Monetary Utility Functions

Prof.dr.em. Freddy DELBAEN
Department of Mathematics
ETH Zurich
http://www.math.ethz.ch/~delbaen.

Extended 2011 Version of the Lecture Notes:
University of Osaka, Osaka 2008
This version March 20, 2020
This is a preliminary draft

Contents

Chapter 1

Introduction

These notes find their origin in courses I taught at a Cattedra Galileiana of the Scuola Normale Superiore di Pisa (March 2000), the University of Pretoria (August 2003), Tokyo University, ETH (2005) and the University of Osaka (2008). The aim of these lectures was to translate problems from Risk Management into mathematics and back. In some sense these notes form another illustration of the fact that problems from applied mathematics and real life, when properly translated are not that far from pure mathematics. In the present notes we will mainly use functional analysis and stochastic calculus to solve problems from Risk Management and Risk Measurement. Because of its close relation with utility functions, I changed the set-up from the Pisa lecture notes ([40]) to a set-up that uses the terminology from utility theory. I also included relations with decision theory as it was developed by Gilboa and Schmeidler, see e.g. [75] for the main paper, Maccheroni, Marinacci and Rustichini [102], [103], Machina and Schmeidler [104], Chateauneuf and Wakker [28], Wakker [129]. In statistics the theory is known as robustness and many papers are published on this topic, see e.g. Huber's book as well as the references given there, [80]. In insurance mathematics, risk measures can be seen as premium calculation principles. We cite Bühlmann, [25] and Gerber, [73] for mathematical definitions of premium principles. So called convex premium principles were introduced by Deprez and Gerber, [51] and this paper contains concepts that are almost the same as the concepts used here. The concept of risk measurement as presented in this book is related to capital requirement for financing institutions. As such it is also related to risk averseness. A risk measure that goes in the direction of risk averseness can be found in Aumann-Serrano, [14]. This measure is related to the exponential utility but is different from what we present. It is impossible to cite all the references that are related to monetary utility functions. I apologise for not mentioning or better for forgetting references that are considered as basic. The multiperiod case is much more complex and involves new concepts. Utility theory in this context goes back to Koopmans [89, 90, 91], Epstein [62], Epstein and Zin [64], Duffie and Epstein [53], Duffie and Skiadas, [54], El Karoui, Quenez, Peng ([61]), Artzner, Delbaen, Eber, Heath, Ku ([5]) and probably many others. In these notes I only introduce some of the basic material and for instance relations with Backward Stochastic Differential Equations (BSDE), see e.g. El Karoui, Quenez, Peng[61], are not treated at all. Utility theory for stochastic processes is another topic that is not covered. The reader can find an introduction to these problems in Artzner, Delbaen, Eber, Heath, Ku [5], Cheridito, Delbaen, Kupper, [29], Delbaen [41].

Part of the courses was devoted to an analysis of Value at Risk and its relation to quantiles. A detailed discussion of this can be found in two papers by Artzner, Delbaen, Eber and Heath, [3] and [4]. It will not be repeated here. We will rather concentrate on the mathematics behind the concept of coherent risk measures or coherent utility functions. They were introduced in the two mentioned papers and the mathematical theory was further developed in Delbaen (1999), [39] and [40]. Further use of coherent risk measures can be found in the papers (and their references) by Kalkbrener, Lotter and Overbeck, [87], Jaschke, [82] and Jaschke and Küchler [83], Tasche, Acerbi and Tasche [11] and Föllmer-Schied, [68]. Since their introduction around 1995, many researchers have extended the theory and giving a complete bibliography is almost impossible. The reader should consult the web-sites to find several papers dealing with this subject. The paper by Föllmer and Schied, [69],[**?**], introduces a generalisation of coherent risk measures to convex risk measures. A mathematical trick will allow us to reduce the characterisation of convex risk measures to the same problem for coherent risk measures. This trick does not contribute to the presentation but it allows an easier use of theorems from functional analysis. The name monetary utility functions was introduced by Föllmer and Schied. Also the word "niveloid" was used for the same property. The expression money based utility function was also used.

In chapter 2 we introduce the notation and recall some basic facts from functional analysis. The reader can consult Diestel's book, [52], for proofs. The Krein-Smulian theorem can be found, as an exercise, in Rudin's book, [120] or for a generalisation (the so-called Banach-Dieudonné theorem) and a full proof, see Grothendieck, [74]. We also give a summary of the results on atomless spaces. These results are well known and are standard (but not always easy) exercises in advanced probability courses.

Chapter 3 gives a short description of Value at Risk. We give a precise definition of what is usually called VaR. It is pointed out that VaR is not sub-additive. Being sub-additive is the mathematical equivalent of diversification. Since we changed the concept pf risk measures into the concept of utility functions we will deal with the a property called super-additivity. For risk adjusted values or utility functions that are not super-additive it may happen that diversified portfolios require more regulatory capital than less diversified portfolios. This observation was made when [3] was prepared (1993), but it was not made concrete. The first who observed that VaR posed a problem in practical problems and especially in the area of credit risk was Albanese, [1]. Especially in the area of Credit Risk the super-additivity property plays a fundamental role. This was shown in a paper by Bonti, Kalkbrener, Lotz and Stahl [22] (paper appeared in 2006 but the results were presented already around 2000). This paper refers to real life data from Deutsche Bank and it shows that capital allocation methods based on VaR could produce a negative amount of required capital and later an amount of economic capital that exceeded the exposure.

Chapter 4 introduces the concept of coherent risk measures and of coherent utility functions. Basically we only deal with coherent utility functions satisfying the Fatou property. Roughly speaking, a coherent utility function is defined via the infimum over a family of expected values. The probabilities used to calculate these expectations form a convex closed set, sometimes referred to as the set of "scenarios" or test probabilities. Stress testing simply means that the set of scenarios contains probability measures that are concentrated on "extreme movements in the market". Examples are given and relations with weak compact sets of L^1 are pointed out. The example on Credit Risk shows that tail expectation (sometimes called Worst Conditional Mean, shortfall, CV@R or TailVaR) is better behaved than VaR. The reader should carefully read the proof given in that chapter. For practical calculations of TailVaR or CVaR, we refer to Rockafellar and Uryasev, [119]. We do not discuss more risk averse utility functions, although we could have given practical examples that show that tail expectation is not yet good enough. Since there is no best risk measure, I did not pursue this discussion. The characterisation theorem permits to give many other examples of coherent utility functions. The interested reader can have a look at Delbaen (1999), [39], to see how Orlicz space theory [93], can be used in the construction of coherent risk measures. The relation with Orlicz space theory became the subject of new research, see Biagini-Frittelli [18], Cheridito-Li, [31], [32]. We also show how convex analysis can be used. The reader familiar with Rockafellar's book, [118], and with Phelps's monograph, [112], can certainly find much more points in common than the ones mentioned here.

In Chapter 5 we characterise the utility functions that only depend on the law of the uderlying random variables. For coherent risk measures this result is due to Kusuoka, [96]. In [72], Frittelli and Rosazza-Gianin could characterise convex law invariant risk measures. See also Jouini-Schachermeyer-Touzi,[84] and Tsukahara, [128]. They proved that these law-invariant measures are necessary Fatou and introduced one-parameter families. The reader should have a look at these papers.

Chapter 6 explains some basic operations on monetary utility functions. The two most important operations are the minimum of two utility functions and the convolution. The latter is usually called "inf-convolution" for convex functions or convex convolution. For concave functions one needs to change the "sign". Operations on utility functions imply, by duality, operations on the penalty functions or on the scenario sets. We give examples where the topological properties are not always preserved. The basic mathematical ingredients can be found in the already cited books, [118] and [112].

In Chapter 7 we mention the connection with convex game theory. The basic references here are Shapley, [126], Rosenmüller, [117], Schmeidler, [123], [124], and Delbaen, [37]. The important relation with commonotonicity (Schmeidler's theorem) is proved in a way that is different from [124]. We should also point out that distorted probability measures were first used by Yaari in decision theory [133] and were used by Denneberg, see [48], to describe premium calculation principles. The characterisation of extreme points is based on selection theorems. It is shown that a distorted game is a convex combination of unanimity games. This result is already present in [76] but the context here is infinite dimensional and hence more complicated. This structural result allows to get the extreme points of the core, a result that goes back to Carlier and Dana [26] and Marinacci et al [7] as well as the references given therein. The earlier results of Ryff, [121], form the mathematical basis of many proofs. We reprove these results in a different way and give some extensions based on results from functional analysis. Many developments on game theory were done at the same time by different authors and in different degrees of generality. I apologise if the references are not always to the original papers.

Chapter 8 shows how coherent utility functions are related to VaR. The main result is that tail expectation is the biggest coherent utility function, only depending on the distribution of the underlying random variable, that is dominated by VaR. Kusuoka, [96] gives another proof of this result. At the same time one can prove that VaR is the hull of all coherent utility functions that are smaller than VaR. The two results are not contradictory since to get

VaR as the hull of coherent utility functions we need utility functions that are not just law determined. Since VaR is not concave this shows that the sup of all coherent utility functions, smaller than VaR, is not concave.

Chapter 9 deals with the problem of capital allocation. In my view one of the most important applications of the theory of monetary utility functions. In our earlier papers, we emphasized that applications to performance measurement and capital allocation were among the driving forces to develop the theory. Denault, [47], looks for axiomatics regarding this problem and wants to characterise the capital allocation via the Shapley value. I tried to give other approaches, especially the use of Aubin's result on fuzzy games, [12] finds a nice interpretation and automatically leads to the introduction of the subgradient. Here again the duality theory plays a fundamental role. The main difference between the two approaches is that the Shapley value leads to a scenario that is in the "middle" of the set of scenarios, whereas our approach leads to extreme points of the set of scenarios. We also point out relations with a paper of Deprez and Gerber [51] that relates properties of coherent utility functions and the derived capital allocation methods with premium calculation principles. In their paper it is argued that the premium to be asked for a new insurance contract cannot be handled independently of the already existing portfolio. Diversification — or the absence of diversification — plays a fundamental role in this philosophy. The paper advocates that premium principles should therefore be defined on random variables and not only on distributions of these random variables. As a probabilist I cannot agree more with their statement. However it is not clear on which spaces the premium principle is defined and especially the existence of derivatives or subgradients is therefore not treated. As an example we will show in chapter xxx that the bid-price in an incomplete market is nowhere differentiable. In infinite dimensions the existence of a derivative leads to non-trivial problems from functional analysis. In our presentation we cannot avoid to use somewhat more theory, I apologise for it. In this chapter we show that differentiability of monetary utility functions automatically leads to weak compactness of the set of scenarios. The proof uses automatic continuity results, which go back to the work of Banach [15] and which was developed by Christensen, [35]. Kalkbrener, [86] gives another set of axioms on capital allocation.

Chapter 10 deals with the definition of coherent utility functions on the space of all random variables. This extension is not obvious and poses some mathematical problems. The approach given here is much simpler than the original approach. We also show that if there a concave utility function is defined on a rearrangement invariant space, then this space must be included in L^1. In particular this shows that it is impossible to define a consistent theory of utility functions on spaces that include Pareto distributed random variables.

Chapter 11 introduces for the two period model concepts such as time consistency. Time consistency was introduced by Koopmans, [89, 90, 91]. It has numerous consequences on the structure of the utility function. We give a presentation that shows that time consistent utility functions (as a functional defined on L^∞) are completely determined by their knowledge (as a functional) at time 0. The two period model is studied in detail since it contains the basic facts for applications in finite discrete time, handled in Chapter 12, and even for applications in continuous time which is not treated in this monograph. In Chapter 12, we show that time consistent utility functions are a concatenation of one period utility functions. This result allows for calcualtions based on dynamic programming principles. The case of coherent utility functions deserves a special attention. Time consistency or recursive utility theory is – as mentioned above – covered by many authors. We cite Detlefsen and Scandalo, [46], Epstein and Schneider, [63], Frittelli and Rosazza-Gianin, [71], Maccheroni, Marinacci and Rusticini, [102], [103], Riedel, [114], Föllmer and Penner, [67], Roorda, Schumacher and Engwerda, [115]. We do not make a connection to the theory of Backward Stochastic Equations and g-expectations. We plan to do that in another text.

I would like to use this occasion to express my thanks to the Scuola Normale Superiore for inviting me to hold the Cattedra Galileiana and to give a series of lectures in 2000. The "Pisa lecture notes" were the start of a more profound mathematical development. I also would like to thank the "Departement voor Wiskunde" of the University of Pretoria. Special thanks go to Professors Johan Swart, Barbara Swart (now at UNISA) and Anton Ströh, at that time chairman of the department. The discussions with the students from Pretoria contributed a lot to the presentation. Pisa and Pretoria were the first to undergo a presentation of the theory. Later the presentations changed and new topics were introduced. Topics such as BSDE were introduced when the author was visiting Fudan University in Shanghai, Shandong University in Weihai and Jinan. Here I could benefit from discussions with the specialists in BSDE and g-expectations. At a later stage I gave a similar course at the "Université de Franche-Comté à Besançon". Here I had a lot of discussions with Professors Kabanov and Stricker, these discussions also contributed to a better understanding of the problems. When visiting Tokyo University (Todai), I had the opportunity to discuss with Professor Kusuoka and Dr. Morimoto. Todai organised a series of lectures with an audience that was a mixture of practitioners and academic researchers. At the University of Osaka (Handai), I had the pleasure to discuss with Prof. Nagai, Prof Sekine (Kyoto University, now in Osaka University) and the other members of the departments of Osaka and Kyoto. The last guinea pigs were students and staff members of Ajou University in Suwon (South-Korea). The series of seminars given there are at the basis of this text. The hospitality in all these institutions is greatly appreciated. Readers familiar with the older "Pisa Notes" can see that the theory has changed a lot. I also presented a one semester course on coherent measures at the ETH. The many discussions with the students, researchers and colleagues are greatly appreciated.

I also want to thank all those who contributed to these lecture notes and made a lot of comments on previous versions. In this respect I cannot underestimate the value of discussions with Akahori, Bao, Barrieu, Ben-Artzi, S. Biagini, Carmona, Chen, Cheridito, Coculescu, Dana, Ekeland, El Karoui, Embrechts, Filipovic, Frittelli, Hu, Koch, Koo, Ku, Kupper, Kusuoka, Lüthi, Madan, Maignan, Nagai, Nikeghbali, Peng, Pratelli, Rosazza-Gianin, Schachermayer, Sekine, Schweizer, Sung, Tang, Tsukahara, Yan, Zariphopoulou ... and I apologise to the many others I forgot to mention.

These lectures would never have existed without the many discussions with the "partners in crime": Artzner, Eber and Heath. When we started the theory around 1993–1994, we had no feeling about the impact it would have. But gradually the theory developed and got more and more attention. Maybe not always there where it should have gotten attention but that is the price we must pay when something new is developed.

As always, lectures only make sense if there is an active audience. I thank the (guinea-pig) students of the Scuola Normale Superiore di Pisa, of the University of Pretoria, of Todai, of Handai , of Ajou University (Suwon) and of ETH, as well as the many practitioners for their interest in the subject and for the many questions they asked.

During the years 1995-2008 I got a grant from Credit Suisse to develop finance activities at the Department of Mathematics of ETH. This financial support allowed to appoint researchers and allowed to develop the theory presented here. Without this grant this work would not have been possible and I thank Credit Suisse for this important support. In particular I want to thank Dr. H.U. Doerrig and Dr. H. Stordel who always expressed a firm interest in the development of these concepts. Of course the work only reflects my personal viewpoint.

Zurich, March 20, 2020

Freddy Delbaen

Chapter 2

Mathematical Preliminaries

2.1 Interpretation of the mathematical concepts

We consider a very simple model in which only two dates ("today" and "tomorrow") matter. The multiperiod model will be treated in later chapters, where we will need the results of the one-period case. For simplicity we also suppose that all (random) amounts of money available tomorrow, have already been discounted. This practice is well known in finance (and in insurance for more than 300 years) and it avoids a lot of notational problems. The discounting can take place with an arbitrary asset, provided the price is strictly positive. So we can use a "sure" bank account with known interest rate at time 0. But we could also use an asset with a return that is only known at date 1. After discounting, the interest rate disappears from the calculations and hence the discounting is equivalent to assume that the interest rate is zero. The reader can consult [6] for a discussion and for a solution on the choice of the numéraire. The procedure of discounting is well understood and we will not comment on it anymore, thereby also avoiding the possible problems it creates. E.g. one can only take "maximal elements" (see [44]), but these problems are beyond the scope of this book.

We fix once and for all a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. All random variables will be defined on Ω. The positive part of a function ξ is denoted by $\xi^+ = \max(\xi, 0)$, the negative part $\xi^- = -\min(\xi, 0)$. Of course $\xi = \xi^+ - \xi^-$. A random variable represents the "discounted" value of a portfolio (or a position). Positive outcomes are good, negative outcomes mean that there is a shortage of money, a bankruptcy, ... We emphasize that it represents the outcome and not just the gain (possibly negative) realised with a transaction or an investment strategy. We will represent such values with bounded random variables. There are two reasons. One reason is mathematical: we will need different probability measures and unbounded random variables might cause problems (integrable with respect to one measure but not with respect to the other measure). The other reason is that we feel that every position taken in real life will lead to bounded outcomes. Of course there are positions and losses that are better modelled by unbounded random variables (e.g. Pareto distributed). In a later chapter we will see how to deal with such problems.

2.2 Some notation and definitions from integration theory

The expectation of a random variable ξ with respect to $\mathbb{P}$ will be denoted by $\mathbb{E}[\xi]$. When more than one probability measure is involved, we will explicitly mention it in the integral and we will write $\mathbb{E}_{\mathbb{P}}[\xi]$ or $\mathbb{P}[\xi]$. We also identify random variables that are equal almost surely. So each time we speak about a random variable we mean in fact the equivalence class of random variables with respect to equality "almost sure". This is common practice in probability theory and in most cases it is harmless. In the chapters on dynamic utility functions, we will draw special attention to the regularity of stochastic processes. In these cases there are uncountably many sets of measure zero involved and the situation is then not so harmless.

In finance, replacing a probability with an equivalent one is quite frequent. From a mathematical point of view, we must pay attention since theorems and properties which depend on variance, higher moments and integrability conditions obviously depend on the probability measure one is working with. There are two spaces that do not depend on the particular probability measure chosen. The first one is the space of (almost surely) bounded random variables L^∞ endowed with the norm:

$$\|\xi\|_\infty = \text{ess.sup}_{\omega\in\Omega}|\xi(\omega)|\,,$$

where by ess.sup of a random variable η we denote the number $\min\{r \mid \mathbb{P}[\eta > r] = 0\}$. (The reader can check that there is a minimum and not just an infimum). The second invariant space is L^0, this is the space of all (equivalence classes of) random variables. This space is usually endowed with the topology of convergence in probability that is

$$\xi_n \xrightarrow{\mathbb{P}} \xi \quad \text{iff} \quad \forall\varepsilon > 0 \quad \mathbb{P}[|\xi_n - \xi| > \varepsilon] \to 0\,,$$

or, equivalently, iff

$$\mathbb{E}[|\xi_n - \xi| \wedge 1] \to 0\,,$$

where $a \wedge b = \min(a, b)$ denotes the minimum of a and b.

Many theorems in measure theory refer to convergence almost surely, altough they remain valid when convergence a.s. is replaced by convergence in probability. This is the case for the dominated convergence theorem of Lebesgue, Fatou's lemma (properly formulated), etc. We will use these extensions without further notice.

We will denote by $L^1(\Omega, \mathcal{F}, \mathbb{P})$ (sometimes $L^1(\mathbb{P})$ or simply by L^1) the space of integrable random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. The dual space of L^1 is L^∞ and the duality (L^1, L^∞) will play a special role. The dual space of L^∞ is $\mathbf{ba}(\Omega, \mathcal{F}, \mathbb{P})$ or just $\mathbf{ba}$ if no confusion can arise. It is the space of bounded finitely additive measures μ such that $\mathbb{P}[A] = 0$ implies $\mu(A) = 0$. We will constantly identify measures with their Radon-Nikodym derivatives. So L^1 becomes a subspace of $\mathbf{ba}$. The set of sigma-additive probability measures, absolutely continuous with respect to $\mathbb{P}$ can then be identified with the set $\{f \in L^1 \mid f \geq 0, \mathbb{E}[f] = 1\}$. This set will be denoted by $\mathbf{P}$. Its weak* closure in $\mathbf{ba}$, denoted by $\mathbf{P}^{\mathbf{ba}}$, is the set of all finitely additive probability measures.

2.3 Some results on atomless spaces

In many cases we need that $(\Omega, \mathcal{F}, \mathbb{P})$ is an atomless probability space. The theorem below shows that this is not a very restrictive assumption. However the case of finite sets Ω, important in practical calculations, is not covered by this assumption.

Definition 1 *The probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called atomless (or diffuse) if for every $A \in \mathcal{F}$ with $\mathbb{P}[A] > 0$, there is a $B \subset A$ such that $0 < \mathbb{P}[B] < \mathbb{P}[A]$.*

The following characterisation of atomless spaces holds.

Theorem 1 *For a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the following are equivalent:*

1. *The probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless.*
2. *There is a family $A_t : t \in [0, 1]$ such that for $t \leq s : A_t \subset A_s$ and such that $\mathbb{P}[A_t] = t$ for all t.*
3. *There is a random variable ξ, defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with a continuous distribution, i.e. for each x, $\mathbb{P}[\xi = x] = 0$.*
4. *There is a random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with a uniform $[0, 1]$ distribution.*

5. *If $C \subset B$, there is a family $\{A_t \mid t \in [\mathbb{P}[C], \mathbb{P}[B]]\}$ such that for $t \leq s$: $A_t \subset A_s$, $C = A_{\mathbb{P}[A]}, B = A_{\mathbb{P}[B]}$ and such that $\mathbb{P}[A_t] = t$ for all t.*

6. *If $(A_t)_{t \in I \subset [0,1]}$ is an increasing family of sets with $\mathbb{P}[A_t] = t$, there is an increasing family of sets $(B_t)_{t \in [0,1]}$ with $\mathbb{P}[B_t] = t$ and where for $t \in I$: $B_t = A_t$.*

7. *If $(A_t)_{t \in I \subset [0,1]}$ is an increasing family of sets with $\mathbb{P}[A_t] = t$, then there is a random variable ξ with a uniform $[0,1]$ law and such that $\{\xi \leq t\} = A_t$.*

8. *There is a sequence of independent random variables r_n such that $\mathbb{P}[r_n = +1] = \mathbb{P}[r_n = -1] = 1/2$.*

9. *For an arbitrary non-degenerate probability distribution μ on $\mathbb{R}$, there is a sequence of independent identically distributed random variables $(\xi_n)_n$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and such that each ξ_n has the law μ.*

If $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless and $\mathbb{P}[A] > 0$, then $(A, A \cap \mathcal{F}, \mathbb{P}[. \mid A])$ is also atomless.

The proof of this theorem is a standard exercise in probability theory. However the solution is not at all trivial. The proof is essentially based on the following lemma. We do not include a proof.

