then we can find a subsequence which converges in the Emery topology.

Consider a partition of unity

$$1 = \sum_{n\geq 1} 1_{\{n-1 \leq |H| < n\}} .$$

For a given sequence $m_k \geq k$, $k \geq 1$ the cut-off sums

$$R_k := \sum_{k \leq n \leq m_k} H 1_{\{n-1 \leq |H| < n\}} \to 0$$

as $k \to \infty$. Furthermore by Cauchy property of $((H_n \bullet S))_{n\geq 0}$, we obtain $(R_k \bullet S) \to 0$ with respect to the Emery topology. Hence

$$\Big(\sum_{k \leq n \leq m_k} K_n 1_{\{n-1 \leq |H| < n\}} \bullet S\Big) \to 0$$

as $k \to \infty$, which translates to $(K_n 1_{\{\|H\| \leq n\}} \bullet S) = (K_n \bullet (1_{\{\|H\| \leq n\}} \bullet S)) \to 0$. Since $(1_{\{\|H\| \leq n\}} \bullet S)) \to S$ in the Emery topology we arrive at the result by joint continuity of the stochastic integral.

Vice versa: assume that we have H predictable satisfying the above properties and take the previous partition of unity. Then $(R_k \bullet S) \to 0$ in the Emery topology as $k \to \infty$ for any sequence $m_k \geq k$, for $k \geq 1$. This, however, means that $((H_n \bullet S))_{n\geq 0}$ forms a Cauchy sequence. □

Vector-valued stochastic integration needs some care since we do not have the usual additivity $(\sum \phi^i \bullet S^i) = \sum (\phi^i \bullet S^i)$ in general. A careful, clear and quick introduction is given in [**12**]: with our constructions all necessary requirements like existence of optional quadratic variation processes and the Davis inequality are at hand to access the paper directly.

2. No Arbitrage Theory for discrete models

The purpose of this section is to illustrate the structure of the general theory by means of discrete models. A *discrete model (for a financial market)* is an adapted $(d+1)$ -dimensional stochastic process S with $S := (S^0, \dots, S^d)$ on a *finite* probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration satisfying the usual conditions. We assume furthermore that trajectories jump at stopping times $0 = \tau_0 \leq \tau_1 \leq \ldots \leq \tau_n \leq \tau_{n+1} = T$ and are constant on the stochastic intervals $]\tau_i, \tau_{i+1}[$.

Assumption 2.1. *The process $S^0_t > 0$ almost surely for every $t \in [0,T]$. We shall refer to this asset as risk-less asset, which means here default-free.*

A *trading strategy* or *portfolio strategy* is a predictable stochastic process ϕ with $\phi_t = (\phi^0_t, \dots, \phi^d_t)$ for $t \in [0,T]$. We think of a portfolio formed by an amount of ϕ^0_t in the numeraire and ϕ^i_t units of risky assets, at time t. The value or wealth at time n of such a portfolio is

$$V_t(\phi) = \phi_t S_t := \sum_{i=0}^{d} \phi^i_t S^i_t$$

for $t \in [0,T]$.

The *discounted value process* is given through

$$\widetilde{V}_t(\phi) = \frac{\phi_t S_t}{S^0_t} = \phi_t \widetilde{S}_t$$

for $t \in [0,1]$, where $\widetilde{S}_t = \frac{S_t}{S^0_t}$ denotes the *discounted price process.*

A trading strategy ϕ is called *self-financing* if

$$\widetilde{V}_t(\phi) = \widetilde{V}_0(\phi) + (\phi \bullet \widetilde{S})_t$$

for $t \in [0,T]$. We interpret this condition that the readjustment of the portfolio at time t to new prices S_n is done without bringing in or consuming any wealth in discounted terms.

Proposition 2.2. *Let $S = (S^0, \dots, S^d)$ be a discrete model of a financial market and ϕ a trading strategy, then the following assertions are equivalent:*

(1) The strategy ϕ is self-financing.

(2) The strategy $(\phi^1, \dots, \phi^d) \in L(\widetilde{S})$ and

$$\phi^0_t = \widetilde{V}_0(\phi) + (\phi \bullet \widetilde{S})_{t-} - \sum_{i=1}^{d} \phi^i_t \widetilde{S}^i_{t-} \,.$$

Proof. The proof is immediate from the definition. Notice that ϕ^0 is predictable, therefore we can leave away the last jump and obtain the last formula. □

Definition 2.3. *Let $S = (S^0, \dots, S^d)$ be a discrete model for a financial market, then the model is called* arbitrage-free *if for any trading strategy ϕ the assertion*

$$V_0(\phi) = 0 \text{ and } V_T(\phi) \geq 0, \text{ then } V_T(\phi) = 0$$

holds true. We call a trading strategy ϕ an arbitrage opportunity (arbitrage strategy) *if $V_0(\phi) = 0$ and $V_T(\phi) \gneqq 0$.*

Definition 2.4. *A* contingent claim (derivative) *is an element of $L^2(\Omega, \mathcal{F}, P)$. We denote by $\widetilde{X}$ the discounted price at time T, i.e. $\widetilde{X} = \frac{1}{S^0_T} X$. We call the subspace of $\mathcal{K} \subset L^2(\Omega, \mathcal{F}, P)$*

$$\begin{aligned}\mathcal{K} &:= \{\widetilde{V_T}(\phi) |\ \phi \text{ self-financing trading strategy, } \widetilde{V_0}(\phi) = 0\} \\ &= \{(\phi \bullet \widetilde{S})_T |\ \phi \text{ predictable}\}\end{aligned}$$

the space of contingent claims replicable at price 0. *We call the convex cone*

$$C = \{Y \in L^2(\Omega, \mathcal{F}, P) | \text{ there is } X \in K \text{ such that } X \geq Y\} = \mathcal{K} - L^2_{\geq 0}(\Omega, \mathcal{F}, P)$$

the cone of claims super-replicable at price 0 *or the outcomes with zero investment and consumption. A contingent claim X is called* replicable at price x *and at time T if there is a self-financing trading strategy ϕ such that*

$$\widetilde{X} = x + (\phi \bullet \widetilde{S})_T \in x + \mathcal{K}.$$

A contingent claim X is called super-replicable at price x *and at time T if there is a self-financing trading strategy ϕ such that*

$$\widetilde{X} \leq x + (\phi \bullet \widetilde{S})_T \in x + \mathcal{K},$$

in other words if $\widetilde{X} \in C$.

Remark 2.5. The set $\mathcal{K}$ is a subspace of $L^2(\Omega, \mathcal{F}, P)$ and the positive cone $L^2_{\geq 0}(\Omega, \mathcal{F}, P)$ is polyhedral, therefore by C is closed.

We see immediately that a discrete model for a financial market is arbitrage-free if

$$\mathcal{K} \cap L^2_{\geq 0}(\Omega, \mathcal{F}, P) = \{0\},$$

which is equivalent to

$$C \cap L^2_{\geq 0}(\Omega, \mathcal{F}, P) = \{0\}.$$

Given a discrete model for a financial market, then we call a measure Q equivalent to P an *equivalent martingale measure* with respect to the numeraire S^0 if the discounted price process $\widetilde{S}^i$ are Q-martingales for $i = 1, \ldots, d$. We denote the set of equivalent martingale measures with respect to the numeraire S^0 by $\mathcal{M}^e(\widetilde{S})$. We denote the absolutely continuous martingale measures with respect to the numeraire S^0 by $\mathcal{M}^a(\widetilde{S})$.

Theorem 2.6 (Fundamental theorem of asset pricing). *Let S be a discrete model for a financial market, then the following two assertions are equivalent:*

1. *The model is arbitrage-free.*
2. *The set of equivalent martingale measures is non-empty, $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$.*

Proof. We shall do the proof in two steps. First we assume that there is an equivalent martingale measure $Q \sim P$ for $\widetilde{S}$. We want to show that there is no arbitrage opportunity. Let ϕ be a self-financing trading strategy and assume that

$$V_0(\phi) = 0, \; V_T(\phi) \geq 0,$$

then the discounted value process of the portfolio

$$\widetilde{V}_t(\phi) = (\phi \cdot \widetilde{S})_t$$

is a martingale with respect to Q and therefore

$$E_Q(\widetilde{V}_T(\phi)) = 0\,.$$

Hence we obtain by equivalence $V_T(\phi) = 0$ almost surely with respect to $\mathbb{P}$, since $V_T(\phi) \geq 0$ Q-almost surely, so there is indeed no arbitrage opportunity.

Next we assume that the market is arbitrage-free. Then

$$\mathcal{K} \cap L^2_{\geq 0}(\Omega, \mathcal{F}, P) = \{0\}$$

and therefore we find a linear functional l that separates $\mathcal{K}$ and the compact, convex set

$$\{Y \in L^2_{\geq 0}(\Omega, \mathcal{F}, P) | \; E_P(Y) = 1\},$$

i.e. $l(X) = 0$ for all $X \in \mathcal{K}$ and $l(Y) > 0$ for all $Y \in L^2_{\geq 0}(\Omega, \mathcal{F}, P)$ with $E_P(Y) = 1$. We define

$$Q(A) = \frac{l(1_A)}{l(1_\Omega)}$$

for measurable sets $A \in \mathcal{F}$ with $1_A \neq 0$, and obtain an equivalent probability measure $Q \sim P$, since $l(1_A) > 0$ for sets with $P(A) > 0$. We have in particular from separation

$$E_Q((\phi \cdot \widetilde{S})_T) = 0$$

for any predictable processes ϕ. Therefore $\widetilde{S}$ is a Q-martingale by Doob's optional sampling theorem. □

Now we can formulate a basic pricing theory for contingent claims.

Definition 2.7. *A* pricing rule *for a contingent claim $X \in L^2(\Omega, \mathcal{F}, P)$ at time T is an adapted, càdlàg stochastic process $\pi(X) = (\pi(X)_t)_{t\in[0,T]}$, which determines the price of the claim at time t at time $t \in [0, T]$, i.e. $\pi(X)_T = X$. A pricing rule is* arbitrage-free *if the discrete time model of a financial market*

$$(S^0, S^1, \dots, S^d, \pi(X))$$

is arbitrage-free. We also have the multi-variate analogue.

Lemma 2.8 (arbitrage-free prices). *Let π be an arbitrage-free pricing rule for a set of contingent claims $\mathfrak{X}$, then the discrete model $(S^0, \dots, S^d)$ is arbitrage-free and there is $Q \in \mathcal{M}^e(\widetilde{S})$ such that*

$$\pi(X)_t = E_Q(\frac{S^0_t}{S^0_T} X|\mathcal{F}_t),$$

for all $X \in \mathfrak{X}$. If the discrete time model S is arbitrage-free, then

$$\pi(X)_t = E_Q(\frac{S^0_t}{S^0_T} X|\mathcal{F}_t)$$

is an arbitrage-free pricing rule for all contingent claims $X \in L^2(\Omega, \mathcal{F}, P)$. Hence the only arbitrage-free prices are conditional expectation of the discounted claims with respect to Q and pricing rules are always linear.

Proof. If the market $(S^0, S^1, \dots, S^d, \pi(X))$ is arbitrage-free, we know that there exists an equivalent martingale measure Q such that the discounted prices are Q-martingales. Hence in particular

$$\frac{\pi(X)_t}{S^0_t}$$

is a Q-martingale, so

$$E(\frac{\pi(X)_T}{S^0_T}|\mathcal{F}_t) = E(\frac{X}{S^0_T}|\mathcal{F}_t) = \frac{\pi(X)_t}{S^0_t}$$

which yields the desired equation.

Given an arbitrage-free discrete model S and define the pricing rules by the above relation for one equivalent martingale measure $Q \in \mathcal{M}^e(\widetilde{S})$, then the whole market is arbitrage-free by the existence of at least one equivalent martingale measure, namely Q. □

Remark 2.9. Taking not an equivalent but an absolutely continuous martingale measure $Q \in \mathcal{M}^a(\widetilde{S})$ means that there is at least one measurable set A such that $Q(A) = 0$ and $P(A) > 0$. Hence the claim 1_A with $P(A) > 0$ would have price 0, which immediately leads to arbitrage by entering this contract $X = 1_A$. Therefore only equivalent martingale measures are possible for arbitrage-free pricing.

The set of equivalent martingale measures $\mathcal{M}^e(\widetilde{S})$ is convex and the set $\mathcal{M}^a(\widetilde{S})$ is compact and convex.

THEOREM 2.10. *Let S be a discrete model for a financial market and assume $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$. Then for all $X \in L^2(\Omega, \mathcal{F}, P)$ the following assertions are equivalent:*

(1) $X \in \mathcal{K}$ ($X \in C$).
(2) For all $Q \in \mathcal{M}^e(\widetilde{S})$ we have $E_Q(X) = 0$ (for all $Q \in \mathcal{M}^e(\widetilde{S})$ we have $E_Q(X) \leq 0$).
(3) For all $Q \in \mathcal{M}^a(\widetilde{S})$ we have $E_Q(X) = 0$ (for all $Q \in \mathcal{M}^a(\widetilde{S})$ we have $E_Q(X) \leq 0$).

PROOF. We shall calculate the polar cone of the cone C,

$$C^0 = \{Z \in L^2(\Omega, \mathcal{F}, P) \text{ such that } E_P(ZX) \leq 0\}$$

by definition. For $Q \in \mathcal{M}^a(\widetilde{S})$ we calculate the Radon-Nikodym-derivative $\frac{dQ}{dP}$ and see that

$$E_P(\frac{dQ}{dP}X) = E_Q(X) = E_Q((\phi \bullet \widetilde{S})_T + Y)$$

for $Y \leq 0$, hence – due to the fact that Q is a martingale measure (so the expectation of the stochastic integral vanishes) – we obtain

$$E_P(\frac{dQ}{dP}X) = E_Q(Y) \leq 0.$$

Consequently $\frac{dQ}{dP} \in C^0$. Given now $Z \in C^0$, then by the same reasoning we obtain

$$E_P(ZX) \leq 0$$

for all $X \in C$. Since the model is arbitrage-free, $Z \geq 0$, assume $Z \neq 0$, so

$$E_P(\frac{Z}{E_P(Z)}(\phi \cdot \widetilde{S})_N) \leq 0$$

for all self-financing trading strategies ϕ. Replacing ϕ by $-\phi$ we arrive at

$$E_P(\frac{Z}{E_P(Z)}(\phi \bullet \widetilde{S})_T) = 0,$$

which means that $\frac{Z}{E_P(Z)} \in \mathcal{M}^a(\widetilde{S})$.

This means that the polar cone of C is exactly given by non-negative multiples of $\frac{dQ}{dP}$ for $Q \in \mathcal{M}^a(\widetilde{S})$, hence all the assertion hold by the bipolar theorem. □

REMARK 2.11. Notice that the fundamental theorem of asset pricing can be viewed as the calculation of the polar cone of C.

The last step of the general theory is the distinction between complete and incomplete markets and a renewed description of pricing procedures.

DEFINITION 2.12. *Let S be a discrete model for a financial market and assume $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$. The financial market is called* complete *if $\mathcal{M}^e(\widetilde{S}) = \{Q\}$, i.e. the equivalent martingale measure is unique. The financial market is called* incomplete *if $\mathcal{M}^e(\widetilde{S})$ contains more than one element. In this case $\mathcal{M}^a(\widetilde{S}) = \langle Q_1, \ldots, Q_m \rangle_{convex}$ for linearly independent measures Q_i, $i = 1, \ldots, m$ and $m \geq 2$.*

THEOREM 2.13 (complete markets). *Let S be discrete model of a financial market with $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$. Then the following assertions are equivalent:*

(1) S is complete financial market.
(2) For every claim X there is a self-financing trading strategy ϕ such that the claim can be replicated, i.e.

$$V_T(\phi) = X.$$

(3) *For every claim X there is a predictable process ϕ and a unique number x such that the discounted claim can be replicated, i.e.*

$$\widetilde{X} = \frac{1}{S_T^0} X = x + (\phi \bullet \widetilde{S})_T.$$

(4) *There is a unique pricing rule for every claim X.*

Proof. We can collect all conclusions from the previous results. 2. and 3. are clearly the same by discounting.

1.⇒2.,3.: If S is complete, then there is a unique equivalent martingale measure Q such that the discounted stock prices are Q-martingales. Take a claim X, then we know by Lemma 2.8 that

$$\pi(X)_t = \frac{S_t^0}{S_T^0} E_Q(X|\mathcal{F}_t)$$

is the only arbitrage-free price for X at time t, since there is only one martingale measure Q. The final value of the martingale $(\frac{\pi(X)_t}{S_t^0})_{0 \leq t \leq T}$ can be decomposed into

$$\frac{\pi(X)_T}{S_T^0} = x + (\phi \bullet \widetilde{S})_T$$

Since $E_Q(\frac{\pi(X)_T}{S_T^0} - x) = 0$ means $\frac{\pi(X)_T}{S_T^0} - x \in \mathcal{K}$ by Theorem 2.10. So we have proved 3. and therefore also 2..

2.⇒4.: Given a claim X. If we are given a portfolio ϕ, which replicates the claim X, then we know that

$$\pi(X)_t = V_t(\phi)$$

for $t \in [0, T]$ defines a pricing rule. Therefore the pricing rule is uniquely given by the values of the portfolio, since the values of the portfolio are unique due to FTAP.

4.⇒1.: If we have a unique pricing rule $\pi(X)$ for any claim X, then we know by Lemma 2.8 that we have only one equivalent martingale measure. □

Example 2.14. We write here instead of time points τ_n simply n for the sake of notational simplicity. The Cox-Ross-Rubinstein model is a complete financial market model: The CRR-model is defined by the following relations

$$S_n^0 = (1+r)^n$$

for $n = 0, \dots, N$ and $r \geq 0$ is the bond-process.

$$S_{n+1} := \begin{cases} S_n(1+a) \\ S_n(1+b) \end{cases}$$

for $-1 < a < b$ and $n = 0, \dots, N$. We can write the probability space as $\{1+a, 1+b\}^N$ and think of $1+a$ as "down movement" and $1+b$ as up-movement. Every path is then a sequence of ups and downs. The σ-algebras $\mathcal{F}_n$ are given by $\sigma(S_0, \dots, S_n)$, which means that atoms of $\mathcal{F}_n$ are of the type

$$\{(x_1, \dots, x_n, y_{n+1}, \dots, y_N) \text{ for all } y_{n+1}, \dots, y_N \in \{1+a, 1+b\}\}$$

with $x_1, \dots, x_n \in \{1+a, 1+b\}$ fixed. Hence the atoms form a subtree, which starts after the moves $x_1, \dots, x_n$.

The returns $(T_i)_{i=1,\dots,N}$ are well-defined by

$$T_i := \frac{S_i}{S_{i-1}}$$

for $i = 1, \dots, N$. This process is adapted and each T_i can take two values

$$T_i = \begin{cases} 1+a \\ 1+b \end{cases}$$

with some specified probabilities depending on $i = 1, \dots, N$. We also note the following formula

$$S_n \prod_{i=n+1}^{m} T_i = S_m$$

for $m \geq n$. Hence it is sufficient for the definition of the probability on $(\Omega, \mathcal{F}, P)$ to know the distribution of $(T_1, \dots, T_N)$, i.e.

$$P(T_1 = x_1, \dots, T_N = x_N)$$

has to be known for each $x_i \in \{1+a, 1+b\}$.

PROPOSITION 2.15. *Let $-1 < a < b$ and $r \geq 0$, then the CRR-model is arbitrage-free if and only if $r \in]a, b[$. If this condition is satisfied, then martingale measure Q for the discounted price process $(\frac{S_n}{(1+r)^n})_{n=0,\dots,N}$ is unique and characterized by the fact that $(T_i)_{i=1,\dots,N}$ are independent and identically distributed and*

$$T_i = \begin{cases} 1+a \text{ with probability } 1-q \\ 1+b \text{ with probability } q \end{cases}$$

for $q = \frac{r-a}{b-a}$.

PROOF. First we assume that there is an equivalent martingale measure Q for the discounted price process $(\frac{S_n}{(1+r)^n})_{n=0,\dots,N}$. Then we can prove immediately that for $i = 0, \dots, N-1$

$$E_Q(T_{i+1}|\mathcal{F}_i) = 1 + r$$

simply by

$$E_Q(\frac{S_{i+1}}{(1+r)^{i+1}}|\mathcal{F}_i) = \frac{S_i}{(1+r)^i}$$
$$E_Q(\frac{S_{i+1}}{S_i}|\mathcal{F}_i) = 1 + r.$$

Taking this property we see by evaluation at $i = 0$ that

$$\begin{aligned} E_Q(T_1) &= 1 + r \\ &= Q(T_1 = 1+a)(1+a) + Q(T_1 = 1+b)(1+b), \\ r &= Q(T_1 = 1+a)a + Q(T_1 = 1+b)b, \end{aligned}$$

since $Q(T_1 = 1+a) + Q(T_1 = 1+b) = 1$ and both are positive quantities. Hence $r \in]a, b[$.

On the other hand the only solution of

$$(1-q)(1+a) + q(1+b) = 1 + r$$

is given through $q = \frac{r-a}{b-a}$. Therefore under the martingale measure Q the condition on conditional expectations of the returns T_i reads as

$$E_Q(1_{\{T_{i+1}=1+a\}}|\mathcal{F}_i) = 1 - q,$$
$$E_Q(1_{\{T_{i+1}=1+b\}}|\mathcal{F}_i) = q$$

and consequently the random variables are independent and identically distributed as described above under Q. Therefore the equivalent martingale measure is unique and given as above.

To prove existence of Q we show that the returns satisfy

$$E_Q(T_{i+1}|\mathcal{F}_i) = 1 + r$$

for $i = 0, \dots, N-1$ if we choose Q as above. If the returns are independent, then

$$E_Q(T_{i+1}|\mathcal{F}_i) = E_Q(T_i)$$

which equals $1 + r$ in the described choice of the measure, hence the result is proved. $\square$

EXAMPLE 2.16. We can calculate the limit of a CRR-model. Fix $\sigma > 0$ the time-normalized volatility, i.e. the standard deviation of the return of the stock. Therefore we assume

$$\ln(1+a) = -\frac{\sigma}{\sqrt{N}}$$
$$\ln(1+b) = \frac{\sigma}{\sqrt{N}},$$

which yields i.i.d random variables

$$T_i = \begin{cases} 1+a \text{ with probability } 1-q \\ 1+b \text{ with probability } q \end{cases}$$

with $q = \frac{b}{b-a} = \frac{\exp(\frac{\sigma}{\sqrt{N}})-1}{\exp(\frac{\sigma}{\sqrt{N}})-\exp(-\frac{\sigma}{\sqrt{N}})}$ denotes the building factor of the martingale measure. The stock price in the martingale measure is given by

$$S_n = S_0 \prod_{i=1}^{n} T_i$$
$$= S_0 \exp(\sum_{i=1}^{n} \ln T_i).$$

The random variables $\ln T_i$ take values $-\frac{\sigma}{\sqrt{N}}, \frac{\sigma}{\sqrt{N}}$ with probabilities q and $1-q$, so

$$E_Q(\ln T_i) = \frac{\sigma}{\sqrt{N}} - \frac{\sigma}{\sqrt{N}} \frac{2\exp(\frac{\sigma}{\sqrt{N}}) - 2}{\exp(\frac{\sigma}{\sqrt{N}}) - \exp(-\frac{\sigma}{\sqrt{N}})}$$
$$= \frac{\sigma}{\sqrt{N}} \frac{2 - \exp(\frac{\sigma}{\sqrt{N}}) - \exp(-\frac{\sigma}{\sqrt{N}})}{\exp(\frac{\sigma}{\sqrt{N}}) - \exp(-\frac{\sigma}{\sqrt{N}})}$$
$$E_Q(\ln(T_i)^2) = \frac{\sigma^2}{N}.$$

Therefore the sums $\sum_{i=1}^{n} \ln T_i$ satisfy the requirements of the central limit theorem, namely

$$\sum_{i=1}^{N} \ln T_i = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \sqrt{N} \ln T_i \to N(-\frac{\sigma^2}{2}, \sigma^2)$$

in law for $N \to \infty$, since $E_Q(N \ln T_i) \to -\frac{\sigma^2}{2}$ as $N \to \infty$ and $\sqrt{N} \ln T_i$ take values $-\sigma, \sigma$.

Consequently for every bounded, measurable function ψ on $\mathbb{R}_{\geq 0}$ we obtain

$$E_Q(\psi(\sum_{i=1}^{n} \ln T_i)) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(-\frac{\sigma^2}{2} + \sigma x) e^{-\frac{x^2}{2}} dx.$$

3. Basics of models for financial markets

In this section some preparatory work for general no arbitrage theory is done: goal is to fix notations for models, provide some important counterexamples, show basic structures in continuous time.

The main **ingredients** building blocks of model for financial markets are:

- $T \in (0, \infty)$: time horizon,
- $t \in [0, T]$: trading dates,
- $(\Omega, \mathcal{F}, \mathbb{P})$: probability space,

- $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$: filtration which satisfies the usual conditions (right continuous and complete) w.r.t. $\mathbb{P}$,
- $\mathcal{F}_t$: information up to and including time t.
- $d+1$ assets, where $d \ge 1$, composed of an asset $S^0 = B$, called numeraire, used as denomination basis, and d price processes $S^i = (S^i_t)_{0\le t\le T}, i = 1, \dots d$. From discrete model considerations we learned that it is reasonable to express all prices/values with respect to this numeraire. Whence the assumption: $B_t \equiv 1$. This means that prices S are already expressed in units of the numeraire.
- We assume that prices processes are adapted and càdlàg processes.

EXAMPLE 3.1 (Black-Scholes model, GBM – geometric Brownian motion). Bank account has instantaneous interest rate r, so $\tilde{B}_t = e^{rt}$ (in undiscounted values). We also have a stock price for $t \in [0,T]$

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

where W is a Brownian motion. Switching to discounted values we get.

$$\begin{aligned} B_t &= \frac{\tilde{B}_t}{\tilde{B}_t} = 1 \\ S_t &= \frac{\tilde{S}_t}{\tilde{B}_t} = S_0 \exp\left\{\sigma W_t + \left(\mu - r - \frac{1}{2}\sigma^2\right) t\right\} \end{aligned}$$

Furthermore, applying Itô's formula gives us that $dS_t = S_t((\mu - r)dt + \sigma dW_t)$.

EXAMPLE 3.2 (General Itô process model). We have

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

where the processes b and σ are $\mathbb{R}^d$ and $\mathbb{R}^{d\times n}$ dimensional respectively, predictable and integrable processes.

EXAMPLE 3.3 (Cox-Ross-Rubinstein binomial model). $\tilde{B}_k = (1+r)^k$ and $\frac{\tilde{S}_k}{\tilde{S}_{k-1}}$ are i.i.d. with two possible values $1+u$, $1+d$ with probability p, $1-p$ respectively (usually $u > r > d > -1$).

REMARK 3.4. We can embed discrete into continuous time by making everything piecewise constant.

DEFINITION 3.5. *We call a predictable process* $\varphi = (\eta, \vartheta^1, \dots, \vartheta^d)$ *with* $\vartheta := (\vartheta^1, \dots, \vartheta^d)$ *a trading strategy with value process*

$$V(\varphi) = (V_t(\varphi))_{0\le t\le T}\,, \tag{3.1}$$

where

$$V_t(\varphi) = \sum_{i=1}^{d} \vartheta^i_t S^i_t + \eta_t \cdot 1 = \vartheta^{\mathrm{tr}}_t S_t + \eta_t$$

is the time t value of the current portfolio. The cost of the trading strategy is defined as

$$C_t(\varphi) := V_t(\varphi) - \int_0^t \sum_{i=1}^{d} \vartheta^i_u dS^i_u, \quad 0 \le t \le T\,,$$

meaning the total cost/expense, on $[0,t]$*, from using strategy* φ*. Notice that we need* $\vartheta \in L(S)$ *here.*

- *In discrete time*, S is piecewise constant, so the integral is a sum, and being $\vartheta \in L(S)$ is always satisfied.
- $V(\varphi)$, $C(\varphi)$ and $\int \varphi dS$ are always $\mathbb{R}$-valued. If ϑ, S are $\mathbb{R}^d$ valued, the integral is a "vector stochastic integral". Note that this is "$\int \sum_i \vartheta^i dS^i$" rather than "$\sum_i \int \vartheta^i dS^i$". This can cause technical problems if one is not careful.

Definition 3.6. *Strategy $\varphi = (\eta, \vartheta)$ is self-financing if $C(\varphi) \equiv C_0(\varphi)$, i.e. $C_t(\varphi) = C_0(\varphi)$ $\mathbb{P}$-a.s. for all t.*

Lemma 3.7. *The following hold:*

(1) $\varphi = (\vartheta, \eta)$ is self-financing iff $V(\varphi) = V_0(\varphi) + \int \vartheta dS$.
(2) There is a bijection between self-financing strategies $\varphi = (\vartheta, \eta)$ and pairs (V_0, ϑ), where $V_0 \in L^0(\mathcal{F}_0)$ and ϑ is predictable and S-integrable. Explicitly: $V_0 = V_0(\varphi)$ and $\eta = V_0 + \int \vartheta dS - \vartheta^{\mathrm{tr}} S$.
(3) If we have $\varphi = (\vartheta, \eta)$ self-financing, then also η is predictable.

Proof. The first assertion is immediate from definition of $C(\varphi)$. The second assertion follows from teh first and $V(\varphi) = \vartheta^{\mathrm{tr}} S + \eta$. For the third assertion we consider a càdlàg process $Y = (Y_t)_{0 \le t \le T}$, write $\Delta Y_t := Y_t - Y_{t-}$ for the jump of Y at time t. From stochastic integration theory, $\Delta(\int \vartheta dS)_t = \vartheta_t^{\mathrm{tr}} \Delta S_t = \vartheta_t^{\mathrm{tr}} S_t - \vartheta_t^{\mathrm{tr}} S_{t-}$. So then the second assertion gives $\eta_t = V_0 + \int_0^t \vartheta_u dS_u - \vartheta_t^{\mathrm{tr}} S_t = V_0 + \int_0^{t-} \vartheta_u dS_u - \vartheta_t^{\mathrm{tr}} S_{t-}$, where the last three terms are all predictable. □

Remark 3.8. $\int \vartheta dS = 0 + \int \vartheta dS = V((0, \vartheta))$ is by Lemma 3.7, the value of the self-financing strategy defined by ϑ and $V_0 = 0$. This also gives cumulative gains/loses from ϑ.

Building up the model as we have, we have some *implicit assumptions* in our setup:

- we can trade continuously in time,
- prices for buying and selling shares are given by S: there are no transaction costs and we have frictionless trading,
- ϑ is $\mathbb{R}^d$-valued, so ϑ_t^i can be positive or negative. η is $\mathbb{R}$-valued, so η_t can be negative. So, short sales and borrowing are allowed; more generally: no constraints on strategies,
- asset prices S are given a priori and exogenously, and do not react to trading activities. Our agents are small investors or price takers. Consequence: the "book value" $V(\varphi)$ agrees with the liquidation value.

Example 3.9. Allowing too many self-financing strategies may be bad. Let $d = 1$, $S = \exp(W_t - t/2)$ be an exponential Brownian motion on $[0, \infty]$ (with the understanding that $S_\infty = 0$), and the time horizon be $T = \infty$. Going short in S, i.e. choosing a trading strategy with $\vartheta = -1$ yields $V_\infty = -(S_\infty - S_0) = 1$ with zero initial investment. The problem is that its wealth $V = \int \vartheta dS$ is not bounded from below and so we may experience huge losses before realizing profit. If we had $S_0 - S_t \ge -a$ for some constant, then S_t would be bounded from above which is apparently not the case.

4. Arbitrage and martingale measures

We start with the following basic idea: In reasonable models "money pumps" should not exist. How can one formalize this? What is the appropriate characterization?

We use the standard model as outlined in the previous section, so we have a probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ over time horizon $[0, T]$, a bank account $B \equiv 1$ and S which is adapted, $\mathbb{R}^d$-valued and càdlàg.

By Lemma 3.7 we have that any $\mathbb{R}^d$-valued, predictable, S-integrable ϑ gives (for $V_0 := 0$) a self-financing strategy with value/wealth $V(\vartheta) = \int \vartheta dS = G(\vartheta)$. We now call ϑ *admissible*, $\vartheta \in \Theta_{adm}$, if the process $G(\vartheta)$ is uniformly bounded from below, i.e. if $G_t(\vartheta) \geq -a$ for all t, P-a.s., for some $a \geq 0$. In other words, we have that all debts are bounded. Note that a does not depend on ω, but may depend on ϑ.

Definition 4.1 (Simple strategy). *$\vartheta \in b\mathcal{E}$: $\vartheta = \sum_{i=1}^n h_i 1_{((\tau_{i-1}, \tau_i]]}$, with $n \in \mathbb{N}$, stopping times $0 \leq \tau_0 < \tau_1 < \cdots < \tau_n < T$ and h^i that is $\mathbb{R}^d$-valued, bounded and $\mathcal{F}_{\tau_{i-1}}$-measurable. We write $\vartheta \in b\mathcal{E}_{det}$ if in addition the τ_i (but not the h_i) are deterministic.*

Remark 4.2. For $\vartheta \in b\mathcal{E}$, $\int \vartheta dS$ is well defined for any $\mathbb{R}^d$-valued stochastic process, with $G(\vartheta) = \int_0^T \vartheta_u dS_u = \sum_{i=1}^n h_i (S_{\tau_i} - S_{\tau_{i-1}})$. In a model with finite discrete time, $b\mathcal{E}$ equals all bounded, predictible $\mathbb{R}^d$-valued ϑ.

Definition 4.3 (Simple arbitrage opportunity). *Let $\vartheta \in b\mathcal{E}$ be admissible, with $G_T(\vartheta) \in L^0_+ \backslash \{0\}$, i.e. $G_T(\vartheta) \geq 0$ P-a.s. and $P[G_T(\vartheta) > 0] > 0$. Then we call ϑ a* simple arbitrage opportunity.

