Lectures on the FTAP

Notes complementing Delbaen and Schachermayer's book

"The Mathematics of Arbitrage"

Harry van Zanten

September 25, 2007 (this version)

use at your own risk
this text undoubtedly contains many errors
please report them to `harry@cs.vu.nl`

Preface

These lecture notes treat various versions of the so-called "fundamental theorem of asset pricing". Many students are familiar with statements about models for financial markets saying that "absence of arbitrage" and the existence of an "equivalent martingale measure" are in some sense equivalent. The purpose of these notes is to present and explain mathematical results making such statements precise.

We start with the relatively simple situation of models defined on a finite underlying probability space, cf. Harrison and Pliska (1981). In this case a mathematically precise and economically satisfactory fundamental theorem can be derived using the separating hyperplane theorem. We then move to general continuous time models, treating the theorem of Kreps (1981). The basic idea behind this theorem is in fact similar to the finite case. However, since the problem is now infinite-dimensional, it is technically much more involved and it is necessary to involve topological considerations in the definition of "no arbitrage", called "no free lunch" by Kreps. Although Kreps' theorem is satisfactory from the mathematical perspective, the use of the weak*-topology in the definition of "no free lunch" destroys the economic interpretation of the result. We finally treat the fundamental theorem of Delbaen and Schachermayer (1994), who work in the setting of asset prices modelled by locally bounded semimartingales, and trading strategies modelled by predictable processes. In this setting Kreps' definition of "no free lunch" can be replaced by the condition of "no free lunch with vanishing risk", which does not involve an unnatural topology, and has a clear economic interpretation. The proof of the fundamental theorem of Delbaen and Schachermayer (1994) is technically very involved, and we only sketch the main arguments in these notes. In particular, we explain the connection to the result of Kreps (1981).

The text is mainly based on Chapters 2, 5 and 8 and 9 of Delbaen and Schachermayer (2006). The necessary results from functional analysis are treated in the appendix, and are taken from Conway (1990) or Rudin (1991).

Readers are assumed to be familiar with basic topology, measure theoretic probability theory, martingale theory and stochastic integration theory, the latter up to the level of stochastic integration of locally bounded predictable processes relative to general semimartingales.

Amsterdam, fall 2006
HvZ

Contents

Preface iii

1 A simple example 1
1.1 Exercises 3

2 Finite underlying probability spaces 5
2.1 Description of the model and basic definitions 5
2.2 Fundamental theorem of asset pricing 7
2.3 Single period versus multiperiod models 8
2.4 Completeness 9
2.5 Change of numerair 11
2.6 No-arbitrage pricing 12
2.7 Example: binomial model 13
2.8 Exercises 15

3 The Kreps-Yan theorem 17
3.1 Description of the model 17
3.2 Kreps-Yan theorem 19
3.3 Discussion 20
3.4 Exercises 22

4 The general theorem 23
4.1 Preliminaries on stochastic integration 23
4.2 No free lunch with vanishing risk and the FTAP 26
4.3 Sketch of proof of the fundamental theorem 27
4.3.1 The relatively easy half 27
4.3.2 The much more difficult half 29
4.4 Example: Itô processes 30
4.5 Exercises 32

A Elements of functional analysis 33
A.1 Separating hyperplane theorem 33
A.2 Topological vector spaces 35
A.3 Hahn-Banach theorem 37
A.4 Dual space 39
A.5 Exercises 44

B Elements of martingale theory 45
B.1 Basic definitions 45

B.2 Theorems 46
B.3 Exercises 48

References 48

1

A simple example

Consider a world with two time points, $t = 0$ (today) and $t = 1$ (tomorrow). In this world there exists a bank where money can be deposited or borrowed at zero interest and a stock is traded. The value S_0 of the stock at time 0 equals 1 and the value S_1 at time $t = 1$ equals the value u with probability $p \in (0, 1)$ and $d < u$ with probability $1-p$, respectively. Note that the stock price process can be viewed as a stochastic process defined on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with Ω consisting of just two elements, one corresponding to the stock price going up, one to the price going down, $\mathcal{F}$ the power set of Ω and $\mathbb{P}$ the probability measure that gives probability p to the event of the price going up and $1 - p$ to the event of the price going down.

A trader in this world can form a portfolio today consisting of a number of units of money in the bank, call this φ_0, and a number of stocks, call this ψ_0. Clearly, this portfolio is worth $V_0 = \varphi_0 + \psi_0$ at time 0. When tomorrow comes, the portfolio will have the new, random value $V_1 = \varphi_0 + \psi_0 S_1$.

We say that there exists an *arbitrage opportunity* in this world if there exists a portfolio (φ_0, ψ_0) as above such that the associated value process V satisfies $V_0 = 0$, $V_1 \geq 0$ and $\mathbb{P}(V_1 > 0) > 0$. Clearly, this corresponds to a possibility of making a risk-free profit.

Proposition 1.0.1. *There exist no arbitrage opportunities in this world if and only if $d < 1 < u$.*

Proof. If $d < 1 < u$ then $q = (1 - d)/(u - d)$ belongs to $(0, 1)$ and hence the probability measure $\mathbb{Q}$ under which the stock price moves up or down according to the probability q instead of p is equivalent (i.e. mutually absolutely continuous) to the underlying probability measure $\mathbb{P}$ (see Exercise 1). Now let (φ_0, ψ_0) be a portfolio such that $V_0 = 0$ and $V_1 \geq 0$. Then by construction we have $\mathbb{E}_{\mathbb{Q}} V_1 = (\varphi_0 + \psi_0 u)q + (\varphi_0 + \psi_0 d)(1 - q) = \varphi_0 + \psi_0 = V_0 = 0$. Hence, since $V_1 \geq 0$, we have $\mathbb{Q}(V_1 > 0) = 0$. But since the measures $\mathbb{P}$ and $\mathbb{Q}$ are equivalent, it follows that $\mathbb{P}(V_1 > 0) = 0$ as well.

To prove the converse statement, suppose for instance that $1 \leq d < u$. Then the stock always performs at least as good as money in the bank and there is a positive probability that it performs strictly better. Borrowing money from the bank and investing it in stock then yields an arbitrage. Specifically, consider the portfolio $\varphi_0 = -1$, $\psi_0 = 1$. The corresponding values process satisfies $V_0 = 0$ and V_1 is either $d-1$ or $u-1$, which is strictly positive with positive probability in this case. The case $d < u \leq 1$ can be handled similarly. $\square$

In the proof of the proposition we noted that in the case of no-arbitrage, i.e. when $d < 1 < u$, the new underlying probability measure $\mathbb{Q}$ under which the stock goes up with probability $q = (1-d)/(u-d)$ is equivalent to the original probability measure $\mathbb{P}$. Observe that

$$\mathbb{E}_{\mathbb{Q}} S_1 = uq + d(1-q) = S_0.$$

In other words, the stock price process $S = (S_0, S_1)$ is a martingale under $\mathbb{Q}$ (relative to the trivial filtration $\mathcal{F}_0 = \{\varnothing, \Omega\}$, $\mathcal{F}_1 = \mathcal{P}(\Omega)$). Note that any other equivalent measure $\mathbb{Q}'$ is fully described by specifying a probability $q' \in (0,1)$ with which the stock goes up (Exercise 1 again). For such a measure $\mathbb{Q}'$ we have

$$\mathbb{E}_{\mathbb{Q}'} S_1 = uq' + d(1-q') = q'(u-d) + d.$$

Hence if $\mathbb{Q}'$ has the property that $\mathbb{E}_{\mathbb{Q}'} S_1 = S_0$, then $(1-d)/(u-d) = q' \in (0,1)$ which is the same as saying that $d < 1 < u$. We just proved that this implies absence of arbitrage opportunities.

We conclude that the condition for no-arbitrage can be reformulated in terms of the existence of certain probability measures.

Proposition 1.0.2 (Fundamental theorem of asset pricing O). *There exist no arbitrage opportunities in this world if and only there exists a probability measure $\mathbb{Q}$ equivalent to the original probability measure $\mathbb{P}$ such that the stock price process $S = (S_0, S_1)$ satisfies $\mathbb{E}_{\mathbb{Q}} S_1 = S_0$.*

For obvious reasons a probability measure $\mathbb{Q}$ as in the proposition is called an *equivalent martingale measure*. Using this terminology the result asserts that absence of arbitrage is equivalent to the existence of an equivalent martingale measure. It turns out that (the appropriate version of) this result, often called the *fundamental theorem of asset pricing*, is true in a very general setting. This is of great interest, since it relates a fundamental economic notion (arbitrage) to an important mathematical concept (martingales). As a result, martingale theory plays a central role in the modelling of financial markets and pricing of derivatives. In these notes we discuss the fundamental theorem of asset pricing in increasingly general settings.

1.1 Exercises

1. Show that two probability measures on a finite set are equivalent (i.e. mutually absolutely continuous) if and only if they give positive probability to the same singletons.

2

Finite underlying probability spaces

2.1 Description of the model and basic definitions

Consider a finite probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $\Omega = \{\omega_1, \dots, \omega_n\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ and $p_i = \mathbb{P}(\{\omega_i\}) > 0$ for every i. Suppose that on this probability space we are given a filtration $(\mathcal{F}_t)_{t=0,\dots,T}$ and a $(d+1)$-dimensional adapted stochastic process $S = (S_t^{(0)}, \dots, S_t^{(d)})_{t=0,\dots,T}$, where T is some finite positive integer. Assume that $\mathcal{F}_T = \mathcal{F}$.

We think of the components of S as the price processes of $d+1$ different financial assets, measured relative to the price of the 0-th asset, called a *numerair*. Since we measure prices relative to the price of the numerair, the price process $S^{(0)}$ equals 1 at all times.

A *portfolio* is a $(d+1)$-dimensional, predictable process $\varphi = (\varphi_t^{(0)}, \dots, \varphi_t^{(d)})_{t=1,\dots,T}$. We think of $\varphi_t^{(i)}$ as the number of assets of type i that is in the portfolio in the time interval $(t-1, t]$. The requirement that φ is predictable means that at each time $t-1$, the portfolio is constructed using only the information available up to that time, i.e. the trader can not look into the future. The *value process* associated with a portfolio φ is the process $V = (V_t)_{t=0,\dots,T}$ defined by

$$V_0 = \sum_{i=0}^{d} \varphi_1^{(i)} S_0^{(i)}, \qquad V_t = \sum_{i=0}^{d} \varphi_t^{(i)} S_t^{(i)}, \quad t \geq 1. \tag{2.1}$$

Clearly V is an adapted process. The value V_0 is called the *initial value* of the portfolio.

A special role is played by portfolios that do not involve injections or withdrawals of money after time 0. Consider such a portfolio with value V_{t-1} at time $t-1$. Just after $t-1$ the portfolio is rebalanced, and the new value equals

$$\sum_{i=0}^{d} \varphi_t^{(i)} S_{t-1}^{(i)} = V_t - \sum_{i=0}^{d} \varphi_t^{(i)} (S_t^{(i)} - S_{t-1}^{(i)}).$$

If no money is injected or withdrawn this should equal V_{t-1}, hence

$$V_t - V_{t-1} = \sum_{i=0}^{d} \varphi_t^{(i)}(S_t^{(i)} - S_{t-1}^{(i)}) = \sum_{i=1}^{d} \varphi_t^{(i)}(S_t^{(i)} - S_{t-1}^{(i)})$$

(we use that $S^{(0)}$ equals 1 at all times).

Definition 2.1.1. A portfolio φ is called *self-financing* if its value process V satisfies

$$V_t - V_{t-1} = \sum_{i=1}^{d} \varphi_t^{(i)}(S_t^{(i)} - S_{t-1}^{(i)})$$

for all $t \geq 1$.

We use the usual notation $\Delta f(t) = f(t) - f(t-1)$ for a (possibly vector-valued) function f on the integers. Moreover, we write $\langle v, w\rangle$ for the Euclidean inner product of two vectors in $\mathbb{R}^d$. Using that notation the preceding display reads

$$\Delta V_t = \langle \varphi_t, \Delta S_t\rangle$$

and a self-financing portfolio satisfies the relation

$$V_t = V_0 + \sum_{u=1}^{t} \langle \varphi_u, \Delta S_u\rangle .$$

If we define the process $\varphi \cdot S$ by $(\varphi \cdot S)_0 = 0$ and $(\varphi \cdot S)_t = \sum_{u=1}^{t} \langle \varphi_u, \Delta S_u\rangle$ for $t \geq 1$, we can write

$$V_t = V_0 + (\varphi \cdot S)_t$$

for a self-financing portfolio.

If we compare the definition of a self-financing portfolio with (2.1) we see that for such a portfolio it holds that

$$\varphi_1^{(0)} = V_0 - \sum_{i=1}^{d} \varphi_1^{(i)} S_0^{(i)}$$

and

$$\Delta\varphi_t^{(0)} = -\sum_{i=1}^{d} \Delta\varphi_t^{(i)} S_{t-1}^{(i)}, \quad t \geq 2.$$

Hence, if we specify the initial value V_0 and $(\varphi^{(1)}, \dots, \varphi^{(d)})$, the process $\varphi^{(0)}$ describing the holdings in the numerair asset is completely determined by the requirement that the portfolio is self-financing.

2.2 Fundamental theorem of asset pricing

In this setting the definition of an arbitrage opportunity is as follows.

Definition 2.2.1. An *arbitrage opportunity* is a self-financing portfolio whose value process V satisfies $V_0 = 0$, $V_T \geq 0$ and $\mathbb{P}(V_T > 0) > 0$.

To prepare for the proof of the theorem below it is useful to reformulate this in more geometric terms. Define the collection of random variables

$$K = K(S) = \{(\varphi \cdot S)_T : \varphi \text{ predictable}\}.$$

Note that K is the set of all possible pay-offs of self-financing portfolios with zero initial value. Denoting the collection of all integrable nonnegative random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ by L^∞_+, absence of arbitrage is the same as the requirement that $K \cap L^\infty_+ = \{0\}$. In our setting of finite Ω we can identify collections of random variables with subsets of of $\mathbb{R}^n$: simply identify a random variable X with the vector of possible realizations $(X(\omega_1), \dots, X(\omega_n))$. For instance L^∞_+ corresponds to the set $\{(x_1, \dots, x_n) : x_1 \geq 0, \dots, x_n \geq 0\}$. This way the requirement $K \cap L^\infty_+ = \{0\}$ of no-arbitrage translates into a geometric statement about subsets of $\mathbb{R}^n$.

By L^∞ we denote the collection of all bounded random variables. Since Ω is finite, bounded just means finite-valued, so that L^∞ can be identified with all of $\mathbb{R}^n$.

We will see, as in the preceding chapter, that absence of arbitrage is equivalent to the existence of an equivalent martingale measure, which is defined as follows.