Lemma 1 *Let $B \subset A$ and suppose $\mathbb{P}[B] < t < \mathbb{P}[A]$, then there is a set $C \in \mathcal{F}$ such that $B \subset C \subset A$ and so that $\mathbb{P}[C] = t$.*

Proof. Replacing the set A by $A \setminus B$ allows us to reduce the problem to $B = \emptyset$. □

Proposition 1 *If $\mathbb{P}$ is atomless and $\xi \in L^\infty$, there exists a sequence $(\xi_n)_n$ such that:*

1. *$\xi \leq \xi_n \leq \xi + \frac{1}{n}$;*

2. *$\xi_n \downarrow \xi$;*

3. *each ξ_n has a continuous distribution.*

Proof. The (obvious) details are left to the reader. Let $\{a_k \mid k \in \mathbb{N}\}$ be the discontinuity set of the distribution function $\mathbb{F}_\xi$ of ξ and let U_k stand for the set $\{\xi = a_k\}$. Then $\mathbb{P}[U_k] > 0$ and for each k we can construct a variable $\eta^k : U_k \to [0,1]$ with the uniform distribution under $\mathbb{P}[\cdot|U_k]$. Take now $\xi_n = \xi + \frac{1}{n}\sum_{k\geq 1}\eta^k \mathbf{1}_{U_k}$. It is easily seen that each ξ_n has a continuous distribution and that the sequence $(\xi_n)_n$ has the required properties. □

2.4 Commonotonicity

Let us start by giving a general definition of commonotonicity. The definition is not quite standard. We will show it implies the usual definition and it also has some *mathematical* beauty. We restrict the definition to the case of two random variables. In the literature the reader can find more general cases.

Definition 2 *Two random variables ξ, η, defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are commonotone if on the product space*

$$(\Omega \times \Omega, \mathcal{F} \otimes \mathcal{F}, \mathbb{P} \otimes \mathbb{P})$$

the random variable $Z(\omega_1, \omega_2) = (\xi(\omega_1) - \xi(\omega_2))(\eta(\omega_1) - \eta(\omega_2))$ is a.s. nonnegative.

To make the writing a little bit easier we use the notation $\xi_i(\omega_1, \omega_2) = \xi(\omega_i)$ (same for η_i). The random variable Z can then be written as $Z = (\xi_1 - \xi_2)(\eta_1 - \eta_2)$. We will keep this notation in the following analysis.

Lemma 2 If ξ and η are commonotone and square integrable, the covariance $Cov\,(\xi, \eta)$ is nonnegative.

Proof. By integrating $Z = (\xi_1 - \xi_2)(\eta_1 - \eta_2)$ on the product space we get:

$$0 \leq \int Z \, d(\mathbb{P} \otimes \mathbb{P}) = 2\left(\int \xi\eta \, d\mathbb{P} \; - \; \int \xi \, d\mathbb{P} \int \eta \, d\mathbb{P}\right).$$

□

Example 1 If we take a random variable ξ and two increasing functions $f, g : \mathbb{R} \to \mathbb{R}$, then the variables $f(\xi)$ and $g(\xi)$ are commonotone. The next propositions will show that this is the general situation for commonotone variables. If ξ and η are commonotone then they do not contribute to diversification. Both variables depend in the same way on a common source. So small values are added to small values and large values are added to large values. So it was no surprise that the correlation was nonnegative.

Proposition 2 *If ξ and η are commonotone random variables, then there exists a set Ω' with $\mathbb{P}[\Omega'] = 1$ and such that for all $(\omega_1, \omega_2) \in \Omega' \times \Omega'$ we have $(\xi(\omega_1) - \xi(\omega_2))(\eta(\omega_1) - \eta(\omega_2)) \geq 0$.*

Proof. The first step is to show that for two couples of real numbers $a < b$ and $c < d$, we necessarily have that either $\mathbb{P}[\xi \leq a, \eta \geq d] = 0$ or $\mathbb{P}[\xi \geq b, \eta \leq c] = 0$. Indeed if both are strictly positive then on $\Omega \times \Omega$ we have that $(\mathbb{P} \times \mathbb{P})[\xi_1 \leq a, \xi_2 \geq b, \eta_1 \geq d, \eta_2 \leq c] > 0$. This shows that $(\mathbb{P} \times \mathbb{P})[Z < 0] > 0$, a contradiction to the assumption. Let us now put

$$N' = \cup_{a<b, c<d, \text{rational with } \mathbb{P}[\xi \leq a, \eta \geq d] > 0} \{\xi \geq b, \eta \leq c\}.$$

Because of what is just proved, $\mathbb{P}[N'] = 0$. Let us put

$$N = N' \cup \left(\cup_{a<b, c<d \text{ rational with } \mathbb{P}[\xi \leq a, \eta \geq d] = 0} \{\xi \leq a, \eta \geq d\}\right).$$

Of course we still have $\mathbb{P}[N] = 0$. For $\omega_1 \notin N, \omega_2 \notin N$, we have that $(\xi(\omega_1) - \xi(\omega_2))((\eta(\omega_1) - \eta(\omega_2)) \geq 0$. Indeed suppose that the product is strictly negative. Then there are rational numbers a, b, c, d such that (maybe after interchanging ω_1, ω_2), $\xi(\omega_1) \leq a < b \leq \xi(\omega_2)$ and $\eta(\omega_1) \geq d > c \geq \eta(\omega_2)$. In case $\mathbb{P}[\xi \leq a, \eta \geq d] > 0$ this will imply $\omega_2 \in N' \subset N$, whereas $\mathbb{P}[\xi \leq a, \eta \geq d] = 0$ would imply $\omega_1 \in N$. So we may put $\Omega' = N^c$. □

Theorem 2 *If ξ, η are commonotone then there are two non-decreasing functions f, g such that $\xi = f(\xi + \eta), \eta = g(\xi + \eta)$.*

Proof. Take Ω' as above and look at the set $S = \{(\xi(\omega), \eta(\omega) \mid \omega \in \Omega'\}$. For $(x, y) \in S$ and $(x', y') \in S$ we have $(x - x')(y - y') \geq 0$. So this remains true for the closure of S (denoted by D). We now claim that $\phi \colon D \to \mathbb{R}, (x, y) \to \phi(x, y) = x + y$ is one to one. Indeed if $(x, y), (x', y') \in D$ and $x + y = x' + y'$ then necessarily $x = x', y = y'$ since otherwise $(x - x')(y - y') < 0$. We also claim that if $z_n = \phi(x_n, y_n)$ is a bounded sequence, then necessarily, the sequence (x_n, y_n) is bounded in D. Indeed if x_n is unbounded, we can extract a subsequence – still denoted x_n – such that either x_n is strictly decreasing to $-\infty$ or strictly increasing to $+\infty$. Let us suppose that $x_n \to +\infty$ (the other case is treated in a symmetric way). Since z_n is a bounded sequence we must have that $y_n \to -\infty$ and by taking a subsequence we may suppose the convergence is strictly decreasing. But then we have $(x_{n+1} - x_n)(y_{n+1} - y_n) < 0$, a contradiction to the commonotonicity. So we get that (x_n, y_n) is bounded as soon as $x_n + y_n$ is bounded. This is enough to show that the image $\phi(D)$ is closed and that the inverse function $(f,g)\colon \phi(D) \to D; (f,g)(x+y) = (x,y)$ is continuous. Obviously, f and g are non-decreasing and even Lipschitz. This ends the proof. □

2.5 Quantiles and Rearrangements

Definition 3 *Let ξ be a random variable and $\alpha \in (0,1)$.*

- *q is called an α–quantile if:*

$$\mathbb{P}[\xi < q] \le \alpha \le \mathbb{P}[\xi \le q]\,,$$

- *the largest α–quantile is:*

$$q_\alpha(\xi) = \inf\{x \mid \mathbb{P}[\xi \le x] > \alpha\},$$

- *the smallest α–quantile is:*

$$q_\alpha^-(\xi) = \inf\{x \mid \mathbb{P}[\xi \le x] \ge \alpha\}.$$

- *if no confusion is possible we drop the argument ξ and simple write q_x, q_x^-,*

- *in case $\alpha = 0$, we can define q_0 without any problem but we take $q_0^- = \lim_{x\to 0, x>0} q_x^-$. Similarly for $\alpha = 1$ we can define q_1^- in the usual way but we take $q_1 = \lim_{x\to 1, x<1} q_x$. In this way the quantiles q^- and q are defined on the closed interval $[0,1]$.*

- *Quantiles allow to define random variables that have the same probability law as the given function ξ. Indeed $q(\xi)\colon [0,1] \to \mathbb{R}; x \to q_x(\xi)$ is an increasing (better nondecreasing) function that has the same law as ξ. Sometimes $x \to q_x(\xi)$ is called the increasing rearrangement of ξ, the "opposite" $x \to q_{1-x}(\xi)$, is called the decreasing rearrangement.*

As easily seen, $q_\alpha^- \le q_\alpha$ and q is an α–quantile if and only if $q_\alpha^- \le q \le q_\alpha$. The set of points α where $q_\alpha^- < q_\alpha$, is at most countable since the (possibly empty) intervals $]q_\alpha^-, q_\alpha[$ are pairwise disjoint. The function q^- is therefore a.s. equal to the increasing rearrangement q.

Proposition 3 *Suppose that the probability space, $(\Omega, \mathcal{F}, \mathbb{P})$, is atomless, then for $\xi \in L^0(\Omega, \mathcal{F}, \mathbb{P})$ there is a uniformly $[0,1]-$distributed random variable $\upsilon \in L^0(\Omega, \mathcal{F}, \mathbb{P})$ such that $\xi = q_\upsilon(\xi)$.*

Proof. Let $A_{\mathbb{P}[\xi \le y]} = \{\xi \le y\}$, then according to Theorem 1 on atomless spaces there is a $[0,1]-$distributed random variable υ such that for all $y \in \mathbb{R}$: $\{\upsilon \le \mathbb{P}[\xi \le y]\} = A_{\mathbb{P}[\xi \le y]}$. Clearly $\xi = q_\upsilon(\xi)$. □

Remark 1 In the language of the previous section, the two random variables ξ and υ are commonotone. The variable $\xi' = q_{1-\upsilon}(\xi)$ has the same law as ξ but is anti-commonotone with ξ. If η is a uniformly $[0,1]-$distributed random variable, then $q_\eta(\xi)$ has the same law as ξ and every random variable that has the same law as ξ, is of this form, this is another way of formulating the previous proposition. The reasoning also shows:

Proposition 4 *Suppose that the probability space is atomless. If ξ, η are random variables, there exist η', η'' having the same law as η and such that ξ, η' are commonotone and ξ, η'' are anti-commonotone.*

Proof. Take υ as in the previous proposition, i.e. $\xi = q_\upsilon(\xi)$, and define $\eta' = q_\upsilon(\eta)$ and $\eta'' = q_{1-\upsilon}(\eta)$. □

The following lemma is due to Hardy and Littlewood (the reader can see that it is a modification of [78], theorem 378):

Lemma 3 *Suppose that $\xi \in L^\infty, 0 \le \eta \in L^1$, let ξ^* be the increasing rearrangement of ξ and η_* the decreasing rearrangement of η. Then*

$$\int_\Omega \xi\eta \ge \int_{[0,1]} \xi^* \eta_* = \int_0^1 q_x(\xi)\, q_{1-x}(\eta)\, dx.$$

In the same way the increasing rearrangement, η^*, of η satisfies:

$$\int_\Omega \xi\eta \le \int_{[0,1]} \xi^* \eta^* = \int_0^1 q_x(\xi)\, q_x(\eta)\, dx.$$

2.6 Some basic theorems from functional analysis

We will frequently make use of the standard duality theory from functional analysis. When we speak about the dual space we always mean the topological dual, i.e. the space of continuous real-valued functionals. The reader can find the relevant theorems in Dunford-Schwartz, [55], Grothendieck, [74] or in Rudin, [120]. We assume that the reader has an introductory knowledge of this theory and knows how to handle the "separation" theorems. We present some less known theorems that I consider as basic but nontrivial. The proofs are omitted. The following theorem is very useful when checking whether sets in a dual space are weak* closed. The theorem is sometimes called the Banach-Dieudonné theorem, sometimes it is referred to as the Krein-Sḿulian theorem.

Theorem 3 *Let E be a Banach space with dual space E^*. Then a convex set $C \subset E^*$ is weak* closed if and only if for each n, the set $W_n = C \cap \{e^* \mid \|e^*\| \le n\}$ is weak* closed.*

Of course, since convex sets that are closed for the so-called Mackey topology are already weak* closed, it suffices to check whether the sets W_n are Mackey closed. Most of the time, the description of the Mackey topology is not easy, but in the case of L^∞ we can make it more precise. Without giving a proof, we recall that on bounded sets of L^∞, the so-called Mackey topology coincides with the topology of convergence in probability. Checking whether a bounded convex set is weak* closed is then reduced to checking whether is it closed for the convergence in probability. More precisely we have the following lemma, which seems to be due to Grothendieck, [74].

Lemma 4 *Let $\mathcal{A} \subseteq L^\infty$ be a convex set. Then $\mathcal{A}$ is closed for the $\sigma(L^\infty, L^1)$ topology if and only if for each n, the set $W_n = \{\xi \mid \xi \in \mathcal{A}, \|\xi\|_\infty \le n\}$ is closed with respect to convergence in probability.*

Remark 2 We warn the reader that the above theorem allows to check whether a set is weak$^*-$closed. It is of no help in constructing the closure of a set.

We will use some more theorems that play a fundamental role in convex analysis, these are the Bishop-Phelps theorem and James's characterisation of weakly compact sets (see Diestel's book, [52], for a proof of these highly non-trivial results).

Theorem 4 *(Bishop-Phelps) Let $B \subset E$ be a bounded closed convex set of a Banach space E. The set $\{e^* \in E^* \mid e^* \text{ attains its supremum on } B\}$ is norm dense in E^*.*

Theorem 5 *(James) Let $B \subset E$ be a bounded closed convex set of a Banach space E. The set B is weakly compact if and only if each $e^* \in E^*$ attains its* *supremum on B. More precisely for each $e^* \in E^*$ there is $b_0 \in B$ such that $e^*(b_0) = \sup_{b\in B} e^*(b)$.*

Theorem 6 *Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless. Take $\alpha \in \mathbb{R}$. The set $\{\eta \mid 0 \le \eta \le 1, \mathbb{E}_{\mathbb{P}}[\eta] = \alpha\}$ is $\sigma(L^\infty, L^1)-$compact and convex. Its extreme points are all of the following form: $\eta = \mathbf{1}_B$ with $\mathbb{P}[B] = \alpha$. Consequently the weak$^*-$closed convex hull of these indicators is the set $\{\eta \mid 0 \le \eta \le 1, \mathbb{E}_{\mathbb{P}}[\eta] = \alpha\}$.*

For an elegant proof of this result we refer to Lindenstrauss [100].

2.7 The Fenchel-Legendre transform

To define the Fenchel-Legendre transform we need two vector spaces that are in duality (see [74]). Most of the time these spaces will be a Banach space E with its topological dual E^*. Of course the special cases of $E = \mathbb{R}$ or $E = \mathbb{R}^d$ are among the most important ones. If $f\colon E \to \mathbb{R} \cup \{+\infty\}$ is a convex function, that is lower semi continuous, then we define the Fenchel-Legendre transform as

$$g(x^*) = \sup_{x\in E} \left(x * (x) - f(x)\right).$$

Of course g, being a supremum of a family of affine functionals (parametrised by $x \in E$), is then convex. If f has a nonempty domain ($= \{x \mid f(x) < \infty\}$), then g is not $-\infty$. So g is a function

$$g\colon E^* \to \mathbb{R} \cup \{+\infty\}.$$

If E^* is equipped with say the weak* topology $\sigma(E^*, E)$, then g is lower semi-continuous as well. Of course it is a priori not excluded that $g(x^*) = +\infty$ for all x^*. It is beyond the scope of this book to give a thorough study of convex functions We refer to [118] and [112].

The following theorem is a consequence of the Hahn-Banach theorem. We do not give the general proof, which can be found in [118]. In the special case of $E = L^\infty$, $E^* = L^1$, we will give a proof adapted to the case of monetary utility functions.

Theorem 7 *Suppose that E, E^* are in a separating duality. Suppose that $f\colon E \to \mathbb{R} \cup \{+\infty\}$ is convex and lower semi-continuous for the topology $\sigma(E, E^*)$. Suppose that f is not identically $+\infty$, then the Fenchel-Legendre* *transform has the same properties $g\colon E^* \to \mathbb{R}\cup\{+\infty\}$, is convex, lower semi-continuous for the topology $\sigma(E^*, E)$ and it is not identically $+\infty$. Furthermore the Fenchel-Legendre transform of g is f:*

$$f(x) = \sup_{x^*\in E^*} \left(x^*(x) - g(x^*)\right).$$

Definition 4 *If f is a lower semi-continuous convex function $f\colon E \to \mathbb{R}\cup\{+\infty\}$, if $f(x) < \infty$, then we define*

$$\partial_x f = \{x^* \in E^* \mid \text{ for all } y \in E : f(y) - f(x) \geq x^*(y - x)\}.$$

$\partial_x f$ is called the subgradient of f in the point x.

Remark 3 The subgradient generalizes the derivative of f. It is possible that the subgradient is empty.

We have the following generalisation of Hölder's inequality, the proof is almost straightforward.

Theorem 8 *If g is the Fenchel-Legendre transform of a convex, lower semi-continuous convex function f, then for all $x^* \in E^*, x \in E$*

$$x^*(x) \leq f(x) + g(x^*),$$

with equality (in $\mathbb{R}$) if and only if $x^* \in \partial_x f$. In that case we also have $x^* \in \partial_{x^*} g$.

Exercise 1 For the following functions, the reader should calculate the Fenchel-Legendre transform and the subgradient. Write down the inequality of the preceding theorem.

1. $E = \mathbb{R}$, $1 \leq p < \infty$ and $f(x) = \frac{1}{p}|x|^p$.
2. $E = \mathbb{R}$, $f(x) = x\log(x)$ for $x > 0$, $f(0) = 0$ and $f(x) = +\infty$ for $x < 0$.
3. $E = \mathbb{R}^d$, $C \subset \mathbb{R}^d$ is a nonempty closed convex set and f is the "indicator" of C defined as $f(x) = 0$ for $x \in C$ and $f(x) = +\infty$ for $x \notin C$.
4. E is a Banach space and $f(x) = \|x\|$.
5. E is a Banach space, $1 \leq p < \infty$ and $f(x) = \frac{1}{p}\|x\|^p$.

6. E is a Hilbert space and $f(x) = \frac{1}{2}\|x\|^2$.

7. E is a Banach space, $C \subset E$ is a nonempty closed convex set and f is the "indicator" function of C defined as $f(x) = 0$ for $x \in C$ and $f(x) = +\infty$ for $x \notin C$.

8. E is a Banach space, E^* its topological dual, $C \subset E^*$ is a nonempty weak*−closed convex set and $f\colon E^* \to \mathbb{R} \cup \{+\infty\}$, is the "indicator" function of C defined as $f(x^*) = 0$ for $x^* \in C$ and $f(x^*) = +\infty$ for $x \notin C$. See also the remark at the end of this section.

Exercise 2 (Orlicz spaces and Young functions) Take a right continuous, non-decreasing (on the continent called increasing) function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$, $\phi(0) = 0$ and $\lim_{x\to+\infty} \phi(x) = +\infty$. Let ψ be the inverse of ϕ defined as $\psi(y) = \inf\{x \mid \phi(x) > y\}$ for $y \geq 0$. $\psi\colon \mathbb{R}_+ \to \mathbb{R}_+$ is right continuous, non-decreasing and $\psi(0) = 0$. Also $\lim_{x\to+\infty} \psi(x) = +\infty$. The function $f(x) = f(|x|) = \int_0^{|x|} \phi(u)\,du$ is then convex, $\lim_{x\to\infty} \frac{f(x)}{|x|} = +\infty$. The Legendre transform of f, here denoted as g, is defined in a similar way as f, namely $g(y) = \int_0^{|y|} \psi(v)\,dv$. The subgradient of f at a point $x > 0$ is the interval $[\lim_{w\to x, w<x} \phi(w), \phi(x)] = [\phi(x-), \phi(x)]$. Sometimes it is good to suppose that $\phi(1) = \psi(1) = 1$ which implies that $f(1) + g(1) = 1$. In this exercise we will make this assumption. With f we can associate the Banach space, L^f, of random variables (defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$):

$$\{\xi \mid \text{ there is } \alpha > 0 \text{ with } \mathbb{E}[f(\xi/\alpha)] < \infty\},$$

where $\|\xi\|_{L^f} = \|\xi\|_f = \inf\{\alpha \mid \mathbb{E}[f(\xi/\alpha)] \leq f(1)\}$. If $\phi(1) = 1$, then $\|a\| = a$ for all constants a. The spaces L^g and L^f form a dual pair. The inequality of Theorem 8, page 25, is then

$$|\mathbb{E}[\xi\eta]| \leq \|\xi\|_f \, \|\eta\|_g,$$

with equality if $\eta \in \partial_\xi f$. Here we need that $f(1) + g(1) = 1$. Essentially it means that $\eta = \phi(\xi)$ or $\xi = \psi(\eta)$.

Exercise 3 Explain why this is not completely correct.

The space
$$L^{(f)} = \{\xi \mid \text{ for all } \lambda > 0 : \mathbb{E}[f(\lambda\xi)] < \infty\}$$
can be strictly smaller than L^f. It is always true that L^g is the dual space of $L^{(f)}$, but the dual space of L^f can be much bigger than L^g. We refer to [93] for more information on Orlicz spaces and Young functions.

Exercise 4 See what happens with $\phi(x) = x^{p-1}$ for $1 < p < \infty$. See what happens for $\phi(x) = \exp(x) - 1$.

Remark 4 Most of the time the norm is defined as $\inf\{\alpha \mid \mathbb{E}[f(\xi/\alpha)] \le 1\}$. This has the disadvantage that the inequality in Theorem 8 above needs additional constants and that $\|a\|_f$ is not equal to a. Defining the norm as we did and asking that $\phi(1) = \psi(1) = 1$ makes live easier – but sometimes calculations more difficult – and allows to see the L^p spaces as special cases.

Remark 5 Suppose that E is a Banach space, E^* its topological dual and E^{**} its bidual (the dual of E^* for the norm topology). If $f\colon E^* \to \mathbb{R} \cup \{+\infty\}$ is a convex function with nonempty domain, we can have different topological notions of lower semi-continuity. In case f is lsc for the weak* topology (i.e. $\sigma(E^*, E)$, we can look at the pair (E^*, E). In that case g is defined on E. There is no guarantee that the subgradient is nonempty. In case f is lsc for the weak topology $\sigma(E^*, E^{**})$, we apply the reasoning to the pair (E^*, E^{**}). The function g is then defined on E^{**}. In that case the Hahn-Banach theorem shows that $\partial_{x^*} f$ is nonempty if $f(x) < \infty$. For convex functions the lsc for $\sigma(E^*, E^{**})$ is the same as for the norm topology on E^*. In the case of utility functions, we will use the duality (L^∞, L^1) but sometimes we will need the duality $(L^\infty, \mathbf{ba})$. We promise and will try not to mix up the two dualities and we ask the reader to be careful when extrapolating results.

2.8 The transform of a concave function

If $h\colon E \to \mathbb{R} \cup \{-\infty\}$ is a $\sigma(E, E^*)$ upper semi continuous concave function, the function $f(x) = -h(-x)$ is a lsc convex function. We define the transform of h as the transform of f. We know that this causes ambiguity when h is affine but we hope that this confusion does not happen and if it happens, that the reader will take care of it. We get

$$g(x^*) = \sup_{x \in E} (x^*(x) - f(x)) = \sup_{x \in E} (x^*(x) + h(-x)) = \sup_{x \in E} (-x^*(x) + h(x))$$

Exercise 5 Let $h : E \to \mathbb{R}$ be a linear function $h \in E^*$. Calculate the Fenchel-Legendre transform of h when seen as a convex function and with the convention above when seen as a concave function. Do the same for an affine function h.