Definition 4.4 (Arbitrage opportunity). *Suppose S is a semimartingale; then an arbitrage opportunity is a strategy ϑ that is predictable, $\mathbb{R}^d$-valued, S-integrable, admissible and with $G_T(\vartheta) \in L^0_{\geq 0} \backslash \{0\}$.*

Definition 4.5 (Absence of arbitrage conditions). *We define the following conditions:*

($\mathbf{NA}_{elem}$): $G_T(b\mathcal{E}) \cap L^0_{\geq 0} = \{0\}$

($\mathbf{NA}^{adm}_{elem}$): $G_T(b\mathcal{E}_{adm}) \cap L^0_{\geq 0} = \{0\}$

(NA): $G_T(\Theta_{adm}) \cap L^0_{\geq 0} = \{0\}$

Lemma 4.6. *If there exists a probability measure $Q \approx P$ such that S is a local Q-martingale, then (NA) and (NA^{adm}_{elem}) hold (and by extension also (NA_{elem}) holds).*

The proof of this lemma requires a result known as the Ansel-Stricker lemma, which we now state.

Lemma 4.7 (Ansel-Stricker lemma). *Suppose S is a semimartingale. If ϑ is predictable and S-integrable, then the stochastic integral $\int \vartheta dS$ is well defined and again a semimartingale. If in addition we require that $\int \vartheta dS$ to be uniformly bounded from below, $\int \vartheta dS$ is again a local martingale (and then, since it is bounded from below, it is a supermartingale by Fatou's Lemma).*

Remark 4.8. If S is a local martingale, then (if S has jumps), $\int \vartheta dS$ can fail to be a local martingale.

Proof of Ansel-Stricker Lemma. We are following a short proof presented by de Donno and Pratelli in [**5**]. We first prove a more general statement: let X be an adapted, càdlàg process and let $(M^n)_{n \geq 0}$ be a sequence of martingales converging uniformly pathwise in probability to X together with a localizing sequence of stopping times $(\eta^k)_{k \geq 0}$ (notice here again so called "stationarity", i.e. $P[\eta_k = \infty] \to 1$ as $k \to \infty$) and integrable random variables $(\theta^k)_{k \geq 0}$. Assume that $X_t^{\eta_k} \geq \theta^k$ for all $k \geq 0$ and that for all stopping times τ the $(\Delta M^n_\tau)^+ \leq (\Delta X_\tau)^+$ and

$(\Delta M^n_\tau)^- \leq (\Delta X_\tau)^-$ holds true, then X is a local martingale. For the proof define stopping times

$$\tau_n := \inf\{t > 0 \mid X_t > n \text{ or } M^n_t > X_t + 1 \text{ or } M^n_t < X_t - 1\} \wedge T$$

for $n \geq 0$. We can assume, by possibly passing to a subsequence, that $\sum \mathbb{P}[\tau_n < 1] < \infty$. We define $\sigma_m := \inf_{n \geq m} \tau_n \wedge \eta^m$ and show now that X^{σ_m} is a martingale. The sequence $(\sigma_m)_{m\geq 0}$ is additionally localizing by the previous construction, since $\sum 1_{\{\tau_n < 1\}}$ is integrable and hence $\mathbb{P}[\inf_{n\geq m} \tau_n = 1] \to 1$ as $m \to \infty$.

At σ_m we can make assertions about the jumps of X by our two further assumptions: let $m \geq 0$ be given, then

$$(\Delta M^n_{t\wedge\sigma_m})^- \leq (\Delta X_{t\wedge\sigma_m})^- \leq m - \theta^m$$

for $n \geq m$ by the second assumption. Since $M^n_t \geq X_t - 1$ for $n \geq m$ (notice that the jumps of M^n are bounded by the jumps of X), we arrive at

$$M^n_{t\wedge\sigma_m} \geq \theta_m - 1 - (m - \theta_m) = 2\theta_m - m - 1\,.$$

This yields by Fatou's Lemma that $X_{t\wedge\sigma_m}$ is integrable since $M^n_{t\wedge\sigma_m} \to X_{t\wedge\sigma_m}$ in probability as $n \to \infty$. For $t = T$ we obtain in particular $X_{t\wedge\sigma_m}$ is integrable, and hence also $\Delta X_{t\wedge\sigma_m}$ by $X_{t\wedge\sigma}$ being bounded from below by an integrable random variable. Again by

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

for $n \geq m$, hence $M^n_{t\wedge\sigma_m} \to X_{t\wedge\sigma_m}$ in $L^1(\mathbb{P})$ for $0 \leq t \leq T$ yielding that X^{σ_m} is a martingale.

We return now to the proof of the Ansel-Stricker Lemma: we assume by Remark 1.21 that S lies in $\mathcal{H}^1$ and let $\vartheta \in L(S)$ be given and define $\vartheta_n := \vartheta 1_{\{\|\vartheta\| \leq n\}}$. Then by definition of the stochastic integral $(\vartheta_n \bullet S) \to (\vartheta \bullet S)$ in the Emery topology, in particular $(\vartheta_n \bullet S) \in \mathcal{H}^1$ for $n \geq 1$. All assumptions of the previous statement are fulfilled due to $(\vartheta \bullet S)$ being bounded from below and jumps of approximations $(\vartheta_n \bullet S)$ being bounded by jumps of $(\vartheta \bullet S)$. $\square$

PROOF OF LEMMA 4.6. $S \in \mathcal{M}_{loc}(Q)$ and $Q \approx P$ give us via Bichteler-Dellacherie that S is a P-semimartingale. We also have that $b\mathcal{E}_{adm} \subseteq \Theta_{adm}$. So it is enough to prove (NA) since this implies (NA$^{adm}_{elem}$). Now, $S \in \mathcal{M}_{loc}(Q)$, take $\vartheta \in \Theta^{adm}$, so ϑ is S-integrable and predictable, so $\int \vartheta dS$ is well defined. Moreover, since ϑ is admissible, $\int \vartheta dS$ is by Ansel-Stricker again in $\mathcal{M}_{loc}(Q)$, hence Q-supermartingale. So $\mathbb{E}_Q[G_T(\vartheta)] \leq \mathbb{E}_Q[G_0(\vartheta)] = 0$.

Whence, if $G_T(\vartheta) \geq 0$ P-a.s., then also (since $Q \approx P$) $G_T(\vartheta) \geq 0$ Q-a.s.; but $\mathbb{E}_Q[G_T(\vartheta)] \leq 0$, so $G_T(\vartheta) = 0$ Q-a.s. and also P-a.s. (since $P \approx Q$).

To prove (NA$_{elem}$) we use that in discrete time, $G(\vartheta) = \int \vartheta dS$ is always a local martingale if S is a local martingale and ϑ is predictable. $\square$

DEFINITION 4.9 (E(L)MM). *An equivalent (local) martingale measure for S is a probability measure $Q \approx P$ such that S is a (local) Q-martingale.*

With this definition, Lemma 4.6 says that if (ELMM) holds for S, then we have (NA).

For the case of finite discrete time, $S = (S_k)_{k=0,1,\ldots T}$, the converse holds, as we will shall see later. In general however, the converse is not true (for a counterexample in infinite discrete time and in continuous time; see [**6**, P5.1.7]) Why does this happen? The key point is that if one can trade infinitely often, one can do "doubling strategies".

To exclude such phenomena, we must forbid not only arbitrage opportunities, but also "limit arbitrage opportunities". For that, we look first at

$$G_T(b\mathcal{E}) - L^\infty_{\geq 0}(P) = \{Y = G_T(\vartheta) - b \mid \vartheta \in b\mathcal{E}, b \in L^\infty_{\geq 0}(P)\},$$

the set of all payoffs starting with wealth 0, doing elementary bounded self-financing trading and discarting a bounded amount b. Intuitively, nothing of that type should be non-negative (except 0), otherwise we again have a “money pump”.

Recall from functional analysis (see any book on functional analysis):

- for $p \in [1, \infty)$, the dual space $(L^p)^*$ of L^p is L^q with $\frac{1}{p} + \frac{1}{q} = 1$. This does not hold for $p = \infty$.
- the pairing between L^p and L^q, for $p \in [1, \infty]$ is given by $(Y, Z) := \mathbb{E}[YZ]$ for $Y \in L^p$, $Z \in L^q$.
- on L^p for $p \in [1, \infty]$ we denote by $\sigma(L^p, L^q)$ the coarsest topology on L^p which makes linear functionals $Y \mapsto (Y, Z)$ continuous for all $Z \in L^q$. Hence $Y_n \to Y$ in $\sigma(L^p, L^q)$ iff $\mathbb{E}[Y_n Z] \to \mathbb{E}[YZ]$, $\forall Z \in L^q$.
- vice versa the dual space of L^p with the $\sigma(L^p, L^q)$-topology is L^q.
- for $p \in [1, \infty[$ the $\sigma(L^p, L^q)$ coincides with the so-called *weak topology*, since L^q is the dual space (with respect to the norm topology) of L^p.
- on dual spaces one often speaks of the *weak-$*$-topology*, i.e. view L^p as the dual of L^q, then the weak-$*$-topology is the coarsest topology on L^p which makes all linear functionals $Y \mapsto (Y, Z)$ continuous for all $Z \in L^q$. Hence, for $1 < p < \infty$, weak and weak-$*$-topology are the same. For $p = 1$, we only have the weak topology $\sigma(L^1, L^\infty)$ (since L^1 is not a dual space), and $Y_n \to Y$ in $\sigma(L^1, L^\infty)$ iff $\mathbb{E}[Y_n Z] \to \mathbb{E}[YZ]$ for all $Z \in L^\infty$. For $p = \infty$, we only have the weak-$*$-topology $\sigma(L^\infty, L^1)$ (since L^1 is not the norm-dual of L^∞); $Z_n \to Z$ in $\sigma(L^\infty, L^1)$ iff $\mathbb{E}[YZ_n] \to \mathbb{E}[YZ]\, \forall Y \in L^1$.
- we shall not use two many words but simply write $\sigma(L^p, L^q)$-topologies.
- The Hahn-Banach theorem for $\sigma(L^p, L^q)$-topologies reads as follows: let $C \subset L^p$ be a $\sigma(L^p, L^q)$-closed, convex cone and $x \notin C$, then there is a $l \in L^q$ such that $l(x) > 0 \geq l(C)$.
- Another important fact: for $p \in [1, \infty)$ a convex subset of L^p is weakly closed (i.e. closed in $\sigma(L^p, L^q)$) if and only if it is (strongly) closed in L^p, i.e. with respect to the norm topology. Hence the case of $\sigma(L^\infty, L^1)$ is of particular interest.

We can quite easily prove the following theorem on the existence of equivalent separating measures.

THEOREM 4.10 (Kreps/Yan). *Fix $p \in [1, \infty]$ and set q conjugate to p. Suppose $C \subseteq L^p$ is a convex cone with $C \supseteq -L^p_{\geq 0}$ and $C \cap L^p_{\geq 0} = \{0\}$. If C is closed in $\sigma(L^p, L^q)$, then there exists $Q \approx P$ with $\frac{dQ}{dP} \in L^q(P)$ and $\mathbb{E}_Q[Y] \leq 0$ for all $Y \in C$.*

SKETCH OF PROOF. Any $x \in L^p_{\geq 0} \backslash \{0\}$ is disjoint from C, so we can by the Hahn-Banach-theorem strictly separate x from C by some $z_x \in L^q$. Then the cone property gives us $\mathbb{E}[z_x Y] \leq 0$, $\forall Y \in C$ and $C \supseteq -L^p_{\geq 0}$ gives $z_x \geq 0$. The strict separation implies $z_x \not\equiv 0$, so that we can normalise to $\mathbb{E}[z_x] = 1$.

We next form the family of sets $\{\Gamma_x := \{z_x > 0\} | x \in L^p_{\geq 0} \backslash \{0\}\}$. Then one can find a countable subfamily $(\Gamma_{x_i})_{i \in \mathbb{N}}$ with $P[\cup_i \Gamma_{x_i}] = 1$. For suitably chosen weights $\gamma_i > 0$, $i \in \mathbb{N}$, one gets that $z := \sum_{i=1}^\infty \gamma_i z_{x_i}$ is $z > 0$ P-a.s., $z \in L^q$ and $\mathbb{E}[zY] \leq 0$, $\forall Y \in C$. Normalise to get $\mathbb{E}[z] = 1$ and then $dQ := zdP$ does the job. $\square$

THEOREM 4.11 (Stricker). *Fix $p \in [1, \infty]$, q conjugate to p and suppose S is an adapted, càdlàg process and that $S_t \in L^p(P)$ for all $t \in [0, T]$. Denote by $\overline{\cdots}$ the closure in L^p for $1 \leq p < \infty$, or the weak-$*$-closure, i.e. the closure in the $\sigma(L^\infty, L^1)$-topology for $p = \infty$. Then are equivalent:*

(1) $\overline{G_T(b\mathcal{E}_{det}) - L^\infty_+(P)} \cap L^p_+(P) = \{0\}$

(2) *The propery (EMM) holds for S, i.e. there exists $Q \approx P$ for S with density $\frac{dQ}{dP} \in L^q(P)$*

Proof of Stricker's theorem. As for direction "2) $\Rightarrow$ 1)": S is a Q-martingale and $\sigma \in b\mathcal{E}_{det}$ is bounded, so $G_\bullet(\vartheta) = \sum_{i=1}^n h^i(S_{t_i \wedge \bullet} - S_{t_{i-1} \wedge \bullet})$ is again a Q-martingale. This gives us that $\mathbb{E}_Q[G_T(\vartheta)] = 0$ and $\mathbb{E}_Q[G_T(\vartheta) - b] \leq 0$ if $b \geq 0$ and bounded. But then, since $\frac{dQ}{dP} \in L^q(P)$, we also get $\mathbb{E}_Q[Y] \leq 0$ for all $Y \in \overline{G_T(b\mathcal{E}_{det}) - L^\infty_{\geq 0}(P)}$. So if also $Y \in L^p_{\geq 0}(P)$, we get $Y = 0$ almost surely.

For "1) $\Rightarrow$ 2)" we have the following consideratin: the set $G_T(b\mathcal{E}_{det})$ is a convex cone in $L^p(P)$, so

$$C := \overline{G_T(b\mathcal{E}_{det}) - L^\infty_{\geq 0}(P)}$$

is again a convex cone, C contains $-L^p_{\geq 0}(P)$ and C is closed in $\sigma(L^p, L^q)$. But also $C \cap L^p_{\geq 0}(P) = \{0\}$. Then the *Kreps-Yan* Theorem gives the existence of the probability measure $Q \approx P$ with $\mathbb{E}_Q[Y] \leq 0$ for all $Y \in C$ and hence $\mathbb{E}[G_T(\vartheta)] \leq 0$ for all $\vartheta \in b\mathcal{E}_{det}$.

We can now take $\vartheta := \pm I_{A_s} I_{(s,t]}$ with $s \leq t$, $A_s \in \mathcal{F}_s$ to get

$$\mathbb{E}_Q[\pm I_{A_s}(S_t - S_s)] \leq 0$$

for all $A_s \in \mathcal{F}_s$. This gives us $\mathbb{E}_Q[S_t - S_s | \mathcal{F}_s] = 0$, which is the martingale property of S under Q. Also, $S_t \in L^1(Q)$ by Hölder, as $S_t \in L^p(P)$, $\frac{dQ}{dP} \in L^q(P)$. ☐

Remark 4.12. Looking back at Stricker's Theorem 4.11 we see that it has the following pros and cons:

- **+:** works for any adapted, càdlàg process S, proof is nice and simple, strategies from $b\mathcal{E}$ are reasonably realistic.
- **−:** need integrability for S ($S_t \in L^p(P)$), strategies in $b\mathcal{E}$ are not admissible in general. The closure with respect to $\sigma(L^\infty, L^1)$ is quite weak and therefore it might be very reasonable to look for alternative hypotheses on the price process.

Example 4.13 (Counterexample in infinite discrete time). We now show that (NA_{elem}) does not imply (EMM) by giving a counterexample. Start with $(Y_n)_{n \in \mathbb{N}}$ under P that are independent, taking values in ± 1, with $P[Y_n = +1] = \frac{1}{2}(1 + \alpha_n)$. Set $S_0 := 1$ and $\Delta S_n := S_n - S_{n-1} = \beta_n Y_n$. Choose $\mathbb{F} = \mathbb{F}^S = \mathbb{F}^Y$.

The only way to get S to be a $(Q, \mathbb{F})$-martingale is to have $Q[Y_n = +1 | \mathcal{F}_{n-1}] = \frac{1}{2}$. So all (Y_n) must be under Q independent and symmetric around 0, i.e. iid under Q with $Q[Y_n = +1] = \frac{1}{2}$. Kakutani's dichotomy theorem (see Williams) then gives us that $Q \approx P$ if and only if

$$\sum_{n=1}^\infty \alpha_n^2 < \infty.$$

Otherwise, we must have $Q \perp P$. So if we take $\sum \alpha_n^2 = +\infty$, then (EMM) does not hold.

What is the role of β_n? It has not been important so far, we just note that $\sum |\beta_n| < \infty$ implies that S is bounded. [Exercise: Show that there exists an arbitrage opportunity in $b\mathcal{E}$ if and only if $\exists$ arbitrage opportunity with ϑ of the form $\vartheta = h 1_{((\sigma,\tau]]}$ for stopping times $\sigma \leq \tau$ and h bounded $\mathcal{F}_\sigma$-measurable (see [**6**, L5.1.5])]. We now choose

$$\beta_n = 3^{-n}$$

so that for each n, we get that $\sum_{k>n}^\infty \beta_k < \beta_n$. A simple consequence of this is that for $m > n$, $\mathrm{sign}(S_m - S_n) = \mathrm{sign}(Y_{n+1})$

We now claim that there does not exist an arbitrage opportunity in $b\mathcal{E}$. Take $\vartheta = h 1_{((\sigma,\tau]]}$ and consider $A_n = \{\sigma = n, \tau > n\} \in \mathcal{F}_n$. Then $G_\infty(\vartheta) = \int_0^\infty \vartheta_u dS_u =$

$h(S_\tau - S_\sigma)$ has $\operatorname{sign}(h(S_\tau - S_\sigma)) = \operatorname{sign}(hY_{n+1})$ on A_n. So if $G_\infty(\vartheta) \geq 0$ P-a.s., we have for all n $\operatorname{sign}(hY_{n+1}) \geq 0$.

But this is not possible: $A_n \in \mathcal{F}_n$, h is $\mathcal{F}_\sigma$-measurable, so hI_{A_n} is $\mathcal{F}_n$-measurable; and Y_{n+1} is independent of $\mathcal{F}_n$ with values ± 1. So we can only have

$$\operatorname{sign}(hI_{A_n} Y_{n+1}) I_{A_n} \geq 0$$

if $hI_{A_n} = 0$. This is for all n, so we get that $h \equiv 0$, $\vartheta \equiv 0$ and as a result $G_\infty(\vartheta) \equiv 0$.

5. No free lunch with vanishing risk (NFLVR)

Suppose S is a semimartingale (with no integration conditions) and recall the space Θ_{adm} of admissible strategies. Condition

(NA): $G_T(\Theta_{adm}) \cap L^0_{\geq 0} = \{0\}$

can be easily shown to be equivalent to $(G_T(\Theta_{adm}) - L^0_{\geq 0}) \cap L^\infty \cap L^0_{\geq 0} = \{0\}$ or equivalently

(NA): $C \cap L^0_{\geq 0} = \{0\}$, with $C := (G_T(\Theta_{adm}) - L^0_{\geq 0}) \cap L^\infty$.

Defined as above, C consists of bounded payoffs one can be dominated by final wealth of an admissible, self-financing strategy with 0 investment capital.

Instead of the hypothesis on $\sigma(L^p, L^q)$-closedness in the Kreps-Yan-Theorem we only speak of intuitive norm closures: notice that for $1 \leq p < \infty$ norm and $\sigma(L^p, L^q)$ closures coincide for convex sets, whereas only in the case $p = \infty$ a (big) gap appears.

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

$$\overline{C}^{L^\infty(\mathbb{P})} \cap L^0_{\geq 0} = \{0\},$$

where $\overline{\cdots}^{L^\infty(\mathbb{P})}$ denotes the norm closure in $L^\infty(\mathbb{P})$.

Proposition 5.2. *For semimartingale S are equivalent:*

(1) (NFLVR)
(2) Any sequence $g_n = G_T(\vartheta^n)$ in $G_T(\Theta_{adm})$ with $G_T^-(\vartheta^n) = g_n^- \to 0$ in L^∞ converges to 0 in L^0.
(3) S satisfies (NA) plus one of the following:
  (a) (NUBPR) (no unbounded profit with bounded risk) The set

$$\mathcal{G}^1 := \{G_T(\vartheta) \mid \vartheta \in \Theta_{adm} \text{ is 1-admissible}\}$$

  is bounded in L^0.
  (b) For every sequence $\epsilon_n \searrow 0$ and every sequence (ϑ^n) of strategies with $G_\bullet(\vartheta^n) \geq -\epsilon_n$, we have $G_T(\vartheta^n) \to 0$ in L^0.

Proof. See [**6**]. We show first (NFLVR)$\Leftrightarrow$(NA)+(NUBPR): if we have (NA), then any $\vartheta \in \Theta_{adm}$ with $G_T(\vartheta) \geq -c$ also has $G_\bullet(\vartheta) \geq -c$. $\square$

Making a short overview of notation, we have for $C := (G_T(\Theta_{adm}) - L^0_{\geq 0}) \cap L^\infty$ and S semimartingale:

(NFLVR): $\overline{C}^{L^\infty(P)} \cap L^0_{\geq 0} = \{0\}$,
(NA): $C \cap L^0_{\geq 0} = \{0\}$,
(NUBPR): The set

$$\mathcal{G}^1 := \{G_T(\vartheta) \mid \vartheta \in \Theta_{adm} \text{ is 1-admissible}\}$$

is bounded in L^0.

The formulation of the next result needs the concept of a *σ-martingale*, which is already familiar to us since it is related to the fact that not every stochastic integral $(\phi \bullet S)$ along a local martingale is a local martingale.

DEFINITION 5.3 (σ-martingale). *An $\mathbb{R}^d$-valued process X is a σ-martingale (under $\mathbb{P}$) if $X = \int \psi dM = \psi \bullet M$ for an $\mathbb{R}^d$-valued local martingale (under $\mathbb{P}$) and an $\mathbb{R}$-valued predictable M-integrable ψ with $\psi > 0$.*

Clearly, X being a martingale implies it is a local martingale, which implies it is a σ-martingale. The converse does not hold in genera, see Michel Emery's famous example: σ-martingales come with the generality of stochastic integration – it can be seen a cumulative effect of re-scaling of infinitesimal increments of martingales.

However, we have the following important remark: suppose X is a σ-martingale and bounded below. Then the Ansel-Stricker theorem gives that X is also a local martingale (and even supermartingale).

EXAMPLE 5.4. See [**6**, Example 7.3.4] for further details: consider a probability space carrying one Bernoulli random variable B and an independent, exponentially distributed random time T with $\mathbb{P}[T \geq x] = \exp(-x)$. Then we can define a stochastic process M via

$$M_t := 1_{\{t \geq T\}} B$$

for $t \geq 0$. We equip the probability space with the natural filtration generated by M. Apparently M is a martingale with respect to its natural filtration, since

$$E[(M_t - M_s) g((M_u)_{u \leq s})] = E[\int_0^\infty B 1_{\{s \leq x < t\}} g((1_{\{u \geq x\}} B)_{u \leq s}) \exp(-x) dx] \tag{5.1}$$

$$= E[B(\exp(-s) - \exp(-t)] = 0 \tag{5.2}$$

for $0 \leq s \leq t$.

Define now $H_t := \frac{1}{t}$, then this deterministic process is M-integrable since the process

$$(H 1_{\{\|H\| \leq n\}} \bullet M) \to X$$

in the semimartingale topology, where

$$X_t = 1_{\{t \geq T\}} \frac{B}{T} \,,$$

for $t \geq 0$. This is true since X is a finite variation process, hence a semimartingale, and the process $(H 1_{\{\|H\| > n\}} \bullet M)$ converges to 0 in the semimartingale topology, since

$$E[|\frac{B}{T}| 1_{\{T \leq 1/n\}}] \to 0$$

as $n \to \infty$. The process X looks a bit like a martingale having again jumps as multiples of B, but there are some integrability issues: first we observe that

$$E[|X_t|] = \int_0^t \frac{1}{x} \exp(-x) dx = \infty \,.$$

This can be easily strengthened since for every stopping time $\tau \neq 0$ with respect to the natural filtration it even holds that $E[|X_\tau|] = \infty$. Hence it also cannot be a local martingale, but X apparently is a σ-martingale.

THEOREM 5.5 (Fundamental theorem of asset pricing). *For semimartingales $S = (S_t)_{0 \leq t \leq T}$ the following statements are equivalent:*

(1) S satisfies (NFLVR),

(2) S admits an equivalent separating measure, i.e. the property (ESM) holds for S: there exists $Q \approx P$ with $\mathbb{E}_Q[G_T(\vartheta)] \leq 0$, for all $\vartheta \in \Theta_{adm}$,

(3) S admits an equivalent σ-martingale measure (EσMM), i.e. the property (EσMM) holds for S: there exists $Q \approx P$ such that S is a Q-σ-martingale.

Remark 5.6. This theorem can be viewed as the "converse" of Lemma 4.6: (NFLVR) implies the existence of an EσMM. Any process S satisfying (EMM) or (ELMM) satisfies (ESM) as shown by using the Ansel-Stricker theorem. Conversely, if S is (locally) bounded, then any equivalent separating measure is an equivalent (local) martingale measure, as seen in the proof of theorem 4.11. But if S is unbounded (i.e. has unbounded jumps, so that it can't be made bounded, even by localizing), an equivalent separating measure need not be an equivalent σ-martingale measure. However, one can show that the set of equivalent σ-martingale measures is dense in the set of all equivalent separating measures, see [**6**] for a proof. We shall see a proof of all this later.

The main mathematical ingredient of Theorem 5.5 is the following important and surprising fact:

Theorem 5.7. *If the semimartingale $S = (S_t)_{0\leq t\leq T}$ satisfies (NFLVR), then the set*

$$C = (G_T(\Theta_{adm}) - L^0_{\geq 0}) \cap L^\infty$$

is weak*-closed in L^∞, i.e. closed in the $\sigma(L^\infty, L^1)$-topology.

Proof of Theorem 5.7. The proof relies on functional analysis and results and techniques from stochastic calculus for general (discontinuous) semimartingales. See [**6**] or [**8**] for the proof. □

Sketch. The direction "3) $\Rightarrow$ 1)" is proven in the same way as Theorem 4.11 by means of the Ansel-Stricker Lemma.

The direction "1) $\Rightarrow$ 2)" can be seen as follows: by Theorem 5.7, (NFLVR) implies that C is closed in $\sigma(L^\infty, L^1)$. As C is also a convex subset of L^∞, and $C \supseteq -L^\infty_{geq0}$, and $C \cap L^\infty_{\geq 0} = \{0\}$, we conclude by Theorem 4.10, that there exists $Q \approx P$ such that $\mathbb{E}_Q[Y] \leq 0$ for all $Y \in C$, i.e. an equivalent separating measure. This easily implies $\mathbb{E}_Q[G_T(\vartheta)] \leq 0$, for all $\vartheta \in \Theta_{adm}$ (use: $G_T(\vartheta) \wedge n \in C$, $n \to \infty$).

Direction "2) $\Rightarrow$ 3)" follows by the previous remark. □

6. No arbitrage in finite discrete time

For the case of finite discrete time, results are easier. Let us denote in this section in a discrete way: $S = (S_k)_{k=0,1,\dots,T}$ be an $\mathbb{R}^d$-valued process adapted to $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,\dots,T}$ and recall that $\mathbb{F}$-predictable processes are simply $\vartheta = (\vartheta_k)_{k=1,\dots,T}$ (or set $\vartheta_0 := 0$) with ϑ_k $\mathcal{F}_{k-1}$-measurable for all k. Then $G_k(\vartheta) = \sum_{j=1}^k \vartheta_j^{\mathrm{tr}}(S_j - S_{j-1}) = \sum_{j=1}^k \vartheta_j^{\mathrm{tr}} \Delta S_j$, $k = 0, 1, \dots, T$. Here:

$$\Theta = \left\{\text{all predictable } \mathbb{R}^d\text{-valued } \vartheta\right\},$$
$$\Theta_{adm} = \{\vartheta \in \Theta | G_\bullet(\vartheta) \geq -a \text{ for some } a \geq 0\}.$$

With the above notation, the classical no arbitrage (NA) condition then becomes

(NA): $G_T(\Theta_{adm}) \cap L^0_{\geq 0} = \{0\}$

Lemma 6.1. *In finite discrete time: (NA) $\Leftrightarrow G_T(\Theta) \cap L^0_{\geq 0} = \{0\}$.*

Proof. The direction "$\Leftarrow$" is clear, since $G_T(\Theta) \supseteq G_T(\Theta_{adm})$. For "$\Rightarrow$" we need to show that any arbitrage from a general $\vartheta \in \Theta$ can also be realized by an admissible $\vartheta' \in \Theta_{adm}$. So we suppose that $\vartheta \in \Theta$ with $G_T(\vartheta) \cap L^0_{\geq 0} \backslash \{0\}$ is not empty. Assume that $G_.(\vartheta) \not\geq 0$, since otherwise we can take $\vartheta = \vartheta'$.

Let $n_0 := \max\{k \in \{0, 1, \dots, T\} | P[G_k(\vartheta) < 0] > 0\}$ be the "last time when ϑ violates 0-admissibility". Then $0 < n_0 < T$ and $A := \{G_{n_0}(\vartheta) < 0\} \in \mathcal{F}_{n_0}$ has $P[A] > 0$. Take $\vartheta' := I_A I_{\{n_0+1,\dots,T\}}\vartheta$, i.e. on A, after n_0, we trade with ϑ. This gives us that $G_k(\vartheta') = I_A I_{\{k>n_0\}} \sum_{j=n_0+1}^{k} \vartheta_j^{\mathrm{tr}} \Delta S_j = I_A I_{\{k>n_0\}}(G_k(\vartheta) - G_{n_0}(\vartheta)) \geq 0$ by definition of n_0, A; so ϑ' is 0-admissible.

Moreover, $G_T(\vartheta') = I_A(G_T(\vartheta) - G_{n_0}(\vartheta))$ is in $L^0_{\geq 0}$ like $G_T(\vartheta)$ and greater than 0 on A with $P[A] > 0$, so $G_T(\vartheta') \in L^0_{\geq 0} \backslash \{0\}$. □

The key mathematical result in this section is

THEOREM 6.2. *In finite discrete time, if S satisfies (NA), the set $C' := G_T(\Theta) - L^0_{\geq 0}$ is closed in L^0.*

In finite discrete time, this translates to the Dalang-Morton-Willinger theorem

THEOREM 6.3 (Dalang/Morton/Willinger). *For an $\mathbb{R}^d$-valued adapted process $S = (S_k)_{k=0,\dots,T}$ in finite discrete time, are equivalent:*

(1) S satisfies (NA), i.e. $G_T(\Theta_{adm}) \cap L^0_+ = \{0\}$,
(2) There exists and equivalent measure $Q \approx P$ such that S is a Q-martingale, i.e. (EMM) holds for S.

PROOF. For the direction "2) ⇒ 1)" see Lemma 4.6. As for direction "1) ⇒ 2)": (NA) is invariant under a change to an equivalent probability measure, so change to $R \approx P$, such that $S_k \in L^1(R)$ for all k. We then drop the R notation and work without loss of generality under the assumption that S is P-integrable. By Lemma 6.1, (NA) is equivalent to $G_T(\Theta) \cap L^0_{\geq 0} = \{0\}$.

Setting $C' := G_T(\Theta) - L^0_{\geq 0}$, (NA) is equivalent to $C' \cap L^0_{\geq 0} = \{0\}$. Set $C := C' \cap L^1$. This set is convex, $\subseteq L^1$, $\supseteq -L^1_{\geq 0}$ and $C \cap L^1_{\geq 0} = \{0\}$. By (NA) and Theorem 6.2, C is closed in L^1 (notice that C' is closed in L^0, which is an even weaker topology), hence also in $\sigma(L^1, L^\infty)$ since it is convex. So the Kreps-Yan theorem gives $Q \approx P$ such that $\mathbb{E}[Y] \leq 0$, for all $Y \in C$. Choose $\vartheta := \pm I_{A_k \times \{k,\dots,l\}}$ with $A_k \in \mathcal{F}_k$ and $k \leq l$ to get $G_T(\vartheta) = \pm I_{A_k}(S_l - S_k)$. As in proof of Theorem 4.11, this shows that S is a Q-martingale. □

REMARK 6.4. In the proof, we could choose for instance

$$dR = \mathrm{const} \cdot \exp\{-\sum_{k=0}^{T} |S_k|\} dP\,.$$

Then $R \approx P$, $\mathbb{E}_R[|S_k|] < \infty$, for all k and $\frac{dR}{dP} \in L^\infty$. Then the Kreps-Yan theorem gives and equivalent martingale measure Q for S with $\frac{dQ}{dR} \in L^\infty$, and so we even have even have an equivalent martingale measure with $\frac{dQ}{dP} \in L^\infty$.

In finite discrete time, we have:

(1) The space $G_T(\Theta) = \{\sum_{j=1}^{T} \vartheta_j^{\mathrm{tr}} \Delta S_j \,|\, \vartheta \text{ predictable } \mathbb{R}^d\text{-valued}\}$ of all final values of stochastic integrals with respect to S is always closed in L^0.
(2) If S satisfies (NA), then $G_T(\Theta) - L^0_{\geq 0}$ is also closed in L^0 (see Theorem 6.2).

Proofs are not difficult, but are notationally involved; use induction over time and dimension of S (when doing induction over dimension, we want to exclude 0 integrals for non-0 strategies).