Definition 2.2.2. A probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$ is called *equivalent martingale measure* if it is equivalent to $\mathbb{P}$ (i.e. $\mathbb{Q} \ll \mathbb{P}$ and $\mathbb{P} \ll \mathbb{Q}$) and S is a (d-dimensional) martingale with respect to $\mathbb{Q}$. The collection of equivalent martingale measures is denoted by $\mathcal{M}^e = \mathcal{M}^e(S)$.

Theorem 2.2.3 (Fundamental theorem of asset pricing I). *There are no arbitrage opportunities in this model if and only if there exists an equivalent martingale measure.*

Proof. Suppose first that there exists a martingale measure $\mathbb{Q}$ and let φ be a self-financing portfolio whose value process V satisfies $V_0 = 0$ and $V_T \geq 0$. Since Ω is finite φ is bounded, hence $V = V_0 + \varphi \cdot S$ is a $\mathbb{Q}$-martingale (see Exercise 1). In particular, $\mathbb{E}_\mathbb{Q} V_T = \mathbb{E}_\mathbb{Q} V_0 = 0$, so $V_T = 0$ with $\mathbb{Q}$-probability one, but then also $\mathbb{P}$-almost surely.

Now assume that no arbitrage opportunities exist, so that $K \cap L^\infty_+ = \{0\}$. Let A be the convex hull of the elements $1_{\{\omega_1\}}, \dots, 1_{\{\omega_n\}}$ in L^∞. This is a convex, compact subset of L^∞, disjoint from K by assumption. Since the latter is a linear subspace it is closed and convex, and hence we can apply the separating hyperplane theorem. This yields a vector $q \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ such that

$$\langle q, f\rangle \leq \alpha < \beta \leq \langle q, h\rangle \tag{2.2}$$

for all $f \in K$ and $h \in A$. Since K is a linear space we can take $\alpha = 0$ in this display, i.e.

$$\langle q, f\rangle \leq 0 < \beta \leq \langle q, h\rangle$$

for all $f \in K$ and $h \in A$ (see Exercise 2). It follows that for every i we have $q_i = \left\langle q, 1_{\{\omega_i\}}\right\rangle \geq \beta > 0$, hence we can renormalize q such that it becomes a vector of strictly positive probabilities adding up to 1. The last display then remains true, but with β suitably normalized. The corresponding probability measure $\mathbb{Q}$, defined by $\mathbb{Q}(\{\omega_i\}) = q_i$, is equivalent to $\mathbb{P}$ and satisfies $\mathbb{E}_{\mathbb{Q}} f \leq 0$ for all $f \in K$. But since K is a linear space this implies that in fact $\mathbb{E}_{\mathbb{Q}} f = 0$ for all $f \in K$. By Exercise 3 if follows that S is a $\mathbb{Q}$-martingale. □

2.3 Single period versus multiperiod models

Recall our setting of a finite underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which we have a filtration $(\mathcal{F}_t)_{t=0,\ldots,T}$ and a d-dimensional adapted stochastic process $S = (S_t^{(1)}, \ldots, S_t^{(d)})_{t=0,\ldots,T}$ describing the discounted asset prices. If $T \geq 2$ this is called a *multiperiod* model. It turns out that that absence of arbitrage in such a model is equivalent to absence of arbitrage in all the single-period sub-models.

Theorem 2.3.1. *There are no arbitrage opportunities in the full model if and only if for every t, the one-period model (S_t, S_{t+1}), with respect to the filtration $(\mathcal{F}_t, \mathcal{F}_{t+1})$, admits no arbitrage opportunities.*

Proof. Suppose all the one-period models are free of arbitrage. Then by the fundamental theorem there exist probability measures $\mathbb{Q}_t$ on $(\Omega, \mathcal{F}_{t+1})$ such that $\mathbb{Q}_t$ is equivalent to $\mathbb{P}$ on $\mathcal{F}_{t+1}$ and $\mathbb{E}_{\mathbb{Q}_t}(S_{t+1} \,|\, \mathcal{F}_t) = S_t$. By the lemma following the theorem we may assume that $\mathbb{Q}_t|_{\mathcal{F}_t} = \mathbb{P}|_{\mathcal{F}_t}$. Now define the process L by $L_0 = 1$ and

$$L_t = \frac{d\mathbb{Q}_0}{d\mathbb{P}} \cdots \frac{d\mathbb{Q}_{t-1}}{d\mathbb{P}},$$

and define the measure $\mathbb{Q}$ by $d\mathbb{Q} = L_T\, d\mathbb{P}$. Then $\mathbb{Q} \in \mathcal{M}^e$ (Exercise 4), hence the full model is free of arbitrage by the fundamental theorem. □

The following lemma applies to the general, multiperiod model, but in the proof of the theorem it is applied only to single period models.

Lemma 2.3.2. *Suppose there is no arbitrage. Let $\mathbb{Q} \in \mathcal{M}^e$ and define $Z_t = \mathbb{E}_{\mathbb{P}}(d\mathbb{Q}/d\mathbb{P} \,|\, \mathcal{F}_t)$ and $L_t = Z_t/Z_0$. Then the measure $\mathbb{Q}^*$ defined by $d\mathbb{Q}^* = L_T\, d\mathbb{P}$ belongs to $\mathcal{M}^e$ and satisfies $\mathbb{Q}^*|_{\mathcal{F}_0} = \mathbb{P}|_{\mathcal{F}_0}$.*

Proof. Since Z is a martingale we have $\mathbb{E}_{\mathbb{P}}(Z_T \,|\, \mathcal{F}_t) = Z_t$ for all $t \le T$. Hence for $A \in \mathcal{F}_t$ and $t \le T$,

$$\mathbb{Q}(A) = \int_A Z_T \, d\mathbb{P} = \int_A Z_t \, d\mathbb{P}.$$

It follows that $Z_t = (d\mathbb{Q}|_{\mathcal{F}_t})/(d\mathbb{P}|_{\mathcal{F}_t})$. The fact that $\mathbb{Q}$ is equivalent to $\mathbb{P}$ now implies that $Z_t > 0$ for all t, and in particular that the process L is well defined. Since Z is a positive $\mathbb{P}$-martingale and Z_0 is $\mathcal{F}_0$-measurable, L is a positive $\mathbb{P}$-martingale as well and therefore $\mathbb{Q}^*$ is a probability measure equivalent to $\mathbb{P}$. The fact that S is a $\mathbb{Q}$-martingale implies that SZ is a $\mathbb{P}$-martingale (check!). Since Z_0 is $\mathcal{F}_0$-measurable, SL is a $\mathbb{P}$-martingale as well. Using the fact that $L_t = (d\mathbb{Q}^*|_{\mathcal{F}_t})/(d\mathbb{P}|_{\mathcal{F}_t})$, we obtain that S is a $\mathbb{Q}^*$-martingale (check!), hence $\mathbb{Q}^* \in \mathcal{M}^e$. Finally, the fact that $L_1 = 1$ implies that $\mathbb{Q}^*|_{\mathcal{F}_0} = \mathbb{P}|_{\mathcal{F}_0}$. □

2.4 Completeness

The FTAP does not say how many equivalent martingale measures there are in the absence of arbitrage. We will see below that this is related to the notion of *completeness*.

Definition 2.4.1. We call $f \in L^\infty$ *attainable* if $f = a + (\varphi \cdot S)_T$ for some $a \in \mathbb{R}$ and predictable process φ. The model is called *complete* if every claim $f \in L^\infty$ is attainable.

So an attainable contingent claim f is a random pay-off at time T that can be realized by following a self-financing strategy requiring some initial capital a.

For the proof of the following theorem it is useful to introduce, in addition to the set K, the set of random variables

$$C = \{f \in L^\infty : \text{there exists a } g \in K \text{ such that } g \ge f\}.$$

This set is a cone[1] containing K and it is easy to see that $K \cap L^\infty_+ = \{0\}$ if and only if $C \cap L^\infty_+ = \{0\}$ (see Exercise 5). Moreover, under no-arbitrage it holds that $K = C \cap (-C)$ (Exercise 6).

Lemma 2.4.2. *For any probability measure $\mathbb{Q}$ we have that S is a $\mathbb{Q}$-martingale if and only if $\mathbb{E}_{\mathbb{Q}} g \le 0$ for all $g \in C$.*

[1] Recall that a subset C of a vector space is called a *cone* if for all $x \in C$ and $a \ge 0$, it holds that $ax \in C$.

Proof. Take $g \in C$, say $g \leq f$ for $f \in K$. Then if S is $\mathbb{Q}$-martingale we have $\mathbb{E}_{\mathbb{Q}} f = 0$ (see Exercise 1) and hence $\mathbb{E}_{\mathbb{Q}} g \leq 0$.

If $\mathbb{E}_{\mathbb{Q}} g \leq 0$ for all $g \in C$ then for all $f \in K$ it holds that $\mathbb{E}_{\mathbb{Q}} f = 0$, since $f \in C$ and $-f \in C$ for $f \in K$. Hence, by Exercise 3, S is a $\mathbb{Q}$-martingale. $\square$

Theorem 2.4.3. *Assume there are no arbitrage opportunities. Then*

$$K = \{f \in L^\infty : \mathbb{E}_{\mathbb{Q}} f = 0 \textit{ for all } \mathbb{Q} \in \mathcal{M}^e\}.$$

Proof. The set C is convex and closed (see Exercise 7) and hence, by the Bipolar Theorem, equals its own bipolar C^{00}. Since C is closed under multiplication with positive scalars we have (see the appendix) $C^0 = \{q \in \mathbb{R}^n : \langle g, q \rangle \leq 0 \text{ for all } g \in C\}$. Hence, by the Lemma preceding the theorem, the collection $\mathcal{M}^a$ of probability measure $\mathbb{Q}$ such that S is a $\mathbb{Q}$-martingale is contained in C^0, hence $\text{cone}(\mathcal{M}^a) \subseteq C^0$. By considering the elements $-1_{\{\omega_i\}} \in C$ we see that every $q \in C^0$ has nonnegative coordinates and hence is a nonnegative multiple of a probability distribution. By the lemma again this probability measure belongs to $\mathcal{M}^a$. We conclude that $\text{cone}(\mathcal{M}^a) = C^0$. Now C^0 is closed under multiplication with positive scalars and hence $C = C^{00} = \{g \in \mathbb{R}^n : \langle g, q \rangle \leq 0 \text{ for all } q \in C^0\}$. Combined with the preceding observations we obtain $C = \{g \in \mathbb{R}^n : \mathbb{E}_{\mathbb{Q}} g \leq 0 \text{ for all } \mathbb{Q} \in \mathcal{M}^a\}$. By Exercise 8 we have that $\mathcal{M}^e$ is dense in $\mathcal{M}^a$ and hence

$$C = \{g \in \mathbb{R}^n : \mathbb{E}_{\mathbb{Q}} g \leq 0 \text{ for all } \mathbb{Q} \in \mathcal{M}^e\}. \tag{2.3}$$

The proof is completed by using the fact that $K = C \cap (-C)$ (Exercise 9). $\square$

Corollary 2.4.4 (Completeness). *Assume there are no arbitrage opportunities.*

(i) The model is complete if and only if the equivalent martingale measure is unique.

(ii) In case of completeness the representation $f = a + f_0$ with $a \in \mathbb{R}$ and $f_0 \in K$ of a claim $f \in L^\infty$ is unique.

Proof. (i). Suppose first that $\mathcal{M}^e = \{\mathbb{Q}\}$ and take $f \in L^\infty$. By Theorem 2.4.3 $f - \mathbb{E}_{\mathbb{Q}} f \in K$, hence f is attainable. Conversely, suppose we have $\mathbb{Q}_1 \neq \mathbb{Q}_2$ in $\mathcal{M}^e$. Then there exists an $f \in L^\infty$ such that $\mathbb{E}_{\mathbb{Q}_1} f \neq \mathbb{E}_{\mathbb{Q}_2} f$. If this f were attainable, there would exist an $a \in \mathbb{R}$ such that $f - a \in K$. By Theorem 2.4.3 this would imply that $\mathbb{E}_{\mathbb{Q}_1} f = a = \mathbb{E}_{\mathbb{Q}_2} f$, a contradiction.

(ii). Exercise 10. $\square$

2.5 Change of numerair

Recall that in our model we have $d+1$ traded assets and the processes $S^{(0)},\ldots,S^{(d)}$ are the prices of these assets relative to the price of the 0-th asset, the so-called *numerair*. Intuitively, absence or presence of arbitrage should not depend on the choice of numerair. In this section we prove that this is indeed the case.

Any asset with a strictly positive price at all times could be taken as a numerair. More generally, we shall allow any self-financing portfolio of assets with a strictly positive value at all times. Let φ be a predictable process and consider the value process $V = 1 + \varphi \cdot S$. Assume that almost surely $V_t > 0$ for all t. We can view this portfolio as a traded asset and use it to express the value of our $d+1$ original assets. The new value process $\tilde{S} = (\tilde{S}^{(0)},\ldots,\tilde{S}^{(d)})$ is given by

$$\tilde{S}^{(i)} = \frac{S^{(i)}}{V}, \qquad i = 0,\ldots,d.$$

Theorem 2.5.1 (Change of numerair). *Suppose the model S admits no arbitrage opportunities. Then the model $\tilde{S}$ admits no arbitrage opportunities either. It holds that $\mathbb{Q} \in \mathcal{M}^e(S)$ if and only the measure $\tilde{\mathbb{Q}}$ defined by $d\tilde{\mathbb{Q}} = V_T\, d\mathbb{Q}$ belongs to $\mathcal{M}^e(\tilde{S})$.*

Proof. We claim that that $K(\tilde{S}) = V_T^{-1}K(S)$. To see this, first observe that, for every i,

$$\Delta\tilde{S}^{(i)}_t = \frac{1}{V_t}\Big(\Delta S^{(i)}_t - \tilde{S}^{(i)}_{t-1}\Delta V_t\Big).$$

It follows that for a given predictable process ψ, we have

$$(\psi\cdot\tilde{S})_T = \sum_t \frac{f_t}{V_t},$$

with f_t an $\mathcal{F}_t$-measurable element of $K(S)$, for every t. By the lemma following the theorem it holds that $f_t/V_t = g_t/V_T$ for certain $g_t \in K(S)$. Hence, we have the inclusion $K(\tilde{S}) \subseteq V_T^{-1}K(S)$. The converse inclusion follows by symmetry, by considering the model $\tilde{S}$ and taking $1/V$ as numerair.