Chapter 3

Value at Risk

3.1 Definition and properties of Quantiles

The philosophy behind the concept of VaR is the following: fix a threshold probability α (say 1%) and define a position as acceptable if and only if the probability to go bankrupt is smaller than α. At first sight this seems to be a good attitude towards risk. However as one can immediately see, the probability alone is not enough to deal with risky situations. Besides the probability of going below zero, the economic agent and especially a supervising authority should also consider what a bankruptcy (if it occurs) means. VaR does not distinguish between a bankruptcy of, say, 1 Euro or a bankruptcy of 1 hundred million Euro. Anyway, VaR is still the most widely (ab)used instrument to "control" risk and in order to study its properties we need more precise definitions. We can understand that limited liability plays a role for the shareholder, but the supervisor should be concerned with the effect of a bankruptcy on Society. By only considering the probability of ruin and not the amount of ruin, a free option is given to the management.

3.2 Definition of VaR

Definition 5 *Given a position* ξ *and a number* $\alpha \in [0,1]$ *we define*

$$\mathrm{VaR}_\alpha(\xi) := -q_\alpha(\xi)$$

and we call ξ *VaR-***acceptable** *if* $\mathrm{VaR}_\alpha(\xi) \leq 0$ *or, equivalently, if* $q_\alpha(\xi) \geq 0$.

We can think of the VaR as the amount of extra-capital that a firm needs in order to reduce the probability of going bankrupt to α. A negative VaR means that the firm would be able to give more money to its managers or to give back some money to its shareholders or that it could change its activities, e.g. it could accept more risk. We can also say that a position ξ, is VaR_α-acceptable if $\mathbb{P}[\xi < 0] \leq \alpha$. So we have two ways to use VaR. Either we say that VaR is the amount of capital to be added in order to become acceptable or we look at the quantile $q_\alpha(\xi)$ as a quantity that describes how good the position ξ is.

Remark 6 VaR_α has the following properties:

1. $\xi \geq 0 \Longrightarrow \text{VaR}_\alpha(\xi) \leq 0$,
2. $\xi \geq \eta \Longrightarrow \text{VaR}_\alpha(\xi) \leq \text{VaR}_\alpha(\eta)$,
3. $\text{VaR}_\alpha(\lambda\xi) = \lambda \text{VaR}_\alpha(\xi)$, $\forall \lambda \geq 0$,
4. $\text{VaR}_\alpha(\xi + k) = \text{VaR}_\alpha(\xi) - k$, $\forall k \in \mathbb{R}$.

In particular, we have $\text{VaR}_\alpha(\xi + \text{VaR}_\alpha(\xi)) = 0$. This simply means that if a position requires some capital, then adding this amount of capital produces a position that becomes acceptable.

Remark 7 In terms of the quantile q_α we can rewrite the preceding as:

1. $\xi \geq 0 \Longrightarrow q_\alpha(\xi) \geq 0$,
2. $\xi \geq \eta \Longrightarrow q_\alpha(\xi) \geq q_\alpha(\eta)$,
3. $q_\alpha(\lambda\xi) = \lambda q_\alpha(\xi)$, $\forall \lambda \geq 0$,
4. $q_\alpha(\xi + k) = q_\alpha(\xi) + k$, $\forall k \in \mathbb{R}$. In particular, we have $q_\alpha(\xi - q_\alpha(\xi)) = 0$.

3.3 Shortcomings

VaR has the nice property that it is defined on the whole space L^0. Therefore it can, in principle, be calculated for every random variable. The problem with such a degree of generality is that VaR_α **necessarily** violates convexity properties. Indeed we know that functionals defined on L^0 never have convexity properties. This result going back to Nikodym ([107]), is the mathematical reason why VaR leads to strange situations. In [3] we discussed some of these issues and warned for the lack of convexity. As an example, consider the case of a bank which has given a \$ 100 loan to a client whose default probability is equal to 0.008. If $\alpha = 0.01$, it is easy to see that $\text{VaR}_\alpha(\xi) \leq 0$. Consider now another bank which has given two loans of \$ 50 each and for both, the default probability is equal to 0.008. In case the default of the two loans are independent, $\mathrm{VaR}_\alpha(\xi)$ is \$ 50. Hence we have that diversification, which is commonly considered as a way to reduce risk, can lead to an increase of VaR. Therefore we argue that VaR is not a good measure of risk. This is the main reason why we are interested in studying other types of risk measures. Contrary to what is believed, examples such as the credit example mentioned, not only arise in theory. They also arise in practice and in a more complicated form, as was presented by Kalkbrenner et al, [87].

Using VaR could even lead to risk appetite (as shown by Leippold, Trojani and Vanini [101]).

Another problem with VaR is that it completely neglects what happens below the threshold. The consequence is that VaR neglects problems coming from avalanche effects or domino effects. For instance, we may have a (tractable) model where the default probability of an agent depends on an economic factor. For some agents the low default probability will remain low when the economy is in bad shape but for others the conditional probability can go up (and even become one) if the economy is in a bad shape. If the economy remains good, nothing serious happens. In case the economy turns the wrong side (say with a probability below the VaR-threshold), a very significant number of agents will go bankrupt, resulting in an extremely high loss. VaR will not detect it since such a bad development got a probability below the threshold. Selling such loans is then encouraged when VaR is used as a risk measure. We did not and did not even intend to use the politically loaded expression *subprime*.

Exercise 6 Build a model based on the previous reasoning.

The calculation of VaR requires the knowledge of the distribution of the profit and loss function. This is true for most risk measures. The calculation of this distribution requires a good model for the dependencies between the credit takers. One cannot take them independent but a complicated dependency leads to impossible calculations. One way out – different from the one sketched above – is then to use Gaussian copulas. These lead to calculations involving transformations of normally distributed variables. The problem is that these dependencies have the deficiency that the global law of the losses is completely determined by the correlation between two agents. This means that once we know the correlation between each pair of agents, one can calculate the law of the number of defaults. This again neglects the possibility of avalanches. The controlling authorities who are concerned by such developments, should not use VaR and should not allow to use VaR, see also [82]. Instead of Gaussian copulas there are many alternatives, see e.g. Schmock, [77].

Chapter 4

Coherent and Concave Utility Functions

4.1 Monetary Utility Functions

Before giving precise definitions, let us recall some terminology from decision theory. We start by recalling that a utility function is simply a function $u\colon L^\infty \to \mathbb{R}$. This is of course a much too general concept. We need to restrict the definition by adding additional properties.

Definition 6 *A utility function $u\colon L^\infty \to \mathbb{R}$ is quasi-concave if for each α the set $\{\xi \mid u(\xi) \geq \alpha\}$ is a convex set.*

Quasi-concavity is seen as a form of risk-averseness. The combination $\frac{\xi+\eta}{2}$ of two payoffs (i.e. random variables ξ and η) for which the economic agent is indifferent (i.e. $u(\xi) = u(\eta)$), is always better than each of the individual payoffs. This is also the mathematical way of saying that diversification is considered as better. The utility function is upper semi-continuous if the previously defined set is closed for each α, of course we need to mention the topology. We will distinguish between two kinds of topologies on L^∞, the topologies compatible with the duality (L^∞, L^1) (e.g. weak* or $\sigma(L^\infty, L^1)$, Mackey or $\tau(L^\infty, L^1)$, compact convergence or $\gamma(L^\infty, L^1)$) and the norm topology defined by $\| \quad \|_\infty$. The utility function is called (weakly) monotone if for each pair of bounded random variables $\xi \geq \eta$ a.s., we have that $u(\xi) \geq u(\eta)$. We remark that our utility functions are defined on spaces of random variables (with identification of random variables that are equal almost surely). In some chapters we will study utility functions that are defined on bigger spaces than L^∞. Sometimes the utility functions will take values in $\mathbb{R} \cup \{-\infty\}$. But we will avoid utility functions that take the value $+\infty$. The latter we consider as unrealistic since it would mean that a future payoff that has a utility equal to $+\infty$ is better than any other claim, it can, from a utility viewpoint, not be improved. A claim that is given a utility equal to $-\infty$ can be part of a realistic model. Indeed such a claim would be highly undesirable, e.g. a claim that cannot be insured because it is too risky.

Definition 7 *We say that a utility function* $u: L^\infty \to \mathbb{R}$ *is monetary if* $u(0) = 0$ *and if for each* $\xi \in L^\infty$ *and each* $k \in \mathbb{R}$ *we have*

$$u(\xi + k) = u(\xi) + k.$$

Remark 8 The term "monetary" was introduced by Föllmer and Schied, [68]. Previously the property was called money based utility function. In [3] and [4] it was called translation invariance. The idea is clear: the utility is measured in money units. Therefore it is numéraire dependent. If we assume, as a normalisation, that $u(0) = 0$, then on the one-dimensional space of constant random variables, the utility function is just the identity, $u(\alpha) = \alpha$. This is in contrast to the von Neumann-Morgenstern utility functions. In the case of monetary utility functions the risk averseness comes from the concavity property of the function when seen as a functional on the whole space L^∞. In the case of von Neumann-Morgenstern functions, the concavity on the space L^∞ is inherited from the concavity of a function on the real line. Here we start with utility functions defined on the space of bounded random variables and not with utility functions defined on the set of real numbers. The basic concepts are the random variables and not the lotteries. We will not discuss the differences with the von Neumann-Morgenstern theory or with its generalisations due to Gilboa-Schmeidler, [75] and Machina and Schmeidler, [104]The knowledge of the monetary utility function on the real line does not give any information on the utility for arbitrary random variables.

Exercise 7 Show that if $k \in \mathbb{R}$, then $u(k) = k$.

Exercise 8 A monetary utility function, u, is characterised by the preferred set $\mathcal{A}$ of 0. Prove that $\mathcal{A} = \{\xi \mid u(\xi) \geq 0\}$ and $u(\xi) = \sup\{\alpha \in \mathbb{R} \mid \xi - \alpha \in \mathcal{A}\}$ and show that the sup is a maximum.

Example 2 Suppose that v is a utility function defined on L^∞. Suppose that $v(0) = 0$ and suppose also that for every ξ and every $\varepsilon > 0$ we have $v(\xi + \varepsilon) > v(\xi)$. With v we associate the monetary utility function that has the same preferred set to 0. More precisely we define $u(\xi) = \sup\{\alpha \mid v(\xi - \alpha) \geq 0\}$. $u(\xi)$ is defined as an implicit function, namely $v(\xi - u(\xi)) = 0$. This procedure is not to be confused with the certainty equivalent. The latter is defined via the equation $v(\xi) = v(\alpha)$.

Example 3 We can use the preceding procedure when v is a von Neumann-Morgenstern utility function. Let us start with a concave, strictly increasing function $v : \mathbb{R} \to \mathbb{R}$. Let us suppose that $v(0) = 0$. The utility function u defined on L^∞ is defined through the relation $\mathbb{E}[v(\xi - u(\xi))] = 0$. The set of acceptable elements is $\mathcal{A} = \{\xi \mid \mathbb{E}[v(\xi)] \geq 0\}$. It is easily seen that the set $\mathcal{A}$ is convex. The function u is not of von Neumann-Morgenstern type.

Example 4 Here we analyse in more detail the preceding example for the exponential utility function $v_\alpha(x) = 1 - e^{-\alpha x}$ where $\alpha > 0$. The function $u_\alpha \colon L^\infty \to \mathbb{R}$ is then defined as

$$u_\alpha(\xi) = -\frac{1}{\alpha} \log \mathbb{E}[e^{-\alpha\xi}].$$

This function is also used as a premium principle, see Bühlmann, [25] and Gerber, [73]. Sometimes it is called the cumulant principle. For $\alpha \to 0$, the utility function $u_\alpha(\xi)$ tends to $u_0(\xi) = \mathbb{E}[\xi]$, whereas for $\alpha \to +\infty$, the function $u_\alpha(\xi)$ tends to $u_\infty(\xi) = \text{ess.inf}\, \xi$ (this is not a trivial exercise). This gives a clear indication why we call the parameter α the "risk-averseness" of the agent. The acceptable set is $\mathcal{A}_\alpha = \{\xi \mid \mathbb{E}[e^{-\alpha x}] \leq 1\} = \{-\eta \mid \mathbb{E}[e^{\alpha\eta}] \leq 1\}$. For the case $\alpha = 0$ we put $\mathcal{A}_0 = \{\xi \mid \mathbb{E}[\xi] \geq 0\}$ whereas with the case $\alpha = +\infty$ could be assigned the set $\mathcal{A}_\infty = \{\xi \mid \xi \geq 0 \; a.s. \}$. The former is too liberal, the latter too restrictive or too severe. One can see that for all $0 \leq \alpha \leq +\infty$, the set $\mathcal{A}_\alpha$ is convex. In order to prepare for duality theory, we will calculate for any $\mathbb{Q} \ll \mathbb{P}$ the quantity

$$c_\alpha(\mathbb{Q}) = \sup\{\mathbb{E}_\mathbb{Q}[-\xi] \mid \xi \in \mathcal{A}_\alpha\} = \sup\{\mathbb{E}_\mathbb{Q}[\eta] \mid \mathbb{E}[e^{\alpha\eta}] \leq 1\}.$$

We will use the following well known inequality proved using elementary calculus or by solving the exercises in chapter 2. For $x \geq 0$ and $y \in \mathbb{R}$ we have

$$xy \leq x \log x - x + e^y.$$

In the case $x > 0$ we have equality if and only $x = e^y$. This inequality shows that $\mathbb{E}[e^{\alpha\eta}] \leq 1$ implies $\mathbb{E}_\mathbb{Q}[\eta] \leq \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$. The equality could be achieved by $\eta = \frac{1}{\alpha} \log \frac{d\mathbb{Q}}{d\mathbb{P}}$ but unfortunately this random variable is not always in L^∞. So we need some truncation argument (left as an exercise) to come to the equality ($0 < \alpha < \infty$):

$$c_\alpha(\mathbb{Q}) = \frac{1}{\alpha} \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right].$$

Let us observe that $c_\alpha(\mathbb{P}) = 0$ meaning that $\xi \in \mathcal{A}_\alpha$ implies $\mathbb{E}[\xi] \geq 0$. For $\alpha \to 0$ we get that $c_\alpha(\mathbb{P}) = 0$ but $c_\alpha(\mathbb{Q})$ tends to $+\infty$ for $\mathbb{Q} \neq \mathbb{P}$. This limit is indeed the function $c_0(\mathbb{Q})$. For $\alpha \to +\infty$ we must be more careful. In this case we have $c_\infty(\mathbb{Q}) = 0$ for all $\mathbb{Q} \ll \mathbb{P}$. However $c_\alpha(\mathbb{Q})$ tends only to 0 for $\mathbb{Q} \ll \mathbb{P}$ with $\mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] < \infty$. This shows that convergence of utility functions and convergence of the penalty function are related but the relation is not always that easy.

The analysis can be carried a little bit further. We restrict it to the case $0 < \alpha < \infty$. From the definition of $c_\alpha(\mathbb{Q})$ it follows that $u_\alpha(\xi) \leq \inf_{\mathbb{Q}} \left(\mathbb{E}_{\mathbb{Q}}[\xi] + c_\alpha(\mathbb{Q})\right)$. If we take ξ such that $u_\alpha(\xi) = 0$, then $d\mathbb{Q} = e^{-\alpha\xi}\, d\mathbb{P}$ defines a probability measure and we get $\mathbb{E}_{\mathbb{Q}}[\xi] + c_\alpha(\mathbb{Q}) = 0 = u_\alpha(\xi)$. So we proved that

$$u_\alpha(\xi) = \inf_{\mathbb{Q}} \left(\mathbb{E}_{\mathbb{Q}}[\xi] + c_\alpha(\mathbb{Q})\right).$$

This equality is also a straightforward consequence of general duality arguments, [68]. We invite the reader to do the same analysis for a more general von Neumann-Morgenstern utility function v. Of course there is no hope to find closed form formulas.

Example 5 For this example we assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless. Let us take the von Neumann-Morgenstern utility function defined on $\mathbb{R}$ as: $v(x) = \beta x$ for $x \leq 0$ and $v(x) = \alpha x$ for $x \geq 0$. In order to be concave we suppose that $0 < \alpha \leq \beta$ where the case of equality leads to a trivial situation. The acceptability set is:

$$\mathcal{A} = \{\xi \mid \mathbb{E}[\alpha\xi^+] \geq \mathbb{E}[\beta\xi^-]\}.$$

The acceptability set does not change if we multiply v by a scalar, so we can, without loss of generality, suppose that $\alpha = 1 \leq \beta$. The acceptable set is a (convex) cone and this implies that $u(\xi)$ is a positively homegeneous function, i.e. for $\lambda \geq 0$ we have $u(\lambda\xi) = \lambda u(\xi)$. The function c therefore only takes two values: 0 and $+\infty$. We claim that the set of scenarios

$$\{\mathbb{Q} \mid \mathbb{E}_{\mathbb{Q}}[\xi] \geq 0 \text{ for all } \xi \in \mathcal{A}\} = \{\mathbb{Q} \mid c(\mathbb{Q}) = 0\}$$

is given by

$$\mathcal{S} = \left\{\mathbb{Q} \mid a \leq \frac{d\mathbb{Q}}{d\mathbb{P}} \leq \beta a \text{ for some } 0 < a\right\}.$$

Indeed if $\mathbb{Q} \in \mathcal{S}$ and $\xi \in \mathcal{A}$, we have $\mathbb{E}_{\mathbb{Q}}[\xi] \geq \mathbb{E}[\xi^+ a - \xi^- a\beta] \geq a\mathbb{E}[\xi^+ - \beta\xi^-] \geq 0$. Therefore $c(\mathbb{Q}) = 0$. Conversely if $\mathbb{Q} \notin \mathcal{S}$ we can find $\varepsilon > 0$ as well as two sets A and B such that: $A \subset \{\frac{d\mathbb{Q}}{d\mathbb{P}} \leq a\}$, $B \subset \{\frac{d\mathbb{Q}}{d\mathbb{P}} \geq b\}$, $b/a \geq \beta + \varepsilon$, $B \cap A = \emptyset$ and $0 < \mathbb{P}[A] = \mathbb{P}[B]/\beta$. Take now $\xi = \mathbf{1}_A - \mathbf{1}_B$. We get that $\mathbb{E}[\xi^+ - \beta\xi^-] = 0$ and hence $\lambda\xi \in \mathcal{A}$ for all $\lambda \geq 0$. But $\mathbb{E}_{\mathbb{Q}}[\xi] \leq a\mathbb{P}[A] - b\mathbb{P}[B] \leq a\mathbb{P}[A] - (\beta+\varepsilon)a\mathbb{P}[A]/\beta \leq -\varepsilon a\mathbb{P}[A]/\beta < 0$. Consequently we have by homogeneity that $c(\mathbb{Q}) = +\infty$.

All these utility functions are concave functions. This is no surprise since we have

Proposition 5 *If $u\colon L^\infty \to \mathbb{R}$ is a quasi-concave, monetary utility function, then u is concave.*

Proof. Let ξ, η be elements in L^∞, and let $\alpha = u(\xi), \beta = u(\eta)$. Then since u is monetary, $u(\xi - \alpha) = 0 = u(\eta - \beta)$. The quasi-concavity then implies that $u\left(\frac{\xi+\eta}{2} - \frac{\alpha+\beta}{2}\right) \geq 0$. Since u is monetary we get that

$$u\left(\frac{\xi+\eta}{2}\right) = u\left(\frac{\xi+\eta}{2} - \frac{\alpha+\beta}{2}\right) + \frac{\alpha+\beta}{2} \geq \frac{\alpha+\beta}{2}.$$

□

Corollary 1 *If u is a monetary utility function with preferred set to zero $\mathcal{A}$, then u is concave if and only if the set $\mathcal{A}$ is convex.*

Proposition 6 *The concave monetary utility function $u\colon L^\infty \to \mathbb{R}$ is weakly monotone (i.e. satisfies $\xi \geq \eta$ implies $u(\xi) \geq u(\eta)$) if and only if the acceptable set $\mathcal{A}$ contains L^∞_+, in other words $\xi \geq 0$ implies $u(\xi) \geq 0$. In this case we have the following properties*

1. *$a \leq \xi \leq b$ implies $a \leq u(\xi) \leq b$,*
2. *$u(\xi - u(\xi)) = 0$*
3. *$|u(\xi) - u(\eta)| \leq \|\xi - \eta\|_\infty$.*

Proof. The equivalence is not difficult but is tricky. If u is weakly monotone then clearly $L^\infty_+ \subset \mathcal{A} = \{\xi \mid u(\xi) \geq 0\}$. Conversely suppose that $\xi \leq \eta$ and suppose that $u(\xi) = 0$. We must show that $u(\eta) \geq 0$. Because u is monetary we only have to deal with the case $u(\xi) = 0$. Take $\varepsilon > 0$ and take $1 \leq \mu \in \mathbb{R}$ so that $\mu(\eta - \xi + \varepsilon) + \xi \geq 0$. This is certainly possible and it implies that $\mu(\eta - \xi + \varepsilon) + \xi \in \mathcal{A}$. Now take $\lambda = 1/\mu \leq 1$ and take the convex combination $\lambda(\xi + \mu(\eta - \xi + \varepsilon) + (1-\lambda)\xi$. This convex combination is equal to $\eta + \varepsilon$ and it belongs to $\mathcal{A}$ since each component belongs to $\mathcal{A}$. Hence by the monetary property $u(\eta) \geq -\varepsilon$. Since ε was arbitrary we get $u(\eta) \geq 0$. The first statement immediately follows from monotonocity, the second property is true because u is monetary. The third property is seen as follows. Clearly $\xi \leq \eta + \|\xi - \eta\|_\infty$ and therefore $u(\xi) \leq u(\eta + \|\xi - \eta\|_\infty) = u(\eta) + \|\xi - \eta\|_\infty$. The other inequality is obtained by interchanging the role of ξ and η. □

Corollary 2 *Under the hypothesis of the proposition we have that u is Lipschitz continuous and $\mathcal{A}$ is a norm closed convex subset of L^∞.*

Remark 9 *From now on we will always assume that concave monetary utility functions are also weakly monotone.*

Definition 8 *The utility function $u : L^\infty \to \mathbb{R}$ is called coherent if it satisfies the following properties*

1. $u(0) = 0$, $u(\xi) \geq 0$ *for* $\xi \geq 0$,
2. $u(\xi + \eta) \geq u(\xi) + u(\eta)$,
3. *for* $\lambda \in \mathbb{R}, \lambda \geq 0$ *we have* $u(\lambda\xi) = \lambda u(\xi)$,
4. *for* $\alpha \in \mathbb{R}$ *we have* $u(\xi + \alpha) = u(\xi) + \alpha$, *this means that u is monetary.*

We remark that the above properties 2 and 3, imply that a coherent utility function is necessarily concave. The difference between coherence and concavity is the homogeneity. Concave monetary utility functions were studied in [?], Coherent utility functions in [3],[4],[39],[40]. We will show later how to reduce the more general concave monetary utility functions to the case of coherent utility functions.

Definition 9 *A coherent risk measure is a function $\rho : L^\infty \longrightarrow \mathbb{R}$ such that*

1. $\xi \geq 0 \Longrightarrow \rho(\xi) \leq 0$,
2. $\rho(\lambda\xi) = \lambda\rho(\xi)$, $\forall\lambda \geq 0$,
3. $\rho(\xi + k) = \rho(\xi) - k$, $\forall k \in \mathbb{R}$,
4. $\rho(\xi + \eta) \leq \rho(\xi) + \rho(\eta)$.