7. No arbitrage in Itô process model

We start with a general probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ with $\mathbb{R}^n$-valued Brownian motion W. Consider the *undiscounted model* with bank account B and d stocks

$S = (S_i)_{i=1,\dots d}$, given by

$$dB_t = B_t r_t dt, \quad B_0 = 1$$
$$dS_t^i = S_t^i \mu_t^i dt + S_t^i \sum_{j=1}^n \sigma_t^{ij} dW_t^j, \quad S_0^i = s_0^i > 0$$

We assume r, μ, σ all predictable and suitably integrable processes. Pass to discounted prices $\tilde{B} := \frac{B}{B} \equiv 1$ and $\tilde{S} := \frac{S}{B}$. These then satisfy

$$d\tilde{S}_t^i = \tilde{S}_t^i (b_t^i dt + \sum_{j=1}^n \sigma_t^{ij} dW_t^j), \ \tilde{S}_0^i = s_0^i > 0, \text{with } b_t^i = \mu_t^i - r_t$$

Compactly we write $d\tilde{S}_t = \tilde{S}_t(b_t dt + \sigma_t dW_t)$ with $b_t \in \mathbb{R}^d$, $\sigma_t \in \mathbb{R}^{d\times n}$, $\tilde{S}_t \in \mathbb{R}^d$ or $\tilde{S}_t = \operatorname{diag}(\tilde{S}_t)$.

Assume $d \leq n$ (so we have more sources of uncertainty than risky assets available for trading) and $\operatorname{rank}(\sigma_t) = d$ P-a.s. for all t. Introduce now $\overline{\lambda}_t := \sigma_t^{\mathrm{tr}}(\sigma_t \sigma_t^{\mathrm{tr}})^{-1} b_t \in \mathbb{R}^n$ to get

$$d\tilde{S}_t = \tilde{S}_t \sigma_t (\overline{\lambda}_t \, dt + dW_t) \, .$$

We call $\overline{\lambda}$ the multi-dimensional instantaneous market price of risk.

What is the structure of martingale measures? We start with some probability measure $Q \approx P$. The density process is defined as $Z^Q = (Z_t^Q)_{0\leq t \leq T}$ with $Z_t^Q = \frac{dQ}{dP}\big|_{\mathcal{F}_t}$, choosing a càdlàg version. Introduce the stochastic logarithm

$$L^Q := \int \frac{1}{Z_-^Q} dZ^Q \in \mathcal{M}_{0,loc}(P)$$

to get $Z^Q = Z_0^Q \mathcal{E}(L^Q)$, $dZ_t^Q = Z_{t-}^Q dL_t^Q$ (which could be discontinuous since we did not assume $\mathbb{F}$ generated by a Brownian motion).

Remark 7.1. Notice that Z^Q is a strictly positive martingale by equivalence of $P \approx Q$, hence $Z_-^Q > 0$ by the Absorption Theorem 1.15. Therefore the stochastic integral is well-defined along the càglàd process Z_- and leads to a local martingale by local boundedness of the integrand.

$\tilde{S}$ is a continuous semimartingale with canonical decomposition $\tilde{S} = \tilde{S}_0 + M + A$ with $M = \int_0^\cdot \tilde{S}_s \sigma_s \, dW_s$ and $A = \int_0^\cdot \tilde{S}_s \sigma_s \overline{\lambda}_s \, ds$. This gives us $\langle M, M\rangle = \langle M^i, M^k\rangle_{i,k=1,\dots,d}$ as $\langle M, M\rangle = \int_0^\cdot \tilde{S}_s \sigma_s \sigma_s^{\mathrm{tr}} \tilde{S}_s \, ds$ and so we see that $A \ll \langle M, M\rangle$ in the sense that $dA_t = d\langle M, M\rangle_t \lambda_t$ with $\lambda_t \in \mathbb{R}^d$:

$$dA_t = \tilde{S}_t b_t \, dt = \tilde{S}_t \sigma_t \sigma_t^{\mathrm{tr}} \tilde{S}_t \tilde{S}_t^{-1} (\sigma_t \sigma_t^{\mathrm{tr}})^{-1} b_t \, dt = d\langle M, M\rangle \lambda_t$$

with

$$\lambda_t := \tilde{S}_t^{-1} (\sigma_t \sigma_t^{\mathrm{tr}})^{-1} b_t \, .$$

The process

$$K = \int \lambda^{\mathrm{tr}} d\langle M, M\rangle \lambda = \int b^{\mathrm{tr}} (\sigma\sigma^{\mathrm{tr}})^{-1} b dt$$

is often called the *mean-variance tradeoff* process. We also have that $K = \int \overline{\lambda}^{\mathrm{tr}} \overline{\lambda} dt = \int |\overline{\lambda}_t|^2 dt$.

$\tilde{S}$ defined as above is called an Itô process model with coefficients b (or μ and r), σ.

Continuous model: $S = S_0 + M + A$ is a continuous semimartingale with its canonical decomposition into a continuous local martingale M and a predictable process A. We say S satisfies the stucture condition (SC') if $A \ll \langle M\rangle$ in the sense that $dA = d\langle M\rangle \lambda$ for some predictable λ. We say that $\tilde{S}$ satisfies (SC) if it (SC') is true and if λ is in $L^2_{loc}(M)$. The last condition means that $\int \lambda^{\mathrm{tr}} d\langle M, M\rangle \lambda$ is finite-valued (i.e. K is finite valued).

REMARK 7.2. Suppose $\tilde{S}$ is a continuous semimartingale. Then $\tilde{S}$ satisfies the structure condition (SC) if and only if $\tilde{S}$ satisfies (NUPBR).

Suppose we have a continuous model and that (SC) holds. If $Q \approx P$ is an equivalent local martingale measure for $\tilde{S}$, what can be said about L^Q?

Since $M \in \mathcal{M}^2_{0,loc}(P)$ (after all a continuous local martingale), we can use the Kunita-Watanabe decomposition to write $L^Q = \int \gamma^Q dM + N^Q$ with $N^Q \in \mathcal{M}_{0,loc}(P)$ and $N^Q \perp M$ (so again, because M is continuous, $\langle N^Q, M\rangle \equiv 0$).

LEMMA 7.3. *$Q \approx P$ is an equivalent local martingale measure for $\tilde{S}$ iff $\gamma^Q = -\lambda$. In the Itô process case we have $\gamma^Q = -\tilde{S}^{-1}(\sigma\sigma^{\mathrm{tr}})^{-1}b$.*

PROOF. By Bayes' rule, we have that $Q \approx P$ an equivalent martingale measure for $\tilde{S}$ iff $Z^Q\tilde{S}$ is in $\mathcal{M}_{loc}(P)$. Using Itô's formula, we compute

$$
\begin{aligned}
d(Z^Q\tilde{S}) &= Z^Q_- d\tilde{S} + \tilde{S} dZ^Q + d\langle Z^Q, \tilde{S}\rangle \\
&= Z^Q_- dM + \tilde{S} dZ^Q + Z^Q_- dA + Z^Q_- d\langle L^Q, \tilde{S}\rangle
\end{aligned}
$$

The first two terms of the right hand side are local martingales, so for $Z^Q\tilde{S}$ to be a martingale in $\mathcal{M}_{loc}(P)$, $A + \langle L^Q, M\rangle$ must be in $\mathcal{M}_{loc}(P)$. Since A and $\langle \bullet \rangle$ are predictable and of finite variation, this is equivalent to saying $A + \langle L^Q, M\rangle \equiv 0$, or $0 \equiv \int d\langle M, M\rangle \lambda + \int d\langle M, M\rangle \gamma^Q = \int d\langle M, M\rangle(\lambda + \gamma^Q)$. ☐

COROLLARY 7.4. *Equivalent local martingale measures Q for $\tilde{S}$ are parametrized via*

$$
\frac{Z^Q}{Z^Q_0} = \mathcal{E}\left(-\int \lambda dM + N^Q\right)
$$

with $N^Q \in \mathcal{M}_{loc,0}(P)$, $N^Q \perp M$ under P as long as the right hand side is a strictly positive martingale.

More precisely, if Q is an equivalent local martingale measure, then Z^Q has the above form with some such N^Q. We also have the converse, so if N^Q is as above, then the corresponding $Z^Q := Z^Q_0\mathcal{E}(-\int \lambda dM + N^Q)$ gives an equivalent local martingale measure, if $Z^Q > 0$ and if we also have that Z^Q is a true P-martingale on $[0, T]$.

REMARK 7.5. The simplest choice of N^Q is $N^Q \equiv 0$. The corresponding process is then (taking $Z^Q_0 := 1$) $\hat{Z} := \mathcal{E}(-\int \lambda dM) = \exp\{-\int \lambda dM - \frac{1}{2}K\}$. If this is a true P-martingale, then the corresponding equivalent local martingale measure $\hat{P}$ is called the *minimal martingale measure.*

REMARK 7.6. Since $N^Q \perp M$, Yor's formula gives $\frac{Z^Q}{Z^Q_0} = \mathcal{E}(-\int \lambda dM + N^Q) = \hat{Z}\mathcal{E}(N^Q)$.

What can we say if $\tilde{S}$ is in addition also an Itô process model?

LEMMA 7.7. *Suppose $\tilde{S}$ is an Itô process model with b, σ. Suppose $\mathbb{F} = \mathbb{F}^W$ and $N \in \mathcal{M}_{0,loc}(P)$. Then $N \perp M$ under P iff $N = \int \gamma dW$ with γ predictable, $\mathbb{R}^n$-valued and $\sigma\gamma \equiv 0$.*

PROOF. $N = \int \gamma dW$ by Itô's representation theorem. $N \perp M$ under P if and only if $\langle N, M\rangle \equiv 0$, i.e. if and only if $0 \equiv \langle \int \gamma dW, \int S\sigma dW\rangle = \int S\sigma\gamma dt$. ☐

Corollary 7.8. *Suppose $\tilde{S}$ is an Itô process model with b, σ. If $\mathbb{F} = \mathbb{F}^W$, then equivalent local martingale measures Q are parametrized via processes γ from the kernel of σ by*

$$Z^Q = \mathcal{E}\left(-\int (\sigma\sigma^{\mathrm{tr}})^{-1} b\sigma dW + \int \gamma dW\right)$$

with $\sigma\gamma \equiv 0$ as long as the right hand side is a strictly positive martingale.

Note: If $d = n$, then there is at most one equivalent local martingale measure for $\tilde{S}$, since $\sigma\gamma \equiv 0$ implies $\gamma \equiv 0$, since σ is now invertible.

A special case of the above is the Black-Scholes model: $d = n = 1$, $\mu, r, \sigma > 0$ are all constants, so we have a unique candidate for the density process of the equivalent local martingale measure: $\hat{Z} = \mathcal{E}(-\int \frac{\mu - r}{\sigma} dW) = \mathcal{E}(-\frac{\mu - r}{\sigma} W)$. Since all coefficients are constant, $\hat{Z}$ is a true P-martingale, so $\hat{P}$ is an equivalent local martingale measure, and $d\tilde{S}_t = \tilde{S}_t \sigma d\hat{W}_t$ is even a true $\hat{P}$-martingale; so $\hat{P}$ is even an equivalent martingale measure.

8. No arbitrage in (exponential) Lévy models

A Lévy process L is a stochastically continuous $\mathbb{R}^d$-valued with stationary, independent increments. Following [**11**] we can choose a càdlàg version of a Lévy process. Additionally we know that the logarithm of the characteristic function of L is of Lévy-Khintchine form.

We analyze how a Lévy process L looks like with respect to an equivalent σ-martingale measure:

Theorem 8.1. *Let L be a one dimensional Lévy process and assume that $S = \exp(L)$ is a σ-martingale, then S is already a martingale*

Proof. By the Ansel-Stricker Lemma a bounded from below σ-martingale is in fact a local martingale, and hence a super-martingale. We therefore have that

$$E[\exp(L_t)] \leq 1\,,$$

for $t \geq 0$, by the super-martingale property. Since L is a Lévy process we know that the Lévy exponent κ is at least well defined on the strip in $\mathbb{C}$ of complex numbers u with real part $0 \leq \mathrm{Re}(u) < 1$ and has Lévy-Khintchine form there. We are interested in showing that $\kappa(u) \to 0$ as $u \nearrow 1$, which then yields the martingale property. Due to κ's Lévy-Khintchine form there are numbers $b \in \mathbb{R}$, $c \geq 0$ and a Radon measure ν on $\mathbb{R} \setminus \{0\}$ such that

$$\kappa(u) = bu + \frac{c^2}{2}u^2 + \int_{\|\xi\| \geq 1} (\exp(u\xi) - 1)\nu(d\,\xi) + \int_{\|\xi\| \leq 1} (\exp(u\xi) - 1 - u\xi)\nu(d\xi)$$

for $0 \leq u < 1$. The first, second and fourth summand are continuous in u as $u \nearrow 1$ by continuity of polynomials and dominated convergence. The third summand can be split in two parts (on the positive and negative real line, respectively), where we can conclude by dominated convergence on the negative real line and by monontone convergence on the positive real line by the fact that $\kappa(u) \leq 0$ as $u \in [0, 1[$ by convexity of the moment generating function. $\square$

A slightly more complicated situation is given when we look at Lévy processes themselves. We can conclude the same result, however, we cannot use the Ansel-Stricker Lemma.

Theorem 8.2. *Let L be a Lévy process. Assume that L is a σ-martingale, then it is a martingale.*

PROOF. A σ-martingale L is a semi-martingale such that there is an increasing sequence of predictable sets $D_n \nearrow \Omega \times [0,1]$ with $(1_{D_n} \bullet L)$ is a local martingale (take the definition of σ-martingales as limit of stochastic integrals of the form $(H1_{\{\|H\|\leq n\}} \bullet M)$ for some predictable strategy $H \in L(M)$). For every n we can hence choose a localizing sequence of stopping times τ_{nm} such that $(1_{D_n} \bullet L^{\tau_{nm}})$ actually are martingales. The compensator (i.e. the predictable process $\tilde{A}$ uniquely associated by the Doob-Meyer decomposition to an increasing, locally integrable finite variation process A making the difference $A - \tilde{A}$ a local martingale) $\tilde{A} = t\nu$ of $A = \sum_{s\leq t} 1_{\{\|\Delta L_s\|\geq 1\}}$ is always well-defined and deterministic due to independent increments and linear in time due to stationarity of increments. We do additionally have that

$$\int_0^{\tau_{nm}} 1_{D_n}(s)\, ds\; 1_{\{\|\xi\|\geq 1\}}\nu$$

is the compensator of $\sum_{s\leq t} 1_{\{\|(\Delta(1_{D_n}\bullet L^{\tau_{nm}})_s\|\geq 1\}}$. If we integrate now $s \mapsto \Delta(1_{D_n} \bullet L)_s$ with respect to this counting measure of the jumps we obtain

$$\sum_{s\leq t} 1_{\{\|(\Delta(1_{D_n}\bullet L^{\tau_{nm}})_s\|\geq 1\}} \Delta(1_{D_n} \bullet L^{\tau_{nm}})_s$$

which in turn is integrable by martingality. Hence we obtain that $\int_{\|\xi\|\geq 1} \xi \nu(d\xi)$ is finite, which proves the martingale property of L. □

9. Pricing and hedging by replication

Assume that we have a standard model of a financial market $(\Omega, \mathcal{F}, \mathbb{F}, P)$ over $[0,T]$ with $B \equiv 1$ and S a $\mathbb{R}^d$-valued semimartingale.

The basic question is: given $H \in L^0(\mathcal{F}_T)$, viewed as a random payoff of a contract at time T, what is its value at $t \leq T$?

We are first going to explain the basic ideas, ignoring all of the (important!) technical details.

DEFINITION 9.1 (Replicating strategy). *A replicating strategy for H is a self-financing φ with $V_T(\varphi) = H$ P-a.s.; we then call H replicable or obtainable by φ.*

THEOREM 9.2 (Valuation of attainable payoffs I). *If $H \in L^0(\mathcal{F}_T)$ is replicable by φ, then its value at any time $t \leq T$ is $V_t(\varphi)$, if there is no arbitrage.*

PROOF. Take $\varphi = (\vartheta, \eta)$ and fix t. Consider on $[t,T]$ self-financing strategy with initial capital $V_t(\varphi)$ and ϑ; see Lemma 3.7. Then we have on (t,T) zero cash-flows (by self-financing) and in T exactly $V_T(\varphi) = H$ P-a.s.. This is exactly the same as one has when simply holding the payoff on $(t,T]$. So values at time t must coincide, too. Otherwise we would have arbitrage. □

REMARK 9.3. How can we compute $V_t(\varphi)$ more easily? Note: H is attainable $\Leftrightarrow$ $\exists$ self financing φ with $V_T(\varphi) = H$ P-a.s. $\Leftrightarrow H = V_0 + \int_0^T \vartheta_u dS_u$ P-a.s., i.e. H is up to V_0 representable as a stochastic integral of S. Moreover, φ self-financing implies by Lemma 3.7 that $V_t(\varphi) = V_0 + \int_0^t \vartheta_u dS_u$, so if Q is an EσMM for S, then $\int \vartheta dS$ is (for sufficiently integrable ϑ) a Q-martingale, and so $V_t(\varphi) = \mathbb{E}_Q[H|\mathcal{F}_t]$, $0 \leq t \leq T$.

THEOREM 9.4 (Valuation of attainable payoffs II). *If $H \in L^0(\mathcal{F}_T)$ is attainable by "reasonable" strategy φ, the value of H at any time $t \leq T$, if there is no arbitrage, is given by $V_t(\varphi) = V_t^H := \mathbb{E}[H|\mathcal{F}_t]$ for any EσMM Q for S.*

EXAMPLE 9.5. Model:

- Bank account $\tilde{B}_t = e^{rt}$
- Stock $\tilde{S}_t = s_0 \exp\{\sigma W_t + (\mu - \frac{1}{2}\sigma^2)t),\ 0 \leq t \leq T$

The discounted stock price $S = \frac{\tilde{S}}{\tilde{B}}$ satisfies the SDE: $dS_t = S_t((\mu - r)dt + \sigma dW_t)$.

A *European call option* with maturity T and strike K has undiscounted payoff $\tilde{H} = (\tilde{S}_T - K)^+$. The discounted payoff is then $H = \frac{\tilde{H}}{\tilde{B}_T} = \left(S_T - Ke^{-rT}\right)^+$. What is its value at $t \leq T$?

Suppose $\mathbb{F} = \mathbb{F}^W$ is generated by Brownian motion (augmented as usual). From Corollary 7.8, there is only one candidate for the density process of an ELMM, namely

$$\hat{Z}_t = \mathcal{E}\left(-\frac{\mu - r}{\sigma}W\right)_T = \exp\left\{-\frac{\mu - r}{\sigma}W_t - \frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 t\right\},\ \hat{Z} > 0$$

is a true P-martingale on $[0, T]$; so $d\hat{P} := \hat{Z}_T dP$ gives $\hat{P} \approx P$ on $\mathcal{F}_T$. By Girsanov, $\hat{W}_t := W_t - \frac{\mu - r}{\sigma}t$, $0 \leq t \leq T$, is then a $\hat{P}$-Brownian motion, and $dS_t = S_t\sigma d\hat{W}_t$ shows that $S_t = s_0\mathcal{E}\left(\sigma\hat{W}\right)_t = s_0\exp\left\{\sigma\hat{W}_t - \frac{1}{2}\sigma^2 t\right\}$ is a true $\hat{P}$-martingale. In other words, $\hat{P}$ is an EMM for S. Also, $\hat{P}$ is the unique equivalent martingale measure.

We suspect that the model is arbitrage free and complete; so guess that H is "attainable" and we also guess its discounted value at time t is

$$\hat{V}_t := \hat{\mathbb{E}}[H|\mathcal{F}_t] = \hat{\mathbb{E}}\left[\left(S_t\exp\left\{\sigma(\hat{W}_T - \hat{W}_t) - \frac{1}{2}\sigma^2(T - t)\right\} - Ke^{-rT}\right)^+ \Big|\mathcal{F}_t\right]$$

S_t is $\mathcal{F}_t$-measurable and $\hat{W}_T - \hat{W}_t$ is independent of $\mathcal{F}_t$ and $\sim \mathcal{N}(0, T - t)$. Set $S \sim \mathcal{N}(0, 1)$ under P so that we get

$$\hat{V}_t = \hat{\mathcal{E}}\left[\left(ae^{bZ - c} - d\right)^+\right] = \hat{v}(t, S_t)$$

with $a = S_t$, $b = \sigma\sqrt{T - t}$, $c = \frac{1}{2}\sigma^2(T - t)$ and $d = Ke^{-rT}$. Here the function $\hat{v}$ can be computed explicitly.

The natural guess for the undiscounted value is then $\tilde{V}_t = \hat{V}_t \cdot \tilde{B}_t = v(t, \tilde{S}_t)$. Doing the computations gives

$$v(t, \tilde{S}_t) = \tilde{S}_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2),\ \text{with}$$
$$d_{1,2} = \frac{\log\frac{\tilde{S}_t}{Ke^{-r(T-t)}} \pm \frac{1}{2}\sigma^2(T - t)}{\sigma\sqrt{T - t}}$$

with Φ being the standard normal cumulative distribution function

$$\Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-\frac{1}{2}x^2}dx$$

The above solution is known as the *Black-Scholes formula*; the derivation was awarded the Nobel prize in economics in 1997.

To justify $\tilde{V}_t$ as a reasonable value for the option at time t, we still need to check whether H is *attainable* (in a good sense). One way exploits Itô's Representation theorem, as follows.

Since $\mathbb{F} = \mathbb{F}^W = \mathbb{F}^{\hat{W}}$, we have that any $H \in L^1(\mathcal{F}_T, \hat{P})$ has a unique representation

$$H = \hat{\mathbb{E}}[H] + \int_0^T \psi_u d\hat{W}_u = \hat{\mathbb{E}}[H] + \int_0^T \vartheta_u dS_u,$$

where $\int\psi d\hat{W} = \int\vartheta dS$ is a $\hat{P}$-martingale; this uses $dS_t = S_t = S_t\sigma d\hat{W}_t$ via $\vartheta_u = \frac{\psi_u}{\sigma S_u}$.

Moreover, if $H \geq 0$ (as for the call option), then $\int\vartheta dS \geq -\hat{\mathbb{E}}[H]$ shows that $\varphi \hat{=} (\hat{\mathbb{E}}[H], \vartheta)$ is 0-admissible. So: every $H \in L^1_+(\mathcal{F}_T, P)$ can be written as final value of some self-financing admissible strategy φ such that $V(\varphi) = \hat{\mathbb{E}}[H] + \int \vartheta dS$ is a $\hat{P}$-martingale. So this justifies calling such H attainable, and so we might say that $(S, \mathbb{F} = \mathbb{F}^W)$ is complete.

An *alternative argument* shows that the call option H is "attainable"; it even works without specifying $\mathbb{F}$ (of course, we must take $\mathbb{F} \supseteq \mathbb{F}^S$ to have S adapted).

We start with the function $v(t,x)$ from the Black-Scholes formula and check by computation that

$$\frac{\partial v}{\partial t} + rx\frac{\partial v}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 v}{\partial x^2} - rv = 0,\ v(T,x) = (x-K)^+.$$

Now $\tilde{S}_t = S_t\tilde{B}_t = e^{rt}S_t$ satisfies $d\tilde{S}_t = \tilde{S}(rdt + \sigma d\hat{W}_t)$, so applying Itô's formula gives:

$$d\tilde{V}_t = dv(t,\tilde{S}_t) = (\dots)dt + (\dots)d\hat{W}_t.$$

Working out the calculations and using that $v_t + v_x rx + \frac{1}{2}v_{xx}\sigma^2 x^2 = rv$ from the PDE in the drift term, we get

$$d\tilde{V}_t = r\tilde{V}_t dt + \frac{\partial v}{\partial t}(t,\tilde{S}_t)\sigma\tilde{S}_t d\tilde{W}_t = \frac{\partial v}{\partial t}(t,\tilde{S}_t)d\tilde{S}_t + (\tilde{V}_t - \frac{\partial v}{\partial t}(t,\tilde{S}_t)\tilde{S}_t)rdt.$$

Because $d\tilde{B}_t = r\tilde{B}_t dt$ we can rearrange the second term and get that $\tilde{\vartheta}_t := \frac{\partial v}{\partial x}(t,\tilde{S}_t)$, $\tilde{\eta}_t := \frac{1}{\tilde{B}_t}(v(t,\tilde{S}_t) - \tilde{\vartheta}\tilde{S}_t)$ to get $d\tilde{V}_t = \tilde{\vartheta}_t d\tilde{S}_t + \tilde{\eta} d\tilde{B}_t$ and also $\tilde{V}_t = v(t, \tilde{S} - t) = \tilde{\vartheta}_t\tilde{S}_t + \tilde{\eta}_t\tilde{B}_t$. This means that $\tilde{\varphi} = (\tilde{\vartheta}, \tilde{\eta})$ is a strategy with undiscounted value process $\tilde{V}(\tilde{\varphi}) = \tilde{V} = v(\bullet, \tilde{S}_\bullet)$, and which is self-financing due to above.

Moreover, $\tilde{V}_T(\tilde{\varphi}) = v(T,\tilde{S}_T) = (\tilde{S}_T - K)^+ = \tilde{H}$ shows that $\tilde{\varphi}$ replicates $\tilde{H}$. Finally, $\tilde{\varphi}$ is even admissible since $v \geq 0$. So in that sense we see again that $\tilde{H}$ is attainable and so its value at t is $v(t,\tilde{S}_t)$.

10. Superreplication and optional decomposition

The basic question we ask ourselves in this section is: How to hedge a non-attainable payoff in an incomplete market?

We use the standard model of $(\Omega, \mathcal{F}, \mathbb{F}, P)$ and S on $[0,T]$. Denote by $\mathbb{P}$ the set of all equivalent σ martingale measures for S and assume $\mathbb{P} \neq \emptyset$; by the fundamental theorem of asset pricing, this guarantees (NFLVR).

Fix a payoff $H \in L^0_{\geq 0}(\mathcal{F}_T)$. Everything would work for $H \geq -$const. as well. We assume H is not attainable, so there is no self-financing strategy φ with $V_T(\varphi) = H$ P-a.s. How do we hedge such an H? Idea: look at strategies that produce at least H and try to find the cheapest one.

Definition 10.1 (Superreplication price). *The super-replication price of $H \in L^0_{\geq 0}(\mathcal{F}_T)$ is*

$$\begin{aligned}\Pi_s(H) &= \inf\{V_0 \in \mathbb{R} \,|\, \exists\, \vartheta \in \Theta_{adm} :\ V_0 + \int_0^T \vartheta_u dS_u \geq H\ P\text{-}a.s.\} \\ &= \inf\{V_0 \in \mathbb{R} \,|\, H - V_0 \in G_T(\Theta_{adm}) - L^0_{\geq 0}\}\end{aligned}$$

The intuition behind this definition is that we can sell H for $\Pi_s(H)$ without risk, because $(\Pi_s(H), \vartheta)$ is a self-financing admissible strategy which produces at least H by time T. We have to be careful, however, since $\Pi_s(H)$ is an infimum; we do not know if it is attained. So we do not know if there exists a $\vartheta \in \Theta_{adm}$ for $V_0 := \Pi_s(H)$.

Lemma 10.2. *Assume that $\mathbb{P} \neq 0$. Then for any payoff $H \in L^0_{\geq 0}(\mathcal{F}_T)$*

$$\Pi_s(H) \geq \sup_{Q \in \mathbb{P}} \mathbb{E}_Q[H].$$

PROOF. Without loss of generality suppose that

$$B := \left\{ V_0 \in \mathbb{R} \,|\, \exists\, \vartheta \in \Theta_{adm} : V_0 + \int_0^T \vartheta_u dS_u \geq H \ P\text{-a.s.} \right\} \neq \emptyset,$$

else $\Pi_s(H) = \infty$. So let $V_0 \in B$ and take some $\vartheta \in \Theta_{adm}$ such that $V_0 + \int_0^T \vartheta_u dS_u \geq H$ P-a.s. Let $Q \in \mathbb{P}$, then $S \in \mathcal{M}_\sigma(Q)$ is a σ-martingale under Q and $G(\vartheta) = \int \vartheta dS$ is bounded below; The Ansel-Stricker Lemma gives us that $G(\vartheta) \in \mathcal{M}_{loc}(Q)$ is a local martingale under Q and in particular a super-martingale. So we get

$$\mathbb{E}_Q[H] \leq V_0 + \mathbb{E}_Q[G_T(\vartheta)] \leq V_0$$

Hence, taking the supremum over Q, infimum over V_0 we get

$$\sup_{Q\in\mathbb{P}} \mathbb{E}_Q[H] \leq \inf B = \Pi_s(H)\,.$$

$\square$

Our goal now is to prove the equality in Lemma 10.2 and also that the infimum for $\Pi_s(H)$ is attained. We fix $H \in L^0_{\geq 0}(\mathcal{F}_T)$ and define the adapted process

$$U_t := \operatorname*{ess\,sup}_{Q\in\mathbb{P}} \mathbb{E}_Q[H|\mathcal{F}_t],\ 0 \leq t \leq T$$

which is the smallest random variable that dominates the set of random variables for any $t \in [0, T]$, i.e. the measurable version of the "supremum". If $\mathcal{F}_0$ is trivial, then $U_0 = \sup_{Q\in\mathbb{P}} \mathbb{E}_Q[H]$.

PROPOSITION 10.3. *Assume $\mathbb{P} \neq \emptyset$ and $H \in L^0_{\geq 0}(\mathcal{F}_T)$. If $\sup_{Q\in\mathbb{P}} \mathbb{E}_Q[H] < \infty$ then U is a Q-supermartingale for every $Q \in \mathbb{P}$, which allows for a càdlàg version.*

PROOF. We argue that U has the supermartingale property: let $s \leq t$, we want to show that $\mathbb{E}_Q[U_t|\mathcal{F}_s] \leq U_s$ for any $Q \in \mathbb{P}$. We fix $Q \in \mathbb{P}$ and introduce for $t \in [0, T]$

$$\begin{aligned} \zeta_t &:= \{Z \,|\, Z \text{ is the density process w.r.t } Q \text{ of some } R \in \mathbb{P}, \text{ and } Z_s = 1 \text{ for } s \leq t\} \\ &= \{Z \,|\, Z \text{ is the density process w.r.t } Q \text{ of some } R \in \mathbb{P}, \text{ with } R = Q \text{ on } \mathcal{F}_t\} \end{aligned}$$

Taking $R = Q$ shows that $1 \in \zeta_t$, so it is not empty; and $\zeta_t \subseteq \zeta_s$ for $0 \leq s \leq t \leq T$. Moreover we claim $\zeta_t = \left\{ \frac{Z^R_{t\vee\bullet}}{Z^R_t} \,|\, Z^R \text{ is density process w.r.t. } Q \text{ of some } R \in \mathbb{P} \right\}$

"$\subseteq$": Take $Z \in \zeta_t$ with corresponding $R \in \mathbb{P}$. Then $Z_t = 1$ and so:

$$Z_\bullet = I_{\{\bullet\leq t\}} + Z_\bullet I_{\{\bullet>t\}} = \frac{Z_{t\vee\bullet}}{Z_t}.$$

"$\supseteq$": Take $R \in \mathbb{P}$ with Q density process Z^R. Let $Z_\bullet = Z^R_{t\vee\bullet}/Z^R_t$. Then $Z > 0$, $Z_s = 1$ for $s \leq t$ and Z is like Z^R a Q-martingale. Moreover, both S and SZ^R are both local Q-martingales (the first one since $Q \in \mathbb{P}$, the second by the Bayes rule because $R \in \mathbb{P}$). So

$$S_\bullet Z_\bullet = S_\bullet I_{\{\bullet\leq t\}} + \frac{S_\bullet Z^R_\bullet}{Z^R_t} I_{\{\bullet>t\}}$$

is also a local Q-martingale. So $dR' := Z_T dQ$ gives $R' \in \mathbb{P}$ with Q density Z.

Now we use Bayes rule again to write

$$U_t = \operatorname*{ess\,sup}_{R\in\mathbb{P}} \mathbb{E}_R\left[H|\mathcal{F}_t\right] = \operatorname*{ess\,sup}_{R\in\mathbb{P}} \mathbb{E}_Q\left[\frac{HZ_T^R}{Z_t^R}|\mathcal{F}_t\right] =$$
$$= \operatorname*{ess\,sup}_{Z\in\zeta_t} \underbrace{\mathbb{E}_Q[HZ_T|\mathcal{F}_t]}_{=:\Gamma_t(Z)}.$$

We claim that the family $\{\Gamma_t(Z)|Z \in \zeta_t\}$ is an *upwards directed set*: if Z and Z' are in ζ_t and $A \in \mathcal{F}_t$, then apparently $ZI_A + Z'I_{A^C}$ is again in ζ_t. So with $A := \{\Gamma_t(Z) \geq \Gamma_t(Z')\} \in \mathcal{F}_t$, we get

$$\max\{\Gamma_t(Z), \Gamma_t(Z')\} = \Gamma_t(Z)I_A + \Gamma_t(Z')I_{A^C} = \mathbb{E}_Q[H(Z_tI_A + Z_t'I_{A^C})|\mathcal{F}_t] = \Gamma_t(\overline{Z})$$

with $\overline{Z} := ZI_A + Z'I_{A^C} \in \zeta_t$. This is useful because the essential supremum of an upward directed family of random variables can be obtained as a monotone increasing limit of a sequence in that family.