It now follows from the definition of arbitrage that the model S is free of arbitrage if and only if this holds for $\tilde{S}$. To complete the proof, take an equivalent probability measure $\mathbb{Q}$. By Lemma 2.4.2 and Exercise 6 it holds that $\mathbb{Q} \in \mathcal{M}^e(S)$ if and only if $\mathbb{E}_{\mathbb{Q}}f = 0$ for all $f \in K(S)$. By the first part of the proof, this holds if and only if $\mathbb{E}_{\mathbb{Q}}V_T f$ for all $f \in K(\tilde{S})$, which is the same as saying that the measure $\tilde{\mathbb{Q}}$ defined by $d\tilde{\mathbb{Q}} = V_T d\mathbb{Q}$ belongs to $\mathcal{M}^e(\tilde{S})$. $\square$

Lemma 2.5.2. *For f an $\mathcal{F}_t$-measurable element of $K(S)$ and $t \leq T$, it holds that $(V_T/V_t)f \in K(S)$.*

Proof. Observe that

$$\frac{V_T}{V_t} f = f + \sum_{s=t+1}^{T} \frac{f}{V_t} \Delta V_s = f + (\psi \cdot S)_T,$$

where

$$\psi_s = \frac{f}{V_t} \varphi_s 1_{\{s \geq t+1\}}.$$

Since f is $\mathcal{F}_t$-measurable the process φ is predictable, and it follows that $(V_T/V_t)f \in K(S)$. □

2.6 No-arbitrage pricing

Suppose that in our market we can buy the claim $f \in L^\infty$ at price a at time $t = 0$. Then the collection of all claims that we can attain with 0 initial cost changes from K to

$$K^{f,a} = \operatorname{span}(K \cup \{f - a\}).$$

In case of no-arbitrage it should hold, as before, that $K^{f,a} \cap L_+^\infty = \{0\}$. This leads us to the following definition.

Definition 2.6.1. We call $a \in \mathbb{R}$ an *arbitrage free price* for the claim $f \in L^\infty$ if $K^{f,a} \cap L_+^\infty = \{0\}$.

Observe that if the claim $f \in L^\infty$ is attainable at price a, i.e. $f - a \in K$, then $K^{f,a} = K$. In the absence of arbitrage we have $K \cap L_+^\infty = \{0\}$ and hence in this case a is an arbitrage-free price for f. Moreover, any other value $b \neq a$ is not an arbitrage-free price for f (Exercise 11).

Theorem 2.6.2 (No-arbitrage pricing). *Assume absence of arbitrage and let $f \in L^\infty$. Define $I = \{\mathbb{E}_\mathbb{Q} f \mid \mathbb{Q} \in \mathcal{M}^e\}$. The set I is the collection of all arbitrage-free prices for f. Either $I = \{a\}$, in which case f is attainable at price a, or I is a bounded, open interval, in which case f is not attainable.*

Proof. Define

$$\underline{\pi}(f) = \inf\{\mathbb{E}_\mathbb{Q} f \mid \mathbb{Q} \in \mathcal{M}^e\}, \qquad \overline{\pi}(f) = \sup\{\mathbb{E}_\mathbb{Q} f \mid \mathbb{Q} \in \mathcal{M}^e\}.$$

Suppose $\underline{\pi}(f) = \overline{\pi}(f) = a$. Then by Theorem 2.4.3, $f - a \in K$, which means that f is attainable at price a. Hence, by the remarks preceding the theorem, a is the unique arbitrage-free price for f.

Now assume $\underline{\pi}(f) < \overline{\pi}(f)$. Since f is bounded and $\mathcal{M}^e$ is convex, $I = \{\mathbb{E}_{\mathbb{Q}} f \,|\, \mathbb{Q} \in \mathcal{M}^e\}$ is a bounded subinterval of $\mathbb{R}$. Suppose $a \in I$. Then there is a $\mathbb{Q} \in \mathcal{M}^e$ such that $\mathbb{E}_{\mathbb{Q}}(f-a) = 0$. This implies that $K^{f,a} \cap L^\infty_+ = \{0\}$ (check). Hence, a is an arbitrage-free price for f. Conversely, suppose that $K^{f,a} \cap L^\infty_+ = \{0\}$. Then by repeating the proof of Theorem 2.2.3 with $K^{f,a}$ in the place of K we find a $\mathbb{Q} \in \mathcal{M}^e$ such that $\mathbb{E}_{\mathbb{Q}} g = 0$ for all $g \in K^{f,a}$. In particular, we see that $\mathbb{E}_{\mathbb{Q}}(f-a) = 0$, hence $a \in I$. It remains to show that I is an open interval. Set $a = \overline{\pi}(f)$, the right boundary point of I, and consider the claim $f - a$. By definition we have $\mathbb{E}_{\mathbb{Q}}(f-a) \leq 0$ for all $\mathbb{Q} \in \mathcal{M}^e$. Since we have the representation (2.3) for the set C, it follows that $f - a \in C$. Hence there exists a $g \in K$ such that $g \geq f - a$. Now suppose that $a \in I$. Then there exists a $\mathbb{Q}^* \in \mathcal{M}^e$ such that $\mathbb{E}_{\mathbb{Q}^*} f = a$ and hence $\mathbb{E}_{\mathbb{Q}^*}(g - (f-a)) = 0$, so that $g = f - a$. But this means that $f - a \in K$, i.e. f is attainable at price a. Theorem 2.4.3 implies that I reduces to a singleton in that case, which leads to a contradiction. We conclude that the right endpoint of I does not belong to I. The left endpoint can be handled similarly (or by considering the claim $-f$). □

2.7 Example: binomial model

Consider a world with n time points, $t = 0, \ldots, n$. In this world there exists a bank where money can be deposited or borrowed at interest rate $r > 0$ and a stock is traded. We denote the bank account process by B, so $B_t = (1+r)^t$. The value X_0 of the stock at time 0 equals 1 and given the stock has value X_t at time t, the value X_{t+1} at time $t+1$ equals uX_t with probability $p \in (0,1)$ and dX_t with probability $1-p$, respectively, where $d < u$ are certain given numbers. This model is called the *binomial model*.

Let us choose the bank account process B as numerair and denote the discounted price processes by $S^{(0)}$ and $S^{(1)}$, so $S^{(0)} \equiv 1$ and $S^{(1)}_t = X_t/B_t$. Then under the objective probability measure $\mathbb{P}$ described above we have $S^{(1)}_0 = 1$ and given $S^{(1)}_t$ we have that $S^{(1)}_{t+1}$ equals $(u/(1+r))S^{(1)}_t$ with probability p and $(d/(1+r))S^{(1)}_t$ with probability $1-p$.

We want to investigate the existence of arbitrage opportunities in this model. According to Theorem 2.3.1, it suffices to consider the one-period model we studied in Chapter 1. Proposition 1.0.1 (applied with $d/(1+r)$ in the place of d and $u/(1+r)$ in the place of u) then implies that the binomial model is free of arbitrage if and only if $d < 1+r < u$. The considerations following Proposition 1.0.1 show that the one-period model admits a unique martingale measure, described by changing the probability with which the stock moves up from p to $q = (1+r-d)/(u-d)$. This implies that the full, multi-period model also admits only one martingale measure $\mathbb{Q}$. Hence, by Corollary 2.4.4, the model is complete.

In this complete model every claim $f \in L^\infty$ is attainable and by Theorem 2.6.2 its arbitrage-free price is given by the expectation of f under the martingale measure. Recall that all of this is still relative to the chosen numerair, the bank account process B. Hence, a claim f should be interpreted as a pay-off of at time T of f units of the bank account process. In ordinary money terms, this is a pay-off of $fB_T = f(1+r)^T$ euros at time T. At time 0 this has the value of $\mathbb{E}_{\mathbb{Q}} f$ units of the bank account process. But B_0 equals one euro, and hence a pay-off of $f(1+r)^T$ euros at time T is worth $\mathbb{E}_{\mathbb{Q}} f$ euros at time 0. In other words, a pay-off of f euros at time T is worth $\mathbb{E}_{\mathbb{Q}}(1+r)^{-T} f$ euros at time 0.

Putting this together we obtain the following result.

Proposition 2.7.1. *The binomial model is free of arbitrage if and only $d < 1 + r < u$. In this case the model is complete and the arbitrage-free value at time 0 of a claim paying $f \in L^\infty$ euros at time T is $\mathbb{E}_{\mathbb{Q}}(1+r)^{-T} f$, where $\mathbb{Q}$ is the probability measure obtained by changing the probability with which the stock price moves up from p to $q = (1+r-d)/(u-d)$.*

2.8 Exercises

1. If S is a (multi-dimensional) martingale and φ is a (multi-dimensional) bounded, predictable process, then $\varphi \cdot S$ is a martingale.

2. In the proof of Theorem 2.2.3, show that the fact that K is a linear space implies that we can take $\alpha = 0$ in (2.2).

3. If $\mathbb{E}_{\mathbb{Q}}(\varphi \cdot S)_T = 0$ for every bounded, predictable process φ, then S is a $\mathbb{Q}$-martingale.

4. In the proof of Theorem 2.3.1, show that $\mathbb{Q} \in \mathcal{M}^e$.

5. Show that $K \cap L^\infty_+ = \{0\}$ if and only if $C \cap L^\infty_+ = \{0\}$.

6. Show that under no-arbitrage, $K = C \cap (-C)$.

7. Let K be a linear subspace of $\mathbb{R}^n$ and $L_+ = \{x \in \mathbb{R}^n : x_1 \geq 0, \ldots, x_n \geq 0\}$.

   (a) For $v_1, \ldots, v_m \in \mathbb{R}^n$, show that the convex cone $C = \{\sum c_i v_i : c_1 \geq 0, \ldots, c_m \geq 0\}$ is closed. (Hint: Write $C = \{ax : a \geq 0, x \in H\}$, where $H = \{\sum c_i v_i : c_1 \geq 0, \ldots, c_m \geq 0, \sum c_i = 1\}$ is the convex hull of the v_i, and first prove that H is compact.)

   (b) Show that $K + L_+$ is closed. (Hint: Say $\dim(K) = k$ and let $f_1, \ldots, f_n$ be an orthonormal basis of $\mathbb{R}^n$ such that $f_1, \ldots, f_k$ is an orthonormal basis of K. Using (a), show that the coordinates of the points of $K + L_+$ relative to this basis form a closed subset of $\mathbb{R}^n$.)

8. Show that under no-arbitrage, $\mathcal{M}^e$ is dense in $\mathcal{M}^a$. (Here we identify probability measures on Ω with points in $\mathbb{R}^n$ again.)

9. Supply the details of the last part of the proof of Theorem 2.4.3.

10. Prove Corollary 2.4.4.(ii).

11. Assume absence of arbitrage. Show that for a claim $f \in L^\infty$ that is attainable at price a, the value a is the unique arbitrage-free price.

12. Consider a one-period model with a bank with zero interest and a stock which has value 1 at time 0 and value s_1, s_2 or s_3, respectively, at time 1, with probabilities p_1, p_2 or p_3, respectively. Assume that $s_1 > s_2 > s_3$ and the p_i's are non-zero and add up to 1.

    (i) Give conditions on the values s_1, s_2, s_3 characterizing absence of arbitrage.

    (ii) In case of absence of arbitrage, investigate whether the model is complete or not.

3

The Kreps-Yan theorem

3.1 Description of the model

Let $S = (S_t)_{t\in[0,T]}$ be a $(d+1)$-dimensional, cadlag, adapted, locally bounded stochastic process, with $T > 0$ a fixed time horizon, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with a filtration $(\mathcal{F}_t)_{t\in[0,T]}$ satisfying the usual conditions. As in the preceding chapter, we think of $S = (S^{(0)}, \dots, S^{(d)})$ as describing the value of $d+1$ financial assets, expressed relative to a chosen numerair, which is the 0-th asset. In particular, we assume that $S^{(0)} \equiv 1$.

By the classical theorems on regularity of continuous-time martingales, the usual conditions on the filtration imply that (local) martingales have a cadlag modification. Since the basic theorems of martingale theory are valid for right-continuous martingales, the usual conditions ensure that we may apply the classical theorems to the martingales that we encounter.

Observe that our setup includes the discrete-time case, simply take the process S (and the filtration) to be constant on the intervals $[t-1, t)$ for integers t. If the underlying probability space is finite the process S is necessarily uniformly bounded, and hence the present setup includes the setting of the preceding chapter.

A first important question is which trading strategies we want to allow in this model. At the very least we shall allow strategies in which the asset portfolio is only rebalanced at a finite number of stopping times, in a predictable manner.

Definition 3.1.1. A d-dimensional process φ is called a *simple trading strategy* if it is of the form

$$\varphi_t = \sum_{i=1}^{n} \varphi_i 1_{(\tau_{i-1},\tau_i]}(t),$$

where $0 = \tau_0 \leq \cdots \leq \tau_n \leq T$ are finite stopping times and the φ_i are d-dimensional, $\mathcal{F}_{\tau_{i-1}}$-measurable random variables.

The strategy is called *admissible* if, in addition, the stopped process S^{τ_n} and the random variables $\varphi_1, \dots, \varphi_n$ are uniformly bounded.

The interpretation of the definition is clear: $\varphi_i^{(j)}$ is the number of assets of type j in the portfolio between times τ_{i-1} and τ_i. As in the preceding chapter we define the stochastic process $\varphi \cdot S$ by

$$(\varphi \cdot S)_t = \sum_{i=1}^{n} \left\langle \varphi_i, S_{\tau_i \wedge t} - S_{\tau_{i-1} \wedge t} \right\rangle, \qquad t \in [0, T].$$

As before, $\varphi \cdot S$ should be interpreted as the value process of a self-financing portfolio starting with 0 initial capital and following the trading strategy φ, the adjustments of the positions in assets 1 to d being financed by taking the appropriate amount from the "bank account", modelled by the numerair asset 0.

Our first notion of no-arbitrage in this setting is the requirement that we can not make a risk-free profit by following a simple, admissible strategy. We proceed analogously to the finite case and first define the space

$$K^s = \{(\varphi \cdot S)_T \,|\, \varphi \text{ simple, admissible}\}$$

of pay-offs that can be achieved with 0 initial capital, following a simple, admissible self-financing trading strategy.

Definition 3.1.2. We say that S satisfies the condition of *no-arbitrage with simple strategies* if $K^s \cap L_+^\infty(\Omega, \mathcal{F}, \mathbb{P}) = \{0\}$.

A sufficient condition for the absence of arbitrage with simple strategies is the existence of a so-called *equivalent local martingale measure*. This is, by definition, a probability measure $\mathbb{Q}$ equivalent to the objective measure $\mathbb{P}$ such that S is a local martingale under $\mathbb{Q}$.

Proposition 3.1.3. *If there exists an equivalent local martingale measure, S admits no arbitrage with simple strategies.*

Proof. Let $\mathbb{Q}$ be an equivalent local martingale measure. We first show that for every simple, admissible strategy φ it holds that

$$\mathbb{E}_{\mathbb{Q}}(\varphi \cdot S)_T = 0.$$

By Definition 3.1.1 it suffices to show that if $0 \leq \sigma \leq \tau \leq T$ are stopping times such that S^τ is bounded and X is a bounded, d-dimensional, $\mathcal{F}_\sigma$-measurable random variable, then

$$\mathbb{E}_{\mathbb{Q}} \langle X, S_\tau - S_\sigma \rangle = 0.$$

This can be derived from the optional stopping theorem (Exercise 1).