As easily seen, ρ is a coherent risk measure if and only if $u = -\rho$ is a coherent utility function. Point *4* (sub-additivity) is the one which is not satisfied by VaR, even if it seems to be a reasonable assumption. In fact, subadditivity of a risk measure is a mathematical way to say that diversification leads to less risk. See [3] and [4] for a discussion of the axiomatics. One interpretation of a risk measure is the following. If the future financial position is described by the random variable ξ, then $\rho(\xi)$ is the amount of capital (positive or negative) that has to be added in order to become acceptable. A position ξ is acceptable if it does not require extra capital or in terms of utility functions: if the utility $u(\xi) \geq 0$. Although the monetary property is criticised by many economists, it is a natural property when dealing with capital requirement.

4.2 Characterisation of coherent risk measures

Because coherent utility functions are monetary, the utility function is completely described by the set of random variables that are preferred to zero. The following theorem describes how to construct examples of coherent utility functions. We first recall that **ba** is the dual space of the Banach space L^∞. The space L^∞ itself is the dual of L^1 but **ba** is much bigger than L^1.

Theorem 9 *With each coherent utility function u, we can associate a convex, $\sigma(\mathbf{ba}, L^\infty)-$compact set, $\mathcal{S}^{\mathbf{ba}}$ of normalised, finitely additive, nonnegative measures (also called finitely additive probability measures), such that*

$$u(\xi) = \inf\{\mu(\xi) \mid \mu \in \mathcal{S}^{\mathbf{ba}}\} = \min\{\mu(\xi) \mid \mu \in \mathcal{S}^{\mathbf{ba}}\}.$$

Conversely a set of finitely additive probability measures $\mathcal{S}^{\mathbf{ba}}$ defines via the relation $u(\xi) = \inf\{\mu(\xi) \mid \mu \in \mathcal{S}^{\mathbf{ba}}\}$, a coherent utility function.

Proof. This is standard duality theory. The polar of the normed-closed cone $\mathcal{A} = \{\xi \mid u(\xi) \geq 0\}$ is the $\sigma(\mathbf{ba}, L^\infty)-$closed cone $\mathcal{A}^o = \{\mu \mid \mu(\xi) \geq 0 \text{ for all } \xi \in \mathcal{A}\}$. Since $\mathcal{A} \supset L^\infty_+$ we get that $\mathcal{A}^o$ only contains nonnegative measures. Therefore $\mathcal{A}^o$ is generated by its "base" $\mathcal{S}^{\mathbf{ba}} = \{\mu \mid \mu(\Omega) = 1 \text{ and } \mu \in \mathcal{A}^o\}$. The bipolar theorem says that $\mathcal{A} = (\mathcal{A}^o)^o$. In other words $\xi \in \mathcal{A}$ if and only if for all $\mu \in \mathcal{S}^{\mathbf{ba}}$ we have $\mu(\xi) \geq 0$. The relation $u(\xi) = \sup\{a \mid \xi - a \in \mathcal{A}\}$ can therefore be rewritten as:

$$u(\xi) = \inf\{\mu(\xi) \mid \mu \in \mathcal{S}^{\mathbf{ba}}\}.$$

□

Remark 10 The set $\mathcal{S}^{\mathbf{ba}}$ is uniquely defined if we require it to be weak* compact and convex. There is a one-to-one correspondence between coherent utility functions and non-empty weak* compact convex subsets of $\mathbf{P}^{\mathbf{ba}}$. *The set $\mathcal{S}^{\mathbf{ba}}$ will always denote a weak* compact and convex set.*

The previous theorem allows us to give examples of coherent utility functions. By choosing the set $\mathcal{S}^{\mathbf{ba}}$ in a special way we get interesting examples. For a discussion of such examples, we prefer to wait since the more appealing examples are given by sets which are subsets of L^1 and not just subsets of **ba**.

4.3 The Fatou Property

To make things more constructive (in the analytic sense), we add a continuity axiom to the definition of a utility function.

Definition 10 *(The Fatou property.) We say that a utility function* $u\colon L^\infty \to \mathbb{R}$ *satisfies the Fatou property – we will say that* u *is Fatou – if for each uniformly bounded sequence* $(\xi_n)_{n\geq 1}$, $\sup_n \|\xi_n\|_\infty < \infty$,

$$\xi_n \xrightarrow{\mathbb{P}} \xi \quad \textit{implies} \quad u(\xi) \geq \limsup u(\xi_n)$$

It is possible to show (in a way similar as in the proof of Fatou's lemma) that, at least for monotone utility functions, the Fatou property is equivalent to a monotonicity property:

$$\sup_n \|\xi_n\|_\infty < \infty, \quad \xi_n \downarrow \xi \text{ a.s. implies } u(\xi_n) \downarrow u(\xi).$$

For completeness, let us sketch the details. Let $(\xi_n)_{n\geq 1}$, $\sup_n \|\xi_n\|_\infty < \infty$ be a sequence such that $\xi_n \to \xi$ a.s.. Then $\eta_n = \sup_{k\geq n} \xi_k$ decreases to ξ. The property above implies that $u(\eta_n)$ tends to $u(\xi)$ and since $\eta_n \geq \xi_n$, we get that $\limsup_n u(\xi_n) \leq \lim_n u(\eta_n) = u(\xi)$ as desired.

Exercise 9 Show that the reduction to a.s. convergent subsequences was allowed.

We can strengthen the previous monotonicity result in the following way.

Proposition 7 *For a coherent utility function* u*, the Fatou property is equivalent to the following statement: for each* $\xi \in L^\infty$ *and each sequence of decreasing sets* $A_n \in \mathcal{F}$*, with* $\lim_n \mathbb{P}[A_n] = 0$*, we have that* $u(\xi + \mathbf{1}_{A_n}) \to u(\xi)$.

Proof. Let $1 \geq \xi_n \geq \xi \geq 0$ be a decreasing sequence of random variables such that $\xi_n \downarrow \xi$ a.s.. Take $\varepsilon > 0$ and let $A_n = \{\xi_n > \xi + \varepsilon\}$. Clearly the sequence A_n is decreasing and $\mathbb{P}[A_n] \downarrow 0$. Since obviously $\xi_n \leq \xi + \varepsilon + \mathbf{1}_{A_n}$

we have that $u(\xi_n) \leq u(\xi + \mathbf{1}_{A_n}) + \varepsilon$ and therefore $\lim u(\xi_n) \leq u(\xi) + \varepsilon$. Since this is true for every $\varepsilon > 0$, the Fatou property follows. □

Remark 11 We warn the reader that it is not sufficient to require the monotonicity only for the case $\xi_n \downarrow 0$, i.e. for $\xi = 0$. This problem will be investigated after the characterisation theorem for Fatou coherent utility functions.

Definition 11 *We say that property (WC) is satisfied if for sequences of random variables*

$$\sup \|\xi_n\|_\infty < \infty, \quad \xi_n \uparrow 0 \textit{ implies } u(\xi_n) \uparrow 0.$$

Proposition 8 *Property (WC) implies the Fatou property.*

Proof. Let $-1 \leq \eta \leq \eta_n \leq 1$, $\eta_n \downarrow \eta$ a.s., then $u(\eta) \geq u(\eta_n) + u(\eta - \eta_n)$ implies $u(\eta) \geq \limsup u(\eta_n) + \lim u(\eta - \eta_n)$. By property (WC) the second term tends to zero and the Fatou property holds. □

Since the superadditivity inequality (used in the proof of the proposition), does not hold in the other direction, we get that property (WC) might be strictly stronger than the Fatou property (and as will be shown later, this is indeed the case).

4.4 Some Examples

Example 6 Let us take a family $\mathcal{S}$ of probability measures $\mathbb{Q}$, all absolutely continuous with respect to $\mathbb{P}$. We identify $\mathbb{Q}$ and $\frac{d\mathbb{Q}}{d\mathbb{P}}$, the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$. We can therefore identify $\mathcal{S}$ with a subset of $L^1(\Omega, \mathcal{F}, \mathbb{P})$. If we define

$$u_{\mathcal{S}}(\xi) = \inf\{\mathbb{E}_{\mathbb{Q}}[\xi] \mid \mathbb{Q} \in \mathcal{S}\}$$

then this $u_{\mathcal{S}}$ is a coherent utility function with the Fatou property and with acceptance cone

$$\mathcal{A} = \{\xi \mid \text{for all } \mathbb{Q} \in \mathcal{S} : \mathbb{E}_{\mathbb{Q}}[\xi] \geq 0\}.$$

Later we will show that **every coherent Fatou utility function has this form**.

Proof. By Theorem 9 only the Fatou property needs to be verified. If $\xi_n \xrightarrow{\mathbb{P}} \xi$ and $\|\xi_n\|_\infty \le 1$ then for every $\mathbb{Q} \in \mathcal{S}$ we have:

$$\mathbb{E}_{\mathbb{Q}}[\xi] \ge \limsup_n \mathbb{E}_{\mathbb{Q}}[\xi_n] \ge \limsup_n u_{\mathcal{S}}(\xi_n)$$

and therefore $u_{\mathcal{S}}(\xi) \ge \limsup_n u_{\mathcal{S}}(\xi_n)$. □
In working with a family $\mathcal{S}$, we can replace it with its convex L^1-closed hull, so that, from now on, we will take $\mathcal{S}$ to be convex and L^1-closed.

Example 7 We consider $\mathcal{S} = \{\mathbb{P}\}$. In this case, $u_{\mathcal{S}}(\xi) = \mathbb{E}_{\mathbb{P}}[\xi]$. A position ξ is then acceptable iff its average $\mathbb{E}_{\mathbb{P}}[\xi]$ is nonnegative. Clearly, such a risk attitude is too tolerant.

Example 8 Let us consider $\mathcal{S} = \{\mathbb{Q} \mid \text{probability on } (\Omega, \mathcal{F}), \mathbb{Q} \ll \mathbb{P}\} = \mathbf{P}$. In this case $u_{\mathcal{S}}(\xi) = \text{ess.inf}(\xi)$ and $u_{\mathcal{S}}(\xi) \ge 0$ if and only if $\xi \ge 0$ a.s. . Hence a position is acceptable if and only if it is nonnegative a.s. . The family $\mathcal{S}$ is too large and therefore $u_{\mathcal{S}}$ is too risk averse. Anyway this $u_{\mathcal{S}}$ provides an example of a coherent risk measure that satisfies the Fatou property but does not verify property (WC). If we consider $\xi_n = -e^{-nx}$ defined on $[0,1]$ with the Borel σ-algebra and the Lebesgue measure, we have $\xi_n \uparrow 0$, almost surely, while $\text{ess.inf}(\xi_n) = -1$.

Example 9 (TailVaR) Let us now see what happens for the convex closed set $\mathcal{S}_k = \{\mathbb{Q} \mid \frac{d\mathbb{Q}}{d\mathbb{P}} \le k\}$. Obviously we only need to investigate the case $k > 1$; indeed, $\frac{d\mathbb{Q}}{d\mathbb{P}} \le 1$ implies that $\mathbb{Q} = \mathbb{P}$, i.e. $\mathcal{S}_1$ reduces to the singleton $\{\mathbb{P}\}$. To avoid technicalities we first deal with the case where the law of ξ is continuous, this means that the distribution function $\mathbb{F}(x) = \mathbb{P}[\xi \le x]$ is continuous. The case where $\mathbb{F}$ might have jumps is done at the end.

Theorem 10 *If ξ has a continuous distribution function and $\alpha = 1/k$, then*

$$u_{\mathcal{S}_k}(\xi) = \mathbb{E}_{\mathbb{P}}[\xi \mid \xi \le q_\alpha(\xi)] \le q_\alpha(\xi) = -\text{VaR}_\alpha(\xi).$$

Proof. Since ξ has a continuous distribution, we get $\mathbb{P}[\xi \le q_\alpha(\xi)] = \alpha = 1/k$. Define now $\mathbb{Q}_0$ such that $\frac{d\mathbb{Q}_0}{d\mathbb{P}} = k1_A$ with $A = \{\xi \le q_\alpha(\xi)\}$. Since $\mathbb{Q}_0 \in \mathcal{S}_k$ and $\mathbb{E}_{\mathbb{Q}_0}[\xi] = \mathbb{E}_{\mathbb{P}}[\xi \mid A]$ we have $u_{\mathcal{S}_k}(\xi) \le \mathbb{E}_{\mathbb{P}}[\xi \mid \xi \le q_\alpha(\xi)]$. By considering now an arbitrary $\mathbb{Q} \in \mathcal{S}_k$, we have

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[\xi] - \mathbb{E}_{\mathbb{Q}_0}[\xi] &= \mathbb{E}_{\mathbb{Q}}[\xi - q_\alpha] - \mathbb{E}_{\mathbb{Q}_0}[\xi - q_\alpha] \\ &= \mathbb{E}\left[(\xi - q_\alpha)\left(\frac{d\mathbb{Q}}{d\mathbb{P}} - \frac{d\mathbb{Q}_0}{d\mathbb{P}}\right)\right] \\ &= \int_A (\xi - q_\alpha)\left(\frac{d\mathbb{Q}}{d\mathbb{P}} - k\right) d\mathbb{P} + \int_{A^c} (\xi - q_\alpha)\,\frac{d\mathbb{Q}}{d\mathbb{P}}\, d\mathbb{P} \\ &\geq 0,\end{aligned}$$

where the last inequality follows because both terms are nonnegative. This ends the proof. □

In Chapter 8 we will give a relation between this utility function and Value at Risk.

In case the distribution of ξ has a discontinuity at q_α, the probability measure $\mathbb{Q}_0$ such that $\frac{d\mathbb{Q}_0}{d\mathbb{P}} = k1_{\{\xi<q_\alpha\}} + \beta 1_{\{\xi=q_\alpha\}}$ (with a suitably chosen $\beta, 0 \leq \beta \leq 1$) does the job. It implies that

$$u_{\mathcal{S}_k}(\xi) = \frac{1}{\alpha}\left(\int_{\xi<q_\alpha} \xi\, d\mathbb{P} + (\alpha - \mathbb{P}[\xi < q_\alpha])\, q_\alpha\right).$$

Using the increasing rearrangement of ξ this can also be written as

$$u_{\mathcal{S}_k}(\xi) = \int_0^\alpha q_u(\xi)\, du.$$

A similar calculation as above shows that for general ξ and for $\{\xi < q_\alpha\} \subset A \subset \{\xi \leq q_\alpha\}$ with $\mathbb{P}[A] = \alpha = 1/k$ we have that $u_{\mathcal{S}_k}(\xi) = k\int_A \xi\, d\mathbb{P}$. But the calculation shows something more. In case $\mathbb{P}[\xi \leq q_\alpha] = \alpha$, the set A is uniquely defined and we have for $\mathbb{Q} \neq \mathbb{Q}_0, \mathbb{Q} \in \mathcal{S}_k$ that the inequality $\int \xi\, d\mathbb{Q}_0 < \int \xi\, d\mathbb{Q}$ is strict! Indeed on A^c we have that $\xi > q_\alpha$ and hence equality would imply that $\mathbb{Q}[A^c] = 0$. Together with $\frac{d\mathbb{Q}}{d\mathbb{P}} \leq k$ this gives $\frac{d\mathbb{Q}}{d\mathbb{P}} = k\mathbf{1}_A$. The measure $\mathbb{Q}_0$ is the unique element in $\mathcal{S}_k$ that gives the quantity $u_{\mathcal{S}_k}(\xi)$. In case we have $\mathbb{P}[\xi < q_\alpha] = \alpha$ we get the same result: the measure $\mathbb{Q}_0$ defined as $\frac{d\mathbb{Q}_0}{d\mathbb{P}} = k\mathbf{1}_A$ with $A = \{\xi < q_\alpha\}$ is the unique measure for which $\mathbb{Q}_0[\xi] = u_{\mathcal{S}_k}(\xi)$. In these cases we say that $\mathbb{Q}_0$ is exposed and that ξ is an exposing functional. This has consequences regarding the differentiability of u.

In case $\mathbb{P}[\xi \leq q_\alpha] > \alpha > \mathbb{P}[\xi < q_\alpha]$ we have more elements in $\mathcal{S}_k$ where the infimum is attained. The set where the minimum is attained is given by

$$\left\{ k\mathbf{1}_{\{\xi<q_\alpha\}} + h\mathbf{1}_{\{\xi=q_\alpha\}} \mid h \in L^\infty; 0 \le h \le k; \int_{\{\xi=q_\alpha\}} h = \alpha - \mathbb{P}[\xi < q_\alpha] \right\}.$$

In the case where the space is atomless we can do more. The extreme points of this set are the elements where $h = k\mathbf{1}_B$ where $B \subset \{\xi = q_\alpha\}$ and $\mathbb{P}[B] = \alpha - \mathbb{P}[\xi < q_\alpha]$. In case $\mathbb{P}[\xi \le q_\alpha] > \alpha > \mathbb{P}[\xi < q_\alpha]$ there are infinitely many choices for the set B.

Example 10 (taken from [40]). This example (with an interpretation in Credit Risk) shows a bad performance of VaR against $u_{\mathcal{S}_k}$. Let us imagine there is a bank which lends \$1 to 150 clients. The clients are supposed to be independent with the same default probability p of 1.2%. For each client i let us put $Z_i = 0$ if he/she does not default and $Z_i = -1$ if he/she defaults. So we suppose $(Z_i)_i$ are independent Bernoulli random variables with $\mathbb{P}[Z_i = -1] = 1.2\%$. The number $Z = \sum_i Z_i$ represents the total number of defaults and therefore the bank's losses. It has the binomial distribution $0 \le k \le 150$:

$$\mathbb{P}[Z = -k] = \binom{150}{k} p^k (1-p)^{150-k}.$$

With $\alpha = 1\%$ we have $VaR_\alpha = 5$ and tail expectation $u_{1/\alpha}(Z) = -6.287$.

If we modify the example and suppose that the clients are **dependent**, things change. A simple way of obtaining a well-behaved dependence structure is by replacing $\mathbb{P}$ with a new probability measure $\mathbb{Q}$ defined as:

$$d\mathbb{Q} = c\, e^{\varepsilon Z^2} d\mathbb{P}\,,$$

where Z and $\mathbb{P}$ are the same as before, ε is positive and c is a normalising constant. Now $\mathbb{Q}[Z_i = -1]$ increases with ε: if we take ε so that $\mathbb{Q}[Z = -1] = 1.2\%$ (taking $p = 1\%$ and $\varepsilon = 0.03029314$), we obtain $VaR_\alpha = 6$ and tail expectation $u_{1/\alpha}(Z) = -14.5$.

We notice that VaR is not able to detect the difference between the two cases, which are better differentiated by TailVaR.

This can be explained as follows. VaR only looks at a quantile, it does not tell us how big the losses are. However, TailVaR takes an average over the worst cases and therefore takes into account the tail distribution of the losses. The probability $\mathbb{Q}$ allows to introduce loans whose defaults are dependent on a common economic factor. It reflects the situation that if a substantial percentage defaults, the conditional probability that others default as well, is very high. For other dependence structures and the relation with copula theory and Dirichlet distributions we refer to work of Schmock et al, [77].

Example 11 We could also consider the following family (where $k > 1$ and $p > 1$):

$$\mathcal{S}_{p,k} = \left\{ \mathbb{Q} \mid \left\| \frac{d\mathbb{Q}}{d\mathbb{P}} \right\|_p \leq k \right\}$$

The following theorem holds:

Theorem 11 *There exists a constant $c = 1 \wedge (k-1)$ such that for all $\xi \in L^\infty$, $\xi \geq 0$ we have:*

$$c\|\xi\|_q \leq -u_{\mathcal{S}_{p,k}}(-\xi) \leq k\|\xi\|_q$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. This proof comes from [39]. For each $h \in \mathcal{S}_{p,k}$ we have

$$\mathbb{E}[\xi h] \leq \|h\|_p \, \|\xi\|_q \leq k\|\xi\|_q.$$

This shows that

$$-u_{\mathcal{S}_{p,k}}(-\xi) = \sup_{h \in \mathcal{S}_{p,k}} \mathbb{E}[h\xi] \leq k\|\xi\|_q.$$

The inequality on the left is more difficult. We suppose that ξ is not identically zero, since otherwise there is nothing to prove. We then define $\eta = \frac{\xi^{q-1}}{\|\xi\|_q^{q-1}}$. As well known and easily checked, we have $\|\eta\|_p = 1$. The random variable η satisfies $\mathbb{E}[\eta] \leq 1$ and $\mathbb{E}[\xi\eta] = \|\xi\|_q$. We now distinguish two cases:

Case 1: $(1 - \mathbb{E}[\eta]) \leq k - 1$. In this case we put $h = \eta + 1 - \mathbb{E}[\eta]$. Clearly $\mathbb{E}[h] = 1$ and $\|h\|_p \leq \|\eta\|_p + 1 - \mathbb{E}[\eta] \leq 1 + k - 1 = k$. We also have $\mathbb{E}[h\xi] \geq \mathbb{E}[\eta\xi] = \|\xi\|_q$.

Case 2: $(1 - \mathbb{E}[\eta]) \geq k - 1$. (This implies that $k \leq 2$). We now take

$$h = \alpha\eta + 1 - \alpha\mathbb{E}[\eta] \text{ where } \alpha = \frac{k-1}{1 - \mathbb{E}[\eta]}.$$

Clearly $h \geq 0$, $\mathbb{E}[h] = 1$ and $\|h\|_p \leq \alpha + 1 - (k-1)\mathbb{E}[\eta]/(1 - \mathbb{E}[\eta]) \leq k$. But also $\mathbb{E}[\xi h] \geq \alpha\|\xi\|_q \geq (k-1)\|\xi\|_q$, since $1 - \mathbb{E}[\eta] \leq 1$. □

Remark 12 If k tends to 1, c tends to 0 and the family $\mathcal{S}_{p,k}$ shrinks to $\{\mathbb{P}\}$. That c tends to zero has to be expected since the L^p and the L^1 norms are not equivalent.

Remark 13 Actually, if $p = q = 2$ we have:

$$\left\| \frac{d\mathbb{Q}}{d\mathbb{P}} - 1 \right\|_2^2 = E\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right)^2 \right] - 1 \leq k^2 - 1$$

so that the densities go to 1 in L^2 as k tends to 1. If $p \geq 2$ we can use the same argument (remember that $\|.\|_2 \leq \|.\|_p$) and if $p < 2$, Clarkson's inequality for L^p-norms must be used.