So for each $t \in [0,T]$ there is an increasing sequence $\left(Z^{(n)}\right)_{n\in\mathbb{N}} \subset \zeta_t$ with

$$U_t = \lim_{n\to\infty} \mathbb{E}_Q[HZ_T^{(n)}|\mathcal{F}_t],$$

hence we obtain

$$\mathbb{E}_Q[U_t|\mathcal{F}_s] = \lim \mathbb{E}_Q\left[\mathbb{E}_Q[HZ_T^{(n)}|\mathcal{F}_t]|\mathcal{F}_s\right] \leq \operatorname*{ess\,sup}_{Z\in\zeta_s} \mathbb{E}_Q[HZ_T|\mathcal{F}_s] = U_s,$$

where the inequality follows from $Z^{(n)} \in \zeta_t \subseteq \zeta_s$. By a similar argument we obtain that $t \mapsto E_Q[U_t]$ is càdlàg, hence there is a càdlàg version of U by martingale regularisation. □

So we have that $U = (U_t)_{0\leq t\leq T}$ is a Q-supermartingale for any $Q \in \mathbb{P}$. One concrete example of such a process is as follows: take $x \in \mathbb{R}$, ϑ an $\mathbb{R}^d$-valued, predictable, S-integrable process and C an increasing càdlàg, adapted process with $C_0 = 0$. Define

$$V^{x,\vartheta,C} := x + \int \vartheta dS - C$$

and interpret this as the value process of a generalised strategy (x, ϑ, C); x is the initial value, ϑ describes the trading and C_t is the amount spent for consumption on $[0,t]$. Note that $C \geq 0$ and

$$V^{x,\vartheta,C} + C = x + \int \vartheta dS,$$

so if $V^{x,\vartheta,C}$ is bounded below, then $\vartheta \in \Theta_{adm}$.

Whenever $\vartheta \in \Theta_{adm}$, $\int \vartheta dS$ is by Ansel-Stricker a Q-supermartingale for all $Q \in \mathbb{P}$. The same is then true for $V^{x,\vartheta,C}$ if this process is uniformly (in t,ω) bounded below; note that

$$0 \leq C \leq \text{const} + \int \vartheta dS$$

shows that C is Q-integrable. Hence each $V^{x,\vartheta,C}$ with $V^{x,\vartheta,C} \geq$ const. is a Q-supermartingale, for all $Q \in \mathbb{P}$. This is the *only* such example.

THEOREM 10.4 (Optional decomposition, Kramkov). *Suppose $\mathbb{P} \neq \emptyset$. Suppose $U = (U_t)_{0\leq t\leq T}$ is an adapted, càdlàg process $U_t \geq 0$ with the property that U is a Q-supermartingale for all $Q \in \mathbb{P}$. Suppose $\mathcal{F}_0$ is trivial. Then there is some $x \in \mathbb{R}$, $\vartheta \in \Theta_{adm}$ and an adapted, increasing, càdlàg process C with $C_0 = 0$ such that*

$$U = V^{x,\vartheta,C} = x + \int \vartheta dS - C$$

(In fact, $x = U_0$.)

REMARK 10.5. If $\mathcal{F}_0$ is non-trivial, we must allow $x \in L^0_{\geq 0}(\mathcal{F}_0)$.

An immediate consequence of the above theorem is the hedging duality:

THEOREM 10.6. *Suppose $\mathbb{P} \neq \emptyset$ and $\mathcal{F}_0$ is trivial. For any $H \in L^0_{\geq 0}(\mathcal{F}_T)$ we then have*

$$\Pi_s(H) = \inf\{V_0 \in \mathbb{R} | H - V_0 \in G_T(\Theta_{adm}) - L^0_{\geq 0}\} = \sup_{Q \in \mathbb{P}} \mathbb{E}_Q[H].$$

Moreover, the infimum is attained as a minimum if $\sup_{Q \in \mathbb{Q}} \mathbb{E}[H] < \infty$.

PROOF. **"$\geq$":** Follows from Lemma 10.2

"$\leq$": is trivial if $RHS = +\infty$. So suppose that $\sup_{Q \in \mathbb{P}} \mathbb{E}[H] < \infty$; with $U_0 := \operatorname{ess\,sup}_{Q \in \mathbb{P}} \mathbb{E}_Q[H|\mathcal{F}_0]$, this means by Proposition 10.3 that U is a Q-supermartingale, $\forall Q \in \mathbb{P}$, so $U = U_0 + \int \vartheta dS - C$ by Theorem 10.4 with $\vartheta \in \Theta_{adm}$, $C \nearrow$, null at 0. So $C_T \geq 0$ and so $H - U_0 = U_T - U_0 = \int_0^T \vartheta dS - C_T \in G_T(\Theta_{adm}) - L^0_{\geq 0}$ shows (by using the definition of $\Pi_s(H)$) that $V_0 \leq U_0 = \sup_{Q \in \mathbb{P}} \mathbb{E}_Q[H]$; the argument also shows that the infimum is attained by $V_0 = U_0$.

□

REMARK 10.7. Here, $C = \hat{B}$ is predictable; in general (i.e. discountinuous $\mathbb{F}$), C is only optional. As $\hat{B} \neq C$ - we add an extra term from $\langle N, \hat{N} \rangle$ to it.

Recall the hedging duality:

$$\Pi_s(H) = \inf\left\{V_0 \in \mathbb{R} | H - V_0 \in G_T(\Theta_{adm}) - L^0_{\geq 0}\right\} = \sup_{Q \in \mathbb{P}} \mathbb{E}[H]$$

with the infimum obtained if the right hand side is finite.

REMARK 10.8.
- Super-replication as a conceptual approach is natural, nice, mathematically beautiful; it also comes up as an auxiliary tool in other problems.
- As an approach to hedging/pricing, it is rather extreme: seller charges enough to reduce his own risk to 0 (because he achieves $V_T(\varphi) \geq H$ P-a.s. with φ admissible and self-financing). All the risk in the deal is with the buyer.
- $\Pi_s(H)$ is a nice price for the seller, but buyer might be unhappy; e.g. can have H bounded ≥ 0 with $\Pi_s(H) = \|H\|_{L^\infty}$ (if S is driven by a Brownian motion and H comes from Poisson jumps), or $\Pi_s((S_T - K)^+) = S_0$ (for certain stochastic volatility models). So buyers might be hard to find.
- We can similarly define a buyer price $\Pi_b(H) = -\Pi_s(-H)$, at least if H is bounded.
- Combining above results shows that reasonable (arbitrage-free) prices/values for H form an interval between $\inf_{Q \in \mathbb{P}} \mathbb{E}_Q[H]$ and $\sup_{Q \in \mathbb{P}} \mathbb{E}_Q[H]$. More precisely: if H is traded at any x from the *open* interval, this gives no arbitrage in the market extended by (x, H). Trading at any price outside the *closed* interval will introduce arbitrage. The behaviour for "boundary prices" depends on S.
- All of the above is OK for H bounded; unbounded H need technical care.
- Up to suitable choices of sign, the superreplication price gives an example of a so-called *convex risk measure*.
- We can use the optional decomposition to characterize attainable payoffs. We call $H \in L^0_{\geq 0}(\mathcal{F}_T)$ attainable, if $H = V_T(V_0, \vartheta) = V_0 + G_T(\vartheta)$ $\mathbb{P}$-a.s. for some $V_0 \in \mathbb{R}$, $\vartheta \in \Theta_{adm}$ such that $G(\vartheta) = \int \vartheta dS$ is a martingale under Q^* for some EMM Q^* (it is always a Q-supermartingale, for all $Q \in \mathbb{P}$).

Then, for $\mathcal{F}_0$ trivial, one can show that H is attainable (in the above sense) if and only if $\sup_{Q\in\mathbb{P}} \mathbb{E}_Q[H] < \infty$ and is attained in some Q^*.

11. American Options

With a European option, the time of the payoff is fixed (usually T). With an American option, the owner/holder can also choose the time of the payoff. How can we model, value and hedge such a product?

We use the usual model of $(\Omega, \mathcal{F}, P)$, $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$, $B \equiv 1$ and $S = (S_t)_{0\le t\le T}$ an $\mathbb{R}^d$-valued semimartingale. We impose absence of arbitrage via $\mathbb{P} \neq \emptyset$.

An *American option* is described by its payoff process $U = (U_t)_{0\le t\le T}$ (discounted as usual); U is $\mathbb{F}$-adapted, càdlàg, ≥ 0. Then U_τ is the payoff due at time τ if the owner decides to exercise the option at τ. The owner/holder chooses τ, but it must be a stopping time to exclude prophets and clairvoyance, with values $\tau \in [0, T]$.

Notation: $\mathcal{S}_{t,T}$ is the set of all stopping times τ with values in $[t, T]$.

Consider the seller/writer of an American option at time $t \in [0, T]$. What can she do?

- If option has already been exercised: nothing.
- Otherwise: suppose the owner chooses to exercise at τ. Then the seller faces a payoff (at τ) of U_τ. To be safe, the seller would like to be able to super-replicate this, from t on; so he needs $\operatorname{ess\,sup}_{Q\in\mathbb{P}} \mathbb{E}_Q[U_\tau|\mathcal{F}_t]$. But the seller does not know τ, so to be safe, he will also need to maximise over $\tau \in \mathcal{S}_{t,T}$. This prepares him for the worst case.

So, the natural selling price at t is:

$$\overline{V}_t := \operatorname*{ess\,sup}_{Q\in\mathbb{P},\tau\in\mathcal{S}_{0,T}} I_{\{\tau\ge t\}}\mathbb{E}_Q[U_\tau|\mathcal{F}_t] = \operatorname*{ess\,sup}_{Q\in\mathbb{P},\tau\in\mathcal{S}_{t,T}} \mathbb{E}_Q[U_\tau|\mathcal{F}_t],\ 0 \le t \le T\,.$$

Proposition 11.1. *Suppose $\mathbb{P} \neq \emptyset$ and $\mathcal{F}_0$ is trivial. If*

$$\overline{V}_0 = \sup_{Q\in\mathbb{P},\tau\in\mathcal{S}_{0,T}} \mathbb{E}[U_\tau] < \infty\,,$$

then $\overline{V}$ is a Q-super-martingale for all $Q \in \mathbb{P}$. Moreover, it is the smallest of all càdlàg processes $V' \ge U$ such that V' is Q-super-martingale, for all $Q \in \mathbb{P}$.

Proof. Similar to Proposition 10.3: fix $Q \in \mathbb{P}$ and set

$$\zeta_t := \{\text{all density processes } Z \text{ w.r.t. } Q \text{ of some } R \in \mathbb{P}, \text{ with } R = Q \text{ on } \mathcal{F}_t\}$$

Then get as in proof of Proposition 10.3 that $\overline{V}_t = \operatorname*{ess\,sup}_{Z\in\zeta_t,\tau\in\mathcal{S}_{t,T}} \underbrace{\mathbb{E}_Q[Z_\tau U_\tau|\mathcal{F}_t]}_{=:\Gamma_t(Z,\tau)}$.

Moreover, the family $\{\Gamma_t(Z,\tau)\,|\,Z \in \zeta_t, \tau \in \mathcal{S}_{t,T}\}$ is upward directed: For $\Gamma_t(Z^i, \tau_i)$, set $A := \{\Gamma_t(Z^1, \tau_1) \ge \Gamma_t(Z^2, \tau_2)\} \in \mathcal{F}_t$, so $\overline{Z} : Z^1 I_A + Z^2 I_{A^C}$ is in ζ_t (see Proposition 10.3) and $\overline{\tau} := \tau_1 I_A + \tau_2 I_{A^C}$ is in $\mathcal{S}_{t,T}$, and then $\max(\Gamma_t(Z^1, \tau_1), \Gamma_t(Z^2, \tau_2)) = \Gamma_t(\overline{Z}, \overline{\tau})$.

So for $s \le t$, we get:

$$\overline{V}_t = \operatorname*{ess\,sup}_{Z\in\zeta_t,\tau\in\mathcal{S}_{t,T}} \Gamma_t(Z,\tau) = \lim_{n\to\infty} \mathbb{E}_Q[Z^n_{\tau_n} U_{\tau_n}|\mathcal{F}_t],$$

and so (by using, in the first equality, monotone convergence due to the set being upward directed)

$$\mathbb{E}_Q[\overline{V}_t|\mathcal{F}_s] = \lim_{n\to\infty} \mathbb{E}_Q\left[\mathbb{E}_Q[Z^n_{\tau_n} U_{\tau_n}|\mathcal{F}_t]|\mathcal{F}_s\right] \le \operatorname*{ess\,sup}_{Z\in\zeta_t,\tau\in\mathcal{S}_{s,T}} \mathbb{E}_Q[Z_\tau U_\tau|\mathcal{F}_s] = \overline{V}_s\,,$$

which gives us the super-martingale property, and also $\overline{V} \ge 0$ and then $\mathbb{E}_Q[\overline{V}_t] \le \overline{V}_0 < \infty$.

We now prove the minimality of $\overline{V}$: Since $t \in \mathcal{S}_{t,T}$, we get $\overline{V} \geq U$ in the sense that $\overline{V}_t \geq U_t$ P-a.s., for all $t \in [0,T]$. If V' satisfies this as well and is a Q-super-martingale, for all $Q \in P$, and càdlàg, then $V'_t \geq \mathbb{E}_Q[V'_\tau|\mathcal{F}_t] \geq \mathbb{E}_Q[U_\tau|\mathcal{F}_t]$, for all $Q \in \mathbb{P}$, for all $\tau \in \mathcal{S}_{t,T}$, where the first inequality follows from the stopping theorem and the second one since $V' \geq U$ and both are càdlàg. So we get that $V'_t \geq \operatorname{ess\,sup}_{Q\in\mathbb{P},\tau\in\mathcal{S}_{t,T}} \mathbb{E}_Q[U_\tau|\mathcal{F}_t] = \overline{V}_t$ P-a.s., for all t. □

REMARK 11.2. One has to show that $\overline{V}$ has version which is càdlàg. This is important for the comparison between V' and $\overline{V}$. This is also important since we want $\overline{V}_\tau \geq U_\tau$, for all $\tau \in \mathcal{S}_{0,T}$.

We now look at generalised strategies with consumption, $x \in \mathbb{R}, \vartheta \in \Theta_{adm}, C$ adapted, increasing càdlàg, null at 0, with $V^{x,\vartheta,C} = x + \int \vartheta dS - C$. We also introduce for the American option the *super-replication price* at 0 as:

$$\Pi_s(U) := \inf \{V_0 \in \mathbb{R} \,|\, \exists \vartheta \in \Theta_{adm} \text{ with } V_0 + G(\vartheta) \geq U\} \,.$$

Note that we want $V_0 + G_\tau(\vartheta) \geq U_\tau$ a.s. for all stopping times; which is well defined as $G(\vartheta)$, U are both càdlàg.

THEOREM 11.3. *Suppose $\mathbb{P} \neq \emptyset$ and $\mathcal{F}_0$ trivial. If*

$$\overline{V}_0 = \sup_{Q\in\mathbb{P},\tau\in\mathcal{S}_{0,T}} \mathbb{E}_Q[U_\tau] < \infty\,,$$

then it holds that

(1) *there exists a generalized strategy with consumption (x, ϑ, C) with $V^{x,\vartheta,C} \geq U$ and (x, ϑ, C) is minimal in the sense that for any (x', ϑ', C') with $V^{x',\vartheta',C'} \geq U$, we have $V^{x,\vartheta,C} \leq V^{x',\vartheta',C'}$. Moreover, we can take $x = \overline{V}_0 = \sup_{Q\in\mathbb{P},\tau\in\mathcal{S}_{0,T}} \mathbb{E}_Q[U_\tau]$.*
(2) *the super-replication price is $\Pi_s(U) = \overline{V}_0 = \sup\{\mathbb{E}_Q[U_\tau] | Q \in \mathbb{P}, \tau \in \mathcal{S}_{0,T}\}$.*

PROOF. By Proposition 11.1, $\overline{V} \geq U$ is a Q-supermartingale, for all $Q \in \mathbb{P}$. So existence of (x, ϑ, C) is immediate from the optional decomposition Theorem 10.4, and also $x = \overline{V}_0$. The minimality: $V' := V^{x',\vartheta',C'}$ is a Q-supermartingale for all $Q \in \mathbb{P}$. So if also $V' \geq U$, then $\overline{V} \leq V'$ by Proposition 11.1.

If $V^{x,\vartheta,0} = x + G(\vartheta) \geq U$, then for any $Q \in \mathbb{P}$: $\mathbb{E}_Q[U_\tau] \leq x + \mathbb{E}_Q[G_\tau(\vartheta)] \leq x$, for all $\tau \in \mathcal{S}_{0,T}$, as $G(\vartheta)$ is Q-supermartingale. So $\Pi_s(U) \geq \overline{V}_0$. For the "$\leq$" part, take (x, ϑ, C) from part 1) with $x = \overline{V}_0$ to get $\vartheta \in \Theta_{adm}$ with $x + G(\vartheta) = V^{x,\vartheta,0} \geq V^{x,\vartheta,C} \geq U$ by 1), and so $\Pi_s(U) \leq x = \overline{V}_0$. □

Interpretation: The initial capital $x = \overline{V}_0 = \sup_{Q\in\mathbb{P},\tau\in\mathcal{S}_{0,T}} \mathbb{E}_Q[U_\tau]$ allows construction of self-financing strategy (x, ϑ) whose value process $V(x, \vartheta) = x + \int \vartheta dS \geq V^{x,\vartheta,C} \geq U$ always lies above U, so following (x, ϑ) keeps the option seller safe and allows him to make the payoff U_τ, no matter which τ is chosen by the option holder. Depending on the τ, the option seller might make a profit of: $x + G_\tau(\vartheta) - U_\tau = V^{x,\vartheta,C}_\tau - U_\tau + C_\tau \geq C_\tau$.

The same reasoning holds at any time t instead of 0; then starting with $\overline{V}_t$ at t leads to profit of $C_\tau - C_t \geq 0$ for $\tau \in \mathcal{S}_{t,T}$, since C is decreasing.

If $\mathbb{P} = \{Q^*\}$ is a singleton (so that, as we know from finite discrete time, we have a complete market), then

$$\overline{V}_t = \operatorname*{ess\,sup}_{\tau\in\mathcal{S}_{t,T}} \mathbb{E}_{Q^*}[U_\tau|\mathcal{F}_t],\ 0 \leq t \leq T.$$

Finding this is the classical *optimal stopping problem*. If one has a Markov structure, this further reduces to the *free boundary problem*, which is a PDE problem with an unknown boundary.

For general $\mathbb{P}$, finding $\overline{V}$ is usually difficult. A frequent approach, especially in the Lévy setting is to start with a P-Lévy model for S and then look for a $Q \in \mathbb{P}$ such that S (or $\log S$) is also Q-Lévy. Then we try to work out $V_t^Q := \operatorname{ess\,sup}_{\tau \in \mathcal{S}_{t,T}} \mathbb{E}_Q[U_\tau|\mathcal{F}_t]$, $0 \le t \le T$.

The next step is to use V^Q as the price process of U. This is partly all right, since it gives no arbitrage; usually, however, there is no hedging strategy to guarantee that one can stay above U in a self-financing way.

For *finite discrete time*, the results are more explicit, since we can construct $\overline{V}$ by backward recursion. For $Q \in \mathbb{P}$, denote by $Z^Q = \left(Z_k^Q\right)_{k=0,\dots,T}$ the density process of Q w.r.t. P. Define the process J recursively backward by $J_T = U_T$ and for $k = 0, 1, \dots, T-1$:

$$J_k = \max\left\{U_k, \operatorname*{ess\,sup}_{Q \in \mathbb{P}} \mathbb{E}_Q[J_{k+1}|\mathcal{F}_k]\right\}$$

Note: by Bayes' rule we obtain $\mathbb{E}_Q[J_{k+1}|\mathcal{F}_k] = \mathbb{E}_P\left[J_{k+1}\frac{Z_{k+1}^Q}{Z_k^Q}|\mathcal{F}_k\right]$ and this needs only the one-step transition probabilities of Q between k and $k+1$.

THEOREM 11.4. *Assume $\mathbb{P} \neq \emptyset$ and final discrete time. Then $J = \overline{V}$, so that $\overline{V}$ has a recursive representation.*

PROOF. All the conditional expectations are well defined in $[0, \infty]$, and we get from $\overline{V}$ the supermartingale property (for each Q) and the minimality as in proposition 11.1, even without integrability.

- **"$\geq$":** By construction, $J \geq U$ and for each $Q \in \mathbb{P}$, $J_k \geq \mathbb{E}_Q[J_{k+1}|\mathcal{F}_k]$, i.e., J has the Q-supermartingale property for all $Q \in \mathbb{P}$. But $\overline{V}$ is minimal, so $J \geq \overline{V}$.
- **"$=$":** Induction: $J_T = U_T = \overline{V}_T$, and if $J_{k+1} \le \overline{V}_{k+1}$, we get for all $Q \in \mathbb{P}$ that $\mathbb{E}_Q[J_{k+1}|\mathcal{F}_k] \le \mathbb{E}_Q[\overline{V}_{k+1}|\mathcal{F}_k] \le \overline{V}_k$ by Proposition 11.1; so $J_k = \max\left\{U_k, \operatorname{ess\,sup}_{Q \in \mathbb{P}} \mathbb{E}_Q[J_{k+1}|\mathcal{F}_k]\right\} \le \max(U_k, \overline{V}_k) = \overline{V}_k$.

□

If the market is complete, so we have $\mathbb{P} = \{Q^*\}$, the recursion becomes

$$\overline{V}_k = \max\{U_k, \mathbb{E}_{Q^*}[\overline{V}_{k+1}|\mathcal{F}_k]\}\,.$$

Financial interpretation: At time k, the option holder can either exercise the option (and get U_k) or he can continue to hold the option for at least one time step. Then the value at time $k+1$ will be $\overline{V}_{k+1}$, and viewed as a time $k+1$ payoff, that has a time k value of $\operatorname{ess\,sup}_{Q \in \mathbb{P}} \mathbb{E}_Q[\overline{V}_{k+1}|\mathcal{F}_k]$. As the option holder is free to choose his decision at k, the value of the contract for him at k is the maximum of the two possibilities.

REMARK 11.5. In the complete market case $\mathbb{P} = \{Q^*\}$, the optional decomposition of $\overline{V}$ is given by the Doob-Meyer decomposition of the Q^*-supermartingale $\overline{V}$. Indeed, $\overline{V}$ is a Q^*-supermartingale, so by Doob-Meyer $\overline{V} =$"Q^*-(local) martingale" $-$ "increasing predictable process"; and since $\mathbb{P} = \{Q^*\}$, S has the martingale representation property, so the above Q^*-martingale is a stochastic integral of S, which gives us the optional decomposition and even C predictable.

EXAMPLE 11.6 (American call option). Suppose $\mathbb{P} = \{Q^*\}$ and $S = \frac{\tilde{S}}{\tilde{B}}$ is a true Q^*-martingale. Consider $\tilde{U}_t = (\tilde{S}_t - K)^+$, $0 \le t \le T$. Then: if $\tilde{B}$ is increasing (i.e. the interest rates are non-negative), then

$$\tilde{V}_t = \tilde{B}_t \mathbb{E}_{Q^*}\left[\frac{(\tilde{S}_T - K)^+}{\tilde{B}_T}|\mathcal{F}_t\right] = \tilde{B}_t \mathbb{E}_{Q^*}\left[\frac{\tilde{U}_T}{\tilde{B}_T}|\mathcal{F}_t\right].$$

So: American call option has the same value as a European call option.

PROOF. An important point is that S is a Q^*-martingale and $x \mapsto (x-K)^+$ is convex; so we get a submartingale, and it is never optimal to stop a submartingale early. More precisely:

$$\frac{\tilde{U}_\tau}{\tilde{B}_\tau} = \left(S_\tau - \frac{K}{\tilde{B}_\tau}\right)^+ \geq \left(S_\tau - \frac{K}{\tilde{B}_t}\right)^+$$

and so

$$\mathbb{E}_{Q^*}[\frac{\tilde{U}_\tau}{\tilde{B}_\tau}|\mathcal{F}_t] \geq \mathbb{E}_{Q^*}[(S_\tau - \frac{K}{\tilde{B}_t})^+|\mathcal{F}_t] \geq (\mathbb{E}_{Q^*}[S_\tau - \frac{K}{\tilde{B}_t}|\mathcal{F}_t])^+ = (S_t + \frac{K}{\tilde{B}_t})^+ = \frac{\tilde{U}_t}{\tilde{B}_t}\,.$$

So $\frac{\tilde{U}}{\tilde{B}}$ is a Q^*-submartingale, so that $\mathbb{E}_{Q^*}\left[\frac{\tilde{U}_T}{\tilde{B}_T}|\mathcal{F}_\tau\right] \geq \frac{\tilde{U}_\tau}{\tilde{B}_\tau}$, for all $\tau \in \mathcal{S}_{t,T}$, hence by conditioning on $\mathcal{F}_t$, $\mathbb{E}_{Q^*}\left[\frac{\tilde{U}_T}{\tilde{B}_T}|\mathcal{F}_t\right] \geq \operatorname{ess\,sup}_{\tau\in\mathcal{S}_{t,T}} \mathbb{E}_{Q^*}\left[\frac{\tilde{U}_\tau}{\tilde{B}_\tau}|\mathcal{F}_t\right]$, whence we get the desired equality. $\square$

Now we replace the call by the put, i.e. $(x-K)^+$ by $(K-x)^+$. Then one might naively expect (since we again have a convex function) that the same result holds for the American put as well, but this is not so (the problem is that $\tilde{B}$ being increasing no longer helps us in the proof). One can even show (e.g. for the binomial tree): if the interest rate r is positive, then for some K the American put has a strictly higher value than a European put. However, if we model dividends by negative rates, we end up with the same phenomenon in the case of the American put.

CHAPTER 2

Utility Optimization

1. Utility optimization in discrete models

We consider the complete and incomplete case in a one period model with a general utility function and some particular examples. This section is preparatory and should provide a feeling for the type of problem, which we are going to treat.

Definition 1.1. *A real valued function* $u : I \to \mathbb{R}$ *is called* utility function *if* $I =]0, \infty[$ *or* $I =]-\infty, \infty[$ *and* u *is an increasing, strictly concave* C^2*-function. We shall denote* $\mathrm{dom}(u) := I$ *and we define* $u(x) = -\infty$ *for* $x \notin dom(u)$*. Furthermore we shall assume that* $\lim_{x \downarrow 0} u(x) = -\infty$ *if* $\mathrm{dom}(u) =]0, \infty[$.

Remark 1.2. In the sequel we shall impose further conditions on utility functions guaranteeing the existence of optimal solutions. For the presentation of the problem this is not necessary.

We consider a financial market $(S_n^0, \ldots, S_n^d)_{n=0,1}$ on $(\Omega, \mathcal{F}, P)$ with one period and aim to solve the following optimization problem for a given utility function $u : \mathrm{dom}(u) \to \mathbb{R}$ and $x \in \mathrm{dom}(u)$.

$$E_P(u(\frac{1}{S_1^0} V_1(\phi)) \to \max,$$
$$V_0(\phi) = x,$$

where ϕ is running over all self-financing trading strategies. This leads to the following one dimensional optimization problem

$$a \mapsto E_P(u(x + a(\widetilde{S}_1 - \widetilde{S}_0))),$$

which can be solved by classical analysis. We see immediately that the existence of an optimal strategy $\widehat{a}(x)$ for a fixed $x \in \mathrm{dom}(u)$ leads to

$$E_P(u'(x + \widehat{a}(x)(\widetilde{S}_1 - \widetilde{S}_0))(\widetilde{S}_1 - \widetilde{S}_0)) = 0.$$

This is in turn means that the vector can be normalized to a probability measure Q, i.e.

$$\frac{dQ}{dP} = \frac{1}{\lambda} u'(x + \widehat{a}(x)(\widetilde{S}_1 - \widetilde{S}_0)),$$

which is a martingale measure since $E_Q(\widetilde{S}_1 - \widetilde{S}_0) = 0$. Therefore the existence of an optimizer leads to arbitrage-free markets.

Next we consider the general situation in discrete models, i.e. finite Ω. Given a financial market $(S_n^0, \ldots, S_n^d)_{n=0,\ldots,N}$ on $(\Omega, \mathcal{F}, P)$ and a utility function u, then we define the *utility optimization problem* as determination of $U(x)$ for $x \in \mathrm{dom}(u)$, i.e.

$$\sup_{\substack{\phi \text{ trading strategy} \\ \phi \text{ self financing} \\ V_0(\phi)=x}} E(u(\frac{1}{S_N^0} V_N(\phi)) =: U(x).$$

We say that the utility optimization problem at $x \in \mathrm{dom}(u)$ is *solvable* if $U(x)$ is finitely valued and if we find an optimal self financing trading strategy $\widehat{\phi}(x)$ for $x \in \mathrm{dom}(u)$ such that

$$U(x) = E(u(\frac{1}{S_N^0} V_N(\widehat{\phi}(x))),$$
$$V_0(\widehat{\phi}(x)) = x.$$

We shall introduce three methods for the solution of the utility optimization problem, where the number of variables involved differ.

We assume that $\mathcal{F} = 2^\Omega$ and $P(\omega) > 0$ for $\omega \in \Omega$. We then have three characteristic dimensions: the dimension of all random variables $|\Omega|$ (the number of paths), then the dimension of discounted outcomes at initial wealth 0, denoted by $\dim \mathcal{K}$, and the number of martingale measures m. We have the basic relation

$$m + \dim \mathcal{K} = |\Omega|.$$

- the pedestrian method is an unconstraint extremal value problem in $\dim \mathcal{K}$ variables.
- the Lagrangian method yields an unconstraint extremal value problem in $|\Omega| + m$ variables.
- the duality method (martingale approach) yields an unconstraint extremal value problem in m variables. Additionally one has to transform the dual value function to the original, which is a one dimensional extremal value problem.

In financial mathematics usually $\dim \mathcal{K} \gg m$, which means that the duality method is of particular importance.

1.1. Pedestrian's method. We can understand utility optimization as unrestricted optimization problem. Define $\mathcal{S}$ the vector space of all predictable strategies $(\phi_n)_{n=0,\ldots,N}$, then the utility optimization problem for $x \in \mathrm{dom}(u)$ is equivalent to solving the following problem

$$F_x : \begin{cases} \mathcal{S} \to \mathbb{R} \cup \{-\infty\} \\ (\phi_n)_{n=0,\ldots,N} \mapsto E(u(x + (\phi \cdot \widetilde{S})_N)) \end{cases}$$
$$\sup_{\phi \in \mathcal{S}} F_x(\phi) = U(x)$$

This is an ordinary extremal value problem for every $x \in \mathrm{dom}(u)$. Let $(\widehat{\phi_n})_{n=0,\ldots,N}$ be an optimal strategy, then necessarily

$$\mathrm{grad}\, F_x((\widehat{\phi_n})_{n=0,\ldots,N}) = 0$$

and therefore we can in principle calculate the optimal strategy. From this formulation we take one fundamental conclusion.

Theorem 1.3. *Let the utility optimization problem at $x \in \mathrm{dom}(u)$ be solvable and let $(\widehat{\phi_n})_{n=0,\ldots,N}$ be an optimal strategy, so*

$$\sup_{\phi \in \mathcal{S}} F_x(\phi) = U(x) = F_x(\widehat{\phi}),$$

then $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$.

Proof. We calculate the directional derivative with respect to 1_A for $A \in \mathcal{A}(\mathcal{F}_{i-1})$ for $i = 1, \ldots, N$,

$$\frac{d}{ds}|_{s=0} E(u(x + (\widehat{\phi} \cdot \widetilde{S})_N + s 1_A \Delta S_i))$$
$$= E(u'(x + (\widehat{\phi} \cdot \widetilde{S})_N) 1_A \Delta S_i).$$

Since $(\widehat{\phi_n})_{n=0,\dots,N}$ is an optimizer we necessarily have that the directional derivatives in direction of the elements $1_A \Delta S_i$ vanish. We define

$$\lambda := E(u'(x + (\widehat{\phi} \cdot \widetilde{S})_N)) > 0$$

since $u'(y) > 0$ for $y \in \operatorname{dom}(U)$. Consequently

$$\frac{dQ}{dP} := \frac{1}{\lambda} u'(x + (\widehat{\phi} \cdot \widetilde{S})_N)$$

defines a probability measure equivalent to P. Hence we obtain from the gradient condition that

$$E_Q(1_A(S_i - S_{i-1})) = 0$$

for all $A \in \mathcal{A}(\mathcal{F}_{i-1})$ and $i = 1, \dots, N$, which means

$$E(S_i | \mathcal{F}_{i-1}) = S_{i-1}$$

for $i = 1, \dots, N$, therefore $Q \in \mathcal{M}^e(\widetilde{S})$. □

Besides baby examples the pedestrian's method is not really made for the solution of the utility optimization problem, since equations become very complicated and the internal structure does not really get clear. Nevertheless the above conclusion is of high importance, since it will be a basic assumption from now on.

Condition 1.4. *We shall always assume $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$.*

Furthermore we can easily formulate a basis existence and regularity result by the pedestrian's method (which allows to make nice general conclusions).

Proposition 1.5. *Assume $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$ and $\lim_{x \to \infty} u'(x) = 0$ if $\operatorname{dom}(u) = \mathbb{R}$, then the utility optimization problem for $x \in \operatorname{dom}(u)$ has a unique solution $\widehat{X}(x) \in x + \mathcal{K}$, which is also the unique local maximum, and $x \mapsto \widehat{X}(x)$ is C^1 on $\operatorname{dom}(u)$. If $x \notin \operatorname{dom}(u)$, then $\sup_{\phi \in \mathcal{S}} F_x(\phi) = -\infty$.*

Proof. The functional $X \mapsto E_P(u(X))$ is C^2, strictly concave and increasing. Assume that there are two optimizers $\widehat{X}_1(x) \neq \widehat{X}_2(x) \in x + \mathcal{K}$, then

$$E_P(u(t\widehat{X}_1(x) + (1-t)\widehat{X}_2(x))) > tE_P(u(\widehat{X}_1(x))) + (1-t)E_P(u(\widehat{X}_2(x))) = U(x)$$

for $t \in]0, 1[$, which is a contradiction. The argument also yields that two local maxima have to coincide. Therefore the optimizer is also the unique local maximum.