Now suppose we have $f \in K^s$, $f \geq 0$, say $f = (\varphi \cdot S)_T$ for a simple, admissible φ. Then by the first part of the proof we have $\mathbb{E}_{\mathbb{Q}} f = 0$. Since f is nonnegative it follows that f vanishes $\mathbb{Q}$-a.s. and hence also $\mathbb{P}$-a.s., since $\mathbb{P}$ and $\mathbb{Q}$ are equivalent. $\square$

The converse of this proposition is, unfortunately, not true. To have the existence of a (local) martingale measure in this general continuous-time setting, the absence of simple arbitrages is not strong enough.

Example 3.1.4. Consider a process $S = (S_t)_{t\in[0,1]}$ with $S_0 = 1$ and that is constant except for jumps at times $t_n = 1 - (n+1)^{-1}$ for $n = 1, 2, \ldots$. At time t_n the process S has a jump of magnitude $3^{-n}Z_n$, where $Z_1, Z_2, \ldots$ are independent random variables with $\mathbb{P}(Z_n = 1) = 1 - \mathbb{P}(Z_n = -1) = 1/2 + \varepsilon_n$ for certain numbers $\varepsilon_n \in (-1/2, 1/2)$. Since the process S is uniformly bounded, it is a martingale under a measure $\mathbb{Q}$ if it local martingale (check!). But there is only one probability measure under which S is a martingale, namely the measure $\mathbb{Q}$ under which $\mathbb{Q}(Z_n = 1) = 1 - \mathbb{Q}(Z_n = -1) = 1/2$. Hence, by Example B.2.4, there exists no equivalent local martingale measure if we choose the ε_n such that $\sum \varepsilon_n^2 = \infty$. However, this model does satisfy the condition of no-arbitrage with simple strategies. To see that, first observe that if a there is a simple admissible arbitrage strategy, then there is simple arbitrage strategy of the form $\varphi = h1_{(\sigma,\tau]}$, for stopping times $\sigma \le \tau \le 1$ and a bounded $\mathcal{F}_\sigma$-measurable random variable h (Exercise 2). Moreover, we only have to consider stopping times that take values in the collection of t_n's. Such a strategy has pay-off $V = h(S_\tau - S_\sigma)$. Now observe that on the event $A_n = \{\sigma = t_{n-1}, \tau \ge t_n\}$ we have $\text{sign}(S_\tau - S_\sigma) = \text{sign}(Z_n) = Z_n$, so $\text{sign}(V) = \text{sign}(h)Z_n$. By assumption $\text{sign}(V) \ge 0$, so we get that

$$\text{sign}(h)1_{A_n} Z_n \ge 0.$$

Note that $A_n \in \mathcal{F}_{t_{n-1}}$ and hence, by definition of $\mathcal{F}_\sigma$, $\text{sign}(h)1_{A_n}$ is $\mathcal{F}_{t_{n-1}}$-measurable. So $\text{sign}(h)1_{A_n}$ and Z_n are independent and in view of the preceding display, it follows that $\text{sign}(h)1_{A_n} = 0$ a.s. (check) . Hence $h = 0$ on every event A_n and therefore $h = 0$ on the event $\{\tau > \sigma\} = \cup_n A_n$. ■

3.2 Kreps-Yan theorem

As in Chapter 2 we can introduce the cone

$$C^s = \{f \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}) : \text{there exists a } g \in K^s \text{ such that } g \ge f\}.$$

Then (see Exercise 5 of Chapter 2) no-arbitrage with simple integrands is equivalent to $C^s \cap L^\infty_+(\Omega, \mathcal{F}, \mathbb{P}) = \{0\}$. In the preceding section we remarked that this condition is not strong enough to guarantee the existence of an equivalent martingale measure. It turns out we have to replace the cone C^s by its closure $\overline{C}$ with respect to the weak*-topology on $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, the latter space viewed as the dual of $L^1(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 3.2.1. We say that S satisfies the condition of *no free lunch* if $\overline{C} \cap L^\infty_+(\Omega, \mathcal{F}, \mathbb{P}) = \{0\}$.

Clearly this condition is stronger than the condition of no-arbitrage with simple strategies. It gives us the following version of the fundamental theorem of asset pricing.

Theorem 3.2.2 (Fundamental theorem of asset pricing II). *The process S satisfies the condition of no free lunch if and only if there exists an equivalent local martingale measure.*

Proof. Suppose first that there exists an equivalent local martingale measure $\mathbb{Q}$. Then by the first part of the proof of Proposition 3.1.3 it holds that $\mathbb{E}_{\mathbb{Q}} f \leq 0$ for all $f \in C^s$. Since the map $f \mapsto \mathbb{E}_{\mathbb{Q}} f$ is weak*-continuous (check), the inequality $\mathbb{E}_{\mathbb{Q}} f \leq 0$ holds in fact for every $f \in \overline{C}$. It follows that if $f \in \overline{C}$, $f \geq 0$, then $\mathbb{E}_{\mathbb{Q}} f = 0$, so that f vanishes $\mathbb{Q}$-a.s. and hence also $\mathbb{P}$-a.s..

Now suppose that S satisfies the condition of no free lunch. For $\delta \in (0,1)$, define $B_\delta = \{f \in L^\infty : 0 \leq f \leq 1,\ \mathbb{E} f \geq \delta\}$. The set B_δ is a weak*-closed subset of the unit ball of L^∞, the latter space viewed as dual of L^1 (Exercise 4). Hence, by Alaoglu's theorem it is weak*-compact. Clearly, it is also convex. By the separation theorem there exists a $g_\delta \in L^1$ and $\alpha, \beta \in \mathbb{R}$ such that

$$\sup_{f \in \overline{C}} \mathbb{E} g_\delta f \leq \alpha < \beta \leq \inf_{h \in B_\delta} \mathbb{E} g_\delta h.$$

Since $0 \in C$ we have $\alpha \geq 0$. Since $\overline{C}$ is a cone, it follows that $\mathbb{E} g_\delta f \leq 0$ for all $f \in \overline{C}$, hence

$$\sup_{f \in \overline{C}} \mathbb{E} g_\delta f \leq 0 < \inf_{h \in B_\delta} \mathbb{E} g_\delta h.$$

Since C contains all negative functions in L^∞ we have $g_\delta \geq 0$ a.s.. Also observe that $1 \in B_\delta$, so that $\mathbb{E} g_\delta > 0$, and therefore g_δ does not vanish almost surely. We renormalize g_δ such that $\mathbb{E} g_\delta = 1$.

For every positive integer n we now define the probability measure $\mathbb{Q}_n$ by $d\mathbb{Q}_n = g_{2^{-n}}\, d\mathbb{P}$ and $\mathbb{Q} = \sum a_n \mathbb{Q}_n$, for a sequence a_n of positive numbers such that $\sum a_n = 1$. Note that if $\mathbb{P}(A) > 2^{-n}$ then $1_A \in B_{2^{-n}}$ and hence $\mathbb{Q}_n(A) > 0$. It follows that $\mathbb{P}$ is absolutely continuous with respect to $\mathbb{Q}$ (check). The fact that $\mathbb{Q}$ is absolutely continuous relative to $\mathbb{P}$ is immediate, and hence $\mathbb{P}$ and $\mathbb{Q}$ are equivalent. To complete the proof observe that if $f \in C$, then $\mathbb{E}_{\mathbb{Q}} f = \sum a_n \mathbb{E}_{\mathbb{Q}_n} f = \sum a_n \mathbb{E} g_{2-n} f \leq 0$. This implies that S is a local martingale under $\mathbb{Q}$ (Exercise 3). □

3.3 Discussion

We saw in this chapter that in the general continuous-time setting, absence of arbitrage with simple strategies is not sufficient for the existence of a (local) martingale measure. The collection of pay-offs that can be attained with simple strategies is somehow to small, and it turned out to be necessary to take the weak*-closure of this set. While this is completely satisfactory from a mathematical point of view, we should observe that it destroys the economic meaning of Theorem 3.2.2. Since taking the weak*-closure is not a very intuitive operation, the weak*-closure of a collection of pay-offs of simple strategies is *not* a set that can for instance be interpreted as a collection of pay-offs of "more complex" strategies.

To obtain an economically meaningful result, we would prefer to replace the weak*-topology by a stronger, more intuitive one. It turns out that this is possible if we are willing to restrict ourselves to asset prices that are semimartingales. Then the class of simple strategies can be enlarged in a natural way, using the theory of integration with respect to semimartingales. Taking the closure with respect to weak*-topology can then be replaced by taking the closure in the norm-topology of L^∞, which is much more satisfactory from the economic perspective.

3.4 Exercises

1. Complete the first part of the proof of Proposition 3.1.3.
2. Show that if there exists a simple arbitrage strategy, there also exists an arbitrage strategy of the form $\varphi = h1_{(\sigma,\tau]}$ with $\sigma \le \tau \le 1$ stopping times and h a bounded, $\mathcal{F}_\sigma$-measurable random variable. (Hint: use induction on the number of stopping times in the given simple strategy.)
3. Show that if the equivalent measure $\mathbb{Q}$ satisfies $\mathbb{E}_{\mathbb{Q}} f \le 0$ for all $f \in C$, then it is an equivalent local martingale measure.
4. Show that the set B_δ defined in the proof of Theorem 3.2.2 is a weak*-closed subset of the unit ball of $(L^1)^*$.

4

The general theorem

4.1 Preliminaries on stochastic integration

In this chapter we will assume that the asset price process S is a one-dimensional semimartingale, defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying the usual conditions. We will consider all processes on a finite time interval $[0, T]$, with the time horizon $T > 0$ fixed. As before, S is interpreted as the value of an asset relative to a numerair asset, the numerair asset itself has the constant value 1.

We will interpret a predictable process φ as a trading strategy, φ_t denoting the number of non-numerair assets that we hold at time t. If φ is a simple process of the form

$$\varphi = \sum \varphi_i 1_{(t_{i-1}, t_i]}(t),$$

for $0 = t_0 < \cdots < t_n = T$ a deterministic partition of $[0, T]$ and φ_i bounded and $\mathcal{F}_{t_{i-1}}$-measurable, then, as explained in Chapter 2, the process

$$(\varphi \cdot S)_t = \sum \varphi_i (S_{t_i \wedge t} - S_{t_{i-1} \wedge t}), \qquad t \in [0, T],$$

can be interpreted as the value process of a self-financing portfolio starting with 0 initial capital and following the trading strategy φ, the adjustments of the position of the non-numerair assets being financed by taking the appropriate amount from the "bank account", modelled by the numerair asset.

As the notation suggests, $\varphi \cdot S$ is exactly the stochastic integral process of the locally bounded, predictable process φ relative to the semimartingale S. Hence, using stochastic integration theory we can now go beyond simple strategies. However, to retain an economically meaningful theory we have to verify that for non-simple predictable processes, $\varphi \cdot S$ can still be interpreted as the value process associated to the trading strategy φ. Since we are now considering strategies that can involve continuous trading, we should consider approximations to make this precise. For a fine enough partition $0 = t_0 < \cdots <$

$t_n = T$ of $[0,T]$, the strategy φ is well approximated by the simple strategy $\sum \varphi_{t_{i-1}} 1_{(t_{i-1},t_i]}$ and the value process associated to this simple strategy equals

$$\sum \varphi_{t_{i-1}}(S_{t_i\wedge\cdot} - S_{t_{i-1}\wedge\cdot}).$$

Hence, we want the latter process to be a good approximation of the integral process $\varphi \cdot S$. Thanks to the continuity property of the stochastic integral this is indeed the case.

Lemma 4.1.1. *Let φ be a left-continuous process. Then if $0 = t_0^n < \cdots < t_{k_n}^n = T$ is a sequence of partitions of $[0,T]$ with mesh tending to 0, we have*

$$\sup_{t\in[0,T]} \Big| \sum \varphi_{t_{i-1}^n}(S_{t_i^n\wedge t} - S_{t_{i-1}^n\wedge t}) - (\varphi\cdot S)_t \Big| \xrightarrow{\mathbb{P}} 0.$$

Proof. For every n, define the process φ^n by

$$\varphi^n = \sum \varphi_{t_{i-1}^n} 1_{(t_{i-1}^n, t_i^n]}.$$

Then φ^n is left-continuous and adapted, hence locally bounded and predictable. Since φ is left-continuous it holds that $\varphi^n \to \varphi$ on $[0,T]\times\Omega$, and we have

$$|\varphi_t^n| \leq \sup_{s\leq t} |\varphi_s|.$$

The process on the right-hand side of the display is adapted and left-continuous, hence predictable and locally bounded (cf. Exercise 5.53 in Van der Vaart's notes). The conclusion of the lemma now follows from Lemma 5.52 of Van der Vaart's notes. □

Unfortunately, the definition of the stochastic integral of locally bounded predictable processes with respect to semimartingales is still not general enough for the next version of the FTAP. To extend the integral we first endow the space of semimartingales with a topology. We denote by $\mathcal{S}(\mathbb{P})$ the space of all $\mathbb{P}$-semimartingales on our fixed filtered probability space. For $X \in \mathcal{S}(\mathbb{P})$ we define

$$\|X\|_{\mathcal{S}(\mathbb{P})} = \sup_{|H|\leq 1} \mathbb{E}(|(H\cdot X)_T| \wedge 1),$$

where the supremum is over all predictable processes H that are uniformly bounded by 1. It is easy to see that $\|\cdot\|_{\mathcal{S}(\mathbb{P})}$ satisfies the triangle inequality (check). It follows that we can define a metric d on $\mathcal{S}(\mathbb{P})$ by setting $d(X,Y) = \|X-Y\|_{\mathcal{S}(\mathbb{P})}$. The topology that this metric generates on $\mathcal{S}(\mathbb{P})$ is called the *semimartingale topology*. Observe that $X^n \to X$ in this topology if and only if

$$(H\cdot X^n)_T \xrightarrow{\mathbb{P}} (H\cdot X)_T$$

for all uniformly bounded predictable processes H (Exercise 1).

We can now use the semimartingale topology to extend the definition of the stochastic integral. For a predictable process H and a positive integer n the process $H1_{\{|H|\leq n\}}$ is bounded and predictable. Hence if X is a semimartingale the stochastic integral process $H1_{\{|H|\leq n\}} \cdot X$ is well defined in the sense of integration of locally bounded predictable processes (see for instance Van der Vaart's notes). If $X \in \mathcal{S}(\mathbb{P})$, the sequence of processes $H1_{\{|H|\leq n\}} \cdot X$ belongs to $\mathcal{S}(\mathbb{P})$ as well. If the sequence has a limit in $\mathcal{S}(\mathbb{P})$ (relative to the semimartingale topology) we say that H is X*-integrable* and we denote the limit semimartingale by $H \cdot X$. Observe that by Lemma 5.52 of Van der Vaart, the new definition of $H \cdot X$ coincides with the old one if H is locally bounded (check).