Example 12 This example is related to work of T. Fischer, see [66]. He suggested, among other constructions, the following coherent utility function. For $\xi \in L^\infty$ we define

$$u(\xi) = \mathbb{E}[\xi] - \alpha \|(\xi - \mathbb{E}[\xi])^-\|_p.$$

The reader can verify that for $0 \leq \alpha \leq 1$ and $1 \leq p \leq \infty$, the above formula defines a coherent utility function. This measure can also be found using a set of probability measures. So let

$$\mathcal{S} = \{1 + \alpha(g - \mathbb{E}[g]) \mid g \geq 0; \|g\|_q \leq 1\}.$$

Here of course $q = p/(p-1)$, with the usual interpretation if $p = 1, \infty$. Clearly the set $\mathcal{S}$ is a convex L^1-closed set of functions h that have expectation equal to 1. We still have to check the positivity of such functions. This is easy since, by $g \geq 0$ and $\alpha \in [0, 1]$, we have

$$1 + \alpha(g - \mathbb{E}[g]) \geq 1 - \alpha\, \mathbb{E}[g] \geq 1 - \|g\|_q \geq 0.$$

We will check that

$$u(\xi) = \inf \{\mathbb{E}[h(\xi)] \mid h \in \mathcal{S}\}.$$

To see this, take $h = 1 + \alpha(g - \mathbb{E}[g])$ where $g = \frac{((\xi - \mathbb{E}[\xi])^-)^{(p-1)}}{\|(\xi - \mathbb{E}[\xi])^-\|_p^{(p-1)}}$. This is the standard way to obtain the $p-$norm by integrating against a function with $q-$norm equal to 1. In case $p = 1$ and therefore $q = \infty$, we take for g the indicator function of the set where $\xi < \mathbb{E}[\xi]$. For this choice of g and h we get:

$$\begin{aligned} \mathbb{E}[h\,\xi] &= \mathbb{E}[\xi] + \mathbb{E}[h\,(\xi - \mathbb{E}[\xi])] \\ &= \mathbb{E}[\xi] + \mathbb{E}[(h - 1 - \alpha\,\mathbb{E}[g])(\xi - \mathbb{E}[\xi])] = \mathbb{E}[\xi] - \alpha\,\|(\xi - \mathbb{E}[\xi])^-\|_p. \end{aligned}$$

For an arbitrary $1+\alpha(g-\mathbb{E}[g]) = h \in \mathcal{S}$ we have, by Hölder's inequality:

$$\mathbb{E}[h\,\xi] \geq \mathbb{E}[\xi] - \|h - 1 - \alpha\,\mathbb{E}[g]\|_q \,\|(\xi-\mathbb{E}[\xi])^-\|_p \geq \mathbb{E}[\xi] - \alpha\,\|(\xi-\mathbb{E}[\xi])^-\|_p.$$

4.5 Characterisation of coherent utility functions with the Fatou property

Let u be a coherent utility function, $u : L^\infty \longrightarrow \mathbb{R}$ and let us assume that the Fatou property holds. Let $\mathcal{A}$ be the set of the acceptable positions, i.e. $\mathcal{A} = \{\xi \mid u(\xi) \geq 0\}$. We note that $\mathcal{A}$ is a convex cone. The next theorem focuses on the relations between u and $\mathcal{A}$:

Theorem 12 *If u satisfies the Fatou property, then $\mathcal{A}$ is closed for the weak* topology $\sigma(L^\infty, L^1)$. Conversely, if $\mathcal{A}$ is a convex cone, closed in the $\sigma(L^\infty, L^1)$ topology and containing L^∞_+, then $\tilde{u}(\xi)$ defined as $\tilde{u}(\xi) = \sup\{\alpha \mid \xi - \alpha \in \mathcal{A}\}$ is a coherent utility function with the Fatou property.*

Moreover if u is a coherent utility function satisfying the Fatou property, there is a convex closed set of probability measures $\mathcal{S} \subset L^1$ such that $u(\xi) = \inf_{\mathbb{Q}\in\mathcal{S}} \mathbb{E}_{\mathbb{Q}}[\xi]$.

Proof. Let us call W the intersection of $\mathcal{A}$ with the unit ball of L^∞. By the Krein-Smulian theorem, if W is closed in the weak* topology, then $\mathcal{A}$ is also closed. We now take a sequence $(\xi_n)_n \in W$ such that $\xi_n \xrightarrow{\mathbb{P}} \xi$. But then $u(\xi) \geq \limsup u(\xi_n) \geq 0$ hence $\xi \in W$, that is W is closed under convergence in probability. In order to show the representation formula, we consider the following:

$$\mathcal{A}^o = \{f \mid f \in L^1 \text{ and } \forall \xi \in \mathcal{A} : \mathbb{E}[\xi f] \geq 0\}$$

which is, by definition, the polar cone of $\mathcal{A}$, taken in L^1. $\mathcal{A}^o$ is L^1 closed and (because $\mathcal{A} \supseteq L^\infty_+$) it is contained in L^1_+. We define $\mathcal{S}$ to be the closed convex set $\{f \in \mathcal{A}^o \,|\, \mathbb{E}[f] = 1\}$, which is, by the way, a basis of the cone $\mathcal{A}^o$. This means that $\mathcal{A}^o = \cup_{\lambda\geq 0}\lambda\mathcal{S}$. The bipolar theorem guarantees that:

$$\begin{aligned}\mathcal{A} &= \{\xi \mid \forall f \in \mathcal{A}^o : \ \mathbb{E}[\xi f] \geq 0\}\\ &= \{\xi \mid \forall f \in \mathcal{S} : \ \mathbb{E}[\xi f] \geq 0\}\end{aligned}$$

and therefore:

$$\begin{aligned}u(\xi) &= \sup\{\alpha \mid \xi - \alpha \in \mathcal{A}\}\\ &= \sup\{\alpha \mid \forall f \in \mathcal{S} : \quad \mathbb{E}[(\xi-\alpha)f] \geq 0\}\\ &= \sup\{\alpha \mid \forall f \in \mathcal{S} : \quad \mathbb{E}[(\xi)f] \geq \alpha\}\\ &= \inf\{\mathbb{E}[\xi\, f] \mid f \in \mathcal{S}\}\end{aligned}$$

□

Exercise 10 Suppose that u is a coherent utility function defined on L^∞. The function u is upper semi continuous for the weak* topology $\sigma(L^\infty, L^1)$ if and only if it satisfies the Fatou property. In that case the Fenchel-Legendre transform of u is the indicator function of the set $\mathcal{S}$ of Theorem 12.

Remark 14 We have in fact established a one-to-one correspondence between:
(a) convex closed sets $\mathcal{S}$ consisting of probabilities which are absolutely continuous with respect to $\mathbb{P}$,
(b) $\sigma(L^\infty, L^1)$-closed convex cones $\mathcal{A}$ containing L^∞_+,
(c) coherent utility functions u with the Fatou property.

4.6 The relation between $\mathcal{S}$ and $\mathcal{S}^{\mathbf{ba}}$.

For coherent utility functions with the Fatou property we now have two representations. One with finitely additive measures, the other one with $\sigma-$additive measures. There must be a relation between these two representations. This relation is described in the following proposition.

Proposition 9 *Let $u : L^\infty \to \mathbb{R}$ be a coherent utility function with the Fatou property. Let $\mathcal{S}$ be the closed convex subset of L^1 such that $u(\xi) = \inf\{\mathbb{E}[f\,\xi] \mid f \in \mathcal{S}\}$. Let u also be represented by the weak* closed convex set $\mathcal{S}^{\mathbf{ba}}$ of $\mathbf{ba}$. Then $\mathcal{S}$ is $\sigma(\mathbf{ba}, L^\infty)$ dense in $\mathcal{S}^{\mathbf{ba}}$.*

Proof. This is a trivial consequence of the Hahn-Banach theorem. Indeed we have, for each $\xi \in L^\infty$:

$$\inf\{\mu(\xi) \mid \mu \in \mathcal{S}^{\mathbf{ba}}\} = u(\xi) = \inf\{\mathbb{E}[f\,\xi] \mid f \in \mathcal{S}\}.$$

□

Corollary 3 *Let u be a coherent utility function represented by $\mathcal{S}^{\mathbf{ba}} \subset \mathbf{ba}$, then u has the Fatou property if and only if $\mathcal{S}^{\mathbf{ba}} \cap L^1$ is weak*$-$dense in $\mathcal{S}^{\mathbf{ba}}$.*

In a previous section we have shown that the Fatou property is equivalent to a convergence property for decreasing sequences, Section 4.3. We have warned the reader that it is not sufficient to require the property for sequences that decrease to 0. The following theorem makes this result precise.

Theorem 13 *For a coherent utility function, $u\colon L^\infty \to \mathbb{R}$, the following are equivalent*

1. *For every decreasing sequence of sets $(A_n)_{n\geq 1}$ with empty intersection, we have that $u(\mathbf{1}_{A_n})$ tends to zero.*
2. $\sup\left\{\|\mu_a\| \mid \mu \in \mathcal{S}^{\mathbf{ba}}\right\} = 1$, *(where $\mu = \mu_a + \mu_p$ is the Yosida–Hewitt decomposition).*
3. *The distance from $\mathcal{S}^{\mathbf{ba}}$ to L^1, defined as $\inf\{\|\mu - f\| \mid \mu \in \mathcal{S}^{\mathbf{ba}}, f \in L^1(\mathbb{P})\}$, is zero. A particular case is $\mathcal{S}^{\mathbf{ba}} \cap L^1 \neq \emptyset$.*

Proof. We start the proof of the theorem with the implication that (2) $\Rightarrow$ (1). So we take $(A_n)_{n\geq 1}$ a decreasing sequence of sets in $\mathcal{F}$ with empty intersection. We have to prove that for every $\varepsilon > 0$ there is n and $\mu \in \mathcal{P}_{\mathbf{ba}}$, such that $\mu(A_n) \leq \varepsilon$. In order to do this we take $\mu \in \mathcal{S}^{\mathbf{ba}}$ such that $\|\mu_a\| \geq 1 - \varepsilon/2$. Then we take n so that $\mu_a(A_n) \leq \varepsilon/2$. It follows that $\mu(A_n) \leq \varepsilon/2 + \|\mu_p\| \leq \varepsilon$.

The fact that 1 implies 2 is the most difficult one and it is based on the following lemma, whose proof is given after the proof of the theorem.

Lemma 5 *If K is a closed, weak* compact, convex set of finitely additive probability measures, such that $\delta = \inf\{\|\nu_p\| \mid \nu \in K\} > 0$, then there exists a non-increasing sequence of sets A_n, with empty intersection, such that for all $\nu \in K$, and for all n, $\nu(A_n) > \delta/4$.*

If (2) were false, then

$$\inf\left\{\|\mu_p\| \mid \mu \in \mathcal{S}^{\mathbf{ba}}\right\} > 0.$$

We can therefore apply the lemma in order to get a contradiction to (1).

The proof that (2) and (3) are equivalent is almost trivial and is left to the reader. □

In the proof of the lemma, we will need a minimax theorem. Since there are many forms of the minimax theorem, let us recall the one we need. It is not written in its most general form, but this version will do. For a proof, a straightforward application of the Hahn–Banach theorem together with the Riesz representation theorem, we refer to [45], page 404.

Theorem 14 (Minimax Theorem) *Let K be a compact convex subset of a locally convex space F. Let L be a a convex set of an arbitrary vector space E. Suppose that ϕ is a bilinear function $\phi\colon E \times F \to \mathbb{R}$. For each $l \in L$ we* *suppose that the partial (linear) function $\phi(l,.)$ is continuous on F. We then have*

$$\inf_{l\in L}\sup_{k\in K}\phi(l,k)=\sup_{k\in K}\inf_{l\in L}\phi(l,k).$$

Proof of Lemma 5 If λ is purely finitely additive, nonnegative, then the Yosida–Hewitt theorem implies the existence of a decreasing sequence of sets, say B_n (depending on λ!), with empty intersection and such that $\lambda(B_n)=\|\lambda\|$. Given $\mu\in K$, it follows that for every $\varepsilon>0$, there is a set, A (depending on μ), such that $\mathbb{P}[A]\le\varepsilon$ and such that $\mu(A)\ge\delta$. For each $\varepsilon>0$ we now introduce the convex set, F_ε, of functions, $f\in L^\infty$ such that f is nonnegative, $f\le 1$ and $\mathbb{E}_\mathbb{P}[f]\le\varepsilon$. The preceding reasoning implies that

$$\inf_{\mu\in K}\sup_{f\in F_\varepsilon}\mathbb{E}_\mu[f]\ge\delta.$$

Since the set K is convex and weak* compact, we can apply the minimax theorem and we conclude that

$$\sup_{f\in F_\varepsilon}\inf_{\mu\in K}\mathbb{E}_\mu[f]\ge\delta.$$

It follows that there is a function $f\in F_\varepsilon$, such that for all $\mu\in K$, we have $\mathbb{E}_\mu[f]\ge\delta/2$. We apply the reasoning for $\varepsilon=2^{-n}$ in order to find a sequence of nonnegative functions f_n, such that for each $\mu\in K$ we have $\mathbb{E}_\mu[f_n]\ge\delta/2$ and such that $\mathbb{E}_\mathbb{P}[f_n]\le 2^{-n}$. We replace the functions f_n by $g_n=\sup_{k\ge n}f_k$ in order to obtain a decreasing sequence g_n such that, of course, $\mathbb{E}_\mu[g_n]\ge\delta/2$ and such that $\mathbb{E}_\mathbb{P}[g_n]\le 2^{-n+1}$. If we now define $A_n=\{g_n\ge\delta/4\}$, then clearly A_n is a decreasing sequence, with a.s. empty intersection and such that for each $\mu\in K$ we have $\mu(A_n)\ge\delta/4$. □

Example 13 This example shows that the equivalent properties of the preceding theorem do not imply the Fatou property. Take $(\Omega,\mathcal{F},\mathbb{P})$ big enough to support purely finitely additive probabilities, i.e. $L^\infty(\Omega,\mathcal{F},\mathbb{P})$ is supposed to be infinite dimensional. Take $\mu\in\mathbf{ba}$, purely finitely additive, and let $\mathcal{S}^{\mathbf{ba}}$ be the segment (the convex hull) joining the two points μ and $\mathbb{P}$. Obviously the equivalent properties of the preceding theorem are satisfied. Indeed there is a $\sigma-$additive measure in $\mathcal{S}^{\mathbf{ba}}$. But clearly the coherent measure cannot satisfy the Fatou property since $\mathcal{S}\cap L^1=\{\mathbb{P}\}$ is not dense in $\mathcal{S}^{\mathbf{ba}}$. To find "explicitly" a sequence of functions that contradicts the Fatou property, we proceed as follows. The measure μ is purely finitely additive and therefore, by the Yosida-Hewitt decomposition theorem (see [135]), there is a countable partition of Ω into sets $(B_n)_{n\geq 1}$ such that for each n, we have $\mu(B_n) = 0$. Of course we may suppose that $\mathbb{P}[B_k] > 0$ for all k. Now we define $A_n = B_1 \cup (\cup_{j\geq n} B_j)$. Clearly $A_n \downarrow B_1$. For $\xi_n = \mathbf{1}_{A_n}$ we then get: $\xi_n \to \mathbf{1}_{B_1}$, $\mu(A_n) = 1$, $u(\xi_n) = \mathbb{P}[A_n] \to \mathbb{P}[B_1]$, $u(\mathbf{1}_{B_1}) = \mu(B_1) = 0$. This violates the Fatou property.

Example 14 In the previous example, $\mathcal{S}^{\mathbf{ba}}$ contained a $\sigma-$additive probability measure. The present example is so that the equivalent properties of the preceding theorem still hold, but there is no $\sigma-$additive probability measure in $\mathcal{S}^{\mathbf{ba}}$. In the language of the theorem, this simply means that the supremum is not a maximum. The set Ω is simply the set of natural numbers. The σ-algebra is the set of all subsets of Ω and $\mathbb{P}$ is a probability measure on Ω charging all the points in Ω. The space L^∞ is then ℓ^∞ and L^1 can be identified with ℓ^1. The set $\mathbf{F}$ denotes the convex weak*-closed set of all purely finitely additive probabilities μ. That the set $\mathbf{F}$ is weak*$-$closed is clear since such measures can be characterised as finitely additive probability measures such that $\mu(\{n\}) = 0$ for all $n \in \Omega$. With δ_n we denote the probability measure (in L^1) that puts all its mass at the point n, the so–called Dirac measure concentrated in n. The set $\mathcal{S}^{\mathbf{ba}}$ is the weak* closure of the set

$$\left\{\left(\sum_{n\geq 1}\frac{\lambda_n}{(n+1)^2}\right)\nu + \sum_{n\geq 1}\lambda_n\left(1-\frac{1}{(n+1)^2}\right)\delta_n \mid \lambda_n \geq 0, \sum_{n\geq 1}\lambda_n = 1, \nu\in\mathbf{F}\right\}.$$

The set is clearly convex and it defines a coherent utility function, u. Since obviously $\sup\left\{\|\mu_a\| \mid \mu \in \mathcal{S}^{\mathbf{ba}}\right\} = 1$, the properties of the theorem hold. The difficulty consists in showing that there is no σ–additive measure in the set $\mathcal{S}^{\mathbf{ba}}$. Take an arbitrary element $\mu \in \mathcal{S}^{\mathbf{ba}}$. By the definition of the set $\mathcal{S}^{\mathbf{ba}}$ there is a generalised sequence, also called a net, μ^α tending to μ and such that

$$\mu^\alpha = \left(\sum_{n\geq 1}\frac{\lambda_n^\alpha}{(n+1)^2}\right)\nu^\alpha + \sum_{n\geq 1}\lambda_n^\alpha\left(1-\frac{1}{(n+1)^2}\right)\delta_n,$$

where each $\nu^\alpha \in \mathbf{F}$, where $\sum_n \lambda_n^\alpha = 1$ and each $\lambda_n^\alpha \geq 0$. We will select subnets, still denoted by the same symbol α, so that

1. the sequence $\sum_n \lambda_n^\alpha\,\delta_n$ tends to $\sum_n \kappa_n\,\delta_n$ for the topology $\sigma(\ell^1, c_0)$. This is possible since ℓ^1 is the dual of c_0. This procedure is the same as selecting a subnet such that for each n, the net λ_n^α tends to κ_n. Of course $\kappa_n \geq 0$ and $\sum_n \kappa_n \leq 1$.

2. from this it follows, by taking subnets, that there is a purely finitely additive, nonnegative measure ν' such that

$$\sum_n \lambda_n^\alpha \left(1 - \frac{1}{(n+1)^2}\right) \delta_n$$

tends to

$$\sum_n \kappa_n \left(1 - \frac{1}{(n+1)^2}\right) \delta_n + \nu'$$

for the topology $\sigma(\mathbf{ba}, L^\infty)$.

3. By taking a subnet we may also suppose that the generalised sequence ν^α converges for $\sigma(\mathbf{ba}, L^\infty)$, to a, necessarily purely finitely additive, element $\nu \in \mathbf{F}$.

4. Of course $\sum_n \frac{|\lambda_n^\alpha - \kappa_n|}{(n+1)^2}$ tends to 0.

As a result we get

$$\mu = \sum_n \frac{\kappa_n}{(n+1)^2} \nu + \nu' + \sum_n \kappa_n \left(1 - \frac{1}{(n+1)^2}\right) \delta_n.$$

If this measure were σ–additive, then necessarily for the non absolutely continuous part, we would have that $\nu' + \sum_n \frac{\kappa_n}{(n+1)^2} \nu = 0$. But, since these measures are nonnegative, this requires that all $\kappa_n = 0$ and that $\nu' = 0$. This would then mean that $\mu = \nu' = 0$, a contradiction to $\mu(\Omega) = 1$.

Example 15 This example shows that in order to represent coherent utility functions via expected values, some control measure is needed. We will construct a utility function on a space of bounded measurable functions that satisfies a continuity property similar to the Fatou property. At the same time we will see that this utility function cannot be described by a set of σ–additive probability measures. We start with the measurable space $([0,1], \mathcal{F})$, where $\mathcal{F}$ is the Borel σ-algebra. A set N is of first category if it is contained in the countable union of closed sets with empty interior (relative to $[0,1]$). The class of Borel sets of first category, denoted by $\mathcal{N}$, forms a σ-ideal in $\mathcal{F}$. For a bounded, Borel measurable function ξ defined on $[0,1]$, we define $u(\xi)$ as the “essential” infimum of ξ. More precisely we define (the reader should prove that there is indeed a maximum in the next formula):

$$u(\xi) = \max\{m \mid \{\xi < m\} \in \mathcal{N}\}.$$

It is clear that $u(\xi)$ defines a coherent utility function. It even satisfies the Fatou property in the sense that $u(\xi) \geq \limsup u(\xi_n)$, where $(\xi_n)_{n\geq 1}$ is a uniformly bounded sequence of Borel functions tending pointwise to ξ. If u were of the form

$$u(\xi) = \inf_{\mathbb{Q}\in\mathcal{S}_\sigma} \mathbb{E}_\mathbb{Q}[\xi],$$

where $\mathcal{S}_\sigma$ is a family of (σ-additive) probability measures, then elements $\mathbb{Q}$ of the family $\mathcal{S}_\sigma$ should satisfy:

$$\mathbb{Q}(N) = 0 \text{ for each set } N \text{ of first category.}$$

Indeed for each set of first category N we have $u(\mathbf{1}_{N^c}) = 1$, hence we have $\mathbb{Q}[N^c] = 1$ for each set N of first category. But if $\mathbb{Q}$ is a Borel measure that is zero on the compact sets of first category, then it is identically zero. Indeed let A be a Borel set and suppose that $\mathbb{Q}[A] \geq \varepsilon > 0$. Let $\{q_n \mid n \geq 1\}$ be an enumeration of the rationals. Because $\{q_n\}$ is of first category, we have $\mathbb{Q}[\{q_n\}] = 0$. So we can choose $\varepsilon_n > 0$ so that $\mathbb{Q}[]q_n-\varepsilon_n, q_n+\varepsilon_n[] \leq \varepsilon 2^{-n-2}$. Let $O = \cup_n]q_n - \varepsilon_n, q_n + \varepsilon_n[$. By the choice of ε_n we have $\mathbb{Q}[O] \leq \varepsilon/4$. Because Borel measures are Radon measures, i.e. regular, there is a compact set $K \subset A$ such that $\mathbb{Q}[K] \geq \varepsilon/2$. Now we put $N = K \setminus O$. This is a set of first category (it is closed and has empty interior) and hence $\mathbb{Q}[N] = 0$. But $\mathbb{Q}[N] \geq \mathbb{Q}[K] - \mathbb{Q}[O] \geq \varepsilon/4$, a contradiction. However we can prove, in the same way as for the representation property of coherent utility functions, that

$$u(\xi) = \inf_{\mu\in\mathcal{S}} \mathbb{E}_\mu[\xi],$$

where $\mathcal{S}$ is a convex set of finitely additive probabilities on $\mathcal{F}$. The set $\mathcal{S}$ does not contain any σ–additive probability measure, although u satisfies some kind of Fatou property. Even worse, every element μ in $\mathcal{S}$ is purely finitely additive and satisfies $\mu(N) = 0$ for $N \in \mathcal{N}$.

Remark 15 The example can be presented using the Baire σ–algebra defined as

$$\mathcal{B} = \{A \subset [0,1] \mid A = O\triangle N, N \text{ is of first category and } O \text{ is an open set}\}.$$

We did not do this because the requirement to be σ–additive on the Baire sets is stronger than the requirement to be σ–additive on the Borel sets. Indeed it is well known (and easily proved using monotone class arguments) that $\mathcal{F} \subset \mathcal{B}$.

4.7 Weak compactness of $\mathcal{S}$

We start with a version of the Dunford-Pettis theorem that includes some excursion to Orlicz-space theory. The theorem is a basic theorem in $L^1 - L^\infty$ duality theory.

Theorem 15 *For closed convex sets $\mathcal{S} \subset L^1$ of probabilities, the following are equivalent:*

1. *$\mathcal{S}$ is weakly compact;*
2. *$\mathcal{S}$ is weakly sequentially compact;*
3. *the set $\{\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathbb{Q} \in \mathcal{S}\}$ is uniformly integrable;*
4. *(de la Vallée-Poussin's criterion for uniform integrability) there exists a function $\Phi : \mathbb{R}^+ \to \mathbb{R}$, increasing, convex, $\Phi(0) = 0$ such that $\lim_{x\to\infty} \frac{\Phi(x)}{x} = +\infty$ and $\sup_{\mathbb{Q}\in\mathcal{S}} \mathbb{E}\left[\Phi\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] < \infty$.*

The following families of weakly compact sets will play a role.
(a) $\Phi(x) = x^p$, $p > 1$; together with point *4* this implies that $\mathcal{S}_{p,k}$ is a weakly compact family; we also have that the set $\mathcal{S}_k$ is weakly compact. This could correspond to the function $\Phi(x) = 1$ for $x < 1$ and $\Phi(x) = +\infty$ for $x \geq 1$. A little bit of liberal thinking is required.
(b) $\Phi(x) = (x + 1)\log(x + 1) - x$; this is another example that can be used in connection with Orlicz space theory, [93]. See Delbaen, [39] on how to use this function in risk measure theory. See also [31, 32] and [18]. for more recent developments. The idea in these papers is that when u is given by a weakly compact set of measures $\mathcal{S}$, then u can be extended to a utility function defined on an Orlicz space L^Ψ. The function Ψ is given by the Legendre transform of Φ where Φ is obtained out of $\mathcal{S}$ via the criterion of de la Vallée-Poussin.