Since $\widetilde{S}$ is a martingale, the space $\mathcal{K}$ of outcomes with zero investment has the property that for $X \in L^2(\Omega, \mathcal{F}, P)$

$$X \in \mathcal{K} \Longleftrightarrow E_Q(X) = 0$$

for all $Q \in \mathcal{M}^a(\widetilde{S})$. Given an equivalent martingale measure $Q \in \mathcal{M}^e(\widetilde{S})$, then we prove that for any $x \in \operatorname{dom}(u)$

$$\lim_{\substack{Y \in \mathcal{K} \\ E_Q(|Y|) \to \infty}} E_P(u(x + Y)) = -\infty.$$

Assume that it were bounded from below by M, so we can find $Y_n \in \mathcal{K}$ such that $E_P(u(x + Y_n)) \geq M$ and $E_Q(|Y|) \geq n$. Since $Y_n \in \mathcal{K}$ we have

$$E_Q(Y_n) = 0$$

and Y_n has positive and negative components. Hence

$$E_Q((Y_n)_+) \geq \frac{n}{2}, \, E_Q((Y_n)_-) \geq \frac{n}{2}.$$

We can choose the sequence Y_n such that the smallest components form a sequence decreasing to $-\infty$ and the sequence of largest components form a sequence increasing to ∞. We have

$$|\frac{\max Y_n}{\min Y_n}| \leq M_1 < \infty$$

for all $n \geq 1$. If $\mathrm{dom}(u) =]0,\infty[$, the assertion is trivial since $-\infty$ is reached after finitely many steps. If $\mathrm{dom}(u) = \mathbb{R}$, then

$$E_P(u(x+Y)) \leq E_P(\max u(Y_n)) - E_P(u(Y_n)_-) \leq u(a_n) - b_n u(c_n)$$

with $a_n \uparrow \infty$ (largest component of Y_n), $c_n \downarrow -\infty$ (smallest component of Y_n), $b_n \in]\epsilon, 1]$ (probability $Q(Y_n = \min Y_n) > 0$) and $|\frac{a_n}{c_n}| \leq M_1$. Hence we obtain the result, since u' increases in negative direction strictly more than in positive direction.

Consequently the function $Y \mapsto E_P(u(x+Y))$ has a maximum on $\mathcal{K}$.

If $x \notin \mathrm{dom}(u)$, then for any $Y \in \mathcal{K}$, there are negative components and therefore $E_P(u(x+Y)) = -\infty$.

For the regularity assertion we take a basis of $\mathcal{K}$ denoted by $(f_i)_{i=1,\ldots,\dim\mathcal{K}}$ and calculate the derivative with respect to this basis at the unique existing optimizer $\widehat{Y}(x) = \widehat{X}(x) - x$,

$$E_P(u'(x+\widehat{Y}(x))f_i) = 0$$

for $i = 1,\ldots,\dim\mathcal{K}$. Calculating the second derivative we obtain the matrix

$$(E_P(u''(x+Y)f_i f_j))_{i,j=1,\ldots,\dim\mathcal{K}}$$

which is invertible for any $Y \in \mathcal{K}$, since u'' is strictly negative. Therefore $x \mapsto \widehat{X}(x)$ is C^1 on $\mathrm{dom}(u)$. □

1.2. Duality methods. Since we have a dual relation between the set of martingale measures and the set $\mathcal{K}$ of claims attainable at price 0, we can formulate the optimization problem as constraint problem: for any $X \in L^2(\Omega,\mathcal{F},P)$

$$X \in \mathcal{K} \Longleftrightarrow E_Q(X) = 0$$

for $Q \in \mathcal{M}^a(\widetilde{S})$ and for any probability measure Q

$$Q \in \mathcal{M}^a(\widetilde{S}) \Longleftrightarrow E_Q(X) = 0$$

for all $X \in \mathcal{K}$. Therefore we can formulate the problem as constraint optimization problem and apply the method of Lagrangian multipliers.

First we define a function $H : L^2(\Omega,\mathcal{F},P) \longrightarrow \mathbb{R}$ via

$$H(X) := E_P(u(X))$$

for a utility function u. For $x \in \mathrm{dom}(u)$ we can formulate the constraints

$$U_x := \mathcal{K} + x = \{X \in L^2(\Omega,\mathcal{F},P) \text{ such that } E_Q(X) = x \text{ for } Q \in \mathcal{M}^a(\widetilde{S})\}.$$

Consequently the utility optimization problem reads

$$\sup_{X \in U_x} E_P(u(X)) = U(x)$$

for $x \in \mathrm{dom}(u)$. Hence we can treat the problem by Lagrangian multipliers, i.e. if $\widehat{X} \in U_x$ is an optimizer, then

$$u'(\widehat{X}) - \sum_{i=1}^{m} \widehat{\eta}_i \frac{dQ_i}{dP} = 0 \qquad \text{(LM)}$$

$$E_{Q_i}(\widehat{X}) = x$$

for $i = 1, \dots, m$, $\mathcal{M}^a(\widetilde{S}) = \langle Q_1, \dots, Q_m \rangle$ and some values $\widehat{\eta}_i$. This result is obtained by taking the gradient of the function

$$X \mapsto E_P(u(X) - \sum_{i=1}^m \eta_i (\frac{dQ_i}{dP} X - x))$$

with respect to some basis. We can choose the $\widehat{\eta}_i$ positive, since $u'(\widehat{X})$ represents a positive multiple of an equivalent martingale measure. Notice that by assumption $u'(x) > 0$ for all $x \in \mathrm{dom}(u)$, and $u'(\widehat{X})$ is finitely valued.

LEMMA 1.6. *If $(\widehat{X}, \widehat{\eta_1}, \dots, \widehat{\eta_m})$ is a solution of the Lagrangian multiplier equation (LM), then the multipliers $\widehat{\eta}_i > 0$ are uniquely determined and $\sum_{i=1}^m \widehat{\eta}_i > 0$. Given $x \in \mathrm{dom}(u)$, the map $x \mapsto (\widehat{\eta}_i(x))_{i=1,\dots,m}$ is C^1.*

PROOF. The coefficients $\widehat{\eta}_i$ are uniquely determined and the inverse function theorem together with the previous result yields the C^1-dependence. □

The Lagrangian $\widetilde{L}$ is given through

$$\widetilde{L}(X, \eta_1, \dots, \eta_m) = E_P(u(X)) - \sum_{i=1}^m \eta_i (E_{Q_i}(X) - x)$$

for $X \in L^2(\Omega, \mathcal{F}, P)$ and $\eta_i \geq 0$. We introduce $y := \eta_1 + \dots + \eta_m$ and $\mu_i := \frac{\eta_i}{y}$ (we can assume $y > 0$ since the value for η_i we are looking for has to satisfy $y > 0$). Therefore

$$L(X, y, Q) = E_P(u(X)) - y(E_Q(X) - x)$$

for $X \in L^2(\Omega, \mathcal{F}, P)$, $Q \in \mathcal{M}^a(\widetilde{S})$ and $y > 0$. We define

$$\Phi(X) := \inf_{\substack{y>0 \\ Q \in \mathcal{M}^a(\widetilde{S})}} L(X, y, Q)$$

for $X \in L^2(\Omega, \mathcal{F}, P)$ and

$$\psi(y, Q) = \sup_{X \in L^2(\Omega, \mathcal{F}, P)} L(X, y, Q)$$

for $y > 0$ and $Q \in \mathcal{M}^a(\widetilde{S})$. We can hope for

$$\sup_{X \in L^2(\Omega, \mathcal{F}, P)} \Phi(X) = \inf_{y>0} \inf_{Q \in \mathcal{M}^a(\widetilde{S})} \psi(y, Q) = U(x).$$

by a mini-max consideration.

REMARK 1.7. Where does the minimax consideration stem from? Look at $X \mapsto \widetilde{L}(X, \eta_1, \dots, \eta_m)$ for fixed $\eta_1, \dots, \eta_m$, then we obtain something strictly concave as sum of two concave functions, where one is strictly concave. Look at $(\eta_1, \dots, \eta_m) \mapsto \widetilde{L}(X, \eta_1, \dots, \eta_m)$ for fixed $X \in L^2(\Omega, \mathcal{F}, P)$, then we obtain something affine.

LEMMA 1.8. *Let u be a utility function and $(S_n^0, S_n^1, \dots, S_n^d)_{n=0,\dots,N}$ be a financial market, which is arbitrage-free, then*

$$\sup_{X \in L^2(\Omega, \mathcal{F}, P)} \Phi(X) = U(x).$$

PROOF. We can easily prove the following facts:

$$\Phi(X) = -\infty \text{ if } E_Q(X) > x$$

for at least one $Q \in \mathcal{M}^a(\widetilde{S})$. Furthermore

$$\Phi(X) = E_P(u(X)) \text{ if } E_Q(X) \leq x$$

for all $Q \in \mathcal{M}^a(\widetilde{S})$. Consequently

$$\sup_{X \in L^2(\Omega,\mathcal{F},P)} \Phi(X) = \sup_{\substack{X \in L^2(\Omega,\mathcal{F},P) \\ E_Q(X) \le x \text{ for } Q \in \mathcal{M}^a(\widetilde{S})}} E_P(u(X)) = U(x)$$

since u is increasing. $\square$

For the proof of the minimax statement we need to calculate ψ, which is done in the next lemma. Therefore we assume the generic conditions for conjugation as stated in the Appendix.

LEMMA 1.9. *Given an arbitrage-free financial market $(S^0, \ldots, S^d)$, the function*

$$\psi(y, Q) = \sup_{X \in L^2(\Omega,\mathcal{F},P)} L(X, y, Q)$$

can be expressed by the conjugate function v of u,

$$\psi(y, Q) = E_P(v(y\frac{dQ}{dP})) + yx.$$

PROOF. By definition we have

$$\begin{aligned} L(X, y, Q) &= E_P(u(X)) - y(E_Q(X) - x) \\ &= E_P(u(X) - y\frac{dQ}{dP}X) + yx. \end{aligned}$$

If we fix $Q \in \mathcal{M}^a(\widetilde{S})$ and $y > 0$, then the calculation of the supremum over all random variables yields

$$\begin{aligned} & E_P(u(X) - y\frac{dQ}{dP}X) \\ &= E_P(\sup_{X \in L^2(\Omega,\mathcal{F},P)} u(X) - y\frac{dQ}{dP}X) \\ &= E_P(v(y\frac{dQ}{dP})) \end{aligned}$$

by definition of the conjugate function. $\square$

DEFINITION 1.10. *Given the above setting we call the optimization problem*

$$V(y) := \inf_{Q \in \mathcal{M}^a(\widetilde{S})} E_P(v(y\frac{dQ}{dP}))$$

the dual problem and V the dual value function for $y > 0$.

Next we formulate that the dual optimization problem has a solution.

LEMMA 1.11. *Let u be a utility function under the above assumptions and assume $\mathcal{M}^e(\widetilde{S}) \neq \emptyset$, then there is a unique optimizer $\widehat{Q}(y)$ such that*

$$V(y) = \inf_{Q \in \mathcal{M}^a(\widetilde{S})} E_P(v(y\frac{dQ}{dP})) = E_P(v(y\frac{d\widehat{Q}(y)}{dP})).$$

Furthermore

$$\inf_{y>0}(V(y) + xy) = \inf_{\substack{y>0 \\ Q \in \mathcal{M}^a(\widetilde{S})}} (E_P(v(y\frac{dQ}{dP})) + xy).$$

PROOF. Since v is strictly convex, C^2 on $]0, \infty[$ and $v'(0) = -\infty$ we obtain by compactness the existence of an optimizer $\widehat{Q}(y)$ and by $v'(0) = -\infty$ that the optimizer is an equivalent martingale measure (since one can decrease the value of

$v(y\frac{dQ}{dP})$ by moving away from the boundary). By strict convexity the optimizer is also unique. The gradient condition for $\widehat{Q}(y)$ reads as follows

$$E_P(v'(\widehat{Q}(y))(\frac{d\widehat{Q}(y)}{dP} - \frac{dQ}{dP})) = 0$$

for all $Q \in \mathcal{M}^a(\widetilde{S})$. The function V shares the same qualitative properties as v and therefore we can define the concave conjugate. Fix $x \in \mathrm{dom}(u)$ and take the optimizer $\widehat{y} = \widehat{y}(x) > 0$, then

$$\begin{aligned}\inf_{y>0}(V(y) + xy) = V(\widehat{y}) + x\widehat{y} &\leq \inf_{Q\in\mathcal{M}^a(\widetilde{S})} E_P(v(y\frac{dQ}{dP})) + xy \\ &\leq E_P(v(y\frac{dQ}{dP})) + xy\end{aligned}$$

for all $Q \in \mathcal{M}^a(\widetilde{S})$ and $y > 0$, so

$$\inf_{y>0}(V(y) + xy) \leq \inf_{\substack{y>0 \\ Q\in\mathcal{M}^a(\widetilde{S})}} (E_P(v(y\frac{dQ}{dP})) + xy).$$

Take $y_1 > 0$ and $Q_1 \in \mathcal{M}^e(\widetilde{S})$ for some $\epsilon > 0$ such that

$$\begin{aligned}\inf_{y>0}(V(y) + xy) + 2\epsilon &\geq V(y_1) + xy_1 + \epsilon \\ &\geq E_P(v(y_1\frac{dQ_1}{dP})) + xy_1 \\ &\geq \inf_{\substack{y>0 \\ Q\in\mathcal{M}^a(\widetilde{S})}} (E_P(v(y\frac{dQ}{dP})) + xy).\end{aligned}$$

Since this holds for every $\epsilon > 0$ we can conclude. □

THEOREM 1.12. *Let $(S^0, \ldots, S^d)$ be an arbitrage-free market and u a utility function with the above properties, then*

$$U(x) = \inf_{\substack{y>0 \\ Q\in\mathcal{M}^a(\widetilde{S})}} (E_P(v(y\frac{dQ}{dP})) + xy)$$

and the mini-max assertion holds.

PROOF. Fix $x \in \mathrm{dom}(u)$ and take an optimizer $\widehat{X}$, then there are Lagrangian multipliers $\widehat{\eta_1}, \ldots, \widehat{\eta_m} \geq 0$ such that $\widehat{y} := \sum_{i=1}^m \widehat{\eta_i} > 0$ and

$$\widetilde{L}(\widehat{X}, \widehat{\eta_1}, \ldots, \widehat{\eta_m}) = U(x),$$

and the constraints are satisfied so $E_{Q_i}(\widehat{X}) = x$ and $\widehat{X}$ is an optimizer. We define a measure $\widehat{Q}$ via

$$u'(\widehat{X}) = \widehat{y}\frac{d\widehat{Q}}{dP}.$$

Since

$$u'(\widehat{X}) - \widehat{y}\sum_{i=1}^m \frac{\widehat{\eta_i}}{\widehat{y}}\frac{dQ_i}{dP} = 0$$

by the Lagrangian multipliers method, we see that

$$\widehat{y}\frac{d\widehat{Q}}{dP} = \widehat{y}\sum_{i=1}^m \frac{\widehat{\eta_i}}{\widehat{y}}\frac{dQ_i}{dP}$$

and therefore $\widehat{Q} \in \mathcal{M}^e(\widetilde{S})$ (its Radon-Nikodym derivative is strictly positive). Furthermore

$$E_P(v(\widehat{y}\frac{d\widehat{Q}}{dP})) + x\widehat{y} = \inf_{Q \in \mathcal{M}^a(\widetilde{S})} (E_P(v(\widehat{y}\frac{dQ}{dP})) + x\widehat{y}),$$

since $v'(y) = -(u')^{-1}(y)$ and $Q_* \in \mathcal{M}^e(\widetilde{S})$ is a minimum if and only if

$$E_P(v'(y\frac{dQ_*}{dP})(\frac{dQ_*}{dP} - \frac{dQ}{dP})) = 0$$

for all $Q \in \mathcal{M}^a(\widetilde{S})$. This is satisfied by $\widehat{y}$ and $\widehat{Q}$. By definition of v we obtain

$$\begin{aligned} E_P(v(\widehat{y}\frac{d\widehat{Q}}{dP})) + x\widehat{y} &= \sup_{X \in L^2(\Omega,\mathcal{F},P)} L(X, \widehat{y}, \widehat{Q}) \\ &= L(\widehat{X}, \widehat{y}, \widehat{Q}), \end{aligned}$$

since $u'(\widehat{X}) = \widehat{y}\frac{d\widehat{Q}}{dP}$, $v(y) = u((u')^{-1}(y) - y(u')^{-1}(y)$, so $v(\widehat{y}\frac{d\widehat{Q}}{dP}) = u(\widehat{X}) - \frac{d\widehat{Q}}{dP}\widehat{y}\widehat{X}$. However $L(\widehat{X}, \widehat{y}, \widehat{Q}) = U(x)$ by assumption on optimality of $\widehat{X}$. Therefore

$$E_P(v(\widehat{y}\frac{d\widehat{Q}}{dP})) + x\widehat{y} = U(x)$$

and $\widehat{y}$ is the minimizer since

$$E_P(v'(\widehat{y}\frac{d\widehat{Q}}{dP})\frac{d\widehat{Q}}{dP}) = -x$$

by assumption. Calculating with the formulas for v yields

$$\begin{aligned} \inf_{\substack{y>0 \\ Q \in \mathcal{M}^a(\widetilde{S})}} (E_P(v(y\frac{dQ}{dP}))) + xy) &= \inf_{y>0} (E_P(v(y\frac{d\widehat{Q}}{dP})) + xy) \\ &= U(x) \\ &= E_P(u(\widehat{X})) \end{aligned}$$

by definition. □

This Theorem enables us to formulate the following duality relation. Given a utility optimization problem for $x \in \operatorname{dom}(u)$

$$\sup_{Y \in \mathcal{K}} E_P(u(x+Y)) = U(x),$$

then we can associate a dual problem, namely

$$\inf_{Q \in \mathcal{M}^a(\widetilde{S})} E_P(v(y\frac{dQ}{dP})) = V(y)$$

for $y > 0$. The main assertion of the minimax considerations is that

$$\inf_{y>0} (V(y) + xy) = U(x),$$

so the concave conjugate of V is U and since V shares the same regularity as U, also U is the convex conjugate of V. First we solve the dual problem (which is much easier) and obtain $y \mapsto \widehat{Q}(y)$. For given $x \in \operatorname{dom}(u)$ we can calculate $\widehat{y}(x)$ and obtain

$$\begin{aligned} V(\widehat{y}(x)) + x\widehat{y}(x) &= U(x) \\ u'(\widehat{X}(x)) &= \widehat{y}(x)\frac{d\widehat{Q}(\widehat{y}(x))}{dP}. \end{aligned}$$

2. Some ideas from optimal stochastic control

Recall the basic problem: maximise $\mathbb{E}[u(V_T(x,\vartheta))]$ over all $\vartheta \in \Theta^x_{adm}$. Here, we have that $\Theta^x_{adm} = \{$predictable S-integrable $\mathbb{R}^d$-valued ϑ with $\int \vartheta dS \geq -x\}$. With no loss of generality we can also impose that $(u(V_T(x,\vartheta)))^- \in L^1(P)$.

We now fix $t \in [0,T]$, $\vartheta \in \Theta^x_{adm}$ and define

$$\Theta(t,\vartheta) := \{\psi \in \Theta^x_{adm} | \psi = \vartheta \text{ on } [0,t]\}\,.$$

The key idea now is to look at all the conditional problems to maximize $\mathbb{E}[u(V_T(x,\psi))|\mathcal{F}_t]$ over all $\psi \in \Theta(t,\vartheta)$ (for every $\vartheta \in \Theta^x_{adm}$). So we define the maximal conditional expected utility, given the initial wealth and an initial strategy ϑ, i.e.

$$J_t(\vartheta) := \operatorname*{ess\,sup}_{\psi\in\Theta(t,\vartheta)} \underbrace{\mathbb{E}[u(V_T(x,\psi))|\mathcal{F}_t]}_{=:\Gamma_t(\psi)}\,.$$

If $\mathcal{F}_0$ is trivial, then for all $\vartheta \in \Theta^x_{adm}$ we have

$$J_0(\vartheta) = J_0 = \sup_{\psi\in\vartheta^x_{adm}} \mathbb{E}[u(V_T(x,\psi))] = U(x)\,,$$

where this U corresponds to the one from the previous chapter.

Remark 2.1. One should be careful with the conditions on u and ϑ to ensure in the sequel that there are no integrability problems, e.g. $u \geq 0$ or u bounded above might be useful assumptions. We do not take care of the exact details here.

The main result is then the following version of the martingale optimality principle from stochastic calculus (dynamic programming principle):

Theorem 2.2 (Martingale Optimality Principle (**MOP**) - with suitable integrability). *The following hold:*

(1) *For every $\vartheta \in \Theta^x_{adm}$, the process*
$$(J_t(\vartheta))_{0\leq t\leq T}$$
is a P-supermartingale.

(2) *A strategy $\vartheta^* \in \Theta^x_{adm}$ is optimal, i.e.*
$$\mathbb{E}[u(V_T(x,\vartheta^*))] = \sup_{\vartheta\in\Theta^x_{adm}} \mathbb{E}[u(V_T(x,\vartheta))]$$
if and only if $(J_t(\vartheta^*))_{0\leq t\leq T}$ is a P-martingale.

Proof. First we check that $\{\Gamma_t(\psi)|\psi \in \Theta(t,\vartheta)\}$ is upward directed: for $t \in [0,T]$, $A \in \mathcal{F}_t$, $\psi^1,\psi^2 \in \Theta(t,\vartheta)$, we have $\psi^1 I_A + \psi^2 I_{A^c} \in \Theta(t,\vartheta)$ so with $A := \{\Gamma_t(\psi^1) \geq \Gamma_t(\psi^2)\} \in \mathcal{F}_t$, we get $\max\{\Gamma_t(\psi^1),\Gamma_t(\psi^2)\} = \Gamma_t(\psi^1 I_A + \psi^2 I_{A^c})$.

So there exists an sequence $(\psi^n)_{n\in\mathbb{N}}$ in $\Theta(t,\vartheta)$ with $J_t(\vartheta) = \nearrow -\lim_{n\to\infty}\Gamma_t(\psi^n)$ and so monotone convergence holds:

$$\begin{aligned}\mathbb{E}[J_t(\vartheta)|\mathcal{F}_s] &= \lim_{n\to\infty}\mathbb{E}[\Gamma_t(\psi^n)|\mathcal{F}_s] = \lim_{n\to\infty} \overbrace{\mathbb{E}[u(V_T(x,\psi^n))|\mathcal{F}_s]}^{=\Gamma_s(\psi^n) \text{ and } \psi^n\in\Theta(t,\vartheta)\subseteq\Theta(s,\vartheta)} \\ &\leq \operatorname*{ess\,sup}_{\psi\in\Theta(s,\vartheta)} \Gamma_s(\psi) = J_s(\vartheta)\,.\end{aligned}$$

Integrability of $J(\vartheta)$ goes analogously; one needs control on J_0, e.g. $U \geq 0$ or $J_0 = U(x) < \infty$ work.

Now we take $\vartheta^* \in \Theta^x_{adm}$; then $J(\vartheta^*)$ is a P-supermartingale by 1). So $J(\vartheta^*)$ is a P-martingale if and only if it has constant expectation; and on $[0,T]$ this is equivalent to:

$$\mathbb{E}[u(V_T(x,\vartheta^*))] = \mathbb{E}[J_T(\vartheta^*)] = J_0 = \sup_{\psi\in\Theta^x_{adm}} \mathbb{E}\left[u(V_T(x,\psi))\right].$$

This means that ϑ^* is optimal. ∎

REMARK 2.3. Note that 2) includes the condition $\vartheta^* \in \Theta^x_{adm}$. So if we just exhibit some predictible S-integrable $\overline{\vartheta}$ s.t. $J(\overline{\vartheta})$ is a P-martingale, we can only conclude optimality of $\overline{\vartheta}$ after we check that $\overline{\vartheta} \in \Theta^x_{adm}$. [This is quite often not handled properly in applications.]

Now we want to exploit theorem 2.2 to get more information on ϑ^*. First, we can prove that $J(\vartheta)$ has a càdlàg version; we use that and decompose uniquely (by Doob-Meyer) as $J(\vartheta) = J_0 + M(\vartheta) - B(\vartheta)$ with $M(\vartheta) \in \mathcal{M}_{0,loc}$, $B(\vartheta)$ predictable, increasing, null at $t = 0$. Can we say even more?

We look at

$$J_t(\vartheta) = \operatorname*{ess\,sup}_{\psi \in \Theta(t,\vartheta)} \mathbb{E}\left[u(V_T(x,\psi))|\mathcal{F}_t\right] = \operatorname*{ess\,sup}_{\psi \in \Theta(t,\vartheta)} \mathbb{E}\left[u(V_t(x,\vartheta) + \int_t^T \psi_u dS_u)|\mathcal{F}_t\right].$$

We expect that each of the conditional expectations, and hence also $J_t(\vartheta)$ is an $\mathcal{F}_t$-measurable functional of $V_t(x,\vartheta)$. So we also expect that $B_t(\vartheta)$ depends on ϑ, $V_t(x,\vartheta)$ in a "nice" way.

From theorem 2.2, $B(\vartheta)$ is always increasing for each ϑ and it is constant (null) for optimal ϑ^*. In other words, the "drift" $b(\vartheta)$" is always ≥ 0, and $\equiv 0$ for ϑ^*. This can be exploited to obtain (non-linear) PDEs for the solution of the optimization problem.

The Merton Problem. Setup: We have a bank account $\tilde{B}$ and a stock $\tilde{S}$ with:

$$\begin{aligned} d\tilde{B}_t &= \tilde{B}_t r dt,\ \tilde{B}_0 = 1 \\ d\tilde{S}_t &= \tilde{S}_t\left(\mu dt + \sigma dW_t\right),\ \tilde{S}_0 > 0 \end{aligned}$$

for $\mu, r \in \mathbb{R}, \sigma \in \mathbb{R}^+$.

For finite time horizon T, we want to maximize the expected utility for final wealth, $\mathbb{E}[u(V_T(x,\vartheta))] = \max_\vartheta$! We do this by re-parametrizing: u is defined on $(0,\infty)$, so $V(x,\vartheta)$ must be > 0, so we can describe a strategy not via number of shares (ϑ) but by fractions of wealth (π).

Call $V(x,\vartheta), \tilde{V}(x,\vartheta)$ the discounted and undiscounted wealth in terms of ϑ, and define $\pi_t := \frac{\vartheta_t \tilde{S}_t}{\tilde{V}_t(x,\vartheta)} = \frac{\vartheta_t S_t}{V(x,\vartheta)}$. π_t is the fraction at time t of total wealth that is invested in stock; the fraction $1 - \pi_t$ is in the bank account.

Call $X^\pi := \tilde{V}(x,\vartheta)$ the undicounted wealth expressed with π, with x fixed. The self-financing condition for X^π is then: $dV(x,\vartheta) = \vartheta dS$, so

$$d\left(\frac{X^\pi}{\tilde{B}}\right) = \frac{\pi X^\pi}{\tilde{B} S} dS = \frac{X^\pi}{\tilde{B}} \pi \frac{dS}{S}$$

and so

$$\begin{aligned} dX_t^\pi &= d\left(\tilde{B}_t \frac{X_t^\pi}{\tilde{B}_t}\right) = \tilde{B}_t d\left(\frac{X_t^\pi}{\tilde{B}_t}\right) + \frac{X_t^\pi}{\tilde{B}_t} d\tilde{B}_t = \pi_t X_t^\pi \overbrace{\frac{dS_t}{S_t}}^{=(\mu-r)dt+\sigma dW} + X_t^\pi r dt \\ &= r X_t^\pi dt + \pi_t X_t^\pi ((\mu - r)dt + \sigma dW_t). \end{aligned}$$

It is our goal to maximize $\mathbb{E}[U(X_T^\pi)]$ over all allowed $\pi = (\pi)_{0 \leq t \leq T}$ in the sequel. For this purpose fix $t \in [0,T]$, strategy π and another strategy ψ with $\psi = \pi$ on

$[0,t]$. Consider

$$\Gamma_t(\psi) = \mathbb{E}[U(X_T^\psi)|\mathcal{F}_t] = \mathbb{E}\left[U(X_t^\pi + \int_t^T dX_u^\psi)|\mathcal{F}_t\right]$$
$$= \mathbb{E}\left[U(X_t^\pi + \int_t^T (rX_u^\psi + \psi_u X_u^\psi(\mu - r))du + \psi_u X_u^\psi \sigma dW_u)|\mathcal{F}_t\right].$$

Our filtration $\mathbb{F}$ is generated by $\tilde{S}$, $\tilde{B}$ or equivalently by W. Recall that W has the Markov property, so "the situation is Markovian": it seems plausible that

- $\Gamma_t(\psi)$ should only depend on the current wealth X_t^π and
- it is sufficient to consider strategies ψ which only depend on current wealth, $\psi_t = g(t, X_t^\psi)$, since the optimal strategy has to be of this type. Notice that this defines a stochastic differential equation for X.

So it is natural to guess that also after optimisation, this persists; we guess that

$$J_t(\pi) = \operatorname*{ess\,sup}_{\psi\in\Theta(t,\vartheta)} \mathbb{E}[U(X_T^\psi)|\mathcal{F}_t] = k(t, X_t^\pi)$$

for some function $k(t,x)$. What do we get then?

Assume k is nice and use Itô's formula. This gives:

$$dJ(\pi) = k_t dt + k_x \underbrace{dX^\pi}_{=\cdots} + \frac{1}{2}k_{xx} \underbrace{d\langle X^\pi\rangle}_{=\pi^2(X^\pi)^2\sigma^2 dt}.$$

So we get:

$$dJ_t(\pi) = \underbrace{k_x(t, X_t^\pi)\pi_t X_t^\pi \sigma dW_t}_{=dM_t(\pi)}$$
$$+ \underbrace{\left(\frac{\partial k}{\partial t} + \frac{\partial k}{\partial x}rx + \frac{\partial k}{\partial x}px(\mu - r) + \frac{1}{2}\frac{\partial^2 k}{\partial x^2}p^2x^2\sigma^2\right)_{(t,x=X_t^\pi,p=\pi_t)} dt}_{=-dB_t(\pi)=-b(t,\pi_t,X_t^\pi)dt}.$$

By the martingale optimality principle, $B(\pi)$ is always increasing and constant at optimal π^*; so $b(\pi)$ (respectively $-b(\pi)$) is always ≥ 0 (≤ 0), and $= 0$ at optimal π^*.

Treating $p \widehat{=} \pi_t$ and $x \widehat{=} X_t^\pi$ as independent variables leads us to guess that $k(t,x)$ should satisfy

$$\sup_{p>0}\left(k_t(t,x) + rxk_x(t,x) + (\mu - r)pxk_x(t,x) + \frac{1}{2}\sigma^2p^2x^2k_{xx}(t,x)\right) = 0.$$

This is the so called *Hamilton-Jacobi-Bellman* (HJB) equation for our control problem. It is a nonlinear PDE. Since $k(T, X_T^\pi) = J_T(\pi) = u(X_T^\pi)$ we impose $k(T,x) = u(x)$ for $x > 0$ as our boundary condition.

The idea now is to try and solve the HJB equation to come up with a candidate for the optimal strategy, π^*.

If we formally maximise over p we get the optimiser $p^*(t,x) = -\frac{\mu-r}{\sigma^2}\frac{k_x(t,x)}{xk_{xx}(t,x)}$. Plugging this in yields the HJB equation in the form:

$$0 = k_t(t,x) + rxk_x(t,x) - \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{(k_x(t,x))^2}{k_{xx}(t,x)},\ k(T,x) = U(x).$$

This is a nonlinear second order PDE for k. Conceptually, we should now to the following:

(1) Find a sufficiently smooth solution $k(t,x)$ to the HJB equation.
(2) Define function $p^*(t,x)$ from k as above.

(3) Consider the SDE: $dX_t = rX_t dt + p^*(t, X_t)X_t((\mu - r)dt + \sigma dW_t)$ obtained by using the "candidate strategy" $p^*(t, X_t)$ for π^* (and writing the self-financing equation), and prove that this has a solution X^*.
(4) Define $\pi_t^* := p^*(t, X_t^*)$ and show that π^* is an allowed strategy. (Then, by 3), $X^{\pi^*} = X^*$.)
(5) Prove that π^* is optimal, either by direct argument (by comparing it to all other allowed π), or by showing that $X^* = X^{\pi^*}$ is such that $(J(\pi_\bullet^*) = k(\bullet, X_\bullet^{\pi^*}) = k(\bullet, X_\bullet^*)$ is a martingale.

The most difficult step is usually the first one.

EXAMPLE 2.4. For power utility $u(x) = \frac{1}{\gamma}x^\gamma$ with $\gamma < 1$, $\gamma \neq 0$, we can solve the PDE explicitly. This goes as follows. the wealth dynamics

$$\frac{dX_t^\pi}{X_t^\pi} = rdt + \pi_t((\mu - r)dt + \sigma dW_t),\ X_0^\pi = x$$

give

$$X_t^\pi = x\mathcal{E}\left(rs + \int \pi_s((\mu - r)ds + \sigma dW_s)\right)_t$$

and so, for $\psi \in \Theta(t, \pi)$,

$$X_T^\psi = X_t^\pi \mathcal{E}\left(rs + \int \psi_s((\mu - r)ds + \sigma dW_s)\right)_{t,T}$$

So,

$$\Gamma_t(\psi) = \mathbb{E}[U(X_T^\psi)|\mathcal{F}_t] = \begin{bmatrix} X_T^\psi = X_t^\pi \frac{X_T^\psi}{X_t^\pi} \\ U(x) = \frac{1}{\gamma}x^\gamma \end{bmatrix} = \frac{1}{\gamma}(X_t^\pi)^\gamma \overbrace{\mathbb{E}[U(\mathcal{E}(\cdots\psi)_{t,T})|\mathcal{F}_t]}^{=:\overline{\Gamma}_t(\psi)}.$$

So of course we set

$$J_t(\pi) = \frac{1}{\gamma}(X_t^\pi)^\gamma \operatorname*{ess\,sup}_{\psi\in\Theta(t,\vartheta)} \overline{\Gamma}_t(\psi)$$

and we guess that $k(t, x) = \frac{1}{\gamma}x^\gamma f(t)$.