Since the FTAP involves changes of measure, it is useful to investigate how stochastic integrals depend on the underlying probability measure. We write $(H \cdot X)^{\mathbb{P}}$ if we want to emphasize the dependence of the integral process on $\mathbb{P}$. For a simple predictable process it is clear that the stochastic integral does not depend on the underlying measure at all. Now let H be a nonnegative, bounded, predictable process and X a $\mathbb{P}$-semimartingale and let $\mathbb{Q}$ be equivalent to $\mathbb{P}$. A general version of Girsanov's theorem says that for equivalent probability measures $\mathbb{P}$ and $\mathbb{Q}$, the spaces $\mathcal{S}(\mathbb{P})$ and $\mathcal{S}(\mathbb{Q})$ of $\mathbb{P}$- and $\mathbb{Q}$-semimartingales coincide, hence X is a $\mathbb{Q}$-semimartingale as well. Now by standard measure theory there exist a sequence of H_n of simple predictable processes, independent of $\mathbb{P}$, such that $H_n \uparrow H$ on $[0,\infty) \times \Omega$. By Lemma 5.52 of Van der Vaart we have that

$$(H_n \cdot X)_t \xrightarrow{\mathbb{P}} (H \cdot X)_t^{\mathbb{P}}$$

for all $t \geq 0$. Hence, there exists a sequence $k_n \to \infty$ such that $(H_{k_n} \cdot X)_t \to (H \cdot X)_t^{\mathbb{P}}$, $\mathbb{P}$-a.s.. Repeating the argument with $\mathbb{Q}$ instead of $\mathbb{P}$ we see that k_n has a further subsequence l_n such that $(H_{l_n} \cdot X)_t \to (H \cdot X)_t^{\mathbb{Q}}$, $\mathbb{Q}$-a.s.. But $\mathbb{P}$ and $\mathbb{Q}$ are equivalent, so $(H \cdot X)_t^{\mathbb{P}} = (H \cdot X)_t^{\mathbb{Q}}$ almost surely (relative to $\mathbb{P}$ or $\mathbb{Q}$). We conclude that for bounded predictable H and $\mathbb{P}$ and $\mathbb{Q}$ equivalent, $(H \cdot X)^{\mathbb{P}}$ and $(H \cdot X)^{\mathbb{Q}}$ are indistinguishable. It can be shown that if $\mathbb{P}$ and $\mathbb{Q}$ are equivalent, the semimartingale topologies induced on $\mathcal{S}(\mathbb{P}) = \mathcal{S}(\mathbb{Q})$ by $\mathbb{P}$ and $\mathbb{Q}$ coincide. For a predictable process H we just observed that $H1_{\{|H|\leq n\}} \cdot X$ does not depend on the probability measure. It follows that whether or not a predictable process H is X-integrable only depends on the equivalence class of the underlying probability measure $\mathbb{P}$, and the same holds for the integral processes $H \cdot X$.

Some care should be taken with integrands that are not locally bounded. For the extended integral it is for instance no longer true that the integral with respect to a local martingale is again a local martingale.

Example 4.1.2. Suppose we have a standard exponential random variable τ and, independent of τ, a standard Bernoulli variable B, i.e. $\mathbb{P}(B = -1) = \mathbb{P}(B = 1) = 1/2$. Define the process M by $M_t = B1_{\{t\geq\tau\}}$. Then M is a martingale relative to its natural filtration $(\mathcal{F}_t)$ (Exercise 3). Now define the deterministic process H by $H_t = 1/t$ for $t > 0$. Then it holds that

$$(H1_{\{|H|\leq n\}} \cdot M)_t = \begin{cases} 0, & t < \tau \\ \dfrac{B}{\tau}1_{\{\tau\geq 1/n\}}, & t \geq \tau. \end{cases}$$

It follows that $H1_{\{|H|\leq n\}} \cdot M$ converges in the semimartingale topology to the process X given by

$$X_t = \begin{cases} 0, & t < \tau \\ \dfrac{B}{\tau}, & t \geq \tau, \end{cases}$$

and $X = H \cdot M$ by definition (check!). Observe however that

$$\mathbb{E}|X_t| = \mathbb{E}\frac{1}{\tau}1_{\{t \geq \tau\}} = \int_0^t \frac{1}{x} e^{-x}\, dx = \infty,$$

hence X is not a martingale. It can be shown that X is not a local martingale either (Exercise 4). ■

4.2 No free lunch with vanishing risk and the FTAP

Replacing the general asset price process of the preceding chapter by a semimartingale will allow us to replace the economically meaningless condition of no free lunch by the condition of *no free lunch with vanishing risk*.

As discussed in the preceding section, we think of a predictable process as describing a self-financing trading strategy. We will assume that a trader has a finite credit line, in the sense that her wealth always stays above some deterministic (but possibly very negative) number. This is formalized by the following definition.

Definition 4.2.1. An *a-admissible strategy* is an S-integrable predictable process φ that satisfies $(\varphi \cdot S) \geq -a$. An *admissible strategy* is a predictable process that is a-admissible for some $a > 0$.

In order to formulate the condition of no free lunch with vanishing risk we introduce the sets

$$K = \{(\varphi \cdot S)_T : \varphi \text{ admissible}\},$$

which is a convex cone in the space L^0 of finite-valued random variables (it is not a linear space in general, since admissibility is a one-sided restriction), and $C = \{f \in L^\infty : \text{there exists a } g \in K \text{ such that } g \geq f\}$. By $\overline{C}$ we denote in this chapter the closure of C with respect to the norm-topology of L^∞.

Definition 4.2.2. We say that S satisfies the condition of *no free lunch with vanishing risk* if $\overline{C} \cap L^\infty_+ = \{0\}$.

Observe that this condition has a clear economic interpretation. If S does not satisfy the condition, there exists for every small enough $\varepsilon > 0$ an admissible strategy φ (depending on ε) such that for the pay-off we have $(\varphi \cdot S)_T > -\varepsilon$ and $\mathbb{P}((\varphi \cdot S)_T > 0) > 0$ (Exercise 2). Hence, if we are willing to take an arbitrarily small, but positive loss, we have a positive probability of receiving a strictly positive pay-off.

We can now formulate the following version of the fundamental theorem of asset pricing. The proof is discussed in the next section.

Theorem 4.2.3 (FTAP IIIa). *If $S = (S_t)_{t\in[0,T]}$ is a bounded, real-valued semimartingale, then there exists an equivalent martingale measure if and only if S satisfies the condition of no free lunch with vanishing risk.*

If S is only locally bounded, the martingale measure has to be replaced by a local martingale measure.

Corollary 4.2.4 (FTAP IIIb). *If $S = (S_t)_{t\in[0,T]}$ is a locally bounded, real-valued semimartingale, then there exists an equivalent local martingale measure if and only if S satisfies the condition of no free lunch with vanishing risk.*

Proof. The sufficiency of no free lunch with vanishing risk follows from the preceding theorem. Indeed, suppose it holds and let $\tau_n \uparrow \infty$ be stopping times such that $|S^{\tau_n}| \le K_n$, with K_n deterministic numbers. Define the new process $\tilde{S}$ by

$$\tilde{S} = S1_{[0,\tau_1]} + \sum_{n\ge 2} 2^{-n}\frac{1}{K_n}1_{(\tau_{n-1},\tau_n]}\cdot S.$$

Then $\tilde{S}$ is a bounded semimartingale. Moreover, it satisfies the condition of no free lunch with vanishing risk (Exercise 5). Hence, by the theorem, there exists an equivalent probability measure $\mathbb{Q}$ such that $\tilde{S}$ is a $\mathbb{Q}$-martingale. But then the original process S is a $\mathbb{Q}$-local martingale (Exercise 6).

The converse statement is proved as in Theorem 4.2.3, see the next section. $\square$

4.3 Sketch of proof of the fundamental theorem

4.3.1 The relatively easy half

We noted above that if M is a local martingale and H is M-integrable, then $H \cdot M$ is not necessarily a local martingale anymore. The proof of the fact that the existence of a martingale measure is sufficient for no free lunch with vanishing risk uses a characterization of the martingality of $H \cdot M$. To explain this characterization it is useful to first note that a local martingale M is in fact locally uniformly integrable. Indeed, let τ_n be a localizing sequence for M, i.e. $\tau_n \uparrow \infty$ a.s. and every M^{τ_n} is a martingale. Then $\sigma_n = \tau_n \wedge n$ also satisfies

$\sigma_n \uparrow \infty$ a.s. and every M^{σ_n} is a uniformly integrable martingale. Moreover, if we now consider the stopping time $\pi_n = \sigma_n \wedge \inf\{t : |M_t| > n\}$ we have that

$$\sup_{t \leq \pi_n} |M_t| \leq n + |M_{\pi_n}|.$$

Since $\pi_n \leq \sigma_n$ and M^{σ_n} is UI, the right-hand side of the display is integrable. So we have proved the following useful lemma.

Lemma 4.3.1. *A local martingale M is locally uniformly integrable. Moreover, there exists a localizing sequence τ_n such that*

$$\mathbb{E} \sup_{t \leq \tau_n} |M_t| < \infty$$

for all n.

Now consider a local martingale M and an M-integrable predictable process H and suppose that $H \cdot M$ is a local martingale. For a localizing sequence τ_n such that $\mathbb{E} \sup_{t \leq \tau_n} |(H \cdot M)_t| < \infty$ we have $\sup_t |\Delta(H \cdot M)_t^{\tau_n}| \leq 2 \sup_{t \leq \tau_n} |(H \cdot M)_t|$, so $\Delta(H \cdot M)^{\tau_n} \geq Z_n$, where $Z_n = -2 \sup_{t \leq \tau_n} |(H \cdot M)_t|$. So we see that there exists a localizing sequence τ_n and integrable random variables Z_n such that $\Delta(H \cdot M)^{\tau_n} \geq Z_n$ for all n. It turns out that the converse is true as well.

Theorem 4.3.2. *Let M be a local martingale and H a predictable processes that is M-integrable. Then $H \cdot M$ is a local martingale if and only if there exists a localizing sequence τ_n and integrable random variables Z_n such that $\Delta(H \cdot M)^{\tau_n} \geq Z_n$ for all n.*

We can now prove that the existence of an equivalent martingale measure implies there is no free lunch with vanishing risk. Suppose there exists an equivalent martingale measure $\mathbb{Q}$ and let φ be an a-admissible strategy. By the observations in the preceding section the integral process $\varphi \cdot S$ does not depend on the underlying measure. Under $\mathbb{Q}$ the process S is a local martingale. Now consider the stopping times $\tau_n = \inf\{t : (\varphi \cdot S)_t \geq n\}$. Then $\tau_n \uparrow \infty$ and by the admissibility of φ we have

$$\Delta(\varphi \cdot S)_t^{\tau_n} = (\varphi \cdot S)_t^{\tau_n} - (\varphi \cdot S)_{t-}^{\tau_n} \geq -(n + a).$$

Hence, by the preceding theorem, $\varphi \cdot S$ is a $\mathbb{Q}$-local martingale. Together with the a-admissibility this implies that $\varphi \cdot S$ is in fact a $\mathbb{Q}$-supermartingale (Exercise 7). It follows that for every $f \in C$ we have $\mathbb{E}_{\mathbb{Q}} f \leq 0$ and then also $\mathbb{E}_{\mathbb{Q}} f \leq 0$ for every $f \in \overline{C}$. This implies that every $f \in \overline{C} \cap L_+^\infty$ vanishes $\mathbb{Q}$-a.s., but then also $\mathbb{P}$-a.s..

4.3.2 The much more difficult half

The essential step in the proof of the fact that no free lunch with vanishing risk is sufficient for the existence of a martingale measure is the following theorem.

Theorem 4.3.3. *If the bounded semimartingale S satisfies the condition of no free lunch with vanishing risk, then the cone C is weak*-closed.*

Indeed, if we have this result we can argue as in the proof of Theorem 3.2.2 to find a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that $\mathbb{E}_{\mathbb{Q}} f \leq 0$ for all $f \in C$. It then follows that $\mathbb{E}_{\mathbb{Q}} f = 0$ for all $f \in K$. Since S is bounded it is integrable. Moreover, for $s \leq t$ and $A \in \mathcal{F}_s$ the predictable process $\varphi = 1_{A \times (s,t]}$ is admissible, so $(\varphi \cdot S)_T \in K$ and hence

$$\mathbb{E}_{\mathbb{Q}} 1_A (S_t - S_t) = \mathbb{E}_{\mathbb{Q}} (\varphi \cdot S)_T = 0,$$

which shows that S is a $\mathbb{Q}$-martingale.

The proof of Theorem 4.3.3 is long and difficult. The first difficulty that arises is that the weak*-topology is in general not metrizable. This implies that to show that a set is weak*-closed, it is in general not enough to consider converging sequences (in fact, one should consider *nets*). However, it turns out that in the present context the situation is not that complicated, since we can use the following consequence of the so-called *Krein-Smulian theorem*.

Theorem 4.3.4. *Let $(E, \mathcal{E}, \mu)$ be a finite measure space and $C \subseteq L^\infty(\mu)$ a convex cone. Suppose that for each uniformly bounded sequence f_n in C that converges in measure to a function f, it holds that $f \in C$. Then C is weak*-closed.*

So to prove Theorem 4.3.3 it suffices to consider a sequence h_n in C such that $|h_n| \leq 1$ for all n and $h_n \xrightarrow{\text{as}} h$ for some $h \in L^\infty$, and show that h belongs to C. To prove that $h \in C$ we have to find an $f_0 \in K$ such that $h \leq f_0$. To that end it turns out to be useful to consider the set

$$D = \{f : \text{there exist 1-admissible } \varphi_n \text{ such that } (\varphi_n \cdot S)_T \xrightarrow{\text{as}} f, \ f \geq h\}.$$

It can be shown that this set contains a maximal element f_0. So the random variable f_0 dominates h and is the almost sure limit of elements $(\varphi_n \cdot S)_T$ of K, φ_n 1-admissible.

The remaining task is to show that f_0 belongs to K itself. The first step is the observation that convergence of $\varphi_n \cdot S$ at the terminal time T in fact implies convergence for all time points. To see this one first shows that as $n, m \to \infty$,

$$\sup_{t \in [0,T]} |(\varphi_n \cdot S)_t - (\varphi_m \cdot S)_t| \xrightarrow{\mathbb{P}} 0. \tag{4.1}$$

The proof of this fact uses the 1-admissibility of the φ_n and the maximality of f_0. Next we want to apply some results on the semimartingale topology, but the preceding display does not imply that $\varphi_n \cdot S$ is a Cauchy sequence in the semimartingale topology. A long and technical proof however shows that for every n there exits a ψ_n in the convex hull of the processes $\varphi_n, \varphi_{n+1}, \ldots$ such that $\psi_n \cdot S$ is a Cauchy sequence in the semimartingale topology. The semimartingale topology can be shown to be complete, so that we now have $\psi_n \cdot S \to Z$ for some semimartingale Z. Moreover, *Mémin's theorem* shows that the semimartingale Z must necessarily be of the form $Z = \psi \cdot S$ for some S-integrable predictable process ψ.