According to the above, for coherent utility functions u, the following are equivalent:

1. the set $\{\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathbb{Q} \in \mathcal{S}\}$ is uniformly integrable;
2. $\mathcal{S}$ is weakly compact;
3. if $(A_n)_{n\geq 1}$ is a family of measurable sets such that $A_n \downarrow \emptyset$, then $u(-\mathbf{1}_{A_n}) \uparrow 0$ or equivalently $\sup_{\mathbb{Q}\in\mathcal{S}} \mathbb{Q}[A_n] \to 0$; this can also be restated as: if $(A_n)_{n\geq 1}$ is a family of measurable sets such that $A_n \uparrow \Omega$, then $u(\mathbf{1}_{A_n}) \uparrow 1$;

4. if $-1 \le \xi_n \le 0$ and $\xi_n \uparrow 0$, then $u(\xi_n) \uparrow 0$.

Remember that point *4* is stronger than the Fatou property! The reader can check that the Example 8 used a non-weakly compact set $\mathcal{S}$. We can give another characterisation of weakly compact sets:

Theorem 16 *$\mathcal{S}$ is weakly compact if and only if every $\xi \in L^\infty$ attains its minimum on $\mathcal{S}$, i.e. there is $\mathbb{Q} \in \mathcal{S}$ such that $u(\xi)$ is exactly $\mathbb{Q}[\xi]$.*

Proof. . This is James's theorem translated to the case of coherent utility functions. □

Theorem 17 *If $\mathcal{S}$ is weakly compact then:*

$$\|\xi_n\|_\infty \le 1\,, \quad \xi_n \xrightarrow{\mathbb{P}} \xi \quad \text{implies} \quad u(\xi) = \lim_{n\to\infty} u(\xi_n).$$

Proof. . A direct application of the property that $\mathcal{S}$ is uniformly integrable. □

Example 16 This example can be seen as an application to a Credit Risk situation. Suppose that $(\xi_n)_n$ are i.i.d and that $\|\xi_n\|_\infty \le 1$. The random variable ξ_i stands for the loss corresponding to the i-th person (*the group is supposed to be independent and identically distributed*). Let $S_n = \xi_1 + \ldots + \xi_n$. The problem is to calculate the total capital needed to face the risk. We need $\rho(S_n) = -u(S_n)$ and the capital or premium we will charge to each person will be $\frac{1}{n}\rho(S_n) = \rho(\frac{S_n}{n})$. Suppose now that $\mathcal{S}$ is weakly compact, for instance the utility function is calculated as in example 9. By the law of large numbers,

$$\frac{S_n}{n} \xrightarrow{a.s.} \mathbb{E}[\xi_1]$$

so that

$$\rho\left(\frac{S_n}{n}\right) \longrightarrow \rho(\mathbb{E}[\xi_1]) \equiv -\mathbb{E}[\xi_1]$$

If we do not have independence, but the correlation coefficients tend to zero when n goes to infinity, the previous result still holds. Indeed if

$$\lim_{k\to\infty} \sup_n |\mathbb{E}[\xi_n \xi_{k+n}] - \mathbb{E}[\xi_n]\mathbb{E}[\xi_{n+k}]| \to 0,$$

then by Bernstein's theorem, $\frac{S_n}{n}$ tends to $\mathbb{E}[\xi_1]$ in probability if $n \to \infty$. We leave the interpretation of this refinement to the reader.

We warn the reader that although the required capital pro capita tends to the expected loss, the total capital can be substantially different from $n\mathbb{E}[-\xi_1]$. This has to do with the speed of convergence. It can be shown – using convexity and inverse martingale arguments (left to the reader) – that $\rho(S_n/n)$ decreases to $\mathbb{E}[-\xi_1]$. The difference $\rho(S_n) + n\mathbb{E}[\xi_1]$ can tend to $+\infty$.

Exercise 11 Fill in the details needed in the previous paragraph. Give an example where effectively $\rho(S_n) + n\mathbb{E}[\xi_1]$ tends to $+\infty$.

Example 17 This example is a modification of the previous example. It shows that if we replace independence by conditional independence, the required capital pro capita changes, even when a large number of agents are participating. So we suppose that the sequence ξ_n is conditionally independent with respect to a sigma-algebra $\mathcal{I}$. And we also suppose that conditional on $\mathcal{I}$ all the random variables have the same law. The sigma-algebra $\mathcal{I}$, could represent the future yet unknown, macro economic situation. Conditionally on the macro economic situation the credit takers are supposed to be independent and identically distributed. However there is a dependence because of the overall economic situation. In this case the law of large numbers reads: S_n/n tends to $\mathbb{E}[\xi_1 \mid \mathcal{I}]$ (almost surely but we only need convergence in probability). In case $\mathcal{S}$ is weakly compact we get $u(S_n/n) \to u(\mathbb{E}[\xi_1 \mid \mathcal{I}])$. Let us see what this means in a credit risk situation. Let us suppose that ξ_n takes values $0, -1$ and suppose conditional independence with respect to $\mathcal{I}$. Then $p = -\mathbb{E}[\xi_1 \mid \mathcal{I}]$ is just the probability of going bankrupt given the information coming from $\mathcal{I}$. This is a random variable. In case we take the utility function TailVar with level α, we see that $u(S_n/n)$ tends to $u(-p) = \mathbb{E}[-p \mid p \geq 1 - q_\alpha(p)]$ (at least when we suppose p to have a continuous distribution, otherwise use the extension as in example 9). This means that the amount of capital needed pro capita is entirely given by the probability law of the macro economic influence.

The example supposed that we had, conditionally on $\mathcal{I}$, identically distributed random variables. Of course we could have refined the example and use some kind of stratification. For different groups the sensitivity to the macro-economic factor could be different (something given by a credit rating, whatever the word means and wherever it comes from). This would then lead to different capital requirements for the individual groups. The total required capital is of course not necessarily the sum of the different individual required capital per group. We will see that for TailVaR, in such models the so-called commonotonicity implies additivity for the required capital. Maybe something to think about when dealing with mortgages, CDO, ... In this context and as said before, the author does not want to use the word *subprime* but the temptation to do so was big.

Exercise 12 We invite the reader to calculate VaR, TailVaR when p has a beta distribution i.e. it has a density (on $[0,1]$) of the form $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$ with $\alpha,\beta>0$. Depending on the kind of default probability one would like to model, one could take any fixed $\varepsilon>0$ and take $\alpha,\beta>0$ so that the total default probability remains equal to ε. One can then calculate the values of VaR and TailVaR as a function of $\beta>0$ (or of α as you wish). One should compare these values with the values for one agent.

4.8 Concave utility functions, duality results

In this section we use convex duality theory in order to get extra information on concave monetary utility functions. In the next section we will present a reduction technique that will allow us to transform the results from the special case of coherent measures. However we find it useful to present also the classical approach. The basic facts can be found in [118]. The definition of Fenchel-Legendre transform has been adapted a little bit, in order to get the sign right. But this is only a cosmetic change. The theory of concave utility functions was developed by Föllmer and Schied, [68].

Definition 12 *If $u\colon L^\infty\to\mathbb{R}$ is a monetary concave utility function, then its Fenchel-Legendre transform (or penalty function) is defined as*

$$\begin{aligned} c\colon \mathbf{ba} &\to \mathbb{R}_+\cup\{\infty\}\\ c(\mu) &= \sup\{-\mu(\xi)+u(\xi)\mid \xi\in L^\infty\}\end{aligned}$$

If we only take the supremum over the constant functions we already get

$$c(\mu)\ge\sup_{a\in\mathbb{R}} a(-\mu(1)+1).$$

This shows that c takes the value $+\infty$ for measures that have total mass different from 1. If $\mu(A)<0$ then we take the functions $n\mathbf{1}_A$ and we get $c(\mu)\ge -n\,\mu(A)+u(n\mathbf{1}_A)\ge -n\,\mu(A)$, yielding that for non positive measures we also get a value $+\infty$. As a result we only need to define the Fenchel-Legendre transform for finitely additive probability measures.

Exercise 13 Show that the definition of the Fenchel-Legendre transform coincides with the one given in chapter 2. See what happens with indicators of convex sets and relate to the exercises given in chapter 2.

Proposition 10 *The function c satisfies the following properties*

1. $c\colon \mathbf{P}^{\mathbf{ba}} \to \mathbb{R}_+ \cup \{+\infty\}$,
2. *c is lower semi-continuous for the weak* topology $\sigma(\mathbf{ba}, L^\infty)$.*
3. *c is convex*
4. $c(\mu) = \sup\{-\mu(\xi) \mid \xi \in \mathcal{A}\} = \sup\{-\mu(\xi) \mid u(\xi) = 0\}$
5. $\min_{\mu\in\mathbf{P}^{\mathbf{ba}}} c(\mu) = 0$.

Proof. Since each function $\mu \to u(\xi) - \mu(\xi)$ is weak* continuous and affine, we get that the supremum is weak* lower semi continuous and convex. Since $u(\xi) - \mu(\xi) = -\mu(\xi - u(\xi))$ and since $u(\xi - u(\xi)) = 0$ we get also item 4. Since the convex set $\mathcal{A}$ contains the positive cone, it has a non empty interior for the norm topology. Since obviously 0 cannot be an interior point of $\mathcal{A}$ (because $u(a) < 0$ for $a < 0$), we can separate the interior of the set $\mathcal{A}$ and the origin. So we get a nonzero functional $\mu \in \mathbf{ba}$ so that for all $\xi \in \mathcal{A}$: $\mu(\xi) \geq 0$. This measure μ is nonnegative and we can normalize it to get an element $\mu \in \mathbf{P}^{\mathbf{ba}}$. Of course we then have $c(\mu) = 0$. □

Remark 16 In case the utility function u is coherent and given by the weak* closed convex set $\mathcal{S}^{\mathbf{ba}}$, the penalty function c only takes the two values $0, +\infty$. Indeed $c(\mu) = 0$ for $\mu \in \mathcal{S}^{\mathbf{ba}}$ and $c(\mu) = +\infty$ for $\mu \notin \mathcal{S}^{\mathbf{ba}}$. This function is called the indicator function of the set $\mathcal{S}^{\mathbf{ba}}$. Conversely when the function c is an indicator function of a set $\mathcal{S}^{\mathbf{ba}}$, then this set is necessarily convex and weak* compact. The utility u is coherent and given by $u(\xi) = \inf_{\mathcal{S}^{\mathbf{ba}}} \mu(\xi)$.

The importance of the function c lies in the fact that by duality we can find the function u back, see [118], [112] for more details on duality. We get:

Theorem 18
$$u(\xi) = \min\{\mu(\xi) + c(\mu) \mid \mu \in \mathbf{P}^{\mathbf{ba}}\}.$$

Proof. That $u(\xi) = \inf\{\mu(\xi) + c(\mu) \mid \mu \in \mathbf{P}^{\mathbf{ba}}\}$ is proved in convex analysis. For completeness we give a proof. Because of the definition of c we have $u(\xi) \leq \mu(\xi) + c(\mu)$ for all $\mu \in \mathbf{P}^{\mathbf{ba}}$. To prove the converse inequality we just have to show that $u(\xi) < 0$ implies the existence of μ with

$\mu(\xi) + c(\mu) < 0$ Suppose that $\xi \notin \mathcal{A}$. By the separation theorem there is a linear functional $\mu \in \mathbf{ba}$ (in case the Fatou property is satisfied we can even take $\mu \in L^1$) such that $\mu(\xi) < \inf\{\mu(\eta) \mid \eta \in \mathcal{A}\}$. Because $\mathcal{A}$ contains L^∞_+, we can already conclude that $\mu \geq 0$. And because μ is not identically zero we may normalise μ so that we can take $\mu \in \mathbf{P^{ba}}$. The definition of $c(\mu)$ can be written as $c(\mu) = -\inf\{\mu(\eta) \mid \eta \in \mathcal{A}\}$ and so we get $\mu(\xi) < -c(\mu)$ or $\mu(\xi) + c(\mu) < 0$. We now prove that the inf is a minimum. Take μ_n a sequence such that $\mu_n(\xi) + c(\mu_n)$ tends to the infimum. The infimum is smaller than $\|\xi\|_\infty$ since $\inf_\mu c(\mu) = 0$. This shows that $c(\mu_n)$ is a bounded sequence. By taking, if necessary, a subsequence we may suppose that both $\mu_n(\xi), c(\mu_n)$ converge. Take a cluster point, say μ, of the sequence μ_n in the compact set $\mathbf{P^{ba}}$. This element μ then satisfies $c(\mu) \leq \lim c(\mu_n)$ and hence $\mu(\xi) + c(\mu) = \inf\{\nu(\xi) + c(\nu) \mid \nu \in \mathbf{P^{ba}}\}$. □

In case the utility function has the Fatou property we can use the duality (L^1, L^∞) and we get that the restriction of c to $\mathbf{P}$ is sufficient. However there is no guarantee that the infimum is a minimum. We get the following theorem, which we give without proof.

Theorem 19 *In case u satisfies the Fatou property we have*

1. *The set $\{(\mathbb{Q}, \beta) \mid \beta \geq c(\mathbb{Q})\} \subset \mathbf{P} \times \mathbb{R}$ is weak* dense in the set $\{(\mu, \beta) \mid \beta \geq c(\mu)\} \subset \mathbf{P^{ba}} \times \mathbb{R}$.*
2. *$\inf_{\mathbb{Q}\in\mathbf{P}} c(\mathbb{Q}) = 0$*
3. *$u(\xi) = \inf\{\mathbb{E}_\mathbb{Q}[\xi] + c(\mathbb{Q}) \mid \mathbb{Q} \in \mathbf{P}\}$.*

Remark 17 The first item shows that for $\mu \in \mathbf{P^{ba}}$ there is a generalized sequence $\mathbb{Q}_\alpha$ with the property that for the topology $\sigma(\mathbf{ba}, L^\infty)$, $\mathbb{Q}_\alpha \to \mu$ and $\lim_\alpha c(\mathbb{Q}_\alpha) = c(\mu)$. There is also a converse to the representation theorem.

Theorem 20 *If $c\colon \mathbf{P} \to \mathbb{R}_+ \cup \{\infty\}$ is a function satisfying the assumptions*

1. *c is lower semi-continuous on L^1.*
2. *c is convex*
3. *$\inf_{\mathbb{Q}\in\mathbf{P}} c(\mathbb{Q}) = 0$,*

then $u(\xi) = \inf\{\mathbb{E}_\mathbb{Q}[\xi] + c(\mathbb{Q}) \mid \mathbb{Q} \in \mathbf{P}\}$ defines a Fatou, monetary concave utility function with penalty function c.

Remark 18 If u is a utility function with penalty function c defined on $\mathbf{P^{ba}}$, then the restriction of c to $\mathbf{P}$ defines – according to the previous theorem – a Fatou utility function u_0. We have $u = u_0$ if and only if u is Fatou.

Example 18 Let $\Phi\colon \mathbb{R}_+ \to \mathbb{R}$, be convex such that $\Phi(1) = 0$. If we put $c(\mathbb{Q}) = \mathbb{E}\left[\Phi\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$ then c defines a convex function, lsc and $c(\mathbb{P}) = 0$. If Φ is strictly convex then this is the only possibility to have $c(\mathbb{P}) = 0$. If $\Phi(x) = x^2 - 1$ or $\Phi(x) = (x-1)^2$, we get the variance of $\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)$. If we take $\Phi(x) = x\log(x)$ we get the entropy. But we could also take $\Phi(x) = \exp(-(x-1)) - 1$ or $\Phi(x) = \sqrt{1+(x-1)^2} - 1$. For strict convex functions Φ, we can show that $c(\mathbb{Q}^n) \to 0$ implies that $\frac{d\mathbb{Q}^n}{d\mathbb{P}} \to 1$ in probability, which implies that $\|\mathbb{Q}^n - \mathbb{P}\| \to 0$. The function c can be extended to the set $\mathbf{P^{ba}}$. There are two ways. The first is to use Theorem 19, item 1 above, the second one is to define the utility function u and then calculate $c(\mu)$. Of course both methods give the same result. In case $\Phi(x)$ satisfies $\lim_{x\to+\infty} \frac{\Phi(x)}{x} = +\infty$ we can easily see that $c(\mu) < \infty$ necessarily implies that $\mu \in \mathbf{P}$. The function $\Phi(x) = \sqrt{1+(x-1)^2} - 1$ has linear growth and for this we can show that $\sup_{\mathbb{Q}} c(\mathbb{Q}) = \sup_\mu c(\mu) < \infty$ (this is an easy exercise). If in these examples we want to calculate $u(\xi)$, we have to solve a convex variational problem. This is well understood and the solution is related to the Legendre transform of Φ. Those who are familiar with the mathematics of optimisation with respect to von Neumann-Morgenstern functions should have no difficulty in finding it. This is beyond the scope of this book.

Exercise 14 For $\Phi(x) = \sqrt{1+(x-1)^2} - 1$ and $c(\mathbb{Q}) = \mathbb{E}\left[\Phi\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$, first show that $c(\mathbb{Q}) < \sqrt{2}$, hence $c(\mu) \le \sqrt{2}$ for all $\mu \in \mathbf{P^{ba}}$; then show that if μ is purely finitely additive $c(\mu) = \sqrt{2}$. Use the convexity of c to show that this characterises the purely finitely additive probability measures. Hint: observe that $\Phi(x) \le x + \sqrt{2} - 1$, with strict inequality for $x \ne 0$.

Example 19 Suppose that the probability space is atomless and let us fix a countable partition of Ω into a sequence of measurable sets A_n with $\mathbb{P}[A_n] > 0$. For $\mu \in \mathbf{P^{ba}}$ we define

$$c(\mu) = \sum_n \mu[A_n]^2.$$

Remark 19 In what follows one can replace the square by any convex function f with the properties $f(0) = 0$, $f(x) > 0$ for $x > 0$ and $f(1) < \infty$.

Proposition 11 *The function c of the previous example is a convex function, $\min_{\mu\in\mathbf{P}^{\mathbf{ba}}} c(\mu)=0$ and it is lower semi-continuous for the weak* topology on $\mathbf{P}^{\mathbf{ba}}$. For each $\mathbb{Q}\in\mathbf{P}$ we have $c(\mathbb{Q})>0$. If $c(\mu)=0$ then μ is purely finitely additive. The utility function u defined by c is Fatou.*

Proof. The first statements are obvious since the mapping $\mu\to\mu[A_n]^2$ is convex and weak* continuous. c is therefore the increasing limit of a sequence of continuous convex functions and hence is lower semi-continuous and convex. The existence of elements in $\mathbf{P}^{\mathbf{ba}}$ such that for all n, $\mu(A_n)=0$, is well known and can be proved using the Hahn-Banach theorem. If $c(\mu)=0$ then for all n: $\mu(A_n)=0$ and this means that μ is purely finitely additive. Of course we have for $\mu\in\mathbf{P}^{\mathbf{ba}}$: $\sum_n\mu(A_n)^2\le\sum_n\mu(A_n)\le 1$. If $\mathbb{Q}\in\mathbf{P}$ we have that at least one of the sets A_n must satisfy $\mathbb{Q}[A_n]>0$, hence $c(\mathbb{Q})>0$. The Fatou property is less trivial. As seen before we must show that for $\mu\in\mathbf{P}^{\mathbf{ba}}$ we can find a generalised sequence or net $\mathbb{Q}_\alpha$ in $\mathbf{P}$, tending to μ and so that $c(\mathbb{Q}_\alpha)$ tends to $c(\mu)$. For this it is sufficient to show the following. Given μ, given $\epsilon>0$ and given a finite partition of Ω in non-zero sets $B_1,\dots,B_N$ we must find $\mathbb{Q}\in\mathbf{P}$ so that $c(\mathbb{Q})\le c(\mu)+\epsilon$ and $\mathbb{Q}(B_j)=\mu(B_j)$ for $j=1,\dots,N$. For a set B_j there are two possibilities: either there is s with $B_j\subset\cup_{n=1}^s A_n$ or there are infinitely many indices n with $\mathbb{P}[B_j\cap A_n]>0$. Since all the sets A_n have a non-zero measure and since the family $(B_j)_j$ forms a partition of Ω the last alternative must occur for at least one index j. So let us renumber the sets B_j and let us select s so that

1. for $j\le N'\le N$ there are infinitely many indices with $\mathbb{P}[A_n\cap B_j]>0$,
2. for $N'<j\le N$ (if any) we have that $B_j\subset\cup_{n=1}^s A_n$.

Fix now an integer $L\ge 1$ so that $1/L\le\epsilon$. We will define the measure $\mathbb{Q}$ by its Radon-Nikodym density. For $j\le N'$ we find indices as follows, we take L indices $s<n_1^1<n_2^1\dots<n_L^1$ so that $\mathbb{P}[A_{n_k^1}\cap B_1]>0$. We then take indices $n_L^1<n_1^2<n_2^2<\dots<n_L^2$ with $\mathbb{P}[A_{n_k^2}\cap B_2]>0$ and so on. We can now define the density of $\mathbb{Q}$ as

$$\frac{d\mathbb{Q}}{d\mathbb{P}}=\sum_{j=1}^N\sum_{k=1}^s\frac{\mu(B_j\cap A_k)}{\mathbb{P}[B_j\cap A_k]}\mathbf{1}_{B_j\cap A_k}+\sum_{j=1}^N\sum_{p=1}^L\frac{\mu(B_j\cap(\cup_{n>s}A_n))}{L\;\mathbb{P}[B_j\cap A_{n_p^j}]}\mathbf{1}_{B_j\cap A_{n_p^j}}.$$

The reader can convince himself that there is no reason to drop the terms with denominator zero. For all $j\le N$ we have that $\mathbb{Q}[B_j]=\mu(B_j)$. Furthermore we have that for $n\le s:\mathbb{Q}[A_n]=\mu(A_n)$. For indices $n>s$ there is at most one of the N sets $B_j \cap A_n$ that is chosen. So we get for $n > s$:

$$\mathbb{Q}[A_{n_p^j}] = \frac{1}{L}\mu(B_i \cap (\cup_{n>s} A_n)) \text{ and for other indices } n \text{ we get } 0.$$

Finally we find

$$\begin{aligned}
c(\mathbb{Q}) &= \sum_n \mathbb{Q}[A_n]^2 = \sum_{n\le s} \mathbb{Q}[A_n]^2 + \sum_{n>s} \mathbb{Q}[A_n]^2 \\
&= \sum_{n\le s} \mu(A_n)^2 + \sum_{n>s} \mathbb{Q}[A_n]^2 \\
&\le c(\mu) + \sum_{j=1}^N \sum_{p=1}^L \frac{1}{L^2}\mu(B_j \cap (\cup_{n>s} A_n))^2 \\
&\le c(\mu) + \frac{1}{L}\sum_{j=1}^N \mu(B_j)^2 \\
&\le c(\mu) + \epsilon.
\end{aligned}$$

□

Example 20 We construct u, a Fatou concave utility function satisfying $c(\mathbb{Q}) > 0$ for all $\mathbb{Q} \in \mathbf{P}$, but $c(\mu) = 0$ does not imply that μ is purely finitely additive. We again take a countable partition into nonnegligible sets, $\{A_n; n \ge 1\}$. We put

$$c(\mathbb{Q}) = \int_{A_1} \left(\frac{d\mathbb{Q}}{d\mathbb{P}} - 1\right)^2 d\mathbb{P} + \sum_{n\ge 2} \mathbb{Q}[A_n]^2.$$

If $\mathbb{Q}$ were a probability measure such that $c(\mathbb{Q}) = 0$, then on A_1,we would have $\frac{d\mathbb{Q}}{d\mathbb{P}} = 1$ whereas on A_1^c we would have $\frac{d\mathbb{Q}}{d\mathbb{P}} = 0$ since $\mathbb{Q}[A_n] = 0$ for $n \ge 2$. This is impossible since it gives for $\mathbb{Q}$ a total mass $\mathbb{P}[A_1] < 1$. The function c is clearly convex, lsc, and $\inf_{\mathbb{Q}} c(\mathbb{Q}) = 0$. The latter can be proved exactly in the same way as in the example 19. In the same way we can show that $c(\mu) < \infty$ implies that μ is sigma addiitive on A_1 and then $c(\mu) = \int_{A_1} \left(\frac{d\mu}{d\mathbb{P}} - 1\right)^2 d\mathbb{P} + \sum_{n\ge 2} \mu[A_n]^2$. We see that $c(\mu) = 0$ if and only if $\mu = \mathbb{P}$ on A_1 and $\mu(A_n) = 0$ for $n \ge 2$. This means that μ is purely finitely additive on A_1^c.