Then $k_t = \frac{1}{\gamma}x^\gamma \dot{f}(t)$, $k_x = x^{\gamma-1}f(t)$, $k_{xx} = (\gamma - 1)x^{\gamma-2}f(t)$ and plugging this into the HJB equation yields

$$0 = \frac{1}{\gamma}x^\gamma\left(\dot{f}(t) + \gamma rf(t) - \frac{1}{2}\frac{(\mu - r)}{\sigma^2}\frac{\gamma}{\gamma - 1}f(t)\right),\ \frac{1}{\gamma}x^\gamma = \frac{1}{\gamma}x^\gamma f(T),\text{ or } f(T) = 1.$$

This ODE for f can be solved explicitly. The explicit candidate for the optimal strategy is then $\pi_t^* = p^*(t, X_t^*) = -\frac{\mu-r}{\sigma^2}\frac{1}{\gamma-1} = \frac{\mu-r}{\sigma^2(1-\gamma)}$ which prescribes to always hold a fixed proportion of total wealth (the so called *Merton proportion*) in the stock (and the rest in the bank account). One can check that this strategy is allowed and optimal.

The strategy π^* being constant still involves trading, because the corresponding ϑ^* (optimal number of shares) is not constant. In case of the Merton problem one could also argue directly that the strategy can neither depend on time nor on current wealth, hence it has to be constant. Given this fact, it is easy to calculate the value of π directly. The solution of the HJB-equation is just making precise this type of reasoning.

3. Utility Optimization for general semi-martingale models

In this section we study the basic problem of an optimal portfolio choice with preferences given by expected utility. We take the standard model with finite time $T < \infty$, $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \le t \le T}, P)$ a filtered probability space satisfying the usual conditions, $B \equiv 1$ the bank account and the discounted asset prices $S = (S_t)_{0 \le t \le T}$, where S is an $\mathbb{R}^d$-valued semimartingale. We impose absence of arbitrage via $\mathbb{P} \neq \emptyset$.

We fix an initial capital $x > 0$ and consider a self-financing strategy (x, ϑ), where ϑ is an $\mathbb{R}^d$-valued predictable S-integrable process. We impose that the strategy ϑ is $-x$-admissible so that the wealth process

$$V(x, \vartheta) = x + (\vartheta \bullet S) \ge 0.$$

Our goal is to find a $-x$-admissible strategy ϑ, so that this strategy maximizes the expected utility from terminal wealth over ϑ, i.e. maximize $\mathbb{E}[U(V_T(x, \vartheta))]$, where U is a utility function on $\mathbb{R}_+$.

Remark 3.1. Note that imposing (x, ϑ) to be a $-x$-admissible strategy ties up with $\mathrm{dom}(U) = \mathbb{R}_+$ and we could have just imposed that $V_T(x, \vartheta) \ge 0$. Moreover, if $\mathrm{dom}(U) = (-a, \infty)$ with $0 < a < \infty$, then we can just translate by a, but if $\mathrm{dom}(U) = \mathbb{R}$, finding a good class of allowed strategies becomes tricky (see Biagini/Frattelli, Biagini/Cerny).

3.1. Basics on utility functions.

For $x > 0$ we introduce

$$v(x) := \{V(x, \vartheta) = x + (\vartheta \bullet S) \mid (x, \vartheta) \text{ 0-admissible, self-financing strategy}\}$$

Definition 3.2. *A utility function is a strictly increasing, strictly concave map $U \in C^1(\mathbb{R}_+; \mathbb{R})$ satisfying the* Inada conditions*:*

(1) $U'(0) := \lim_{x \searrow 0} U'(x) = +\infty$,

(2) $U'(\infty) := \lim_{x \to \infty} U'(x) = 0$.

Suppose U is a utility function and define

$$u(x) := \sup_{V \in v(x)} \mathbb{E}[U(V_T(x, \vartheta))],$$

for which we will assume that $u(x_0) < \infty$ for some $x_0 > 0$.

Remark 3.3. U quantifies the subjective preferences by assigning to a monetary amount z a subjective utility of $U(z)$. The fact that U is increasing means that more is better and the concavity of U captures the idea of risk aversion or the effect that an extra dollar means more to a beggar than to a millionaire.

For a given $x > 0$, $u(x)$ can be interpreted as the maximal expected utility one can obtain via investment from an initial wealth x and the standing assumption implies that the optimisation problem is well-posed for at least one x_0.

Note that U is defined on $\mathbb{R}_+$ and $V_T \ge 0$, but $U(0) \in [-\infty, \infty)$ exists (as a limit $x \to 0$), so that $U(V_T(x, \vartheta))$ is well-defined in $[-\infty, \infty)$.

Moreover we set $\mathbb{E}[U(V_T)] := -\infty$ if $(U(V_T))^- \notin L^1(P)$, since $u(x) \ge U(x) > -\infty$ for any $x > 0$, i.e. we do not lose any information if we exclude such strategies.

If U is unbounded and S allows arbitrage, then $u \equiv +\infty$, so the problem just makes sense in an arbitrage-free model.

The standing assumption, i.e. $u(x_0) < \infty$ for some $x_0 > 0$, implies that $u(x) < \infty$ for any $x > 0$.

For $y > 0$ we introduce the conjugate or Legendre transform of $-U(-\cdot)$ in the sense of convex analysis,

$$J(y) := \sup_{x > 0}(U(x) - xy),$$

(see Rockafellar Chapter 12) and denote by $I := (U')^{-1}$ the inverse of the derivative of U.

LEMMA 3.4 (*Conjugacy relation*). *$J \in C^1(\mathbb{R}_+;\mathbb{R})$ is strictly decreasing, strictly convex, $J'(0) = -\infty$, $J'(\infty) = 0$, $J(0) = U(\infty)$ and $J(\infty) = U(0)$. Moreover for any $x > 0$ we have the conjugacy relation*

$$U(x) = \inf_{y>0}(J(y) + xy),$$

in addition $J' = -I$ and for any $y > 0$ we have

$$J(y) = U(I(y)) - yI(y).$$

PROOF. (SKETCH). J is clearly decreasing and convex, as it is a supremum of convex (even affine) functions. To show that $J \in C^1(\mathbb{R}_+;\mathbb{R})$ we assume that $U \in C^2(\mathbb{R}_+;\mathbb{R})$. Then $I \in C^1(\mathbb{R}_+;\mathbb{R})$ and for a fixed $y > 0$, $\sup_{x>0}(U(x) - xy)$ is attained in $x = I(y)$, so that

$$J(y) = U(I(y)) - I(y)y.$$

This last expression shows that $J \in C^1(\mathbb{R}_+;\mathbb{R})$ and

$$J'(y) = \underbrace{U'(I(y))}_{=\,y} I'(y) - I'(y)y - I(y) = -I(y).$$

$\square$

EXAMPLE 3.5. Classical utility functions on $\mathbb{R}_+$ are

$$U(x) := \log(x),$$

with corresponding conjugate

$$J(y) = \sup_{x>0}(U(x) - xy) \underbrace{=}_{x=\frac{1}{y}} -\log(y) - 1;$$

and for $\gamma \in (-\infty, 1) \setminus \{0\}$,

$$U(x) := \frac{1}{\gamma}(x^\gamma - 1),$$

with Legendre transform

$$J(y) = \sup_{x>0}(U(x) - xy) \underbrace{=}_{x=y^{\frac{1}{\gamma-1}}} \frac{1-\gamma}{\gamma} y^{\frac{\gamma}{\gamma-1}} - \frac{1}{\gamma}.$$

Note that for $\gamma < 0$, U is bounded from above by zero, while for $\gamma > 0$, U is unbounded. Moreover, for $\gamma \to 0$ we obtain the first case.

3.2. Abstract formulation and the dual problem. Let U be a utility function as above and $x > 0$, the primal problem is

$$u(x) = \sup_{V \in v(x)} \mathbb{E}[U(V_T)].$$

Consider the set of positions that can be superreplicated from initial wealth $x > 0$, with $-x$-admissible self-financing strategies, i.e.

$$\mathcal{C}(x) := \left\{ f \in L^0_+(\mathcal{F}_T) \mid \exists V \in v(x) : f \le V_T \right\} = (x + G_T(\Theta^x_{adm}) - L^0_+) \cap L^0_+,$$

where

$$\Theta^x_{adm} := \{\vartheta = (\vartheta_t)_{0\le t\le T} \mid \vartheta \in \Theta_{adm} : (\vartheta \bullet S) \ge -x\}.$$

Note that $v(x)_T \subseteq \mathcal{C}(x)$ and if $f \in \mathcal{C}(x)$ then $\mathbb{E}[U(f)] \le u(x)$, for the latter take some $V \in v(x)$ so that $V_T \ge f$; since U is increasing we have $U(f) \le U(V_T)$ and hence $\mathbb{E}[U(f)] \le \mathbb{E}[U(V_T)] \le u(x)$.

So the primal problem can be written as

$$u(x) = \sup_{f \in \mathcal{C}(x)} \mathbb{E}[U(f)].$$

As we shall see, $\mathcal{C}(x)$ is easier to describe than $v(x)$. Note also that if $f^* \in \mathcal{C}(x)$ is optimal, then there is some $\vartheta^* \in \Theta^x_{adm}$ so that

$$f^* \leq x + G_T(\vartheta^*)$$

and $V(x, \vartheta^*) \in v(x)$ is a solution to the primal problem, because

$$u(x) = \mathbb{E}[U(f^*)] \leq \mathbb{E}[U(V_T(x, \vartheta^*))] \leq u(x).$$

In order to gain more information about the primal problem we want to introduce a suitable dual problem using the conjugacy relation of U and J, and exploiting the absence of arbitrage condition. Take $Q \in \mathbb{P}(\neq \emptyset)$ and denote by Z the density process of Q with respect to P, then $S \in \mathcal{M}_{loc}(Q)$ is a local martingale with respect to Q.

Let $V = V(x, \vartheta) \in v(x)$, then $(\vartheta \bullet S)$ is well-defined and bounded below by $-x$, hence by Ansel-Stricker $(\vartheta \bullet S) \in \mathcal{M}_{loc}(Q)$ is a local martingale with respect to Q, so it is also a Q-super-martingale.

Moreover, since Z is a density process of an equivalent probability measure we have $Z > 0$ and $\mathbb{E}[Z_0] = 1$. So, if $\mathcal{F}_0$ is trivial or if we insist on $Q = P$ on $\mathcal{F}_0$, then $Z_0 \equiv 1$.

This motivates the following set: for $z > 0$ we introduce the family of all $(\mathcal{F}_t)_{t \in [0,T]}$-adapted, positive, RCLL processes Z starting at z such that for any $V \in v(1)$, ZV is a P-supermartingale, i.e.

$$\mathcal{Z}(z) := \{Z \mid Z \geq 0 \text{ adapted, càdlàg} : Z_0 = z, \ \forall V \in v(1) : ZV \ P\text{-super-martingale}\}.$$

Note that for any $x > 0$, $v(x) = xv(1)$; so the last condition is equivalent to saying that for any $V \in v(x)$, ZV is a P-super-martingale.

Remark 3.6. Any $Z \in \mathcal{Z}(z)$ is itself a super-martingale, to see this take $(x, \vartheta) = (1, 0)$, then $V(1, 0) \equiv 1 \in v(1)$ so that $ZV = Z$ is a super-martingale. Moreover, $\mathcal{Z}(z)$ contains all density processes $Q \in \mathbb{P}$ with $Q = P$ on $\mathcal{F}_0$. Finally, $\mathcal{Z}(z) = z\mathcal{Z}(1)$.

This set allows us to derive the dual problem in the following way: let $x, z > 0$, $V \in v(x)$ and $Z \in \mathcal{Z}(z)$, then ZV is a P-super-martingale starting at $Z_0V_0 = zx$, so

$$\mathbb{E}[Z_T V_T] \leq zx.$$

Recall the Legendre transform of U, i.e. for any $y > 0$,

$$J(y) = \sup_{x>0}(U(x) - xy) \geq U(x) - xy,$$

to obtain, using the super-martingale property that

$$\mathbb{E}[U(V_T)] \leq \mathbb{E}[J(Z_T) + V_T Z_T] \leq \mathbb{E}[J(Z_T)] + zx.$$

Taking the supremum over $V \in v(x)$ and the infimum over $Z \in \mathcal{Z}(z)$ yields the following expression

$$u(x) \leq \inf_{Z \in \mathcal{Z}(z)} \mathbb{E}[J(Z_T)] + zx.$$

So, for $z > 0$ it is a natural dual problem to look for

$$j(z) := \inf_{Z \in \mathcal{Z}(z)} \mathbb{E}[J(Z_T)].$$

Remark 3.7. The primal problem maximizes a concave functional, while the dual problem minimizes a convex functional.

In analogy to $\mathcal{C}(x)$, we introduce the set

$$\mathcal{D}(z) := \left\{ h \in L^0_+ \mid \exists Z \in \mathcal{Z}(z) : h \leq Z_T \right\}$$

to get the abstract equivalent version of the dual problem

$$j(z) = \inf_{h \in \mathcal{D}(z)} \mathbb{E}[J(h)],$$

this follows from the following two observations: $\mathcal{Z}(z)_T \subseteq \mathcal{D}(z)$ and if $h \in \mathcal{D}(z)$, then $\mathbb{E}[J(h)] \geq j(z)$.

Moreover, note that if we fix $z > 0$ we obtain that

$$j(z) \geq \sup_{x>0}(u(x) - xz),$$

and if we fix $x > 0$ we get

$$u(x) \leq \inf_{z>0}(j(z) + zx).$$

This is very reminiscent of the conjugacy relation between U and J. We will see that we actually get equalities above, plus solvability of the primal as well as the dual problem at the expense of one extra assumption on U.

3.3. Solving the (abstract) dual problem. The main goal of this section will be to show that if $j(z) < \infty$, then there is a unique optimizer $h^*_z \in \mathcal{D}(z)$ of the abstract dual problem, i.e.

$$\mathbb{E}[J(h^*_z)] = j(z),$$

in other words, the mapping $h \mapsto \mathbb{E}[J(h)]$ on $\mathcal{D}(z)$ attains its infimum in h^*_z.

This would be immediate if we could show that for some topology $\mathcal{D}(z)$ were compact and $h \mapsto \mathbb{E}[J(h)]$ continuous. This does not work, however, we can show that $\mathcal{D}(z)$ is closed and convex in L^0 and the function is convex and lower semicontinuous with respect to the topology of convergence in probability.

One of the key properties of compactness is that for a given sequence we can extract a convergent subsequence. In problems with convexity, one often works with convex combinations. For any sequence $(a_n)_{n\in\mathbb{N}}$ in some real vector space, e.g. L^0, denote by

$$\begin{aligned}&\operatorname{conv}((a_k)_{k\geq n})\\ &:= \left\{ \sum_{k=n}^{\infty} \lambda_k a_k \,\Big|\, \forall k \geq n : \lambda_k \geq 0,\ \lambda_k \neq 0 \text{ for finitely many } k \geq n,\ \sum_{k=n}^{\infty} \lambda_k = 1 \right\}\end{aligned}$$

the convex hull spanned by the subsequence $(a_k)_{k\geq n}$.

LEMMA 3.8 (*Komlós*). *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of positive random variables in some probability space, $X_n \geq 0$ for any $n \in \mathbb{N}$. Then there is a sequence $(\tilde{X}_n)_{n\in\mathbb{N}}$, such that for any $n \in \mathbb{N}$, $\tilde{X}_n \in conv((X_k)_{k\geq n})$, and a random variable X taking values in $[0,\infty]$ such that*

$$\tilde{X}_n \xrightarrow[n\to\infty]{P\text{-a.s.}} X$$

converges P-a.s.

PROOF. See, e.g., [**6**]. $\square$

PROPOSITION 3.9 (*Topological properties of $\mathcal{D}(z)$*). *For any $z > 0$, $\mathcal{D}(z) \subset L^0_+$ is a closed and convex subset.*

PROOF. Note that $\mathcal{D}(z)$ is convex, since $\mathcal{Z}(z)$ is convex. We just need to argue that $\mathcal{D}(z)$ is closed with respect to the topology of convergence in probability. Let $(h_n)_{n\in\mathbb{N}} \subset \mathcal{D}(z)$ such that

$$h_n \xrightarrow[n\to\infty]{L^0} h$$

for some $h \in L^0$. First we notice that $h \geq 0$, since $h_n \geq 0$ for any $n \in \mathbb{N}$; so that we just need to prove that there is some $Z \in \mathcal{Z}(z)$ such that $h \leq Z_T$.

Proving the existence of Z uses the above Komlós-type result: for any $n \in \mathbb{N}$ choose some $Z^n \in \mathcal{Z}(z)$ such that $h_n \leq Z^n_T$. By Lemma 3.8 we can construct out of these sequences new sequences $(\tilde{h}_n)_{n\in\mathbb{N}}$ and $((\tilde{Z}^n_r)_{n\in\mathbb{N}})_{r\in\mathbb{Q}\cap[0,T]}$, where $\tilde{h}_n \in \operatorname{conv}((h_k)_{k\geq n})$ and $\tilde{Z}^n_r \in \operatorname{conv}((Z^n_r)_{k\geq n})$ for any $n \in \mathbb{N}$ and $r \in \mathbb{Q}\cap[0,T]$, such that the convergence to the respective limit random variables

$$\tilde{h}_n \xrightarrow[n\to\infty]{P\text{-}a.s.} h_\infty \quad \textit{and}$$

$$\forall r \in \mathbb{Q}\cap[0,T] : \tilde{Z}^n_r \xrightarrow[n\to\infty]{P\text{-}a.s.} Z^\infty_r$$

holds simultaneously, this can be done using a diagonal argument.

Claim. $h_\infty = h$ P-a.s.

PROOF. We know that $h_n \xrightarrow[n\to\infty]{L^0} h$ and $\tilde{h}_n \in conv\,((h_k)_{k\geq n})$ for any $n \in \mathbb{N}$, so $\tilde{h}_n \xrightarrow[n\to\infty]{P\text{-}a.s.} h$. Hence $\tilde{h}_n \xrightarrow[n\to\infty]{L^0} h_\infty$ and $\tilde{h}_n \xrightarrow[n\to\infty]{L^0} h$, so $h_\infty = h$ P-a.s. □

Claim. $h \leq Z^\infty_T$ P-a.s.

PROOF. For any $n \in \mathbb{N}$, we know that $h_n \leq Z^n_T$. So $\tilde{h}_n \leq \sup_{k\geq n} Z^k_T$, hence $h_\infty \leq \liminf_{n\to\infty} Z^n_T$, and $\liminf_{n\to\infty} Z^n_T = Z^\infty_T$ P-a.s. □

It remains to show that there is some $Z \in \mathcal{Z}(z)$ so that $Z^\infty_T \leq Z_T$ P-a.s. We want to construct Z out of $\tilde{Z}^n$. For this we notice that for any $n \in \mathbb{N}$, $\tilde{Z}^n \in \mathcal{Z}(z)$, since $\mathcal{Z}(z)$ is convex; so $Z^\infty_0 = z$ and if we take rational $r \leq s$ we obtain for any $V \in v(1)$

$$\mathbb{E}[\underbrace{Z^\infty_s V_s}_{\lim\limits_{n\to\infty} \tilde{Z}^n_s V_s} \mid \mathcal{F}_r] \underbrace{\leq}_{\text{Fatou}} \liminf_{n\to\infty} \mathbb{E}[\tilde{Z}^n_s V_s \mid \mathcal{F}_r] \underbrace{\leq}_{\tilde{Z}^n\in\mathcal{Z}(z)} \liminf_{n\to\infty} \tilde{Z}^n_r V_r = Z^\infty_r V_r.$$

So, by taking $V = 1$, we see that $Z^\infty V$ is indeed a P-super-martingale on $\mathbb{Q}\cap[0,T]$. Define Z by setting

$$Z_t := \lim_{\substack{r \searrow t \\ \mathbb{Q}\cap[0,T]}} Z^\infty_r ,$$

so that $Z_T = Z^\infty_T$ and $Z_0 = Z^\infty_0 = z$. Moreover Z is an càdlàg P-supermartingale (see Dellacherie/Meyer Theorem VI.2). We still need to show that for any $V \in v(1)$, ZV is actually a P-supermartingale, so let $V \in v(1)$ and $t \geq s$, with $s \in \mathbb{Q}$ then

$$\mathbb{E}[\underbrace{Z_t V_t}_{\lim\limits_{r\searrow t} \tilde{Z}^\infty_t V_r} \mid \mathcal{F}_s] \underbrace{\leq}_{\text{Fatou}} \liminf_{r\searrow t} \mathbb{E}[\tilde{Z}^\infty_r V_r \mid \mathcal{F}_s] = Z_s V_s \,,$$

which completes the proof. □

PROPOSITION 3.10. *The mapping $h \mapsto \mathbb{E}[J(h)]$ on $\mathcal{D}(z)$ is lower semicontinuous in L^0.*

PROOF. Let $h \mapsto F(h) := \mathbb{E}[J(h)]$ and $(h_n)_{n\in\mathbb{N}} \subset \mathcal{D}(z)$ such that $h_n \xrightarrow[n\to\infty]{L^0} h$, we need to show that

$$F(h) \leq \liminf_{n\to\infty} F(h_n).$$

Decompose $J(h) = (J(h))^+ - (J(h))^-$ into its positive and negative parts and note that by Fatou's lemma we obtain

$$\mathbb{E}[(J(h))^+] \leq \liminf_{n\to\infty} \mathbb{E}[(J(h_n))^+].$$

The result follows from the following observation,

Claim. $\{(J(h))^- \mid h \in \mathcal{D}(z)\}$ is P-uniformly integrable.

PROOF. Without loss of generality we may assume that $J(\infty) = -\infty$, else we have a uniform bound $(J(h))^- \leq -J(\infty)$, since J is decreasing. Define $\varphi := (-J)^{-1} : \mathbb{R} \longrightarrow [0,\infty)$, then $\varphi \geq 0$, φ is increasing and

$$\lim_{x\to\infty} \frac{\varphi(x)}{x} = \lim_{y\to\infty} \frac{\varphi(-J(y))}{-J(y)} = \lim_{y\to\infty} \frac{y}{-J(y)} = \lim_{y\to\infty} \frac{1}{-J'(y)} \underbrace{=}_{\text{Lemma 6.1}} \infty.$$

Using that $(x)^- = \max(-x, 0)$ we obtain

$$\varphi((J(h))^-) \leq \varphi(-J(h)) + \varphi(0) = h + \varphi(0)$$

for any $h \in \mathcal{D}(z)$. Hence,

$$\sup_{h\in\mathcal{D}(z)} \mathbb{E}[\underbrace{\varphi((J(h))^-)}_{\leq h+\varphi(0)}] \leq \sup_{h\in\mathcal{D}(z)} \mathbb{E}[\underbrace{h}_{\leq Z_T,\ Z\in\mathcal{Z}(z)}] + \varphi(0) \leq \sup_{Z\in\mathcal{Z}(z)} \underbrace{\mathbb{E}[Z_T]}_{\leq z} + \varphi(0) < \infty.$$

By the Theorem of de la Vallée-Poussin we conclude that $\{(J(h))^- \mid h \in \mathcal{D}(z)\}$ is P-uniformly integrable, whence the claim. □

□

THEOREM 3.11 (*Solution of the dual problem*). *For any $z > 0$ such that $j(z) < \infty$ there is a unique $h_z^* \in \mathcal{D}(z)$ satisfying $j(z) = \mathbb{E}[J(h_z^*)]$.*

PROOF. We assume $z > 0$ with $j(z) < \infty$. To construct a solution we take any sequence $(h_n)_{n\in\mathbb{N}} \subset \mathcal{D}(z)$ such that for any $n \geq 2$,

$$\infty > \mathbb{E}[J(h_{n-1})] > \mathbb{E}[J(h_n)],$$

and $\mathbb{E}[J(h_n)] \leq j(z) + \frac{1}{n}$ for $n \geq 1$. Note that $h_n \geq 0$, so by Komlos Lemma we can get a sequence $(\tilde{h}_n)_{n\in\mathbb{N}}$, where $\tilde{h}_n \in \operatorname{conv}((h_k)_{k\geq n})$ and $\tilde{h}_n \xrightarrow[n\to\infty]{P\text{-a.s.}} h$. So $\tilde{h}_n \xrightarrow[n\to\infty]{L^0} h$, hence $h \in \mathcal{D}(z)$, since $\mathcal{D}(z)$ is closed in L^0 according to Lemma 3.9. Moreover h has values in $[0,\infty)$, since it can be dominated by some Z_T with $Z \in \mathcal{Z}(z)$.

Moreover, since a) $(\tilde{h}_n)_{n\in\mathbb{N}} \subset \mathcal{D}(z)$, b) J is convex and c) $(\mathbb{E}[J(h_n)])_{n\in\mathbb{N}})$ is decreasing, we obtain

$$j(z) \overset{a)}{\leq} \mathbb{E}[J(\tilde{h}_n)] \overset{b)}{\leq} \sup_{k\geq n} \mathbb{E}[J(h_k)] \overset{c)}{=} \mathbb{E}[J(h_n)] \searrow j(z).$$

So,

$$j(z) \overset{\text{by def}}{\leq} \mathbb{E}[J(h)] \overset{\text{by lower semi-continuity}}{\leq} \liminf_{n\to\infty} \mathbb{E}[J(\tilde{h}_n)] \leq \liminf_{n\to\infty} \mathbb{E}[J(h_k)] = j(z).$$

Finally set $h := h_z^*$ to be the optimizer. □

COROLLARY 3.12. *The mapping j is decreasing and strictly convex on $\{j < \infty\}$, and continuous in the interior of $\{j < \infty\}$.*

PROOF. To show convexity take $z, z' \in \{j < \infty\}$ such that $z \neq z'$. We know that $h_z^* \neq h_{z'}^*$. Take some $t \in (0,1)$, set $w := tz + (1-t)z'$ and $\bar{h} := th_z^* + (1-t)h_{z'}^*$. Furthermore $\bar{h} \in \mathcal{D}(w)$ and therefore

$$j(w) \leq \mathbb{E}[J(\bar{h})] \underbrace{<}_{J(\bar{h}) < tJ(h_z^*) + (1-t)J(h_{z'}^*)} tj(z) + (1-t)j(z'),$$

which gives us convexity of j and also $w \in \{j < \infty\}$, since the right hand side above is finite.

Moreover, any convex function is continuous in the interior of the set on which it is finite *(see Rockafellar Theorem 10.1)*.

To show that j is decreasing, it suffices to observe that $\mathcal{D}(z) = z\mathcal{D}(1) \subseteq \mathcal{D}(z')$ for $z \leq z'$. Since $j(z) = \inf_{h \in \mathcal{D}(z)} \mathbb{E}[J(h)]$, that means that for $z' \geq z$, we are taking the infimum over a bigger set, so $j(z)$ is decreasing. □

3.4. From dual to primal problem: idea and motivation. Fix $x, z > 0$ and let $f \in \mathcal{C}(x)$ and $h \in \mathcal{D}(z)$. By definition of J we first obtain

$$U(f) \leq J(h) + fh.$$

Let $V \in v(x)$ and $Z \in \mathcal{Z}(z)$ such that $f \leq V_T$ and $h \leq Z_T$, then ZV is a P-supermartingale and

$$\mathbb{E}\left[fh\right] \leq \mathbb{E}\left[V_T Z_T\right] \leq xz.$$

So,

$$u(x) \leq j(z) + xz.$$

Each side provides a bound on the other side. Let us try to obtain equality everywhere. Note that by Lemma 3.4 we can write

$$J(y) = U(I(y)) - yI(y)\,,$$

so to get equality in the second inequality we need

$$\mathbb{E}\left[hI(h)\right] = xz\,.$$

Suppose we can achieve this for some $h \in \mathcal{D}(z)$, then for those $h \in \mathcal{D}(z)$ we obtain

$$\mathbb{E}[U(I(h))] = \mathbb{E}[J(h)] + \mathbb{E}[hI(h)] = \mathbb{E}[J(h)] + xz \geq j(z) + xz \geq \inf_{z>0}(j(z) + xz) \geq u(x),$$

Note that if we take $h = h_z^*$, then

$$\mathbb{E}[J(h_z^*)] + xz = j(z) + xz,$$

if we take $z_x > 0$ so that $j'(z_x) = -x$, i.e. the minimizer of the mapping $z \mapsto j(z) + zx$, then

$$j(z_x) + xz_x = \inf_{z>0}(j(z) + xz).$$

Finally, if $I(h_z^*) \in \mathcal{C}(x)$ we get

$$u(x) \geq \mathbb{E}[U(I(h_z^*))]$$

by definition. So,

$$\mathbb{E}\left[U(I(h_z^*))\right] = \mathbb{E}\left[J(h_z^*)\right] + \mathbb{E}\left[h_z^* I(h_z^*)\right] = \mathbb{E}\left[J(h_z^*)\right] + xz = j(z) + xz \geq u(x) \geq \mathbb{E}\left[U(I(h_z^*))\right],$$

that is we get equality everywhere, which means that $I(h_z^*)$ must be optimal for the primal problem.

REMARK 3.13. The primal problem is a resource allocation problem, i.e. how to optimally choose a portfolio to maximize expected utility with respect to the physical measure. It turns out that the optimal portfolio $\hat{X}$ is chosen such that $U'(\hat{X})$ is a multiple of a Radon-Nikodym derivative of an equivalent martingale measure. High positions in the optimal portfolio correspond via U' to relatively low values of the Radon-Nikodym derivative, i.e. prices of Arrow-Debreu assets which are relative to its probability small. In other words: optimal investments avoid states, where the Arrow-Debreu prices are high in comparison to the probability of the state.

We can also interpret the dual problem in a similar way: the dual problem minimizes $h \mapsto E[U(I(h)) + xz - hI(h)]$. If we interpret $I(h)$ as option payoff at time T, then minimization of the previous functional translates into minimization of expected utility of the payoff $I(h)$ for the buyer, where the price of the option should be equal to x. The market acts as "seller" of this option with premium x and minimal expected utility for the buyer. One can see the option payoff as functional, which translates levels of marginal utility proportional to Arrow-Debreu prices into capital according to $I = (U')^{-1}$. Note that $xz - hI(h)$ should also be seen as additional utility of the position: if $E[xz - hI(h)] < 0$, then one can reduce the position further by increasing z, so necessarily the price of the option will be at most x. This is a reasonable interpretation since buying an optimal portfolio always needs a selling counterpart.

By reverse engineering we obtain a recipe for solving the primal problem,

(1) Start with $x > 0$ and define $z_x > 0$ via $-j'(z) = x$.
(2) Solve the dual problem for $z_x > 0$ to get the dual optimizer $h^*_{z_x} \in \mathcal{D}(z)$ and define $f^*_x := I(h^*_{z_x})$.
(3) Show that $\mathbb{E}[h^*_{z_x} I(h^*_{z_x})] = xz_x$.
(4) Show that $f^*_x \in \mathcal{C}(x)$.

If we can achieve this, then the computations above show that $\mathbb{E}[U(f^*_x)] = u(x)$, i.e. f^*_x solves the primal utility maximization problem. In addition we also obtain that

$$u(x) = \inf_{z>0}(j(z) + zx) \text{ and } j(z) = \sup_{x>0}(u(x) - zx),$$

i.e. the conjugacy relation extends is the original conjugacy relation between U and J.

3.5. Auxiliary results. We first want to define $z_x > 0$ given $x > 0$ via

$$-j'(z) = x,$$

so we need to study j, which in turn is linked to h^*_z.

LEMMA 3.14. *The mapping* $(0, \infty) \longrightarrow L^0_+$, $z \mapsto h^*_z$ *is continuous in the interior of* $\{j < \infty\}$.