Observe that since ψ_n is a convex combination of $\varphi_n, \varphi_{n+1}, \ldots$, the process ψ_n is 1-admissible. Since the convergence in the semimartingale topology implies that

$$(\psi_n \cdot S)_t \xrightarrow{\mathbb{P}} (\psi \cdot S)_t$$

for all $t \in [0, T]$, it follows that ψ is 1-admissible as well. The fact that ψ_n is a convex combination of $\varphi_n, \varphi_{n+1}, \ldots$ also implies that the almost sure limit of $(\psi_n \cdot S)_T$ equals the almost sure limit of $(\varphi_n \cdot S)_T$, which is f_0 (Exercise 8). Combined with the previous observations we conclude that $f_0 = (\psi \cdot S)_T$, so indeed $f_0 \in K$.

4.4 Example: Itô processes

Suppose we have, on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying the usual conditions, continuous adapted processes B and X satisfying

$$B_t = \exp\Big(\int_0^t r_s\, ds\Big),$$

$$dX_t = \mu_t X_t\, dt + \sigma_t X_t\, dW_t,$$

where W is a standard Brownian motion and r, μ and σ are locally bounded predictable processes. All processes are indexed by $[0, T]$ for some fixed time horizon $T > 0$. We think of B as describing a bank account process with (continuous, stochastic) interest rate r_t, and X as describing the value of a stock with local return rate μ_t and (possibly stochastic) volatility σ_t.

We use B as numerair, putting $S = X/B$. Integration by parts then gives the stochastic differential equation

$$dS_t = (\mu_t - r_t) S_t\, dt + \sigma_t S_t\, dW_t$$

for the discounted process S (check). Now suppose that the Sharp ratio

$$\theta_t = \frac{\mu_t - r_t}{\sigma_t}$$

is uniformly bounded by a deterministic constant for all $t \in [0, T]$. Then by the classical Girsanov theorem, there exists a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ under which the process

$$B_t = W_t + \int_0^t \theta_s\, ds, \qquad t \in [0, T],$$

is a Brownian motion. Combining the definition of B with the SDE for S we get

$$dS_t = \sigma_t S_t \, dB_t.$$

In particular, we see that S is a $\mathbb{Q}$-local martingale. Hence, by Corollary 4.2.4, this model satisfies the condition of no free lunch with vanishing risk.

Observe that the classical Black-Scholes model corresponds to the special case that r, μ and σ are deterministic and independent of time. The condition on the Sharp ratio is then trivially satisfied, so we recover the well-known fact that the Black-Scholes model is free of arbitrage (in the sense of no free lunch with vanishing risk).

4.5 Exercises

1. Show that $X_n \xrightarrow{\mathbb{P}} X$ if and only if $\mathbb{E}(|X_n - X| \wedge 1) \to 0$.
2. Show that if S does not satisfy the condition of no free lunch with vanishing risk, there exists for every $\varepsilon > 0$ small enough an admissible strategy φ^ε with a pay-off satisfying $(\varphi^\varepsilon \cdot S)_T > -\varepsilon$ and $\mathbb{P}((\varphi^\varepsilon \cdot S)_T > 0) > 0$.
3. Show that the process M defined in Example 4.1.2 is a martingale.
4. Show that the process X in Example 4.1.2 is not a local martingale.
5. Show that the process $\tilde{S}$ in the proof of Corollary 4.2.4 satisfies the condition of no free lunch with vanishing risk.
6. Show that the process S in the proof of Corollary 4.2.4 is a $\mathbb{Q}$-local martingale.
7. Show that a local martingale that is bounded from below by a deterministic number is a supermartingale.
8. Show that in Section 4.3.2, we have the a.s. convergence $(\psi_n \cdot S)_T \to f_0$.

A

Elements of functional analysis

A.1 Separating hyperplane theorem

Let $v \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}$ be given and consider the set $H = \{x \in \mathbb{R}^n : \langle v, x\rangle = \gamma\}$. For $x \in H$ we have

$$\langle v, x - (\gamma/\|v\|^2)v\rangle = 0,$$

so $H = v^\perp + (\gamma/\|v\|^2)v$. The complement of H consists of the two sets $\{x : \langle v, x\rangle < \gamma\}$ and $\{x : \langle v, x\rangle > \gamma\}$ on the two "sides" of the hyperplane.

The following theorem says that for two disjoint, convex sets, one compact and one closed, there exists two "parallel" hyperplanes such that the sets lie strictly one different sides of those hyperplanes.

The assumption that one of the sets is compact can not be dropped (see Exercise 1)

Theorem A.1.1 (Separating hyperplane theorem). *Let K and C be disjoint, convex subsets of $\mathbb{R}^n$, K compact and C closed. There exist $v \in \mathbb{R}^n$ and $\gamma_1, \gamma_2 \in \mathbb{R}$ such that*

$$\langle v, x\rangle < \gamma_1 < \gamma_2 < \langle v, y\rangle$$

for all $x \in K$ and $y \in C$.

Proof. Consider the function $f : K \to \mathbb{R}$ defined by $f(x) = \inf\{\|x - y\| : y \in C\}$, i.e. $f(x)$ is the distance of x to C. The function f is continuous (check) and since K is compact, there exists $x_0 \in K$ such that f attains its minimum at x_0. Let $y_n \in C$ be such that $\|x_0 - y_n\| \to f(x_0)$. By the parallelogram law we have

$$\begin{aligned}\left\|\frac{y_n - y_m}{2}\right\|^2 &= \left\|\frac{y_n - x_0}{2} - \frac{y_m - x_0}{2}\right\|^2\\ &= \tfrac{1}{2}\|y_n - x_0\|^2 + \tfrac{1}{2}\|y_m - x_0\|^2 - \left\|\frac{y_n + y_m}{2} - x_0\right\|^2.\end{aligned}$$

By convexity $(y_n + y_m)/2 \in C$, so that $\|(y_n + y_m)/2 - x_0\| \geq f(x_0)$. Hence, we have

$$\left\| \frac{y_n - y_m}{2} \right\|^2 \leq \tfrac{1}{2}\|y_n - x_0\|^2 + \tfrac{1}{2}\|y_m - x_0\|^2 - f^2(x_0).$$

The right-hand side of this display converges to 0 as $n, m \to \infty$. So the y_n form a Cauchy sequence and hence they converge to some $y_0 \in \mathbb{R}^n$. Since C is closed, $y_0 \in C$. Let $v = y_0 - x_0$. Since K and C are disjoint, $v \neq 0$. It follows that $0 < \|v\|^2 = \langle v, y_0 - x_0 \rangle = \langle v, y_0 \rangle - \langle v, x_0 \rangle$. It remains to show that $\langle v, x \rangle \leq \langle v, x_0 \rangle$ and $\langle v, y_0 \rangle \leq \langle v, y \rangle$ for all $x \in K$ and $y \in C$.

Take $y \in C$. Since C is convex, the line segment $y_0 + \lambda(y - y_0)$, $\lambda \in [0, 1]$, belongs to C. Since y_0 minimizes the distance to x_0, we have

$$\|y_0 - x_0\| \leq \|y_0 - x_0 + \lambda(y - y_0)\|$$

for every λ. By squaring this we find that

$$0 \leq 2\lambda \langle y_0 - x_0, y - y_0 \rangle + \lambda^2 \|y - y_0\|^2.$$

Dividing by λ and then letting $\lambda \to 0$ gives $\langle v, y - y_0 \rangle \geq 0$, as desired.

A similar argument shows that $\langle v, x \rangle \leq \langle v, x_0 \rangle$ for $x \in K$. □

The *polar* C^0 of a set $C \subseteq \mathbb{R}^n$ is defined as

$$C^0 = \{y \in \mathbb{R}^n : \langle x, y \rangle \leq 1 \text{ for all } x \in C\}.$$

Note that in the special case that C is closed under multiplication with positive scalars, we have $C^0 = \{y \in \mathbb{R}^n : \langle x, y \rangle \leq 0 \text{ for all } x \in C\}$ (check). For a given x, the set $C_x^0 = \{z : \langle x, z \rangle \leq 0\}$ is the set of all vectors that lie on the same side of $x^\perp$ as $-x$. The polar is in this case the intersection of all the sets C_x^0 for $x \in C$.

To illustrate the *bipolar theorem* geometrically, consider a V-shaped set: C the union of two rays emanating from the origin. Then one readily sees that the polar of the polar of C precisely equals the convex hull of C. The general result is as follows.

Theorem A.1.2 (Bipolar theorem). *Let $C \subseteq \mathbb{R}^n$ contain 0. Then the bipolar $C^{00} = (C^0)^0$ equals the closed convex hull of C.*

Proof. It is clear that C^{00} is a closed, convex set containing C, so the closed convex hull A of C is a subset of C^{00}. Suppose that the converse inclusion does not hold. Then there exists a point $x_0 \in C^{00}$ that is not in A. By the separating hyperplane theorem there then exists a vector $v \in \mathbb{R}^n$ and $\gamma_1, \gamma_2 \in \mathbb{R}$ such that $\langle x_0, v \rangle > \gamma_1 > \gamma_2 > \langle y, v \rangle$ for all $y \in A$. Since $0 \in C \subseteq A$ we have $\gamma_1 > 0$. Dividing by γ_1 shows there exists a vector $v \in \mathbb{R}^n$ such that $\langle x_0, v \rangle > 1 > \langle y, v \rangle$ for all $y \in A$. The second inequality implies that $v \in C^0$, and then the first one implies that $x_0 \notin C^{00}$, which gives a contradiction. □

A.2 Topological vector spaces

A vector space X is called a *topological vector space* if it is endowed with a topology which is such that every point of X is a closed set and the addition and scalar multiplication operations are continuous.

It is easy to see that translation by a fixed vector and multiplication by a nonzero scalar are homeomorphisms of a topological vector space. This implies in particular that the topology is *translation-invariant*, meaning that a set $E \subseteq X$ is open if and only if each of its translates $x + E$ is open.

Topological vector spaces have nice separation properties. Combined with the fact that points are closed sets, the next theorem implies for instance that they are always Hausdorff.

Theorem A.2.1. *Suppose that K and C are disjoint subsets of a topological vector space X, K compact and C closed. Then there exits a neighborhood V of 0 such that $K + V$ and $C + V$ are disjoint.*

Proof. The continuity of addition implies that for every neighborhood W of 0 there exist neighborhoods V_1 and V_2 of 0 such that $V_1 + V_2 \subseteq W$ (check). Now put $U = V_1 \cap V_2 \cap (-V_1) \cap (-V_2)$. Then U is symmetric (i.e. $U = -U$) and $U + U \subseteq W$. Applying the same procedure to the neighborhood U we see that for every neighborhood W of 0 there exists a symmetric neighborhood U such that $U + U + U \subseteq W$ (etc.).

Pick an $x \in K$. Then $X \backslash C$ is an open neighborhood of x. By translation invariance and the preceding paragraph there exists a symmetric neighborhood V_x of 0 such that $x + V_x + V_x + V_x$ does not intersect C. By the symmetry of V_x this implies that $x + V_x + V_x$ and $C + V_x$ are disjoint (check). Since K is compact, it is covered by finitely many sets $x_1 + V_{x_1}, \dots, x_n + V_{x_n}$. Put $V = V_{x_1} \cap \cdots \cap V_{x_n}$. Then

$$K + V \subseteq \bigcup (x_i + V_{x_i} + V) \subseteq \bigcup (x_i + V_{x_i} + V_{x_i})$$

and none of the terms in the last union intersects $C + V$. □

The following lemma implies that if V is a neighborhood of 0 in a topological vector space X, then for every $x \in X$ it holds that $x \in rV$ if r is large enough. A set V with this property is called *absorbing*.

Lemma A.2.2. *Suppose V is a neighborhood of 0 in a topological vector space X and r_n is a sequence of positive numbers tending to infinity. Then*

$$\bigcup r_n V = X.$$

Proof. Fix $x \in X$. Then since V is open in X and $\lambda \mapsto \lambda x$ from $\mathbb{R}$ to X is continuous, $\{\lambda : \lambda x \in V\}$ is open in $\mathbb{R}$. The set contains 0, and hence it contains $1/r_n$ for n large enough. This completes the proof. □

For an arbitrary absorbing subset A (for instance a neighborhood of 0) of a topological vector space X we define the *Minkowsky functional* $\mu_A : X \to [0, \infty)$ by

$$\mu_A(x) = \inf\{t > 0 : x/t \in A\}.$$

Note that μ_A is indeed finite-valued, since A is absorbing. The following lemma collects properties that we need later.

Lemma A.2.3. *Let A be a convex, absorbing subset of a topological vector space X and let μ_A be its Minkowsky functional.*

(i) $\mu_A(x + y) \leq \mu_A(x) + \mu_A(y)$ *for all* $x, y \in X$.

(ii) $\mu_A(tx) = t\mu_A(x)$ *for all* $x \in X$ *and* $t \geq 0$.

Proof. For $x, y \in X$ and $\varepsilon > 0$, consider $t = \mu_A(x) + \varepsilon$, $s = \mu_A(y) + \varepsilon$. Then by definition of μ_A, $x/t \in A$ and $y/s \in A$. Hence, the convex combination

$$\frac{x+y}{s+t} = \frac{t}{s+t}\frac{x}{t} + \frac{s}{s+t}\frac{y}{s}$$

belongs to A as well. This proves (i). The proof of (ii) is easy. □

For the proof of the following characterization of continuous linear functionals we need the notion of a *balanced neighborhood*. A set $B \subseteq X$ is said to be *balanced* if $\alpha B \subseteq B$ for every scalar $\alpha \in \mathbb{R}$ with $|\alpha| \leq 1$.

Lemma A.2.4. *Every neighborhood of* 0 *contains a balanced neighborhood of* 0.

Proof. Let U be a neighborhood of 0. Since scalar multiplication is continuous, there exists a $\delta > 0$ and a neighborhood V of 0 in X such that $\alpha V \subseteq U$ whenever $|\alpha| < \delta$. Then $W = \cup_{|\alpha|<\delta}\, \alpha V$ is a balanced neighborhood of 0. □

A linear map $\Lambda : X \to \mathbb{R}$ is called a *linear functional* on the space X. A linear functional on X is called *bounded* on a subset $A \subseteq X$ if there exists a number $K > 0$ such that $|\Lambda x| \leq K$ for all $x \in A$.

Theorem A.2.5. *Let Λ be a nontrivial linear functional on a topological vector space X. Then Λ is continuous if and only if Λ is bounded on a neighborhood of* 0.