Example 21 Let us take $0 \le f \in L^\infty$, $\text{ess.inf}(f) = 0$ and define $c(\mathbb{Q}) = \mathbb{E}_\mathbb{Q}[f]$. This function satisfies all the requirements and defines a utility function $u(\xi) = \inf_\mathbb{Q} \mathbb{E}_\mathbb{Q}[\xi + f] = \text{ess.inf}(\xi + f)$. The set $\mathcal{A}$ is a cone but its top is not the element 0 but is $-f$. If $f > 0$ a.s. , then we get another example of a utility function such that for all $\mathbb{Q} \in \mathbf{P}$, $c(\mathbb{Q}) > 0$. The elements $\mu \in \mathbf{P^{ba}}$ with $c(\mu) = 0$ are supported on the tail of the sequence $A_n = \{f \le n^{-1}\}$ — meaning that for all n: $\mu(A_n) = 1$ — hence these are purely finitely additive since $A_n \downarrow \emptyset$. If $f = \mathbf{1}_A$, $0 < \mathbb{P}[A] < 1$, $\xi \in \mathcal{A}$ if and only if $\xi \ge -\mathbf{1}_A$. The function $u(\alpha \mathbf{1}_A)$ is concave on $\mathbb{R}$, it is zero for $\alpha \ge -1$ and is $\alpha + 1$ for $\alpha \le -1$.

4.9 Extension of a Fatou utility function

If $u\colon L^\infty \to \mathbb{R}$ is a Fatou utility function, we can extend it to the cone of measurable functions that are bounded above. The procedure is the same as in measure theory. If η is a random variable that is bounded above we define

$$u(\eta) = \inf\{u(\xi) \mid \xi \in L^\infty, \xi \ge \eta\}.$$

The set of random variables that are bounded above does not form a vector space, it is only a cone. Algebraically we can describe this cone as $L^\infty - L^0_+$. The following properties are obvious

Proposition 12 *u satisifes*

1. $u\colon L^\infty - L^0_+ \to \mathbb{R} \cup \{-\infty\}$
2. *If* $\eta \ge \eta'$ *then* $u(\eta) \ge u(\eta')$
3. u *is concave and monetary*
4. *Exactly as in measure theory one can prove: if* $\eta_n \downarrow \eta$ *and if* η_1 *is bounded above, then* $u(\eta_n) \downarrow u(\eta)$.
5. *For* $\eta \in L^\infty - L^0_+$ *we have* $u(\eta) = \inf_{\mathbb{Q} \in \mathbf{P}} \left(\mathbb{E}_\mathbb{Q}[\eta] + c(\mathbb{Q})\right)$

4.10 Gâteaux differentiability of utility functions, subgradient.

As already seen in Chapter 2, the Fenchel-Legendre transform can also be used to find the derivative of the concave function u. The general results on duality of convex functions can be translated directly. But in Chapter 2, we promised to give full proofs in the case of utility functions. The reader familiar with convex duality will immediately recognise the consequences of the general theory. Let us recall:

Definition 13 *The function u is called Gâteaux differentiable at a point $\xi \in L^\infty$, if for all $\eta \in L^\infty$, the function $x \to u(\xi + x\eta)$ is differentiable at $x = 0$ and the derivative defines a continuous linear function of η. In other words there exists an element $\mu \in \mathbf{ba}$ such that*

$$\mu(\eta) = \lim_{x \to 0} \frac{u(\xi + x\eta) - u(\xi)}{x}.$$

The subgradient of u at ξ is defined as

$$\partial_\xi(u) = \{\mu \in \mathbf{ba} \mid u(\xi + \eta) \le u(\xi) + \mu(\eta) \text{ for all } \eta \in L^\infty\}.$$

The weak-subgradient of u at ξ is defined as*

$$\partial^*_\xi(u) = \left\{f \in L^1 \mid u(\xi + \eta) \le u(\xi) + \mathbb{E}[f\eta] \text{ for all } \eta \in L^\infty\right\}.$$

The set $\partial_\xi(u)$ is not empty as shown in the following theorem, but in the next section we will give a criterion that shows that $\partial^*_\xi(u)$ can be empty. Of course we have $\partial^*_\xi(u) = \partial_\xi(u) \cap L^1$.

Theorem 21 *Let u be a monetary concave utility function. $\mu \in \partial_\xi(u)$ if and only if $u(\xi) = \mu(\xi) + c(\mu)$. Consequently $\partial_\xi(u) \neq \emptyset$.*

Proof. If $\mu \in \partial_\xi(u)$ then we have for all $\eta \in L^\infty$:

$$u(\xi + \eta) \le u(\xi) + \mu(\eta).$$

If we replace η by $-\xi + \eta$ we get for all η: $u(\eta) \le u(\xi) + \mu(-\xi + \eta) \le c(\mu) + \mu(\eta)$. This can be rewritten as $u(\eta) - \mu(\eta) \le u(\xi) + \mu(-\xi) \le c(\mu)$. Taking sup over all η then gives the equality $c(\mu) = u(\xi) + \mu(-\xi)$, as desired. The converse is easier. If $u(\xi) = c(\mu) + \mu(\xi)$, then for all η we have $u(\xi + \eta) \le c(\mu) + \mu(\xi + \eta) = u(\xi) + \mu(\eta)$. □

Proposition 13 *The graph of $\partial(u)$ is closed in the product topology given by the norm topology on L^∞ and the weak* topology on $\mathbf{P}^{\mathbf{ba}}$.*

Proof. Take a generalised sequence $(\xi_n, \mu_n)_n$ such that $\|\xi_n - \xi\|_\infty \to 0$ and $\mu_n \to \mu$, weak*. For all η we have

$$u(\xi + \eta) = \lim_n u(\xi_n + \eta) \le \lim_n u(\xi_n) + \lim_n \mu_n(\eta) = u(\xi) + \mu(\eta),$$

showing that $\mu \in \partial_\xi(u)$ □

Proposition 14 *The monetary utility function u is Gâteaux differentiable at ξ if and only if $\partial_\xi(u)$ is a singleton $\{\mu\}$. In that case the derivative is μ.*

Proof. This is easy and well known. Suppose first that $\mu_1 \neq \mu_2$ are two different elements in $\partial_\xi(u)$ and suppose that u is differentiable at ξ. Take η such that $\mu_1(\eta) < \mu_2(\eta)$. Now let us calculate the derivative

$$\lim_{x\to 0} \frac{u(\xi + x\eta) - u(\xi)}{x} \le \lim_{x\to 0} \frac{x\mu_1(\eta)}{x} = \mu_1(\eta)$$

whereas for $-\eta$ we find

$$\lim_{x\to 0} \frac{u(\xi - x\eta) - u(\xi)}{x} \le \lim_{x\to 0} \frac{x\mu_2(-\eta)}{x} = -\mu_2(\eta)$$

which we can rewrite as

$$\mu_1(\eta) = \lim_{x\to 0} \frac{u(\xi + x\eta) - u(\xi)}{x} = \lim_{x\to 0} \frac{u(\xi - x\eta) - u(\xi)}{-x} \ge \mu_2(\eta).$$

This is a contradiction to the choice of η. Conversely if $\partial_\xi(u) = \{\mu\}$ we have for given η and for each $x \in \mathbb{R}$ the existence of an element μ_x such that $u(\xi + x\eta) = \mu_x(\xi) + c(\mu_x)$. We will show that $\mu_x \to \mu$ as $x \to 0$. Since $\xi + x\eta$ tends to ξ in norm, the previous proposition shows that every cluster point of $(\mu_x)_x$ must be equal to μ. In the compact space $\mathbf{P^{ba}}$ this shows the convergence of μ_x to μ. The rest is now easy.

$$\begin{aligned}
&\limsup_x \frac{u(\xi + x\eta) - u(\xi)}{x} \\
&\le \limsup_x \frac{\mu(\xi + x\eta) + c(\mu) - (\mu(\xi) + c(\mu))}{x} = \mu(\eta), \text{ and} \\
&\liminf_x \frac{u(\xi + x\eta) - u(\xi)}{x} \\
&\ge \liminf_x \frac{c(\mu_x) + \mu_x(\xi + x\eta) - (c(\mu_x) + \mu_x(\xi))}{x} \\
&= \liminf_x \frac{\mu_x(x\eta)}{x} = \lim_x \mu_x(\eta) = \mu(\eta).
\end{aligned}$$

This shows that the limit exists and that it is equal to $\mu(\eta)$. $\square$

Example 22 It is not sufficient to suppose that $\partial^*_\xi(u)$ is a singleton. We will give an example where u is even a coherent utility function u. Let us consider the probability space $(\mathbb{N}, 2^{\mathbb{N}}, \mathbb{P})$ where $\mathbb{N}$ is the set of natural numbers (including 0) and where $\mathbb{P}\{n\} = \frac{1}{2^{n+1}}$. For $\mathcal{S}$ we take the set of **all** probabilities on $\mathbb{N}$. We now define ξ in the following way: $\xi(0) = -1$ and $\xi(n) = -\left(1 - \frac{1}{n}\right)$ if $n \geq 1$. It is immediately seen that $u(\xi) = -1$ and that $\partial^*_\xi(u) = \{\delta_0\}$ (i.e. the Dirac measure in 0). If we define η by: $\eta(0) = 0$ and $\eta(n) = \xi(n) = -\left(1 - \frac{1}{n}\right)$ if $n \geq 1$, we find that $u(\xi + \varepsilon\eta) = -(1 + \varepsilon)$ for all $\varepsilon > 0$. So, $\frac{u(\xi+\varepsilon\eta)-u(\xi)}{\varepsilon} = -1$ whereas $\delta_0[\eta] = 0$. The set $\partial_\xi(u)$ is much bigger, it consists of all the convex combinations of δ_0 and elements $\mu \in \mathbf{P^{ba}}$, satisfying $\mu(n) = 0$ for all n. The latter are probabilities on the Stone-Čech compactification $\beta\mathbb{N}$ of $\mathbb{N}$, more precisely probability measures that are supported by $\beta\mathbb{N} \setminus \mathbb{N}$.

Remark 20 In case the utility function u is coherent (given by the set $\mathcal{S}^{\mathbf{ba}}$) and is Gâteaux differentiable at ξ, then the derivative μ is the unique element in $\mathcal{S}^{\mathbf{ba}}$ that minimises $\nu(\xi)_{\nu \in \mathcal{S}^{\mathbf{ba}}}$. In this case we say that μ is an exposed point of $\mathcal{S}^{\mathbf{ba}}$.

Theorem 22 *Suppose that $u : L^\infty \to \mathbb{R}$ is a monetary utility function with the Fatou property. Suppose that u is Gâteaux differentiable at $\xi \in L^\infty$. Then $\partial_\xi(u) \in L^1$.*

Proof. Because u satisfies the Fatou property, it is Borel measurable with respect to the weak* topology. Indeed for every $k \in \mathbb{R}$, $\{\eta \mid u(\eta) \leq k\}$ is weak* closed. We will now show that $u'(\xi) = \partial_\xi(u)$ is also Borel measurable. This is easy since for every $\eta \in L^\infty_0$ we have

$$u'(\xi)(\eta) = \lim_{n\to\infty} \frac{u(\xi + 1/n\eta) - u(\xi)}{1/n}.$$

As a limit of a *sequence* of Borel measurable functions, $u'(\xi)$ is Borel measurable. The results on automatic continuity, see [35], show that necessarily $u'(\xi) \in L^1$. $\square$

Remark 21 The previous theorem is essential to show that in an incomplete market, the bid price is nowhere differentiable.

4.11 A class of examples

We start with a concave utility function u satisfying the Fatou property. We suppose that it is given by the penalty function $c\colon \mathbf{P} \to \overline{\mathbb{R}_+}$. The more general case where u is not necessarily Fatou is less interesting and is left as an "exercise". The set of acceptable elements is $\mathcal{A} = \{\xi \mid \mathbb{E}_{\mathbb{Q}}[\xi] + c(\mathbb{Q}) \geq 0 \text{ for all } \mathbb{Q} \in \mathbf{P}\}$. The set $\mathcal{A}$ is convex and weak* closed. Now we take an element η with $u(\eta) = 0$. We define a new set $\mathcal{A}^1 = -\eta + \mathcal{A}$ and use this as the acceptance set of a new utility function u^1. Of course $u^1(\xi) = \sup\{a \mid \xi - a \in \mathcal{A}^1\} = \sup\{a \mid \xi + \eta - a \in \mathcal{A}\} = u(\xi + \eta)$. The utility function is still concave and monetary. Because $\mathcal{A}^1$ is weak* closed, u^1 is Fatou. This can also be checked directly. The penalty function c^1 is defined as $c^1(\mathbb{Q}) = \sup\{\mathbb{E}_{\mathbb{Q}}[-\xi] \mid \xi \in \mathcal{A}^1\} = \sup\{\mathbb{E}_{\mathbb{Q}}[-\xi+\eta] \mid \xi \in \mathcal{A}\} = c(\mathbb{Q}) + \mathbb{E}_{\mathbb{Q}}[\eta]$. Clearly c^1 is convex, lower semi continuous, $\inf_{\mathbb{Q}\in\mathbf{P}} c^1(\mathbb{Q}) = 0$. The function u^1 satisfies the weak compactness property if and only if u satisfies the weak compactness property (see the next section for a definition of this property). The function u^1 is Gâteaux differentiable at 0 if u is Gâteaux differentiable at η. If u is coherent and given by the closed convex set $\mathcal{S}$, the penalty function c^1 is given by $c^1(\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}}[\eta]$ if $\mathbb{Q} \in \mathcal{S}$ and $c^1(\mathbb{Q}) = +\infty$ if $\mathbb{Q} \notin \mathcal{S}$. u^1 is no longer coherent. If u is given by the principle as in example 3, the utility function u^1 is not always of the same type (exercise: except when the von Neumann-Morgenstern function is exponential, where after a change of measure u^1 is again given by the same principle, see example 4). This is a good argument why we need a wider class than the von Neumann-Morgenstern functions.

4.12 Concave utility functions, reduction technique, weak compactness

In this section we reduce the study of monetary concave utility functions to the case of coherent utility functions. The geometric theorem is formulated in a rather abstract way so that it can be applied to the general case as well as to the Fatou case. We start with the definition of the recession cone, also called asymptotic cone. To fix the notation we will denote by E a locally convex topological space. This space is either L^∞ with the weak* topology (for the Fatou case) or L^∞ with the norm topology (general case). The topological dual of E is denoted by E^*.

Definition 14 *If $K \subset E$ is a convex set containing the origin, then the set*

$$K^e = \left\{ x \,\middle|\, \begin{array}{l} x \in E, \text{there exists nets (or generalised sequences)} \\ x^\alpha \in K, \lambda^\alpha \in \mathbb{R}_+, \lambda^\alpha \to 0, \text{ such that } \lambda^\alpha x^\alpha \to x \end{array} \right\},$$

is called the recession cone of K.

Proposition 15 *The recession cone K^e of a closed convex set K containing the origin, is a closed convex cone. More precisely*

$$K^e = \cap_{\varepsilon>0}\, \varepsilon K.$$

Proof. This is standard but for completeness we give a proof. First observe that $\cap_{\varepsilon>0}\, \varepsilon K \subset K^e$. Indeed if $x \in \cap_{\varepsilon>0}\, \varepsilon K$ then clearly for all $\varepsilon > 0$ there is $x^\varepsilon \in K$ so that $x = \varepsilon\, x^\varepsilon$, we can apply the definition of K^e. For the converse we first observe that if $0 < \eta < \varepsilon$ then $\eta K \subset \varepsilon K$. To see this let us write $x = \eta y$ with $y \in K$. Then we can write $x = \varepsilon \left(\frac{\eta}{\varepsilon} y + (1 - \frac{\eta}{\varepsilon})\, 0\right)$. By convexity and since $0 \in K$, the expression between brackets is in K. If $x \in K^e$ then $x = \lim_\alpha \lambda^\alpha x^\alpha$ where $x^\alpha \in K$ and $\lambda^\alpha \to 0$. Take now $\varepsilon > 0$. For α big enough, i.e. in a cofinal set, we get that $\lambda^\alpha \leq \varepsilon$ and hence for α big enough we get $\lambda^\alpha x^\alpha \in \varepsilon K$. Since this is true for all α big enough and since K is closed, we get $x \in \varepsilon K$. This proves $x \in \cap_{\varepsilon>0}\, \varepsilon K$. □

We now extend the space E in the following way. We put $F = E \times \mathbb{R}$ and we endow it with the product topology. For a given convex closed set with $0 \in K \subset E$, we put $K_1 = K \times \{1\} \subset F$. The closed convex cone generated by K_1 is the set:

$$K' = \cup_{t>0}(tK \times \{t\}) \cup (K^e \times \{0\}).$$

The set K_1 is a closed convex subset of K', namely $K = \{x \mid (x, 1) \in K'\}$. Since K' is a cone it is easy to characterise it with its polar cone. Now the dual space of F is precisely $E^* \times \mathbb{R}$ with the obvious inproduct defined as $((e^*, \beta), (e, t)) = e^*(e) + \beta t$. If K'^o denotes the dual cone we find that

$$K' = \{(x, t) \mid \forall (x^*, \beta) \in K'^o : x^*(x) + \beta\, t \geq 0\}.$$

In particular, for $t = 1$, we find the set K :

$$K = \{x \mid \forall (x^*, \beta) \in K'^o : x^*(x) + \beta \geq 0\}.$$

Let us apply this for the case of an acceptance set $\mathcal{A}$ that is convex weak* closed and such that $0 \in \mathcal{A} \supset L^\infty_+$. The set $\mathcal{A}$ is the set $\{\xi \mid u(\xi) \geq 0\}$, where

u is a concave monetary utility function defined on L^∞ and satisfying the Fatou property. The space $E = L^\infty$ is equipped with the topology $\sigma(L^\infty, L^1)$. The dual is then $E^* = L^1$. The space $F = E \times \mathbb{R}$ can be seen as the L^∞ space on the probability space $\Omega' = \Omega \cup \{p\}$, where p is an extra point added to Ω, $p \notin \Omega$. On Ω' we put the σ–algebra $\mathcal{F}' = \{A' \subset \Omega' | A' \cap \Omega \in \mathcal{F}\}$ and with the probability $\mathbb{P}'[A'] = (1/2)\mathbb{P}[A' \cap \Omega] + (1/2)\,\mathbf{1}_{A'}(p)$. The probabilities $\mathbb{Q}'$ defined on Ω', absolutely continuous with respect to $\mathbb{P}'$ can be identifed with the pairs (f, β) of nonnegative random variables f defined on Ω and numbers $\beta \geq 0$ such that $\mathbb{E}[f] + \beta = 2$. The construction above gives a weak* closed cone $L^\infty_+(\Omega') \subset \mathcal{A}' \subset L^\infty(\Omega')$ that can be seen as the acceptance cone of a coherent utility function u' defined on $L^\infty(\Omega')$. We can therefore apply the theory of coherent utility functions. The polar of the corresponding set $\mathcal{A}'$ is

$$\mathcal{A}'^o = \{(f, \beta) \mid \forall \xi \in \mathcal{A} : \mathbb{E}[f\,\xi] + \beta \geq 0 \text{ and } \forall \xi \in \mathcal{A}^e : \mathbb{E}[f\,\xi] \geq 0\}.$$

It follows that $(f, \beta) \in \mathcal{A}'^o$ implies that $f \geq 0$ and $\beta \geq 0$. If the element $f = 0$, then β is only restricted to be nonnegative. The representation Theorems 9, 12, then state that

$$u'(\xi, t) = \inf\left\{\frac{1}{2}\left(\mathbb{E}[f\xi] + \beta\, t\right) \mid (f, \beta) \in \mathcal{A}'^o; \mathbb{E}[f] + \beta = 2\right\}.$$

In particular we see that $\xi \in \mathcal{A}$ (or $(\xi, 1) \in \mathcal{A}'$) if and only if for all $(f, \beta) \in \mathcal{A}'^o$, we have $\mathbb{E}[f\xi] + \beta \geq 0$. Of course we only need to use the elements with $f \neq 0$. Let us analyse this a little bit further. Let us look at the closed convex set $\{(f, \beta) \in \mathcal{A}'^o \mid \mathbb{E}[f] = 1\}$. We find a function denoted by c (defined for probability measures $\mathbb{Q} \ll \mathbb{P}$), taking values in $\mathbb{R}_+ \cup \{+\infty\}$ and such that $(\mathbb{Q}, \beta) \in \mathcal{A}'^o$ if and only if $\beta \geq c(\mathbb{Q})$. The function $c(\mathbb{Q})$ can also be found as follows. For given $\mathbb{Q}$ we get that $c(\mathbb{Q}) = \sup\{-\mathbb{E}_\mathbb{Q}[\eta] \mid \eta \in \mathcal{A}\}$. This means that c is up to obvious sign changes, the support functional of $\mathcal{A}$. It is convex and lower semi-continuous in the sense that $\{\mathbb{Q} \mid c(\mathbb{Q}) \leq \alpha\}$ is closed (in L^1) and convex for each $\alpha \in \mathbb{R}$. Putting things together gives

$$\mathcal{A} = \{\xi \in L^\infty \mid \text{ for all } \mathbb{Q} \text{ probability measure, } \mathbb{Q} \ll \mathbb{P} : \mathbb{E}_\mathbb{Q}[\xi] + c(\mathbb{Q}) \geq 0\}.$$

Since the concave utility function is normalised so that $u(0) = 0$, we get that $\inf_\mathbb{Q} c(\mathbb{Q}) = 0$. Since u is monetary we also get

$$u(\xi) = \inf\{\mathbb{E}_\mathbb{Q}[\xi] + c(\mathbb{Q}) \mid \mathbb{Q} \text{ is a probability measure } \mathbb{Q} \ll \mathbb{P}\}.$$

Conversely if c is defined for all probability measures $\mathbb{Q} \ll \mathbb{P}$, takes values in $\mathbb{R}_+ \cup \{+\infty\}$, is lower semi-continuous, convex and satisfies $\inf_{\mathbb{Q}} c(\mathbb{Q}) = 0$, the above equality defines a concave monetary, Fatou utility function on L^∞.