PROOF. Suppose by contradiction that this is not the case, i.e. there is some $z \in (\{j < \infty\})^\circ$, a sequence $(z_n)_{n\in\mathbb{N}} \subset (\{j < \infty\})^\circ$ converging to z and an $\epsilon > 0$ such that

$$\limsup_{n\to\infty} P\left[|h^*_{z_n} - h^*_z| > \epsilon\right] \geq \epsilon.$$

Using Chebychev's inequality we obtain

$$P\left[h^*_z > \frac{1}{\epsilon}\right] \leq \epsilon\mathbb{E}\left[h^*_z\right] \leq \epsilon z$$

and analogously for any $n \in \mathbb{N}$

$$P\left[h^*_{z_n} > \frac{1}{\epsilon}\right] \leq \epsilon\mathbb{E}\left[h^*_{z_n}\right] \leq \epsilon z.$$

So, by shrinking ϵ if necessary, we may assume that

$$\limsup_{n\to\infty} P\left[|h^*_{z_n} - h^*_z| > \epsilon,\ h^*_z + h^*_{z_n} \leq \frac{1}{\epsilon}\right] \geq \epsilon.$$

Define the sequence $(h_n)_{n\in\mathbb{N}} \subset L^0_+$ by

$$h_n := \frac{1}{2}(h^*_{z_n} + h^*_z).$$

Since J is convex we obtain that

$$J(h_n) \leq \frac{1}{2}(J(h^*_{z_n}) + J(h^*_z)).$$

Moreover, since $\limsup_{n\to\infty} P\left[|h^*_{z_n} - h^*_z| > \epsilon,\ h^*_z + h^*_{z_n} \geq \frac{1}{\epsilon}\right] \geq \epsilon$ and J is strictly convex there is some $\eta > 0$ so that

$$\limsup_{n\to\infty} P\left[J(h_n) \leq \frac{1}{2}(J(h^*_{z_n}) + J(h^*_z)) - \eta\right] \geq \eta > 0.$$

For any $n \in \mathbb{N}$ we get

$$\begin{aligned}\mathbb{E}[J(h_n)] &\leq \mathbb{E}\left[\big(\frac{1}{2}(J(h^*_{z_n}) + J(h^*_z)) - \eta\big)\mathbb{1}_{\{J(h_n)\leq\frac{1}{2}(J(h^*_{z_n})+J(h^*_z))-\eta\}}\right] + \\ &\quad + \mathbb{E}\left[J(h_n)\mathbb{1}_{\{J(h_n)>\frac{1}{2}(J(h^*_{z_n})+J(h^*_z))-\eta\}}\right] \\ &\underbrace{\leq}_{J(h_n)\leq\frac{1}{2}(J(h^*_{z_n})+J(h^*_z))} \frac{1}{2}\mathbb{E}\left[J(h^*_{z_n}) + J(h^*_z)\right] - \eta P\left[J(h_n) \leq \frac{1}{2}(J(h^*_{z_n}) + J(h^*_z)) - \eta\right].\end{aligned}$$

Hence, using the continuity of the map j on $\{j < \infty\}$ due to Corollary 3.12, we obtain

$$\begin{aligned}\liminf_{n\to\infty} \mathbb{E}[J(h_n)] &\leq \frac{1}{2}\Big(\underbrace{\liminf_{n\to\infty} j(z_n)}_{=j(z)} + j(z)\Big) - \eta \limsup_{n\to\infty} \underbrace{P[J(h_n) \leq \frac{1}{2}(J(h^*_{z_n}) + J(h^*_z)) - \eta]}_{\geq \eta} \\ &\leq j(z) - \eta^2 .\end{aligned}$$

Since $(h_n)_{n\in\mathbb{N}} \subset L^0_+$ we can use Lemma 3.8 to obtain a sequence $(\tilde{h}_n)_{n\in\mathbb{N}} \in \mathrm{conv}((h_k)_{k\geq n})$ such that $\tilde{h}_n \xrightarrow[n\to\infty]{P\text{-a.s.}} h$, and $h \in L^0_+$.

Claim. For any $\delta > 0$ we have $h \in \mathcal{D}(z + \delta)$.

PROOF. Let $\delta > 0$. Since $z_n \xrightarrow[n\to\infty]{} z$, there is some $N \in \mathbb{N}$ such that for any $n \geq N$, $z_n \leq z + \delta$. Hence, for any $n \geq N$ we have $h_n \in \mathcal{D}(z + \delta)$ and by convexity of $\mathcal{D}(z + \delta)$ we also obtain $\tilde{h}_n \in \mathcal{D}(z + \delta)$. So $h \in \mathcal{D}(z + \delta)$, by closedness of $\mathcal{D}(z + \delta)$. □

So, for any $\delta > 0$,

$$j(z+\delta) \underbrace{\leq}_{h\in\mathcal{D}(z+\delta)} \mathbb{E}[J(h)] \underbrace{\leq}_{\text{Prop. 3.10}} \liminf_{n\to\infty} \mathbb{E}[J(\tilde{h}_n)] \underbrace{\leq}_{J \text{ is convex}} \liminf_{n\to\infty} \mathbb{E}[J(h_n)] \leq j(z) - \eta^2.$$

On the other hand, by continuity of j,

$$j(z) - \eta^2 \geq j(z + \delta) \xrightarrow[\delta\searrow 0]{} j(z),$$

which is a contradiction, since $\eta > 0$. □

As mentioned before, we will need to assume an extra condition on U in order to ensure solvability of both the primal and dual problem.

DEFINITION 3.15. *U has* reasonable asymptotic elasticity (RAE) *if*

$$AE(U) := AE_{+\infty}(U) := \limsup_{x\to\infty} \frac{xU'(x)}{U(x)} < 1.$$

REMARK 3.16. The intuition behind this definition is the following, $U'(x)$ can be interpreted as the marginal increase in utility, since $U'(x) \approx U(x+1) - U(x)$ and $U(x+1) - U(x)$ measures the increase in utility when wealth increases from x to $x+1$. Similarly,

$$\frac{U(x+1)}{x} \approx \frac{1}{x} \sum_{j=2}^{\lfloor x+1 \rfloor} \big(U(j) - U(j-1)\big) + \underbrace{\frac{1}{x} U(1)}_{\mathcal{O}(\frac{1}{x})}$$

measures the average increase of utility when wealth increases successively from 1 to $x+1$. Moreover, U is concave, so

$$U(j) - U(j-1) \approx U'(j-1) \geq U'(x),$$

so that

$$\frac{U(x+1)}{x} \geq U'(x).$$

In fact, one can prove that $AE(U) \leq 1$. Having equality, i.e. $AE(U) = 1$, would mean that for large x, the marginal utility and the average increase of utility are almost the same, so U would behave asymptotically linear, and this is unreasonable.

EXAMPLE 3.17. If $U(x) = \log(x)$ then

$$AE(U) = \limsup_{x\to\infty} \frac{xU'(x)}{U(x)} = \limsup_{x\to\infty} \frac{x\frac{1}{x}}{\log(x)} = 0,$$

so log satisfies (RAE).

If $U(x) = \frac{1}{\gamma} x^\gamma$ with $\gamma \in (-\infty, 1) \setminus \{0\}$ then

$$AE(U) = \limsup_{x\to\infty} \frac{xU'(x)}{U(x)} = \limsup_{x\to\infty} \frac{xx^{\gamma-1}}{\gamma^{-1}x^\gamma} = \gamma < 1,$$

so U satisfies (RAE).

Finally, if $U(x) = \frac{x}{\log(x)}$, then

$$AE(U) = \limsup_{x\to\infty} \frac{xU'(x)}{U(x)} = \limsup_{x\to\infty} \big(1 - \frac{1}{\log x}\big) = 1,$$

so this U does not satisfy (RAE). Note also that this U behaves asymptotically linear.

LEMMA 3.18. *If U satisfies (RAE) then there is some $y_0 > 0$ and a constant $C > 0$ such that*

$$-J'(y) \leq C\frac{J(y)}{y}$$

for any $y \in (0, y_0)$.

PROOF. For any $y > 0$, we can write by Lemma 3.4

$$J(y) = U(I(y)) - yI(y),$$

where $I = (U')^{-1} = J'$. Since U has (RAE) we can find some $x_0 > 0$ and a $\beta < 1$ so that

$$xU'(x) \leq \beta U(x),$$

for any $x \geq x_0$.

Moreover, since both, I and U' are decreasing, taking $x = I(y)$ gives $y \leq y_0 := U'(x_0)$. So

$$I(y)y \leq \beta U(I(y)) = \beta J(y) + \beta y I(y)$$

So,

$$-J'(y) = I(y) \le \frac{\beta}{1-\beta}\frac{J(y)}{y}$$

for any $0 < y \le y_0$. $\square$

Lemma 3.19. *If U has (RAE) then the mapping $(0,\infty) \longrightarrow \mathbb{R}$, $z \mapsto \mathbb{E}[h_z^* I(h_z^*)]$ is continuous in the interior of $\{j < \infty\}$.*

Proof. By Lemma 3.14 we know that the mapping $(0,\infty) \longrightarrow L^0_+$, $z \mapsto h_z^*$ is continuous in the interior of $\{j < \infty\}$. Hence, $z \mapsto h_z^* I(h_z^*)$ is continuous on $(\{j < \infty\})^\circ$, since I is continuous. Let $z \in (\{j < \infty\})^\circ$ and $(z_n)_{n\in\mathbb{N}} \subset (\{j < \infty\})^\circ$ such that $z_n \xrightarrow[n\to\infty]{} z$. Then,

$$\lim_{n\to\infty} \mathbb{E}[h_{z_n}^* I(h_{z_n}^*)] = \mathbb{E}[h_z^* I(h_z^*)]$$

holds, if we can show that $(h_{z_n}^* I(h_{z_n}^*))_{n\in\mathbb{N}}$ is uniformly integrable.

We will argue that $(h_{z_n}^* I(h_{z_n}^*)\mathbb{1}_{\{h_{z_n}^* \ge y_0\}})_{n\in\mathbb{N}}$ and $(h_{z_n}^* I(h_{z_n}^*)\mathbb{1}_{\{h_{z_n}^* < y_0\}})_{n\in\mathbb{N}}$ are uniformly integrable if y_0 is suitably chosen.

Claim. $(h_{z_n}^* I(h_{z_n}^*)\mathbb{1}_{\{h_{z_n}^* \ge y_0\}})_{n\in\mathbb{N}}$ is uniformly integrable for any $y_0 > 0$.

Proof. Let $z' > 0$ such that $|z_n| \le z'$ for any $n \in \mathbb{N}$. Then $(h_{z_n}^*)_{n\in\mathbb{N}} \subset \mathcal{D}(z')$. By Lemma 3.4 we obtain for any $n \in \mathbb{N}$

$$h_{z_n}^* I(h_{z_n}^*)\mathbb{1}_{\{h_{z_n}^* \ge y_0\}} = (U(I(h_{z_n}^*)) - J(h_{z_n}^*))\mathbb{1}_{\{h_{z_n}^* \ge y_0\}} \le U(I(y_0)) + (J(h_{z_n}^*))^-.$$

As in the proof of Proposition 3.10 we can show that $\{(J(h))^- | h \in \mathcal{D}(z')\}$ is uniformly integrable, which gives the result. $\square$

Claim. $(h_{z_n}^* I(h_{z_n}^*)\mathbb{1}_{\{h_{z_n}^* < y_0\}})_{n\in\mathbb{N}}$ is uniformly integrable for a suitably chosen $y_0 > 0$.

Proof. Since U has (RAE) there is $y_0 > 0$ and $C > 0$ such that

$$yI(y) = -yJ'(y) \le CJ(y)$$

for any $0 < y \le y_0$. So

$$\underbrace{h_{z_n}^* I(h_{z_n}^*)\mathbb{1}_{\{h_{z_n}^* < y_0\}}}_{\ge 0} \le C|J(h_{z_n}^*)|.$$

Moreover, $((J(h_{z_n}^*))^-)_{n\in\mathbb{N}}$ is uniformly integrable, as argued above. Hence it is enough to show that $((J(h_{z_n}^*))^+)_{n\in\mathbb{N}}$ is uniformly integrable to conclude that $(|J(h_{z_n}^*)|)_{n\in\mathbb{N}}$ and hence also $(h_{z_n}^* I(h_{z_n}^*)\mathbb{1}_{\{h_{z_n}^* < y_0\}})_{n\in\mathbb{N}}$ are uniformly integrable.

To show this note that $h_{z_n}^* \xrightarrow[n\to\infty]{L^0} h_z^*$ so that

$$J(h_{z_n}^*) \xrightarrow[n\to\infty]{L^0} J(h_z^*),$$

and therefore also for the positive parts

$$(J(h_{z_n}^*))^+ \xrightarrow[n\to\infty]{L^0} (J(h_z^*))^+.$$

Moreover, since j is continuous on $\{j < \infty\}^\circ$ we obtain

$$\mathbb{E}[J(h_{z_n}^*)] = j(z_n) \xrightarrow[n\to\infty]{} j(z) = \mathbb{E}[J(h_z^*)],$$

which implies that

$$\mathbb{E}[(J(h_{z_n}^*))^+] \xrightarrow[n\to\infty]{} \mathbb{E}[(J(h_z^*))^+].$$

By Scheffé's theorem we conclude that

$$(J(h_{z_n}^*))^+ \xrightarrow[n\to\infty]{L^1} (J(h_z^*))^+,$$

which implies that $((J(h^*_{z_n}))^+)_{n\in\mathbb{N}}$ is uniformly integrable. □

Both claims together with $y_0 > 0$ as in the second claim give the desired result. □

REMARK 3.20. A slight variation of above arguments also gives

$$\lim_{n\to\infty} \mathbb{E}[h^*_{z_n} I(\mu_n h^*_{z_n})] = \mathbb{E}[h^*_z I(h^*_z)]$$

if $z_n \xrightarrow[n\to\infty]{} z$ in the interior of $\{z < \infty\}$ and $\mu_n \nearrow_{n\to\infty} 1$.

The key result of this section is the next theorem, for this recall the standing assumptions: $\mathbb{P} \neq \emptyset$, U is a utility function satisfying the Inada conditions and $u(x_0) < \infty$ for some $x_0 > 0$.

THEOREM 3.21. *For any $z > 0$ we have*

$$j(z) = \sup_{x>0}(u(x) - xz),$$

in particular there is some $z_0 > 0$ such that for any $z \geq z_0$, $j(z) < \infty$. Moreover, if U has (RAE) then $j(z) < \infty$ for all $z > 0$.

PROOF. The key idea is the following, by definition of J we have

$$J(y) = \sup_{x>0}(U(x) - xy),$$

so it seem plausible to try to prove that for any $h \in \mathcal{D}(z)$

$$\mathbb{E}[J(h)] = \sup_{f\in L^\infty_+} \mathbb{E}[U(f) - fh].$$

This would yield

$$j(z) = \inf_{h\in\mathcal{D}(z)} \mathbb{E}[J(h)] = \inf_{h\in\mathcal{D}(z)} \sup_{f\in L^\infty_+} \mathbb{E}[U(f) - fh].$$

We want to interchange inf and sup, this requires a minimax theorem. Such results need compactness for at least one of the sets over which we optimize. We use von Neumann's minimax theorem *(see Aubin Theorem 2.7.1)*. Consider L^∞ as the dual space of L^1 with the weak*-topology $\sigma(L^\infty, L^1)$. Fix $n \in \mathbb{N}$, by the Tychonov-Alaoglu theorem we know that

$$B_n := \{f \in L^\infty_+ | f \leq n\} = \{f \in L^\infty | f \leq n\} \cap L^\infty_+$$

is weak*-compact. Moreover, $\mathcal{D}(z)$ is a convex subset of L^1. So the mapping

$$B_n \times \mathcal{D}(z) \longrightarrow \mathbb{R},\ (f,h) \mapsto \mathbb{E}[U(f) - fh],$$

satisfies following conditions,

$$\begin{cases} B_n \text{ is a compact convex subset,} \\ \text{for all } h \in \mathcal{D}(z),\ f \mapsto \mathbb{E}[U(f) - fh] \text{ is concave like } U \end{cases}$$

and

$$\begin{cases} \mathcal{D}(z) \text{ is a convex subset,} \\ \text{for all } f \in B_n,\ h \mapsto \mathbb{E}[U(f) - fh] \text{ is convex, since linear} \end{cases}$$

so von Neumann's minimax theorem gives

$$\sup_{f\in B_n} \inf_{h\in\mathcal{D}(z)} \mathbb{E}[U(f) - fh] = \inf_{h\in\mathcal{D}(z)} \sup_{f\in B_n} \mathbb{E}[U(f) - fh]. \tag{3.1}$$

We want to let $n \to \infty$, so we will first study the right hand side of the equation and show that it is equal to $j(z)$.

Claim. $\lim_{n\to\infty} \inf_{h\in\mathcal{D}(z)} \sup_{f\in B_n} \mathbb{E}[U(f) - fh] = j(z)$.

PROOF. Define

$$J^n(y) := \sup_{0<x\leq n} (U(x) - xy).$$

Note that $J^n \leq J$ and $\sup_{f\in B_n} \mathbb{E}[U(f) - fh] = \mathbb{E}[J^n(h)]$, so

$$\inf_{h\in\mathcal{D}(z)} \mathbb{E}[J^n(h)] \leq j(z).$$

Define

$$j_n(z) := \inf_{h\in\mathcal{D}(z)} \mathbb{E}[J^n(h)].$$

Take a sequence $(h_n)_{n\in\mathbb{N}} \subset \mathcal{D}(z)$ such that

$$\mathbb{E}[J^n(h_n)] \xrightarrow[n\to\infty]{} \lim_{n\to\infty} j_n(z)$$

and use Lemma 3.8 to get a sequence $(\tilde{h}_n)_{n\in\mathbb{N}} \subset \mathcal{D}(z)$ such that $\tilde{h}_n \in \mathrm{conv}((h_k)_{k\geq n})$ and $\tilde{h}_n \xrightarrow[n\to\infty]{P\text{-a.s.}} h$, with $h \in \mathcal{D}(z)$. Note that for any $y \geq I(n)$ we have

$$J^n(y) = J(y),$$

hence also for any $y \geq I(1)(\geq I(n))$ and for any $n \in \mathbb{N}$, $y \mapsto J^n(y)$ is decreasing. As in the proof of Proposition 3.10 we obtain that $((J(\tilde{h}_n))^-)_{n\in\mathbb{N}}$ is uniformly integrable and

$$\lim_{n\to\infty} j_n(z) = \lim_{n\to\infty} \mathbb{E}[J^n(h_n)] \underbrace{\geq}_{J^n \text{ is convex}} \liminf_{n\to\infty} \mathbb{E}[J^n(\tilde{h}_n)] \underbrace{\geq}_{Fatou} \mathbb{E}[\liminf_{n\to\infty} J^n(\tilde{h}_n)] = \mathbb{E}[J(h)] \geq j(z).$$

So

$$\lim_{n\to\infty} \inf_{h\in\mathcal{D}(z)} \sup_{f\in B_n} \mathbb{E}[U(f) - fh] = j(z).$$

□

Next we want to show that the left hand side of equation (3.1) equals $\sup_{x>0}(u(x)-zx)$.

Claim. $\lim_{n\to\infty} \sup_{f\in B_n} \inf_{h\in\mathcal{D}(z)} \mathbb{E}[U(f) - fh] \geq \sup_{x>0}(u(x) - zx)$.

PROOF. Let $f \in \mathcal{C}(x) \cap B_n$ and $x > 0$, then we obtain

$$\sup_{h\in\mathcal{D}(z)} \mathbb{E}[fh] \leq xz,$$

so

$$\mathbb{E}[U(f)] - xz \leq \inf_{h\in\mathcal{D}(z)} \mathbb{E}[U(f) - fh].$$

Taking the sup on the left over $f \in \mathcal{C}(x) \cap B_n$ and the sup over $f \in B_n$ on the right we obtain for any $n \in \mathbb{N}$

$$\sup_{f\in\mathcal{C}(x)\cap B_n} \mathbb{E}[U(f)] - xz \leq \sup_{f\in B_n} \inf_{h\in\mathcal{D}(z)} \mathbb{E}[U(f) - fh].$$

Now let $n \to \infty$ to get

$$\underbrace{\lim_{n\to\infty} \sup_{f\in\mathcal{C}(x)\cap B_n} \mathbb{E}[U(f)]}_{= \sup_{f\in\mathcal{C}(x)} \mathbb{E}[U(f)] = u(x)} - xz \leq \lim_{n\to\infty} \sup_{f\in B_n} \inf_{h\in\mathcal{D}(z)} \mathbb{E}[U(f) - fh].$$

This holds for any $x > 0$, hence

$$\sup_{x>0}(u(x) - xz) \leq \sup_{f\in L^\infty_+} \inf_{h\in\mathcal{D}(z)} \mathbb{E}[U(f) - fh]$$

□

Claim. $\lim_{n\to\infty} \sup_{f\in B_n} \inf_{h\in\mathcal{D}(z)} \mathbb{E}[U(f) - fh] \le \sup_{x>0}(u(x) - zx)$.

PROOF. Let $n \in \mathbb{N}$ and $f \in B_n$. Define

$$x^* := \inf\{x > 0 | f \in \mathcal{C}(x)\} \ (< \infty).$$

Without loss of generality we may assume that $x^* > 0$, else $f \equiv 0$. Let $\epsilon > 0$, then $f \in \mathcal{C}(x^* + \epsilon)$, and by definition of x^* we know that $f \notin \mathcal{C}(x^* - \epsilon)$, hence

$$\sup_{h\in\mathcal{D}(z)} \mathbb{E}[fh] > (x^* - \epsilon)z,$$

and $f \in \mathcal{C}(x^* + \epsilon)$. So

$$\inf_{h\in\mathcal{D}(z)} \mathbb{E}[U(f) - fh] < \mathbb{E}[U(f)] - (x^* - \epsilon)z \le u(x^* + \epsilon) - (x^* - \epsilon)z =$$

$$u(x^* + \epsilon) - (x^* + \epsilon)z + 2\epsilon z \le \sup_{x>0}(u(x) - xz) + 2\epsilon z.$$

Now let $\epsilon \searrow 0$ to get the claim. $\square$

Next we have to argue the existence of some $z_0 > 0$ such that $j(z) < \infty$ for any $z \ge z_0$. For this, we note that u is concave and increasing. By assumption $u(x_0) < \infty$ for some $x_0 > 0$ and hence for all $x > 0$. Moreover, j is decreasing and so $j(z) < \infty$ for all $z \ge z_0$, unless $j \equiv \infty$, so we must argue that this is not the case.

Claim. $j \neq \infty$

PROOF. Take a sequence $z_n \xrightarrow[n\to\infty]{} \infty$ and find another sequence $(x_n)_{n\in\mathbb{N}}$ such that

$$u(x_n) - x_n z_n \ge \min\{n, j(z_n) - \frac{1}{n}\}.$$

Now u is increasing and concave, so the maximizer x_z^* of the map $x \mapsto u(x) - xz$ is decreasing in z, so we can choose the sequence $(x_n)_{n\in\mathbb{N}}$ to be bounded. Let $(x_{n_k})_{k\in\mathbb{N}}$ be a convergent subsequence, i.e. $x_{n_k} \xrightarrow[n\to\infty]{} x_\infty$. By continuity of u we obtain

$$u(x_\infty) = \lim_{k\to\infty} u(x_{n_k}) \ge \limsup_{k\to\infty} \left(x_{n_k} z_{n_k} + \min\{n_k, j(z_{n_k}) - \frac{1}{n_k}\}\right).$$

Suppose by contradiction that $j \equiv \infty$, then

$$u(x_\infty) \ge \limsup_{k\to\infty} \left(x_{n_k} z_{n_k} + n_k\right) = \infty,$$

but $u(x_\infty) < \infty$. $\square$

Finally we argue that if U has (RAE) then $j(z) < \infty$ for all $z > 0$. So suppose that $AE(U) < 1$, then

$$J(\mu y) \le C_\mu J(y)$$

for $\mu \in (0, 1]$ and for $y \le y_0$. Let $z_0 > 0$ so that $j(z_0) < \infty$. Since j is decreasing we only need to control j to the left of z_0, so take some $z < z_0$, set $\mu := \frac{z}{z_0}$ and note that $\mu h_{z_0}^* \in \mathcal{D}(z)$. Then, using the above inequality and the fact that J is decreasing, we obtain

$$J(\mu h_{z_0}^*) \le C_\mu J(h_{z_0}^*) \mathbb{1}_{\{h_{z_0}^* \le y_0\}} + J(\mu y_0),$$

so that

$$j(z) \le \mathbb{E}[J(\mu h_{z_0}^*)] \le C_\mu \underbrace{\mathbb{E}[J(h_{z_0}^*)]}_{=j(z_0)} + J(\mu y_0) = C_\mu j(z_0) + J(\mu y_0) < \infty,$$

i.e. $j(z) < \infty$. $\square$

LEMMA 3.22 (*Smoothness properties of j*). *If U has (RAE) then $j \in C^1((0,\infty))$, j' is increasing and for any $z > 0$ we have*

$$-zj'(z) = \mathbb{E}[h_z^* I(h_z^*)].$$

PROOF. Suppose we know that $j \in C^1((0,\infty))$, then for any $z > 0$ the limit

$$-zj'(z) = \lim_{\lambda\to 1} -z\frac{j(z)-j(\lambda z)}{z-\lambda z} = \lim_{\lambda\to 1}\frac{j(z)-j(\lambda z)}{\lambda - 1}$$

exists. Consider first the right derivative

$$g_r(z) := \lim_{\lambda\searrow 1}\frac{j(z)-j(\lambda z)}{\lambda-1},$$

and suppose we can prove that $g_r(z)$ exists and

$$g_r(z) = \mathbb{E}[h_z^* I(h_z^*)],$$

then

$$-zj_r'(z) = \mathbb{E}[h_z^* I(h_z^*)],$$

where j_r' is the right derivative of j, which always exists, as j is monotonic and convex. But, $z \mapsto \mathbb{E}[h_z^* I(h_z^*)]$ is continuous on $(0,\infty)$, so j_r' is continuous and so $j \in C^1((0,\infty))$ as it is convex *(see Rockafellar Theorem 24.1)*. Also note that j is strictly convex on $(0,\infty)$, and so j' must be strictly increasing. So we only need to compute $g_r(z)$.

Let $\mu > 0$, then $\frac{1}{\mu}h_{\mu z}^* \in \mathcal{D}(z)$ and hence

$$j(z) \le \mathbb{E}\left[J\left(\frac{1}{\mu}h_{\mu z}^*\right)\right],$$

and this gives

$$\limsup_{\lambda\searrow 1}\frac{j(z)-j(\lambda z)}{\lambda-1} \le \limsup_{\lambda\searrow 1}\frac{1}{\lambda-1}\mathbb{E}\underbrace{\left[J\left(\frac{1}{\lambda}h_{\lambda z}^*\right) - J(h_{\lambda z}^*)\right]}_{\int_{\frac{1}{\lambda}h_{\lambda z}^*}^{h_{\lambda z}^*} -J'(w)dw} \le$$

$$\le \limsup_{\lambda\searrow 1}\frac{1}{\lambda-1}\mathbb{E}\left[-J'\left(\frac{1}{\lambda}h_{\lambda z}^*\right)h_{\lambda z}^*(1-\frac{1}{\lambda})\right] = \limsup_{\lambda\searrow 1}\frac{1}{\lambda}\mathbb{E}\left[h_{\lambda z}^* I\left(\frac{1}{\lambda}h_{\lambda z}^*\right)\right] = \mathbb{E}\left[h_z^* I(h_z^*)\right];$$

on the other hand note that $\mu h_z^* \in \mathcal{D}(\mu z)$, so

$$j(\mu z) \le \mathbb{E}[J(\mu h_z^*)]$$

so similarly as above we obtain

$$\liminf_{\lambda\searrow 1}\frac{j(z)-j(\lambda z)}{\lambda-1} \ge \liminf_{\lambda\searrow 1}\frac{1}{\lambda-1}\mathbb{E}\underbrace{\left[J(h_z^*) - J(\lambda h_z^*)\right]}_{\int_{h_z^*}^{\lambda h_z^*} -J'(w)dw} \ge \liminf_{\lambda\searrow 1}\mathbb{E}\left[-J'(\lambda h_z^*)h_z^*)\right] \ge$$

$$\ge \liminf_{\lambda\searrow 1}\mathbb{E}\left[h_z^* I(\lambda h_z^*)\right] = \mathbb{E}\left[h_z^* I(h_z^*)\right],$$

where the last equality in both cases is justified by Lemma 3.19 and the subsequent remark. □

REMARK 3.23. The reverse engineering recipe requires us to find for a given $x > 0$ some $z_x > 0$ such that $j'(z) = -x$. Now j' is continuous and strictly monotonic, so if z_x exists, it must be unique. To prove existence, we first need to understand the range of values of j'.

Lemma 3.24 (*Range of j'*). *$j'(\infty) := \lim_{z\to\infty} j'(z) = 0$ and if U has (RAE) then*

$$j'(0) := \lim_{z\searrow 0} j'(z) = -\infty.$$

Proof. We know that $j(z) < \infty$ for all $z \geq z_0$. The same argument as in the proof of Lemma 3.22 yields that $j \in C^1((0,\infty))$, note that this does not use (RAE). Now, $-j$ is concave and increasing by Corollary 3.12, so $j'(\infty)$ exists by monotonicity and $-j'(\infty) \geq 0$. Moreover, $-J$ is increasing and $-J' = I$ is decreasing with $I(\infty) = 0$, so for any $\epsilon > 0$ there is some $C_\epsilon > 0$ so that for all $y > 0$,

$$-J(y) =\leq \epsilon y + C_\epsilon,$$

So,

$$0 \leq -j'(\infty) = \lim_{z\to\infty} -\frac{j(z)}{z} = \lim_{z\to\infty} \sup_{h\in\mathcal{D}(z)} \mathbb{E}[-J(h)]\frac{1}{z} \leq \lim_{z\to\infty} \sup_{h\in\mathcal{D}(z)} \mathbb{E}[C_\epsilon + \epsilon h]\frac{1}{z} \leq$$

$$\leq \lim_{z\to\infty} \frac{C_\epsilon}{z} + \frac{\epsilon}{z}\underbrace{\mathbb{E}[h]}_{\leq z} \leq \epsilon,$$

finally let $\epsilon \searrow 0$ to obtain the first claim.

By Theorem 3.21 we have for all $z > 0$

$$j(z) = \sup_{x>0}(u(x) - xz),$$

and by Corollary 3.12, j is strictly convex on $\{j < \infty\}$, hence by general duality we obtain that $u \in C^1((0,\infty))$ *(see Rockafellar Theorem V 26.3)*. We already know that $j \in C^1(\{j < \infty\}^\circ)$, so u' and $-j'$ are inverse to each other.

If we now use that U has (RAE) then $\{j < \infty\}^\circ = (0,\infty)$ and so we can use the inverse relation everywhere to get

$$j'(0) = -\frac{1}{u'(\infty)}.$$

But, $AE(U) < 1$ implies that $AE(u) < 1$ *(see Kramkov/Schachermayer)* and that in turn implies $u'(\infty) = 0$, so $j'(0) = -\infty$. ☐

3.6. Solution of the primal problem. Recall the primal problem: for a given $x > 0$, find

$$u(x) = \sup_{f\in\mathcal{C}(x)} \mathbb{E}[U(f)]$$

under the assumption that $\mathbb{P} \neq \emptyset$ and U is a utility function satisfying the Inada conditions with (RAE), such that $u(x_0) < \infty$ for some $x_0 > 0$. Recall the recipe for solving the primal problem,

1. Start with $x > 0$ and define $z_x > 0$ via $-j'(z) = x$. The mapping $j' : (0,\infty) \to (0,\infty)$ is continuous, strictly decreasing and surjective, so this is always possible and $z_x \in (0,\infty)$ is unique.
2. Solve the dual problem for $z_x > 0$ to get the dual optimizer $h^*_{z_x} \in \mathcal{D}(z)$ and define $f^*_x := I(h^*_{z_x})$. We know that for any $z > 0$, $j(z) < \infty$ so we can find the solution of the dual problem for any $z > 0$.
3. Show that $\mathbb{E}[h^*_{z_x} I(h^*_{z_x})] = xz_x$. This is even true for any $z > 0$ and its corresponding h^*_z. But for later use in step 4), we even prove more.

Lemma 3.25. *For all $z > 0$ and for any $h \in \mathcal{D}(z)$ we have*

$$\mathbb{E}[h^*_z I(h^*_z)] = -zj'(z) \geq \mathbb{E}[hI(h^*_z)]$$

Proof. The first equality follows from Lemma 3.22. To show the inequality let $z > 0$ and $h \in \mathcal{D}(z)$. For any $\delta \in (0,1)$ we know that $h_\delta := \delta h + (1-\delta)h_z^* \in \mathcal{D}(z)$ by convexity of $\mathcal{D}(z)$, so that

$$0 \underbrace{\leq}_{h_z^* \text{ is optimal}} \mathbb{E}[J(h_\delta) - J(h_z^*)] \underbrace{=}_{-J'=I} \mathbb{E}\left[\int_{h_\delta}^{h_z^*} I(w)dw\right] \leq \mathbb{E}[I(h_\delta)(h_z^* - h_\delta)] = \mathbb{E}[I(h_\delta)(h_z^* - h)]\delta,$$

for the last inequality we used that on $\{h_\delta \leq h_z^*\}$ we have $I(w) \leq I(h_\delta)$ and on $\{h_\delta > h_z^*\}$ we have $I(w) \geq I(h_\delta)$ but $h_z^* - h_\delta \leq 0$. Since I is decreasing and $h_\delta \geq (1-\delta)h_z^*$ we obtain

$$\mathbb{E}[I(h_\delta)h] \leq \mathbb{E}[I(h_\delta)h_z^*] \leq \mathbb{E}[I((1-\delta)h_z^*)h_z^*].$$

Hence by Fatou's lemma as $\delta \searrow 0$

$$\mathbb{E}[I(h_z^*)h] = \mathbb{E}[\liminf_{\delta \searrow 0} I(h_\delta)h] \leq \liminf_{\delta \searrow 0} \mathbb{E}[I(h_\delta)h]$$

and by monotone convergence

$$\lim_{\delta \searrow 0} \mathbb{E}[I((1-\delta)h_z^*)h_z^*] = \mathbb{E}[I(h_z^*)h_z^*].$$

So

$$\mathbb{E}[I(h_z^*)h] \leq \mathbb{E}[I(h_z^*)h_z^*]$$

□

Recall the last step,

4. Show that $f_x^* \in \mathcal{C}(x)$.

Note that we have for any $h \in \mathcal{D}(z_x)$

$$\mathbb{E}[f_x^* h] = \mathbb{E}[hI(h_{z_x}^*)] \leq -z_x j'(z_x) = z_x x.$$

So in particular for any $h \in \mathcal{D}(1)$ we obtain

$$\mathbb{E}[f_x^* h] \leq x.$$

Hence, step 4) follows immediately from the following observation,

Proposition 3.26. *If $\mathcal{F}_0$ is trivial and $x > 0$, then for any $f \in L_+^0(\mathcal{F}_T)$: $f \in \mathcal{C}(x)$ if and only if $\mathbb{E}[fh] \leq x$ for any $h \in \mathcal{D}(1)$.*

Proof. Suppose that $f \in \mathcal{C}(x)$ and let $h \in \mathcal{D}(1)$. Find some $V \in v(x)$ so that $f \leq V_T$ and choose $Z \in \mathcal{Z}(1)$ such that $h \leq Z_T$. Then ZV is a P-supermartingale, so

$$\mathbb{E}[fh] \leq \mathbb{E}[VZ] \leq V_0 Z_0 = x.$$

Conversely suppose that $\mathbb{E}[fh] \leq x$ for any $h \in \mathcal{D}(1)$. Note that if $Q \in \mathbb{P}$ then $\dfrac{dQ}{dP} \in \mathcal{D}(1)$, so that

$$\sup_{Q \in \mathbb{P}} \mathbb{E}_Q[f] \leq \sup_{h \in \mathcal{D}(1)} \mathbb{E}[hf] \leq x.$$

Define the process Y by

$$Y_t := \operatorname*{ess\,sup}_{Q \in \mathbb{P}} \mathbb{E}_Q[f|\mathcal{F}_t], \quad 0 \leq t \leq T.$$

The process Y is a Q-supermartingale for any $Q \in \mathbb{P}$, moreover we can choose an càdlàg version of Y. By the optional decomposition theorem there is some $\vartheta \in \Theta_{adm}$ and an adapted, increasing, positive process C null at zero such that

$$Y = Y_0 + \int \vartheta dS - C,$$

where $Y_0 = \sup_{Q \in \mathbb{P}} \mathbb{E}_Q[f]$. Note that, since $Y, C \geq 0$ we get

$$\int \vartheta dS = Y + C - Y_0 \geq -Y_0 \geq -x.$$

So we actually have $\vartheta \in \Theta^x_{adm}$. Moreover,

$$f = Y_T = Y_0 + \int_0^T \vartheta_u dS_u - C_T \leq x + \int_0^T \vartheta_u dS_u$$

and $V(x, \vartheta) = x + \int \vartheta dS \in v(x)$, so $f \in \mathcal{C}(x)$. □

Putting all these pieces together we obtain

Theorem 3.27 (*Solution of the primal problem*). *Suppose that $\mathcal{F}_0$ is trivial and U has (RAE). Then, for all $x > 0$, the primal problem of maximizing expected utility from final wealth has a unique solution $f^*_x \in \mathcal{C}(x)$, which is given by*

$$f^*_x = I(h^*_{z_x}),$$

where $h^*_{z_x}$ is the unique solution of the dual problem with $z_x > 0$ defined by $-j'(z_x) = x$.