Proof. Suppose Λ is continuous. Then the null space $N = \{x \in X : \Lambda x = 0\}$ is closed. Since Λ is nontrivial, there exists $x \in X \backslash N$. By Theorem A.2.1 there exists a balanced neighborhood V of 0 such that $x + V$ and N are disjoint. Then $\Lambda(V)$ is a balanced subset of $\mathbb{R}$. Suppose it is not bounded. Then since it is balanced, it most be all of $\mathbb{R}$. In particular, there then exists a $y \in V$ such that

$\Lambda y = -\Lambda x$. But then $x + y \in N$, a contradiction. Hence, $\Lambda(V)$ is bounded, i.e. Λ is bounded on V.

Conversely, suppose that $|\Lambda x| \leq M$ for all $x \in V$. For $r > 0$, put $W = (r/M)V$. Then for $x \in W$, say $x = (r/M)y$ for $y \in V$, we have $|\Lambda x| = (r/M)|\Lambda y| \leq r$. Hence, Λ is continuous at 0. By translation invariance, it is continuous everywhere. □

A.3 Hahn-Banach theorem

The proof of the following version of the Hahn-Banach theorem relies on the axiom of choice, in the form of the *Hausdorff maximality theorem*:

Every nonempty partially ordered set $\mathcal{P}$ contains a totally ordered subset $\mathcal{Q}$ which is maximal with respect to the property of being totally ordered.

A proof of this fact can for instance be found in Rudin (1987), pp. 395–396.

Theorem A.3.1 (Hahn-Banach theorem). *Suppose X is a (real) vector space and $p : X \to \mathbb{R}$ satisfies $p(x + y) \leq p(x) + p(y)$ and $p(tx) = tp(x)$ for $x, y \in X$ and $t \geq 0$. Then if f is a linear functional on a subspace M of X such that $f(x) \leq p(x)$ for all $x \in M$, f extends to a linear functional Λ on the whole space X such that*

$$-p(-x) \leq \Lambda x \leq p(x)$$

for all $x \in X$.

Proof. Suppose M is a proper subspace of X and pick $x_1 \in X \backslash M$. For $x, y \in M$ we have

$$f(x) + f(y) = f(x + y) \leq p(x + y) \leq p(x - x_1) + p(y + x_1),$$

hence $f(x) - p(x - x_1) \leq p(y + x_1) - f(y)$. So there exists an α such that

$$f(x) - \alpha \leq p(x - x_1), \quad f(y) + \alpha \leq p(y + x_1) \tag{A.1}$$

for all $x, y \in M$. Now let M_1 be the vector space spanned by M and x_1. An element of M_1 is of the form $x + \lambda x_1$ for some $\lambda \in \mathbb{R}$. So we can extend f to M_1 by setting $f_1(x + \lambda x_1) = f(x) + \lambda\alpha$. Then f_1 is a well-defined linear functional on M_1 and the inequalities in (A.1) imply that $f_1(x) \leq p(x)$ for all $x \in M_1$ (check).

Let $\mathcal{C}$ be the collection of pairs (M', f'), where M' is a subspace of X containing M and f' is a linear extension of f to M' such that $f \leq p$ on M'. Put an ordering on $\mathcal{C}$ by saying that $(M', f') \leq (M'', f'')$ if $M' \subseteq M''$ and $f''|_{M'} = f'$. This is a partial ordering and $\mathcal{C}$ is not empty. Hence, by the Hausdorff maximality theorem, we can extract a maximal totally ordered subcollection $\mathcal{C}'$. Let $\tilde{M}$ be the union of all M' for which $(M', f') \in \mathcal{C}'$. Then

$\tilde{M}$ is a subspace of X (check). If $x \in \tilde{M}$ then $x \in M'$ for some M' such that $(M', f') \in \mathcal{C}'$. We then put $\Lambda x = f'(x)$. This defines a linear function Λ on $\tilde{M}$ and we have that $\Lambda \leq p$ on $\tilde{M}$ (check). If $\tilde{M}$ were a proper subspace of X the construction of the preceding paragraph would give us a further extension of Λ, contradicting the maximality of $\mathcal{C}'$. Hence, $\tilde{M} = X$. This completes the proof, upon noting that $\Lambda \leq p$ implies that $-p(-x) \leq -\Lambda(-x) = \Lambda x$ for all $x \in X$. □

Before we use the Hahn-Banach theorem to prove the infinite-dimensional version of the separating hyperplane theorem we introduce some more concepts and notation.

A topological vector space X is called *locally convex* if for every neighborhood V of 0 there exists a convex neighborhood U of 0 such that $U \subseteq V$. The space of continuous linear maps from X to $\mathbb{R}$ is denoted by X^*. It is called the *dual* of X, and is treated in more detail in the next section.

Theorem A.3.2 (Separation theorem). *Let A and B be disjoint, nonempty, convex subsets of a topological vector space X.*

(i) If A is open there exist $\Lambda \in X^*$ and $\gamma \in \mathbb{R}$ such that

$$\Lambda x < \gamma \leq \Lambda y$$

for every $x \in A$ and $y \in B$.

(ii) If X is locally convex, A is compact and B is closed, there exist $\Lambda \in X^*$ and $\gamma_1, \gamma_2 \in \mathbb{R}$ such that

$$\Lambda x < \gamma_1 < \gamma_2 < \Lambda y$$

for every $x \in A$ and $y \in B$.

Proof. (i). Pick $a_0 \in A$ and $b_0 \in B$ and put $x_0 = b_0 - a_0$. Define $C = A - B + x_0$ and note that C is a convex, open neighborhood of 0. Let μ_C be the Minkowsky functional of C.

Let M be the linear subspace generated by x_0 and define a linear functional f on M by putting $f(\lambda x_0) = \lambda$. Since A and B are disjoint, $x_0 \notin C$ so we have $\mu_C(x_0) \geq 1$ and hence, for $\lambda \geq 0$, $f(\lambda x_0) = \lambda \leq \lambda \mu_C(x_0) = \mu_C(\lambda x_0)$. For $\lambda < 0$ we have $f(\lambda x_0) < 0 \leq \mu_C(\lambda x_0)$. By Lemma A.2.3 and the Hahn-Banach theorem, Theorem A.3.1, the functional f extends to a linear functional Λ on X, and the extension satisfies $\Lambda x \leq \mu_C(x)$ for all $x \in X$. In particular $\Lambda \leq 1$ on C, so that $|\Lambda| \leq 1$ on the neighborhood $C \cap -C$ of 0. By Theorem A.2.5 this implies that Λ is continuous, i.e. $\Lambda \in X^*$.

Now for $a \in A$ and $b \in B$ we have that

$$\Lambda a - \Lambda b + 1 = \Lambda(a - b + x_0) \leq \mu_C(a - b + x_0) < 1,$$

since $a - b + x_0 \in C$ and C is open (Exercise 2), so $\Lambda a < \Lambda b$. It follows that $\Lambda(A)$ and $\Lambda(B)$ are disjoint, convex subsets of $\mathbb{R}$, the first one lying on the left of the second one. Since A is open and Λ is nonconstant, $\Lambda(A)$ is open as well

(Exercise 3). Letting γ be the right end point of $\Lambda(A)$ completes the proof of (i).

(ii). By Theorem A.2.1 and the local convexity of X there exists a convex neighborhood V of 0 such that $(A+V)\cap B=\varnothing$. By the proof of part (i) there exists $\Lambda\in X^*$ such that $\Lambda(A+V)$ and $\Lambda(B)$ are disjoint, convex subsets of $\mathbb{R}$, the first one lying on the left of the second one, the first one being open. Moreover, $\Lambda(A)$ is a compact subset of $\Lambda(A+V)$. The proof is now easily completed. □

Corollary A.3.3. *If X is a locally convex topological vector space, X^* separates the points of X.*

Proof. given distinct points $x,y\in X$, apply the separation theorem with $A=\{x\}$ and $B=\{y\}$. □

For $x\in X$ and $\Lambda\in X^*$ we define, in analogy with the finite-dimensional situation, $\langle x,\Lambda\rangle=\Lambda x$. The *polar* C^0 of a set $C\subseteq X$ is defined as

$$C^0=\{\Lambda\in X^*:\langle x,\Lambda\rangle\le 1 \text{ for all } x\in C\}.$$

Similarly, the *bipolar* is defined as

$$C^{00}=(C^0)^0=\{x\in X:\langle x,\Lambda\rangle\le 1 \text{ for all } \Lambda\in C^0\}.$$

Theorem A.3.4 (Bipolar theorem). *The bipolar C^{00} of a subset C of a locally convex topological vector space X equals the closed convex hull of C.*

Proof. It is clear that C^{00} is a convex set containing C, so the closed convex hull A of C is a subset of C^{00}. Suppose that the reverse inclusion does not hold. Then there exists a point $x_0\in C^{00}$ that is not in A. By the separation theorem there then exists a functional $\Lambda\in X^*$ such that $\Lambda x_0>1>\Lambda y$ for all $y\in A$ (check). The second inequality implies that $\Lambda\in C^0$, and then the first one implies that $x_0\notin C^{00}$, which is a contradiction. □

A.4 Dual space

The *dual* of a topological vector space X is the space X^* of continuous linear functionals on X. By Theorem A.2.5 this is the same as the space of linear functionals that are bounded on a neighborhood of 0.

It is easy to see that if the topology on X is induced by a *norm* $\|\cdot\|$, a linear functional Λ belongs to X^* if and only if the unit ball in X is mapped into a bounded subset of $\mathbb{R}$. In that case we define the *norm of* Λ by

$$\|\Lambda\|=\sup_{\|x\|\le 1}|\Lambda(x)|$$

and we have the relation $|\Lambda(x)| \leq \|\Lambda\| \, \|x\|$ for every $x \in X$.

Example A.4.1. Let $(E, \mathcal{E}, \mu)$ be a measure space with μ a finite measure, $p \in [1, \infty)$ and $X = L^p(E, \mathcal{E}, \mu)$. Consider a continuous linear functional Λ on X. Then the map $\nu : \mathcal{E} \to \mathbb{R}$ defined by $\nu(B) = \Lambda(1_B)$ is a signed measure (note that the finiteness of μ implies that ν is well-defined). Indeed, if B_n are disjoint elements of $\mathcal{E}$ and $B = \cup B_n$, then $1_{\cup_{k \leq n} B_k} \to 1_B$ in L^p. Since Λ is continuous, this implies that ν is countably additive. If $\mu(B) = 0$ then 1_B vanishes in L^p and hence $\nu(B) = 0$, so $\nu \ll \mu$. Hence, by the Radon-Nikodym theorem, there exists a $g \in L^1$ such that

$$\Lambda(1_B) = \nu(B) = \int_B g \, d\mu$$

for all $B \in \mathcal{E}$. By linearity we then have

$$\Lambda(f) = \int fg \, d\mu \tag{A.2}$$

for all simple functions f. Every bounded measurable function f is the uniform limit of simple functions and since μ is finite, uniform convergence implies convergence in L^p. It follows that (A.2) holds for all $f \in L^\infty$.

Suppose that $p > 1$ and let q be the conjugate exponent. For $E_n = \{x : |g(x)| \leq n\}$ we have, since g is bounded on E_n and Λ is continuous and hence bounded,

$$\int_{E_n} |g|^q \, d\mu = \int_{E_n} |g|^{q-1} \mathrm{sign}(g) g \, d\mu = \Lambda(1_{E_n} |g|^{q-1} \mathrm{sign}(g)) \leq \|\Lambda\| \Big(\int_{E_n} |g|^q \, d\mu \Big)^{1/p}.$$

It follows that

$$\Big(\int_{E_n} |g|^q \, d\mu \Big)^{1/q} \leq \|\Lambda\| \tag{A.3}$$

and letting $n \to \infty$ shows that $g \in L^q$. If $p = 1$ then for every $B \in \mathcal{E}$ we have

$$\Big| \int_B g \, d\mu \Big| = |\Lambda(1_B)| \leq \|\Lambda\| \mu(B).$$

But this implies that $|g| \leq \|\Lambda\|$ a.e. (indeed: if not there would exist an $\varepsilon > 0$ such that the set $B = \{x : |g(x)| > \|\Lambda\| + \varepsilon\}$ has positive μ-measure, leading to a contradiction), hence $g \in L^\infty$.

So in all cases the function g in (A.2) belongs to L^q. We proved already that (A.2) holds for all bounded functions f. Now Λ is continuous on L^p by assumption and Hölders inequality implies that the right-hand side is continuous for $f \in L^p$ as well. This shows that the relation holds in fact for all $f \in L^p$. Uniqueness of g is easy to prove. We conclude that we may identify the dual of L^p with L^q. Moreover, using (A.3) it is easy to see that for $\Lambda \in (L^p)^*$ given by (A.2), we have $\|\Lambda\| = \|g\|_{L^q}$ (Exercise 4). ■

Let X be a topological vector space with dual X^*. Every point $x \in X$ induces a linear functional on X^*, defined by $\Lambda \mapsto \Lambda x$. The *weak*-topology* of X^* is the weakest (i.e. smallest) topology making all these maps continuous.

The following theorem states that X^* with the weak*-topology is a locally convex topological vector space. This implies for instance that we can apply the separation theorem to it. In general, the space X^* endowed with the weak*-topology is not a Banach space. (In fact, it is not even metrizable if X is an infinite-dimensional Banach space.)

Theorem A.4.2. *The dual X^* of a topological vector space X, endowed with the weak*-topology, is a locally convex topological vector space. Its dual is given by $\{\Lambda \mapsto \Lambda x : x \in X\}$.*

Proof. Denote by f_x be the linear functional $\Lambda \mapsto \Lambda x$. If $\Lambda \neq \Lambda'$ in X^*, there exists an $x \in X$ such that $f_x\Lambda \neq f_x\Lambda'$. Hence, in $\mathbb{R}$ there exist disjoint neighborhoods U of $f_x\Lambda$ and U' of $f_x\Lambda'$. Since f_x is continuous, $f_x^{-1}(U)$ and $f_x^{-1}(U')$ are disjoint neighborhoods of Λ and Λ'. This shows that X^* is Hausdorff, and in particular that points are closed.

To show that the weak*-topology is translation invariant, consider an open base set

$$U = \{\Lambda : \Lambda x_1 \in B_1, \dots, \Lambda x_n \in B_n\}$$

and $\Lambda' \in X^*$. Then $\Lambda' + U = \{\Lambda : \Lambda x_1 \in B_1 + \Lambda' x_1 \dots, \Lambda x_n \in B_n + \Lambda' x_n\}$ is an open base set as well. It follows that the topology is translation invariant. Note that the open sets V of the form

$$V = \{\Lambda : |\Lambda x_1| < r_1, \dots, |\Lambda x_n| < r_n\} \tag{A.4}$$

for $x_1, \dots, x_n \in X$ and $r_1, \dots, r_n > 0$ form a local base at 0. Every such set V is convex, balanced and absorbing (check). In particular, X^* is a locally convex space.