Proof. Uniqueness follows from the strict concavity of U. To show that f^*_x is a solution we note that $f^*_x \in \mathcal{C}(x)$ and

$$\mathbb{E}[U(f^*_x)] = \mathbb{E}[U(I(h^*_{z_x}))] = \mathbb{E}[J(h^*_{z_x}) + h^*_{z_x} I(h^*_{z_x})] = j(z_x) + \mathbb{E}[h^*_{z_x} I(h^*_{z_x})] \underbrace{=}_{\text{Lemma 3.22}}$$

$$= j(z_x) + x z_x \geq \inf_{z>0}(j(z) + xz) \geq u(x) = \sup_{f \in \mathcal{C}(x)} \mathbb{E}[U(f)] \geq \mathbb{E}[U(f^*_x)],$$

so f^*_x is optimal. □

Remark 3.28. From Theorem 3.21 we know that for any $z > 0$,

$$j(z) = \sup_{x>0}(u(x) - xz).$$

From the above proof we also know that for any $x > 0$,

$$u(x) \geq \mathbb{E}[U(f^*_x)] \geq \cdots \geq \inf_{z>0}(j(z) + xz) \geq u(x),$$

so that for all $x > 0$

$$u(x) = \inf_{z>0}(j(z) + xz).$$

So u and j satisfy the same conjugacy as the original U and J. Note that we could also have deduced this last expression from the properties of u and j via abstract convex analysis.

The extra condition on U, i.e. the fact that U has (RAE) is optimal in the following sense:

- If this condition is satisfied then we can solve the primal problem for any model S (for S at least satisfying the standing assumptions, i.e. $\mathbb{P} \neq \emptyset$, $u(x_0) < \infty$ for some $x_0 > 0$, etc.).
- If U does not have (RAE) then we can find a model S such that the primal problem there is not solvable, even though $\mathbb{P} = \mathbb{P}(S) \neq \emptyset$ *(see Kramkov/Schachermayer).*

The above approach is very general and, via the dual problem, also gives a lot of extra information. But how about a shorter way if we are only interested in the primal problem? At least under some (minor) extra conditions on U and on $\mathbb{P}$ one can prove directly the existence of a solution to the primal problem, using a Komlós-type argument to get a candidate for the optimizer.

CHAPTER 3

Appendix

1. Methods from convex analysis

In this chapter basic duality methods from convex analysis are discussed. We shall also apply the notions of dual normed vector spaces in finite dimensions. Let V be a real vector space with norm and real dimension $\dim V < \infty$, then we can define the *pairing*

$$\begin{aligned} \langle .,. \rangle : V \times V' &\to \mathbb{R} \\ (v,l) &\mapsto l(v) \end{aligned}$$

where V' denotes the dual vector space, i.e. the space of continuous linear functionals $l : V \to \mathbb{R}$. The dual space carries a natural dual norm namely

$$||l|| := \sup_{||v|| \leq 1} |l(v)|.$$

We obtain the following duality relations:

- If for some $v \in V$ it holds that $\langle v, l \rangle = 0$ for all $l \in V'$, then $v = 0$.
- If for some $l \in V'$ it holds that $\langle v, l \rangle = 0$ for all $v \in V$, then $l = 0$.
- There is a natural isomorphism $V \to V''$ and the norms on V and V'' coincide (with respect to the previous definition).

If V is an euclidean vector space, i.e. there is a scalar product $\langle .,. \rangle : V \times V \to \mathbb{R}$, which is symmetric and positive definite, then we can identify V' with V and every linear functional $l \in V'$ can be uniquely represented $l = \langle ., x \rangle$ for some $x \in V$.

Definition 1.1. *Let V be a finite dimensional vector space. A subset $C \subset V$ is called* convex *if for all $v_1, v_2 \in C$ also $tv_1 + (1-t)v_2 \in C$ for $t \in [0,1]$.*

Since the intersection of convex sets is convex, we can define the convex hull of any subset $M \subset V$, which is denoted by $\langle M \rangle_{conv}$. We also define the closed convex hull $\overline{\langle M \rangle_{conv}}$, which is the smallest closed, convex subset of V containing M. If M is compact the convex hull $\langle M \rangle_{conv}$ is already closed and therefore compact.

Definition 1.2. *Let C be a closed convex set, then $x \in C$ is called* extreme point *of C if for all $y, z \in C$ with $x = ty + (1-t)z$ and $t \in [0,1]$, we have either $t = 0$ or $t = 1$. This is equivalent to saying that there are no two different points $x_1 \neq x_2$ such that $x = \frac{1}{2}(x_1 + x_2)$.*

First we treat a separation theorem, which is valid in a fairly general context and known as Hahn-Banach Theorem.

Theorem 1.3. *Let C be a closed convex set in an euclidean vector space V, which does not contain the origin, i.e. $0 \notin C$. Then there exists a linear functional $\xi \in V'$ and $\alpha > 0$ such that for all $x \in C$ we have $\xi(x) \geq \alpha$.*

Proof. Let r be a radius such that the closed ball $B(r)$ intersects C. The continuous map $x \mapsto ||x||$ achieves a minimum $x_0 \neq 0$ on $B(r) \cap C$, which we denote by x_0, since $B(r) \cap C$ is compact. We certainly have for all $x \in C$ the relation $||x|| \geq ||x_0||$. By convexity we obtain that $x_0 + t(x - x_0) \in C$ for $t \in [0, 1]$ and hence

$$||x_0 + t(x - x_0)||^2 \geq ||x_0||^2.$$

This equation can be expanded for $t \in [0, 1]$,

$$||x_0||^2 + 2t \langle x_0, x - x_0 \rangle + t^2 ||(x - x_0)||^2 \geq ||x_0||^2,$$
$$2t \langle x_0, x - x_0 \rangle + t^2 ||(x - x_0)||^2 \geq 0.$$

Take now small t and assume $\langle x_0, x - x_0 \rangle < 0$ for some $x \in C$, then there appears a contradiction in the previous inequality, hence we obtain

$$\langle x_0, x - x_0 \rangle \geq 0$$

and consequently $\langle x, x_0 \rangle \geq ||x_0||^2$ for $x \in C$, so we can choose $\xi = \langle ., x_0 \rangle$. □

As a corollary we have that each subspace $V_1 \subset V$, which does not intersect with a convex, compact and non-empty subset $K \subset V$ can be separated from K, i.e. there is $\xi \in V'$ such that $\xi(V_1) = 0$ and $\xi(x) > 0$ for $x \in K$. This is proved by considering the set

$$C := K - V := \{w - v \text{ for } v \in V \text{ and } w \in K\},$$

which is convex and closed, since V, K are convex and K is compact, and which does not contain the origin. By the above theorem we can find a separating linear functional $\xi \in V'$ such that $\xi(w - v) \geq \alpha$ for all $w \in K$ and $v \in V$, which means in particular that $\xi(w) > 0$ for all $w \in K$. Furthermore we obtain from $\xi(w) - \xi(v) \geq \alpha$ for all $v \in V$ that $\xi(v) = 0$ for all $v \in V$ (replace v by λv, which is possible since V is a vector space, and lead the assertion to a contradiction in case that $\xi(v) \neq 0$).

THEOREM 1.4. *Let C be a compact convex non-empty set, then C is the convex hull of all its extreme points.*

PROOF. We have to show that there is an extreme point. We take a point $x \in C$ such that the distance $||x||^2$ is maximal, then x is an extreme point. Assume that there are two different points x_1, x_2 such that $x = \frac{1}{2}(x_1 + x_2)$, then

$$||x||^2 = ||\frac{1}{2}(x_1 + x_2)||^2 < \frac{1}{2}(||x_1||^2 + ||x_2||^2)$$
$$\leq \frac{1}{2}(||x||^2 + ||x||^2) = ||x||^2,$$

by the parallelogram law $\frac{1}{2}(||y||^2 + ||z||^2) = ||\frac{1}{2}(y+z)||^2 + ||\frac{1}{2}(y-z)||^2$ for all $y, z \in V$ and the maximality of $||x||^2$. This is a contradiction. Therefore we obtain at least one extreme point.

The set of all extreme points is a compact set, since it lies in C and is closed. Take now the convex hull of all extreme points, which is a closed convex subset S of C and hence compact. If there is $x \in C \setminus S$, then we can separate by a hyperplane l the point x and S such that $l(x) \geq \alpha > l(y)$ for $y \in S$. The set $\{l \geq \alpha\} \cap C$ is compact, convex, nonempty and has therefore an extreme point z, which is also an extreme point of C. So $z \in S$, which is a contradiction. □

Next we treat basic duality theory in the finite dimensional vector space V with euclidean structure. We identify the dual space V' with V by the above representation.

DEFINITION 1.5. *A subset $C \subset V$ is called convex cone if for all $v_1, v_2 \in C$ the sum $v_1 + v_2 \in C$ and $\lambda v_1 \in C$ for $\lambda \geq 0$. Given a cone C we define the polar C^0*

$$C^0 := \{l \in V \text{ such that } \langle l, v \rangle \leq 0 \text{ for all } v \in C\}.$$

The intersection of convex cones is a convex cone and therefore we can speak of the smallest convex cone containing an arbitrary set $M \subset V$, which is denoted by $\langle M \rangle_{cone}$. We want to prove the bipolar theorem for convex cones.

THEOREM 1.6 (Bipolar Theorem). *Let $C \subset V$ be a convex cone, then $C^{00} \subset V$ is the closure of C.*

PROOF. We show both inclusions. Take $v \in \overline{C}$, then $\langle l, v \rangle \leq 0$ for all $l \in C^0$ by definition of C^0 and therefore $v \in C^{00}$. If there were $v \in C^{00} \setminus \overline{C}$, where $\overline{C}$ denotes the closure of C, then for all $l \in C^0$ we have that $\langle l, v \rangle \leq 0$ by definition. On the other hand we can find $l \in V$ such that $\langle l, \overline{C} \rangle \leq 0$ and $\langle l, v \rangle > 0$ by the separation theorem since $\overline{C}$ is a closed cone. By assumption we have $l \in C^0$, however this yields a contradiction since $\langle l, v \rangle > 0$ and $v \in C^{00}$. □

DEFINITION 1.7. *A convex cone C is called* polyhedral *if there is a finite number of linear functionals $l_1, \dots, l_m$ such that*

$$C := \cap_{i=1}^n \{v \in V | \langle l_i, v \rangle \leq 0\}.$$

In particular a polyhedral cone is closed as intersection of closed sets.

LEMMA 1.8. *Given $e_1, \dots, e_n \in V$. For the cone $C = \langle e_1, \dots, e_n \rangle_{con}$ the polar can be calculated as*

$$C^0 = \{l \in V \text{ such that } \langle l, e_i \rangle \leq 0 \text{ for all } i = 1, \dots, n\}.$$

PROOF. The convex cone $C = \langle e_1, \dots, e_n \rangle_{cone}$ is given by

$$C = \{\sum_{i=1}^n \alpha_i e_i \text{ for } \alpha_i \geq 0 \text{ and } i = 1, \dots, n\}.$$

Given $l \in C^0$, the equation $\langle l, e_i \rangle \leq 0$ necessarily holds and we have the inclusion $\subset$. Given $l \in V$ such that $\langle l, e_i \rangle \leq 0$ for $i = 1, \dots, n$, then for $\alpha_i \geq 0$ the equation $\sum_{i=1}^n \alpha_i \langle l, e_i \rangle \leq 0$ holds and therefore $l \in C^0$ by the explicit description of C as $\sum_{i=1}^n \alpha_i e_i$ for $\alpha_i \geq 0$. □

COROLLARY 1.9. *Given $e_1, \dots, e_n \in V$, the cone $C = \langle e_1, \dots, e_n \rangle_{con}$ has a polar which is polyhedral and therefore closed.*

PROOF. The polyhedral cone is given through

$$\begin{aligned} C^0 &= \{l \in V \text{such that } \langle l, e_i \rangle \leq 0 \text{ for all } i = 1, \dots, n\} \\ &= \cap_{i=1}^n \{l \in V | \langle l, e_i \rangle \leq 0\}. \end{aligned}$$

□

LEMMA 1.10. *Given a finite set of vectors $e_1, \dots, e_n \in V$ and the convex cone $C = \langle e_1, \dots, e_n \rangle_{con}$, then C is closed.*

PROOF. Assume that $C = \langle e_1, \dots, e_n \rangle_{con}$ for vectors $e_i \in V$. If the e_i are linearly independent, then C is closed by the argument, that any $x \in C$ can be uniquely written as $x = \sum_{i=1}^n \alpha_i e_i$. Suppose next that there is a non-trivial linear combination $\sum_{i=1}^n \beta_i e_i = 0$ with $\beta \in \mathbb{R}^n$ non-zero. We can write $x \in C$ as

$$x = \sum_{i=1}^n \alpha_i e_i = \sum_{i=1}^n (\alpha_i + t(x)\beta_i) e_i = \sum_{j \neq i(x)} \alpha'_i e_i$$

with

$$\begin{aligned} i(x) &\in \{i \text{ such that } |\frac{\alpha_i}{\beta_i}| = \max_{\beta_j < 0} |\frac{\alpha_j}{\beta_j}|\}, \\ t(x) &= -\frac{\alpha_{i(x)}}{\beta_{i(x)}} \end{aligned}$$

Then $\alpha'_j \geq 0$ by definition. Consequently we can construct by variation of x a decomposition

$$C = \cup_{i=1}^{n'} C_i$$

where C_i are cones generated by $n-1$ vectors from the set $e_1, \ldots, e_n$. By induction on the number of generators n we can conclude by the statement on linearly independent generators. □

PROPOSITION 1.11. *Let $C \subset V$ be a convex cone generated by $e_1, \ldots, e_n$ and $\mathcal{K}$ a subspace, then $\mathcal{K} - C$ is closed convex.*

PROOF. First we prove that $\mathcal{K} - C$ is a convex cone. Taking $v_1, v_2 \in \mathcal{K} - C$, then $v_1 = k_1 - c_1$ and $v_2 = k_2 - c_2$, therefore

$$\begin{aligned} v_1 + v_2 &= k_1 + k_2 - (c_1 + c_2) \in \mathcal{K} - C, \\ \lambda v_1 &= \lambda k_1 - \lambda c_1 \in \mathcal{K} - C. \end{aligned}$$

In particular $0 \in \mathcal{K} - C$. The convex cone is generated by a generating set $e_1, \ldots, e_n$ for C and a basis $f_1, \ldots, f_p$ for $\mathcal{K}$, which has to be taken with $-$ sign, too. So

$$\mathcal{K} - C = \langle -e_1, \ldots, -e_n, f_1, \ldots, f_p, -f_1, \ldots, -f_p \rangle_{con}$$

and therefore $\mathcal{K} - C$ is closed by Lemma 1.10. □

LEMMA 1.12. *Let C be a polyhedral cone, then there are finitely many vectors $e_1, \ldots, e_n \in V$ such that*

$$C = \langle e_1, \ldots, e_n \rangle_{con}.$$

PROOF. By assumption $C = \cap_{i=1}^{p} \{v \in V | \langle l_i, v \rangle \leq 0\}$ for some vectors $l_i \in V$. We intersect C with $[-1,1]^m$ and obtain a convex, compact set. This set is generated by its extreme points. We have to show that there are only finitely many extreme points. Assume that there are infinitely many extreme points, then there is also an adherence point $x \in C$. Take a sequence of extreme points $(x_n)_{n \geq 0}$ such that $x_n \to x$ as $n \to \infty$ with $x_n \neq x$. We can write the defining inequalities for $C \cap [-1,1]^m$ by

$$\langle k_j, v \rangle \leq a_j$$

for $j = 1, \ldots, r$ and we obtain $\lim_{n \to \infty} \langle k_j, x_n \rangle = \langle k_j, x \rangle$. Define

$$\epsilon := \min_{\langle k_j, x \rangle < a_j} a_j - \langle k_j, x \rangle > 0.$$

Take n_0 large enough such that $| \langle k_j, x_{n_0} \rangle - \langle k_j, x \rangle | \leq \frac{\epsilon}{2}$, which is possible due to convergence. Then we can look at $x_{n_0} + t(x - x_{n_0}) \in C$ for $t \in [0,1]$. We want to find a continuation of this segment for some $\delta > 0$ such that $x_{n_0} + t(x - x_{n_0}) \in C$ for $[-\delta, 1]$. Therefore we have to check three cases:

- If $\langle k_j, x_{n_0} \rangle = \langle k_j, x \rangle = a_j$, then we can continue for all $t \leq 0$ and the inequality $\langle k_j, x_{n_0} + t(x - x_{n_0}) \rangle = a_j$ remains valid.
- If $\langle k_j, x \rangle = a_j$ and $\langle k_j, x_{n_0} \rangle < a_j$, we can continue for all $t \leq 0$ and the inequality $\langle k_j, x_{n_0} + t(x - x_{n_0}) \rangle \leq a_j$ remains valid.
- If $\langle k_j, x \rangle < a_j$, then we define $\delta = 1$ and obtain that for $-1 \leq t \leq 1$ the inequality $\langle k_j, x_{n_0} + t(x - x_{n_0}) \rangle \leq a_j$ remains valid.

Therefore we can find δ and continue the segment for small times. Hence x_n cannot be an extreme point, since it is a nontrivial convex combination of $x_{n_0} - \delta(x - x_{n_0})$ and x, which is a contradiction. Therefore $C \cap [-1,1]^m$ is generated by finitely many extreme points $e_1, \ldots, e_n$and so

$$C = \langle e_1, \ldots, e_n \rangle_{con}$$

by dilatation. □

2. Optimization Theory

We shall first consider general principles in optimization theory related to analysis and proceed to special functionals.

DEFINITION 2.1. *Let $U \subset \mathbb{R}^m$ be a subset with $U \subset V$, where V is open in $\mathbb{R}^m$. Let $F : V \to \mathbb{R}$ be a C^2-function. A point $x \in U$ is called local maximum (local minimum) of F on U if there is a neighborhood W_x of x in V such that for all $y \in U \cap W_x$*

$$F(y) \leq F(x)$$

or respectively $F(y) \geq F(x)$.

LEMMA 2.2. *Let $U \subset \mathbb{R}^m$ be a subset with $U \subset V$, where V is open in $\mathbb{R}^m$ and let $F : V \to \mathbb{R}$ be a C^2-function. Given a local maximum (or local minimum) $x \in U$ of F on U and a C^2-curve $c :]-1, 1[\to V$ such that $c(0) = x$ and $c(t) \in U$ for $t \in]-1, 1[$, the following necessary condition holds true,*

$$\frac{d}{dt}|_{t=0} F(c(t)) = \langle \operatorname{grad} F(x), c'(0)\rangle = 0.$$

PROOF. The function $t \mapsto F(c(t))$ has a local extremum at $t = 0$ and therefore the first derivative at $t = 0$ must vanish. □

We shall now prove a version of the Lagrangian multiplier theorem for affine subspaces $U \subset \mathbb{R}^m$. We take a affine subspace $U \subset \mathbb{R}^m$ and an open neighborhood $V \subset \mathbb{R}^m$ such that $U \cap V \neq \emptyset$, where a C^2-function $F : V \to \mathbb{R}$ is defined.

THEOREM 2.3. *Let x be a local maximum (local minimum) of F on $U \cap V$ and assume that there are $k := m - \dim U$ vectors $l_1, \dots, l_k \in \mathbb{R}^m$ and real numbers $a_1, \dots, a_k \in \mathbb{R}$ such that*

$$U = \{x \in V \text{ with } \langle l_i, x\rangle = a_i \text{ for } i = 1, \dots, k\}.$$

Then

$$\operatorname{grad} F(x) \in \langle l_1, \dots, l_k\rangle$$

or in other words there are real numbers $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ such that

$$\operatorname{grad} F(x) = \lambda_1 l_1 + \dots + \lambda_k l_k.$$

PROOF. Take a C^2-curve $c :]-1, +1[\to V$, then c takes values in U if and only if

$$c(0) \in U$$

and

$$\langle l_i, c'(t)\rangle = 0$$

for $i = 1, \dots, k$ and $t \in]-1, 1[$. The proof is simply done by Taylor's formula. Fix $t \in]-1, 1[$ and take

$$c(t) = c(0) + \int_0^t c'(s)ds.$$

By definition $c(t) \in U$ if and only if $\langle l_i, c(t)\rangle = a_i$, but

$$\langle l_i, c(t)\rangle = \langle l_i, c(0)\rangle + \int_0^t \langle l_i, c'(s)\rangle\, ds$$
$$= a_i$$

by assumption for $i = 1, \dots, k$. We denote the span of $l_1, \dots, l_k$ by T and can consequently state that a C^2-curve $c :]-1, +1[\to V$ takes values in U if and only if $c(0) \in U$ and $c'(t) \in T^0$ for all $t \in]-1, 1[$. Furthermore we can say that T^0 is generated by all derivatives of C^2-curves $c :]-1, +1[\to V$ taking values in U at time $t = 0$ (simply take a line with direction a vector in T^0 through some point of U).

By the previous lemma we know that for all C^2-curves $c :]-1, +1[\to V$ with $c(0) = x$ the relation

$$\langle \operatorname{grad} F(x), c'(0) \rangle = 0$$

holds. Therefore $\operatorname{grad} F(x) \in T^{00}$. By the bipolar theorem we know that $T^{00} = T = \langle l_1, \ldots, l_k \rangle$, which proves the result. □

Remark 2.4. This leads immediately to the receipt of Lagrangian multipliers as it is well known from basic calculus: a necessary condition for an extremal point of $F : V \to \mathbb{R}$ subject to the conditions $\langle l_i, x \rangle = a_i$ for $i = 1, \ldots, k$ is to solve the extended problem with the Lagrangian L

$$L(x, \lambda_1, \ldots, \lambda_k) = F(x) - \sum_{i=1}^{k} \lambda_i (\langle l_i, x \rangle - a_i).$$

Taking the gradients leads to the system of equations

$$\begin{aligned} \operatorname{grad} F(x) - \sum_{i=1}^{k} \lambda_i l_i &= 0 \\ \langle l_i, x \rangle &= a_i \end{aligned}$$

for $i = 1, \ldots, k$, which necessarily has a solution if there is an extremal point at x.

Remark 2.5. How to calculate a gradient? The gradient of a C^1-function $F : V \to \mathbb{R}$ on a finite dimensional vector space V is defined through

$$\langle \operatorname{grad} F(x), w \rangle = \frac{d}{ds}|_{s=0} F(x + sw),$$

for $x \in V$ and $w \in \mathbb{R}^n$ (and a scalar product!). This can be calculated with respect to any basis and gives a coordinate representation. The derivative of F is understood as element of the dual space

$$dF(x)(w) := \frac{d}{ds}|_{s=0} F(x + sw)$$

for $x \in V$ and $w \in \mathbb{R}^n$ (even without scalar product!). The derivative can be calculated with respect to a basis $(e_i)_{i=1,\ldots,\dim V}$. That means that it simply represents a collection of directional derivatives of a function, i.e.

$$\operatorname{grad}_{(e_i)} F(x) := (\frac{d}{ds}|_{s=0} F(x + s e_i))_{i=1,\ldots,\dim V}$$

for $x \in V$.

3. Conjugate Functions

Given a concave, increasing function $u : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$, which usual conventions for the calculus with $-\infty$. We denote by $\operatorname{dom}(u)$ the set $\{u > -\infty\}$ and assume that $\operatorname{dom}(u)$ is either $]0, \infty[$ or $\mathbb{R}$. We shall always assume that u is strictly concave and C^2 on $\operatorname{dom}(u)$ and that

$$\lim_{x \downarrow 0} u(x) = -\infty$$

in the case $\operatorname{dom}(u) =]0, \infty[$ and

$$\lim_{x \to -\infty} u(x) = -\infty$$

in the case $\operatorname{dom}(u) = \mathbb{R}$.

In this and more general cases we can define the conjugate function

$$v(y) := \sup_{x\in\mathbb{R}}(u(x) - yx)$$

for $y > 0$.

Since the function $x \mapsto u(x) - yx$ is strictly concave for every $y > 0$, there is some hope for a maximum. If there is one, let's say $\widehat{x}$, then it satisfies

$$u'(\widehat{x}) = y. \tag{3.1}$$

Since the second derivative exists and is strictly negative, $\widehat{x}$ is a local maximum if the above equation is satisfied. By strict concavity the local maximum is unique and global, too.

We need basic assumptions for the existence and regularity of the conjugate function:

(1) If $\mathrm{dom}(u) =]0, \infty[$ (negative wealth not allowed), then we assume

$$\lim_{x\downarrow 0} u'(x) = \infty,$$
$$\lim_{x\to\infty} u'(x) = 0 \text{ (marginal utility tends to 0).}$$

(2) If $\mathrm{dom}(u) = \mathbb{R}$ (negative wealth allowed), then we assume

$$\lim_{x\downarrow -\infty} u'(x) = \infty,$$
$$\lim_{x\to\infty} u'(x) = 0 \text{ (marginal utility tends to 0).}$$

Under these assumptions we can state the following theorem on existence and convexity of v.

Theorem 3.1. *Let $u : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ be a concave function satisfying the above assumptions, then the conjugate function is strictly convex and C^2 on $\mathrm{dom}(v) =]0, \infty[$. Additionally for $\mathrm{dom}(u) =]0, \infty[$ we have*

$$v'(0) := \lim_{y\downarrow 0} v'(y) = -\infty,$$
$$\lim_{y\to\infty} v'(y) = 0$$

and for $\mathrm{dom}(u) = \mathbb{R}$

$$v'(0) := \lim_{y\downarrow 0} v'(y) = -\infty,$$
$$\lim_{y\to\infty} v'(y) = \infty$$

Furthermore the inversion formula

$$u(x) = \inf_{y>0}(v(y) + xy)$$

holds true.

Proof. By formula 3.1 and our assumptions we see that for every $y > 0$ there is exactly one $\widehat{x}$, since u' is strictly decreasing and C^1. We denote the inverse of u' by $(u')^{-1}$. Therefore v is well-defined and at least C^1, since the inverse is C^1. Furthermore

$$\begin{aligned}
v(y) &= u((u')^{-1}(y)) - y \cdot (u')^{-1}(y) \\
v'(y) &= u'((u')^{-1}(y))((u')^{-1})'(y) - (u')^{-1}(y) - y((u')^{-1})'(y) \\
&= -(u')^{-1}(y) \\
v''(y) &= -((u')^{-1})'(y) = -\frac{1}{u''((u')^{-1}(y))} > 0
\end{aligned}$$

Hence v is C^2 on $]0,\infty[$ and a fortiori, by $v''>0$, strictly convex.

We know that u' is positive and strictly decreasing from ∞ to 0 by the previous assumptions, hence the two limiting properties for v, since $v'(y) = -(u')^{-1}(y)$.

Replacing v by $-v$, we can apply the same reasoning for existence of the concave conjugate of v. Take $\widehat{y} > 0$ such that $\inf_{y>0}(v(y) + xy)$ takes the infimum, then necessarily

$$v'(\widehat{y}) = -x,$$

hence $-(u')^{-1}(\widehat{y}) = -x$ and therefore $\widehat{y}(x) = u'(x)$. Inserting yields

$$\begin{aligned} v(u'(x)) + x\widehat{y}(x) &= u((u')^{-1}(u'(x))) - u'(x)(u')^{-1}(u'(x)) + xu'(x) \\ &= u(x), \end{aligned}$$

which is the desired relation. □

4. Exam Questions

For the oral exam I shall choose randomly three questions from the following list, from which you have the right to select two for your exam. The exam is "open book", i.e. you can use the script during the exam. You will have about 12 minutes of time for each question after about 6 minutes of preparation. I expect you to speak about the question like in a seminar, i.e. explaining the structure of the answer and important details such that a good mathematician, who does not know precisely about the topic could in principle follow.

- Explain the notion of a semi-martingale. Why are semi-martingales important in mathematical Finance. What is a good integrator? What does the Bichteler-Dellacherie theorem tell? Give a sketch of its proof?
- Why is every semi-martingale a good integrator? Argue that finite variation processes are good integrators, then argue that L^2-martingales are good integrators, next argue that L^1-martingales are good integrators by means of an inequality and prove this inequality.
- Define the Emery topology (finite time horizon) and show that the set of semi-martingales is a complete metric space with it. Introduce the Burkholder-Davis-Gundy inequalities and provide an idea for the proof based on deterministic inequalities (which you do not need to prove). Introduce and discuss the space $\mathcal{H}^1$ and Davis inequality on it. Argue why a local martingale is locally $\mathcal{H}^1$.
- Describe the No Arbitrage Theory for discrete models and show the proof of Theorem 2.6 (ch. 1). What are arbitrage-free prices of a claim? Why is this a reasonable concept? What does hedging mean? What is a complete model?
- Describe arbitrage opportunities in a general semi-martingale model. Describe and prove the Ansel-Stricker Lemma 4.7. (ch. 1). Prove Lemma 4.6. (ch. 1).
- Describe and prove the Kreps/Yan Theorem 4.10 and Stricker's Theorem 4.11. (ch. 1). What are strong, what are weak sides of those theorems?
- Describe (No Free Lunch with Vanishing Risk) and (No Unbounded Profit with bounded Risk). What is the assertion of the fundamental theorem of asset pricing? What is a σ-martingale (provide an example)?
- Give a sketch of the proof of the fundamental theorem of asset pricing following the article of Yuri Kabanov [**8**]. This means not going into detail but just explaining the main headlines of the lemmas in order to arrive at a separating measure.
- What does No Arbitrage mean in case of Ito process models? Explain the Black-Scholes model.

- What does No Arbitrage mean in case of exponential Lévy process models? Explain hedging and incompleteness in exponential Lévy models.
- What is a superhedging price? Prove Proposition 10.3. (ch. 1). Explain the optional decomposition theorem and prove Theorem 10.6. (ch. 1) with it.
- What is an American option? Prove Theorem 11.3. (ch. 1).
- Explain utility optimization for discrete models. Prove Theorem 1.12. (ch. 2). Explain the dual approach.
- Explain the martingale optimality principle and prove Theorem 2.2. (ch. 2). Explain and solve the Merton problem.
- Explain the primal and dual problem of utility optimization in the general semi-martingale setting. Explain the strategy to solve the primal problem via the dual problem (subsection 3.4 and 3.6 of chapter 2) with highlighting some intermediate results.

Bibliography

[1] Mathias Beiglböck, Walter Schachermayer, and Bezirgen Veliyev. A short proof of the Doob-Meyer theorem. *Stochastic Process. Appl.*, 122(4):1204–1209, 2012.

[2] Mathias Beiglböck and Pietro Siorpaes. Pathwise versions of the burkholder-davis-gundy inequalities. *preprint*, 2013.

[3] Mathias Beiglböck and Pietro Siorpaes. A simple proof of the bichteler-dellacherie theorem. *preprint*, 2013.

[4] Klaus Bichteler. Stochastic integration and L^p-theory of semimartingales. *Ann. Probab.*, 9(1):49–89, 1981.

[5] Marzia De Donno and Maurizio Pratelli. On a lemma by Ansel and Stricker. In *Séminaire de Probabilités XL*, volume 1899 of *Lecture Notes in Math.*, pages 411–414. Springer, Berlin, 2007.

[6] Freddy Delbaen and Walter Schachermayer. *The mathematics of arbitrage*. Springer Finance. Springer-Verlag, Berlin, 2006.

[7] M. Emery. Une topologie sur l'espace des semimartingales. In *Séminaire de Probabilités, XIII (Univ. Strasbourg, Strasbourg, 1977/78)*, volume 721 of *Lecture Notes in Math.*, pages 260–280. Springer, Berlin, 1979.

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

[9] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[10] Kostas Karadaras. On the closure in the emery topology of semimartingale wealth-process sets. *Annals of Aplied Probability*, 23(4):1355–1376, 2013.

[11] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2004. Stochastic Modelling and Applied Probability.

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