For the set V in the preceding display we have $V/2 + V/2 = V$ and hence addition is continuous at $(0, 0)$. As for scalar multiplication, suppose that $\alpha\Lambda \in V$ for some scalar $\alpha \in \mathbb{R}$ and $\Lambda \in X^*$. By Exercise 2, there exists $t > 0$ such that $t < 1/|\alpha|$ and $\Lambda \in tV$. For $\varepsilon > 0$ and $\Lambda' \in tV$ we have that $(\alpha+\varepsilon)\Lambda' \in (\alpha+\varepsilon)tV$. Hence, since V is balanced, $(\alpha+\varepsilon)\Lambda' \in V$ for all ε such that $|\alpha|t+|\varepsilon|t \leq 1$. Since $|\alpha|t < 1$ there is a nonempty interval around 0 of ε satisfying this condition. Hence, scalar multiplication is continuous.

It remains to identify the dual of X^* (endowed with the weak*-topology). If $x \in X$, the linear map $\Lambda \mapsto \Lambda(x)$ is weak*-continuous by definition of the weak*-topology. Conversely, let $f : X^* \to \mathbb{R}$ be weak*-continuous. By Theorem A.2.5, f is bounded on a neighborhood of 0, and hence also on a base set V of the form (A.4). This implies that f vanishes on the set $N = \{\Lambda : \Lambda x_1 = \cdots = \Lambda x_n = 0\}$ (Exercise 5). Now N is the kernel of the linear map $\pi : X^* \to \mathbb{R}^n$ defined by $\pi(\Lambda) = (\Lambda x_1, \dots, \Lambda x_n)$. It follows that the linear map $F : \pi(X^*) \to \mathbb{R}$ given by $F(\pi(\Lambda)) = f(\Lambda)$ is well defined (check). We can extend F to a linear functional on $\mathbb{R}^n$. It is then necessarily of the form $F(z_1, \dots, z_n) = \sum \alpha_i z_i$ for certain real numbers α_i. In particular,

$$f(\Lambda) = F(\Lambda x_1, \dots, \Lambda x_n) = \sum \alpha_i \Lambda x_i.$$

So indeed, $f(\Lambda) = \Lambda x$, with $x = \sum \alpha_i x_i$. □

If X is a Banach space its dual X^* is endowed with a norm, and the unit ball in X^* is the set $\{\Lambda \in X^* : |\Lambda x| \leq \|x\| \text{ for all } x \in X\}$. In the norm-topology this set is not compact in general (think of an infinite-dimensional Hilbert space). In the weak*-topology however, it is always compact.

Theorem A.4.3 (Banach-Alaoglu). *The unit ball of the dual of a Banach space is weak*-compact.*

Proof. Denote the Banach space by X and let B^* be the unit ball in its dual. By Tychonov's theorem, $P = \Pi_{x\in X}[-\|x\|, \|x\|]$ is compact (relative to the product topology). We can view P as a collection of functions on X, with $f \in P$ if and only if $|f(x)| \leq \|x\|$ for all $x \in X$. As such, we have $B^* \subseteq X^* \cap P$. Hence, B^* inherits two topologies: the weak*-topology from X^* and the product topology from P. These two topologies on B^* coincide. To see this, take $\Lambda_0 \in B^*$. The sets of the form

$$V_1 = \{\Lambda \in X^* : |\Lambda x_1 - \Lambda_0 x_1| < r_1, \ldots, |\Lambda x_n - \Lambda_0 x_1| < r_n\}$$

and

$$V_2 = \{f \in P : |f(x_1) - \Lambda_0 x_1| < r_1, \ldots, |f(x_n) - \Lambda_0 x_1| < r_n\}$$

form a local base for the weak*-topology and, respectively, the product topology at Λ_0. Since $B^* \subseteq X^* \cap P$ we have $V_1 \cap B^* = V_2 \cap B^*$ and hence the two relative topologies coincide.

Next we show that B^* is closed in P. Take f_0 in the closure of B^* (with respect to the product topology). For $x, y \in X$, $\alpha, \beta \in \mathbb{R}$ and $\varepsilon > 0$ we have that the set

$$U = \{f \in P : |f(x) - f_0(x)| < \varepsilon, |f(y) - f_0(y)| < \varepsilon, |f(\alpha x + \beta y) - f_0(\alpha x + \beta y)| < \varepsilon\}$$

is an open neighborhood of f_0. Hence, there exist an $f \in U \cap B^*$. Since f is linear we have

$$\begin{aligned} &f_0(\alpha x + \beta y) - \alpha f_0(x) - \beta f_0(y) \\ &\quad = (f_0 - f)(\alpha x + \beta y) - \alpha(f_0 - f)(x) - \beta(f_0 - f)(y) \end{aligned}$$

and hence

$$|f_0(\alpha x + \beta y) - \alpha f_0(x) - \beta f_0(y)| \leq (1 + |\alpha| + |\beta|)\varepsilon.$$

Since ε was arbitrary, it follows that f_0 is linear. By definition of P we have that $|f_0(x)| \leq \|x\|$ for every $x \in X$, so indeed $f_0 \in B^*$.

The proof is now completed upon noting that by the preceding paragraph, B^* is compact with respect to the product topology. But by the first part of the proof, the latter topology coincides on B^* with the weak*-topology. □

Example A.4.4. Although the weak*-topology has some nice properties according to Theorem A.4.2, it is good to note that it is typically "strange". Consider for instance a finite measure μ on the line and view $L^\infty(\mu)$ as the dual of $L^1(\mu)$. Then from the form of the local base at 0 given in the proof of the theorem one sees that a sequence f_n in L^∞ converges in the weak*-topology to 0 if $\int f_n g \, d\mu \to 0$ for every $g \in L^1$. By dominated convergence, this holds for instance for $f_n = 1_{(-n,n)^c}$. This sequence does however not converge to 0 in the ordinary, uniform topology on L^∞. More generally, to say that a function $f \in L^\infty$ belongs to the weak*-closure of a set $C \subseteq L^\infty$ does not necessarily mean that f is well-approximated by elements of C in a uniform or any other intuitively reasonable way. ■

A.5 Exercises

1. Give an example which shows that the separation theorem does not hold in general if the assumption of compactness of one of the sets in dropped.
2. Suppose that C is an open neighborhood of 0 in a topological vector space and let μ_C be its Minkowsky functional. Show that for all $x \in C$ it holds that $\mu_C(x) < 1$.
3. Show that a non-constant linear functional on a topological vector space maps open sets to open sets.
4. In Example A.4.1, show that for the functional Λ on L^p defined by (A.2) we have $\|\Lambda\| = \|g\|_{L^q}$.
5. In the last part of the proof of Theorem A.4.2, show that the functional f vanishes on the set N.

B

Elements of martingale theory

B.1 Basic definitions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A collection of $\mathbb{R}^d$-valued random variables $X = (X_t)_{t \in T}$ indexed by a set $T \subseteq \mathbb{R}$ is called a *(d-dimensional) stochastic process*. We call the process *continuous* (or *cadlag*), it its trajectories $t \mapsto X_t(\omega)$ are continuous (or cadlag). The process is called *bounded* if there exists a finite number K such that a.s. $\|X_t\| \leq K$ for all t.

A *filtration* is a collection $(\mathcal{F}_t)_{t \in T}$ of sub-σ-fields of $\mathcal{F}$ such that $\mathcal{F}_s \subseteq \mathcal{F}_t$ for all $s \leq t$. It is said to satisfy the *usual conditions* if it is *right-continuous*, i.e. $\cap_{s>t} \mathcal{F}_s = \mathcal{F}_t$ for all t and $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets in $\mathcal{F}$. A process X is called *adapted* to $(\mathcal{F}_t)$ is for every t, X_t is $\mathcal{F}_t$-measurable. For a process X and $t \in T$ we define $\mathcal{F}_t^X$ to be the σ-field generated by the collection of random variables $\{X_s : s \leq t\}$. The filtration $(\mathcal{F}_t^X)$ is called the *natural filtration* of the process X. It is the smallest filtration to which it is adapted. A process $X = (X_t)_{t \in [0,T]}$ is called *progressively measurable* relative to the filtration $(\mathcal{F}_t)_{t \in [0,T]}$ if for all t, the map $(\omega, s) \mapsto X_s(\omega)$ on $\Omega \times [0,t]$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t])$-measurable.

A $[0,\infty]$-valued random variable τ is called a *stopping time* relative to the filtration $(\mathcal{F}_t)$ if $\{\tau \leq t\} \in \mathcal{F}_t$ for every t. If τ is a stopping time and X a process, the *stopped process* X^τ is defined by $X_t^\tau = X_{\tau \wedge t}$. A *localizing sequence* is a sequence of stopping times τ_n increasing a.s. to infinity. A process X is said to have a property P *locally* if there exists a localizing sequence τ_n such that for every n, the stopped process X^{τ_n} has the property P.

A process M is called a *martingale* relative to the filtration $(\mathcal{F}_t)$ if every M_t is integrable and for all $s \leq t$ it holds that $\mathbb{E}(M_t \,|\, \mathcal{F}_s) = M_s$ a.s.. In accordance with the previously introduced notation the process M is called a *local martingale* if there exists a localizing sequence τ_n such that for every n, the stopped process M^{τ_n} is a martingale. Every martingale is a local martingale, but not vice versa..

B.2 Theorems

For a filtration $(\mathcal{F}_t)$ and a stopping time τ we define

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \le t\} \in \mathcal{F}_t \text{ for all } t\}.$$

The set $\mathcal{F}_\tau$ is always a σ-field and should be thought of as the collection of events describing the history before time τ.

Theorem B.2.1 (Optional stopping theorem). *Let M be a cadlag, uniformly integrable martingale. Then for all stopping times $\sigma \le \tau$,*

$$\mathbb{E}(M_\tau \mid \mathcal{F}_\sigma) = M_\sigma.$$

Theorem B.2.2 (Kakutani's theorem). *Let $X_1, X_2, \ldots$ be independent nonnegative random variables with mean 1. Define $M_0 = 1$ and $M_n = X_1 X_2 \cdots X_n$. It holds that M is uniformly integrable if and only if $\sum(1 - \mathbb{E}\sqrt{X_n}) < \infty$. If M is not uniformly integrable, then $M_n \to 0$ a.s..*

Corollary B.2.3. *Let $X = (X_1, X_2, \ldots)$ and $Y = (Y_1, Y_2, \ldots)$ be two sequences of independent random variables. Assume X_i has a positive density f_i with respect to a dominating measure μ, and Y_i has a positive density g_i with respect to μ. Then the laws of the sequences X and Y are equivalent probability measures on $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ if and only if*

$$\sum_{i=1}^n \int (\sqrt{f_i} - \sqrt{g_i})^2 \, d\mu < \infty.$$

If the laws are not equivalent, they are mutually singular.

Proof. Let $(\Omega, \mathcal{F}) = (\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ and $Z = (Z_1, Z_2, \ldots)$ the coordinate process on $(\Omega, \mathcal{F})$, so $Z_i(\omega) = \omega_i$. Let $\mathcal{F}_n \subseteq \mathcal{F}$ be the σ-field generated by $Z_1, \ldots, Z_n$. Since the densities f_i and g_i are all positive, the distributions $\mathbb{P}_X$ and $\mathbb{P}_Y$ of the sequences X and Y are equivalent on $\mathcal{F}_n$. For $A \in \mathcal{F}_n$ we have

$$\mathbb{P}_X(A) = \int_A M_n \, d\mathbb{P}_Y,$$

where the Radon Nikodym derivative is defined by $M_n = \prod_{i=1}^n f_i(Z_i)/g_i(Z_i)$. Observe that under $\mathbb{P}_Y$, the process M is a martingale to which the preceding theorem applies. It is readily verified that the measures $\mathbb{P}_X$ and $\mathbb{P}_Y$ are equivalent on the whole σ-field $\mathcal{F}$ if and only if M is uniformly integrable with respect to $\mathbb{P}_Y$ (Exercise 1). Hence, by the preceding theorem, the measures are equivalent if and only if

$$\sum_{i=1}^{n}\Big(1-\int\sqrt{f_i g_i}\,d\mu\Big)<\infty.$$

The proof of the first part is completed by noting that $\int(\sqrt{f_i}-\sqrt{g_i})^2\,d\mu = 2-2\int\sqrt{f_i g_i}\,d\mu$.

We noted that if $\mathbb{P}_X$ and $\mathbb{P}_Y$ are not equivalent, then M is not uniformly integrable relative to $\mathbb{P}_Y$. Hence, by the preceding theorem, $M_n \to 0$, $\mathbb{P}_Y$-a.s.. We can reverse the roles of X and Y, which amounts to replacing M by $1/M$. Then we find that if $\mathbb{P}_X$ and $\mathbb{P}_Y$ are not equivalent, $1/M_n \to 0$, $\mathbb{P}_X$-a.s.. It follows that for the event $A = \{M_n \to 0\}$ we have $\mathbb{P}_Y(A) = 1$ and $\mathbb{P}_X(A) = 0$. □

Example B.2.4. Let $X = (X_1, X_2, \ldots)$ and $Y = (Y_1, Y_2, \ldots)$ be two sequences of independent random variables. Suppose that $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = 1/2$ and $\mathbb{P}(Y_i = 1) = 1 - \mathbb{P}(Y_i) = -1 = 1/2 + \varepsilon_i$ for some $\varepsilon_i \in (-1/2, 1/2)$. By the corollary, applied with μ the counting measure, $f_i(1) = f_i(-1) = 1/2$, $g_i(1) = 1 - g_i(-1) = 1/2 + \varepsilon_i$, the laws of the sequences X and Y are equivalent if and only if

$$\sum\Big((\sqrt{1/2}-\sqrt{1/2+\varepsilon_i})^2+(\sqrt{1/2}-\sqrt{1/2-\varepsilon_i})^2\Big)<\infty.$$

By Taylor's formula the function $h(x) = (\sqrt{1/2}-\sqrt{1/2+x})^2 + (\sqrt{1/2}-\sqrt{1/2-x})^2$ behaves like a multiple of x^2 near $x = 0$ (check!). It follows that the sequences are equivalent if and only if $\sum\varepsilon_i^2 < \infty$.

B.3 Exercises

1. In the proof of Corollary B.2.3, show that the measures $\mathbb{P}_X$ and $\mathbb{P}_Y$ are equivalent on the whole σ-field $\mathcal{F}$ if and only if M is uniformly integrable.

References

Conway, J. B. (1990). *A course in functional analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition.

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Math. Ann.* **300**(3), 463–520.

Delbaen, F. and Schachermayer, W. (2006). *The mathematics of arbitrage*. Springer Finance. Springer-Verlag, Berlin.

Harrison, J. M. and Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Process. Appl.* **11**(3), 215–260.

Kreps, D. M. (1981). Arbitrage and equilibrium in economies with infinitely many commodities. *J. Math. Econom.* **8**(1), 15–35.

Rudin, W. (1987). *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition.

Rudin, W. (1991). *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, second edition.