It is not so easy to describe the utility function u' in terms of the utility function u. We will not analyse this quantitative relation between u and u'. Let us just mention that $u(\xi)$ is not necessarily equal to $u'(\xi, 1)$. We will only need the following relation between u and u':

Lemma 6 *For $\xi \in L^\infty(\Omega)$ we have $u(\xi) = 0$ if and only if $u'(\xi, 1) = 0$.*

Proof. Suppose that $u'(\xi, 1) = 0$. Since $\xi \in \mathcal{A}$ by definition of $\mathcal{A}'$, we already have $u(\xi) \geq 0$. But we actually have $u(\xi) = 0$. Indeed for each $\varepsilon > 0$ we have $(\xi - \varepsilon, 1 - \varepsilon) \notin \mathcal{A}'$ and hence $(\frac{\xi-\varepsilon}{1-\varepsilon}, 1) \notin \mathcal{A}'$. This means that $u\left(\frac{\xi-\varepsilon}{1-\varepsilon}\right) < 0$. Since $\frac{\xi-\varepsilon}{1-\varepsilon}$ converges *uniformly* to ξ we get $u(\xi) \leq 0$. Suppose now that $u(\xi) = 0$, then $(\xi, 1) \in \mathcal{A}$ and hence $u'(\xi, 1) \geq 0$. But since $\xi - \varepsilon \notin \mathcal{A}$ we have $(\xi - \varepsilon, 1) \notin \mathcal{A}'$ and $u'(\xi - \varepsilon, 1) < 0$. If ε converges to zero, the continuity of u' for the uniform convergence implies that $u'(\xi, 1) \leq 0$. □

Remark 22 The generalisation of coherent utility functions to concave utility functions was developed by Föllmer and Schied, see [**?**]. The presentation in this section is different. There is not much advantage coming from this homogenisation technique, except that it allows us to use theorems from functional analysis in an easier way. We leave it up to the reader to rephrase the theory for monetary utility functions that do not necessarily satisfy the Fatou property.

Theorem 23 *Let u be a concave monetary utility function satisfying the Fatou property and represented by the lower semi continuous convex function $c(\mathbb{Q})$. The set*

$$\{\xi \mid \text{there is a probability } \mathbb{Q} \text{ with } \mathbb{E}_{\mathbb{Q}}[\xi] + c(\mathbb{Q}) = u(\xi)\}$$

is norm dense in the space L^∞.

Proof. This statement follows from the Bishop-Phelps theorem. Let us take $\xi \in L^\infty$ such that $\|\xi\|_\infty = 1$. The Bishop-Phelps theorem is now applied to the element $(\xi, 1) \in L^\infty(\Omega')$ and the bounded closed convex set

$$M = \{(\mathbb{Q}, \beta) \mid \mathbb{Q} \text{ a probability, } (\mathbb{Q}, \beta) \in \mathcal{A}'^{o} \text{ and } \beta \leq 6\}.$$

For every $1/2 > \varepsilon > 0$ there is an "inf attaining" element $(\eta, 1-\delta)$ such that $\|\xi - \eta\|_\infty < \varepsilon$, $|\delta| < \varepsilon$. This means that there is a probability measure $\mathbb{Q}$ so that

$$\mathbb{E}_\mathbb{Q}[\eta]+(1-\delta)c(\mathbb{Q}) = \inf_{(\mathbb{Q}',\beta)\in M} \left(\mathbb{E}_{\mathbb{Q}'}[\eta] + (1-\delta)\beta\right) = \inf_{\mathbb{Q}'} \left(\mathbb{E}_{\mathbb{Q}'}[\eta] + (1-\delta)c(\mathbb{Q}')\right).$$

The second equality follows from the fact that $c(\mathbb{Q}') = \inf\{\beta \mid (\mathbb{Q}', \beta) \in \mathcal{A}'^o\}$ and the observation that we do not need elements with $c(\mathbb{Q}') \geq 6$. We can rewrite this as follows

$$\mathbb{E}_\mathbb{Q}\left[\frac{1}{1-\delta}\eta\right] + c(\mathbb{Q}) = \inf_{\mathbb{Q}'} \left(\mathbb{E}_{\mathbb{Q}'}\left[\frac{1}{1-\delta}\eta\right] + c(\mathbb{Q}')\right).$$

In other words $u\left(\frac{1}{1-\delta}\eta\right) = \mathbb{E}_\mathbb{Q}[\frac{1}{1-\delta}\eta] + c(\mathbb{Q})$. But as easily seen $\|\xi - \frac{1}{1-\delta}\eta\|_\infty \leq 3\varepsilon$. □

Remark 23 The Bishop-Phelps theorem has many applications and "family members" of it were used in optimisation theory. The variational principle of Ekeland can be used to get the preceding theorem in a direct way, see [57], [112].

Let us now see what could be the equivalent property of weak compactness in the case of concave utility functions. One useful property is that for uniformly bounded sequences ξ_n, converging in probability to ξ, we should have $\lim u(\xi_n) = u(\xi)$. Another generalisation could be: the basis of the cone $\mathcal{A}'^o$ is weakly compact. We use the same notation as in the section 4.12.

Theorem 24 *For a concave monetary utility function $u : L^\infty \to \mathbb{R}$ the following are equivalent*

1. *u satisfies the property $\lim u(\xi_n) = u(\xi)$ for uniformly bounded sequences $(\xi_n)_n$, converging in probability to a random variable ξ.*

2. *The basis of $\mathcal{A}'^o$ defined as $\{(f, \beta) \in \mathcal{A}'^o \mid \mathbb{E}[f] + \beta = 2\}$ is weakly compact.*

3. *The convex function c satisfies: for each $\infty > \alpha \geq 0$, $\{\mathbb{Q} \mid c(\mathbb{Q}) \leq \alpha\}$ is weakly compact in L^1.*

4. *For each $\xi \in L^\infty$ there is a probability $\mathbb{Q}$ so that $u(\xi) = \mathbb{E}_\mathbb{Q}[\xi] + c(\mathbb{Q})$.*

Proof. The proof follows the lines $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 2 \Rightarrow 1$. We will show that the coherent risk measure u' defined on $L^\infty(\Omega')$ satisfies property (WC) of section 4.3. By Theorem 15 this will show the weak compactness of the basis of $\mathcal{A}'^o$. Let us take a uniformly bounded sequence (ξ_n, t_n) that increases to $(0,0)$. This means that ξ_n increases to 0 and $t_n \uparrow 0$ For $\varepsilon > 0$ we get that $t_n + \varepsilon \geq \varepsilon/2$ for n big enough. Since the variables are increasing we can pass to a subsequence and hence we may suppose that $t_n + \varepsilon \geq \varepsilon/2$ for all n. The sequence $\frac{\xi_n+\varepsilon}{t_n+\varepsilon}$ tends to 1 and remains uniformly bounded. This implies that $u\left(\frac{\xi_n+\varepsilon}{t_n+\varepsilon}\right)$ tends to 1. In other words $\frac{\xi_n+\varepsilon}{t_n+\varepsilon} \in \mathcal{A}$ for n big enough. By definition of the cone $\mathcal{A}'$ this implies that $\left(\frac{\xi_n+\varepsilon}{t_n+\varepsilon}, 1\right) \in \mathcal{A}'$ and hence also $(\xi_n + \varepsilon, t_n + \varepsilon) \in \mathcal{A}'$. In other words $u'((\xi_n + \varepsilon, t_n + \varepsilon)) \geq 0$ and hence $u'(\xi_n, t_n) \geq -\varepsilon$. This shows that $u'(\xi_n, t_n)$ tends to zero. In other words the basis of $\mathcal{A}'^o$ is weakly compact.

Let us now suppose that the set $B = \{(f, \beta) \in \mathcal{A}'^o \mid \mathbb{E}[f] + \beta = 2\}$ is weakly compact. If $f = \frac{d\mathbb{Q}}{d\mathbb{P}}$, then the element $\left(\frac{2f}{1+c(\mathbb{Q})}, \frac{2c(\mathbb{Q})}{1+c(\mathbb{Q})}\right) \in B$ (with the obvious modifications if $c(\mathbb{Q}) = \infty$). Conversely if $(f, \beta) \in B$ and $\beta < 2$ then $c(\frac{f}{2-\beta}) = \frac{\beta}{2-\beta}$. In this correspondence the elements $\mathbb{Q}$ with $c(\mathbb{Q}) = \infty$ are mapped onto the element $(0, 2) \in B$. It is now clear that the set $\{\mathbb{Q} \mid c(\mathbb{Q}) \leq \alpha < \infty\}$ is coming from the set $\{(f, \beta) \mid \beta \leq \beta_0 = \frac{2\alpha}{1+\alpha} < 2\}$, which as a closed set of B is again weakly compact. Since $\beta_0 < 2$, the multiplication of the first coordinate f is by a uniformly bounded real number and hence the set $\{\mathbb{Q} \mid c(\mathbb{Q}) \leq \alpha < \infty\}$ is weakly compact as the image of a weakly compact set.

Let us now suppose 3 and prove 4. For given $\xi \in L^\infty$ let $\mathbb{Q}_n$ be a sequence such that $u(\xi) = \lim\left(\mathbb{E}_{\mathbb{Q}_n}[\xi] + c(\mathbb{Q}_n)\right)$. Since the sequence $c(\mathbb{Q}_n)$ is eventually bounded, we find that the sequence $\mathbb{Q}_n$ is taken in a weakly compact set. So we may suppose that it converges weakly to an element $\mathbb{Q}$. Since $c(\mathbb{Q}) \leq \liminf c(\mathbb{Q}_n)$ we get that $u(\xi) = \mathbb{E}_\mathbb{Q}[\xi] + c(\mathbb{Q})$.

Now we prove that 4 implies 2. By James' theorem we have to show that every element (ξ, t) attains its minimum on the basis $B \subset KK^o$. Of course may suppose that $u'(\xi, t) = 0$. This implies that $t \geq 0$ and hence we distinguish two cases $t = 0$ and $t = 1$ the latter can be obtained by normalising the element (ξ, t). The homogeneity of u' guarantees that we still have an element $u'(\xi, 1) = 0$. We first treat the case of $(\xi, 0)$. As observed above, we have that $(0, 2) \in B$ and hence there is an element in B that realises the infimum of $(\xi, 0)$ on the set B. The case $(\xi, 1)$ is easier. There is a probability $\mathbb{Q}$ so that $\mathbb{E}_\mathbb{Q}[\xi] + c(\mathbb{Q}) = u(\xi)$. But $u(\xi) = 0$ as shown in lemma 6. So we get $\mathbb{E}_{\mathbb{Q}}[\xi] + c(\mathbb{Q}) = 0$. If we divide this expression by $2/(1 + c(\mathbb{Q})$ we get an element $(f, \beta) \in B$ that realises the minimum.

We now show that 2 implies 1. This is also easy. We suppose that ξ_n is a uniformly bounded sequence that converges to ξ in probability. Suppose that $u(\xi) > 0$. We have to show that $u(\xi_n)$ becomes eventually nonnegative. Since $(\xi_n, 1)$ tends to $(\xi, 1)$, we have $u'(\xi_n, 1)$ tends to $u'(\xi, 1)$ as a consequence of weak compactness. But $u'(\xi, 1) > 0$, as seen from lemma 6. For n big enough this means $u'(\xi_n, 1) \geq 0$ and therefore $\xi_n \in \mathcal{A}$. Hence $u(\xi_n) \geq 0$ □

Remark 24 The above theorem was first proved by Jouini-Schachermayer-Touzi [84]. Their proof was more involved and was based on the (rather complicated) proof of James's theorem.

4.13 The one-sided derivative

Because of concavity, monetary concave utility functions have a one-sided derivative at every point $\xi \in L^\infty$. It is defined as

$$\varphi_\xi(\eta) = \lim_{\varepsilon \downarrow 0} \frac{u(\xi + \varepsilon\eta) - u(\xi)}{\varepsilon}.$$

If $\xi = 0$ we get

$$\varphi(\eta) = \lim_{\epsilon \downarrow 0} \frac{u(\epsilon\eta)}{\epsilon}.$$

Proposition 16 *The function φ is the smallest coherent utility function bigger than u.*

Proof. Take ψ coherent and $\psi \geq u$. Since for each $\varepsilon > 0$ we have $\psi(\eta) = \psi(\epsilon\eta)/\epsilon \geq u(\epsilon\eta)/\epsilon \geq u(\eta)$, we get $\psi(\eta) \geq \varphi(\eta) \geq u(\eta)$. □

The acceptance cone for φ is easily obtained via the acceptance set $\mathcal{A}$.

Proposition 17 *The acceptance cone of φ, $\mathcal{A}_\varphi$, is given by the $\|.\|_\infty$ closure of the union $\cup_n n\mathcal{A}$*

Proof. Suppose first that $\eta \in n\mathcal{A}$, then for $\epsilon \leq 1/n$ we have by convexity of $\mathcal{A}$ that $u(\epsilon\eta) \geq 0$. This shows that $\varphi(\eta) \geq 0$. It follows that $\cup_n n\mathcal{A} \subset \mathcal{A}_\varphi$. Since the latter set is norm closed, it also has to contain the norm-closure of this union. If $\varphi(\eta) > 0$, we have that for ϵ small enough $u(\epsilon\eta) > 0$ and hence for n big enough we have $\eta \in n\mathcal{A}$. This shows the opposite inclusion. □

The scenario set – or the polar cone – that defines the coherent utility function φ is given by the following proposition

Proposition 18 *With the notation introduced above we have*

$$\varphi(\eta) = \inf_{\mu \in \mathcal{S}^{\mathbf{ba}}} \mu(\eta),$$

where the set $\mathcal{S}^{\mathbf{ba}} = \{\mu \in \mathbf{P}^{\mathbf{ba}} \mid c(\mu) = 0\}$.

Proof. Because of the previous proposition, $\mu \in \mathcal{S}^{\mathbf{ba}}$ if and only if $\mu(\eta) \geq 0$ for all $\eta \in \mathcal{A}$. This is equivalent to saying that $c(\mu) = 0$. □

Corollary 4 *The one-sided derivative φ of u at 0 is Fatou if and only if* $\{\mathbb{Q} \in \mathbf{P} \mid c(\mathbb{Q}) = 0\}$ *is weak* dense in* $\{\mu \in \mathbf{P}^{\mathbf{ba}} \mid c(\mu) = 0\}$.

Remark 25 For the derivative at a point ξ we use the transformation $u_\xi(\eta) = u(\xi + \eta) - u(\xi)$. It follows that the derivative at a point ξ is given by

$$\varphi_\xi(\eta) = \inf\{\mu(\eta) \mid c(\mu) + \mu(\xi) = u(\xi)\}.$$

The example 19 shows that for $c(\mu) = \sum_n \mu(A_n)^2$ the scenario set of φ is $\mathcal{S}^{\mathbf{ba}} = \{\mu \in \mathbf{P}^{\mathbf{ba}} \mid \mu(A_n) = 0 \text{ for all } n\}$. Since this set only contains purely finitely additive probability measures, the one-sided derivative cannot be Fatou. The results on automatic continuity [35] do not apply to concave functions.

The coherent utility function φ was the smallest coherent utility function dominating u. Is there also a largest coherent utility function, ψ such that $\psi \leq u$? The answer is given by the following proposition.

Proposition 19 *The recesssion cone, $\mathcal{A}^e$ of $\mathcal{A}$ defines a coherent utility function ψ that is the largest coherent utility function smaller than u. ψ is Fatou if u is Fatou.*

Proof. If ψ is a coherent function smaller that u then its acceptance cone is contained in $\mathcal{A}$. Since $\mathcal{A}^e$ is the largest cone contained in $\mathcal{A}$, the proposition follows. If u is Fatou we get that $\mathcal{A}^e = \cap_n \frac{1}{n}\mathcal{A}$ and hence the weak* closedness of $\mathcal{A}$ implies that also $\mathcal{A}^e$ is weak* closed. We recall that this recession cone was also used in the homogenisation technique. □

4.14 Relevance: Halmos-Savage theorem

Definition 15 *A monetary utility function u is called relevant if $\xi \in L^\infty$, $\xi \le 0$ and $\mathbb{P}[\xi < 0] > 0$ imply $u(\xi) < 0$.*

Theorem 25 *For a Fatou monetary concave utility function, the following are equivalent:*

1. *u is relevant, i.e. for $\xi \ge 0, \mathbb{P}[\xi > 0] > 0$ we have $u(-\xi) < 0$.*
2. *$A \in \mathcal{F}$, $\mathbb{P}[A] > 0$, $\varepsilon > 0$ imply $u(-\varepsilon \mathbf{1}_A) < 0$;*
3. *For all $\xi \ge 0$, $\mathbb{E}[\xi] > 0$ there is $\mathbb{Q}$ with $c(\mathbb{Q}) - \mathbb{E}_\mathbb{Q}[\xi] < 0$.*
4. *For all $\delta > 0$ there exist $\mathbb{Q}$ and $\eta > 0$ such that for all $1 \ge \xi \ge 0$, $\mathbb{E}[\xi] \ge \delta$: $c(\mathbb{Q}) - \mathbb{E}_\mathbb{Q}[\xi] \le -\eta < 0$. Here η is determined by δ, also the measure $\mathbb{Q}$ depends on δ.*
5. *For all $\delta > 0$ there exist $\mathbb{Q} \sim \mathbb{P}$ such that for all $1 \ge \xi \ge 0$, $\mathbb{E}[\xi] \ge \delta$: $c(\mathbb{Q}) - \mathbb{E}_\mathbb{Q}[\xi] < 0$.*

Proof. $1 \Rightarrow 2$ is trivial. That $2 \Rightarrow 1$ is easy. Let ξ be given and take $\varepsilon > 0$ so that $\mathbb{P}[\xi > \varepsilon] > 0$. Now $\xi \ge \varepsilon \mathbf{1}_{\{\xi > \varepsilon\}}$, hence $u(-\xi) \le u(-\varepsilon \mathbf{1}_{\{\xi>\varepsilon\}}) < 0$ by assumption 2. $1 \Leftrightarrow 3$ since $u(-\xi) = \inf_\mathbb{Q}(c(\mathbb{Q} - \mathbb{E}_\mathbb{Q}[\xi])$. Now comes the serious work. $1, 3 \Rightarrow 4$. This will follow from the separation theorem. Let $K_\delta = \{\xi \mid 1 \ge \xi \ge 0; \mathbb{E}[\xi] \ge \delta\}$. This set is weak* compact in L^∞. The set $-K_\delta$ is disjoint from the weak* closed set $\mathcal{A}$. The separation theorem yields a linear functional (taken in L^1) that separates the two sets. Because $\mathcal{A}$ contains the positive cone, the functional is nonnegative and we can normalise it to a probability, $\mathbb{Q}$. This yields

$$\sup_{\xi \in K_\delta} \mathbb{E}_\mathbb{Q}[-\xi] < \inf_{\beta \in \mathcal{A}} \mathbb{E}_\mathbb{Q}[\beta] = -c(\mathbb{Q}).$$

This means that there is $\mathbb{Q}$ and $\eta > 0$ such that

$$\sup_{\xi \in K_\delta} (c(\mathbb{Q}) - \mathbb{E}_\mathbb{Q}[\xi]) \le -\eta.$$

Let us now show that $4 \Rightarrow 5$. We have to find $\mathbb{Q} \sim \mathbb{P}$ but we do not require the uniform bound with some $\eta > 0$. The proof is as in the Halmos-Savage theorem, [79]. Define the class:

$$\mathcal{B} = \left\{ \left\{ \frac{d\mathbb{Q}}{d\mathbb{P}} > 0 \right\} \middle| \sup_{\xi \in K_\delta} (c(\mathbb{Q}) - \mathbb{E}_\mathbb{Q}[\xi]) < 0 \right\}.$$

This class is stable for countable unions. Indeed if

$$\sup_{\xi \in K_\delta} (c(\mathbb{Q}^n) - \mathbb{E}_{\mathbb{Q}^n}[\xi]) < 0$$

for a sequence $\mathbb{Q}^n$, then the measure $\mathbb{Q} = \sum_n 2^{-n}\mathbb{Q}^n$ satisfies, by the convexity of c, the same *strict* inequality. So there is a maximal element, $B \in \mathcal{B}$ and let this be given by $\mathbb{Q}$, $B = \left\{\frac{d\mathbb{Q}}{d\mathbb{P}} > 0\right\}$. If $A = B^c$ would be nonnegligible, we distiguish two cases. First let us suppose that $\mathbb{P}[A] \ge \delta$, then $\mathbf{1}_A \in K_\delta$ and hence $c(\mathbb{Q}) - \mathbb{Q}[\mathbf{1}_A] < 0$ meaning that $A \cap B \neq \emptyset$. So we are left with the second case $\mathbb{P}[A] < \delta$. Let us now apply the assumption 4 with $\delta' = \mathbb{P}[A]$. This gives a measure $\mathbb{Q}^A$ such that

$$\sup_{\xi \in K_{\delta'}} (c(\mathbb{Q}^A) - \mathbb{E}_{\mathbb{Q}^A}[\xi]) < 0.$$

The class $K_{\delta'}$ contains K_δ and hence $\left\{\frac{d\mathbb{Q}^A}{d\mathbb{P}} > 0\right\} \in \mathcal{B}$. But this time we have $c(\mathbb{Q}^A) - \mathbb{Q}^A[A] < 0$, showing that again $A \cap B \neq \emptyset$, a contradiction to the maximality of B. This shows that $B = \Omega$ or $\mathbb{Q} \sim \mathbb{P}$. That $5 \Rightarrow 1$ is straightfoward. Indeed 3 is equivalent to 1 and 5 is obviously stronger. □

Remark 26 In assumption 4 we used the condition $1 \ge \xi$ only to get a weak* compact set. Due to monotonicity the condition is equivalent to

For all $\delta > 0$ there exist $\mathbb{Q}$ and $\eta > 0$ such that for all $\xi \ge 0$, $\mathbb{E}[\xi] \ge \delta$: $c(\mathbb{Q}) - \mathbb{E}_{\mathbb{Q}}[\xi] \le -\eta < 0$.

Example 23 Take again Example 21 with $1/2 \ge f > 0$ and $\text{ess.inf}\, f = 0$. The function $c(\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}}[f]$ defines the utility $u(\xi) = \text{ess.inf}(\xi + f)$. For $\xi = \mathbf{1}_A$, $\mathbb{P}[A] > 0$, we find $u(-\mathbf{1}_A) < 0$. If we take $A = \{f > \varepsilon\}$ where $\mathbb{P}[f > \varepsilon] > 0$, we get $u(-\varepsilon\mathbf{1}_A) \ge 0$. This shows that in the theorem we cannot restrict to indicators but we have to take multiples of indicators.

Proposition 20 *If u is relevant then for all $\varepsilon > 0$ there is $\mathbb{Q} \sim \mathbb{P}$ with $c(\mathbb{Q}) \le \varepsilon$. Consequently for every $\mathbb{Q} \in \mathbf{P}$ there is a sequence $\mathbb{Q}^n$ of equivalent measures such that $c(\mathbb{Q}) = \lim_n c(\mathbb{Q}^n)$.*

Proof. Let $\varepsilon > 0$. By item 5 there is $\mathbb{Q} \sim \mathbb{P}$ such that $c(\mathbb{Q}) \le \inf\{\mathbb{E}_{\mathbb{Q}}[\xi] \mid \xi \ge 0; \mathbb{E}_{\mathbb{P}}[\xi] \ge \varepsilon\} \le \varepsilon$. To prove the last assertion, take $\mathbb{Q}^0 \sim \mathbb{P}$ with $c(\mathbb{Q}^0) < \infty$. Then take $\mathbb{Q}^n = \frac{1}{n}\mathbb{Q}^0 + \frac{n-1}{n}\mathbb{Q}$. Clearly $\mathbb{Q}^n \sim \mathbb{P}$.

Since c is lower semi continuous we have $c(\mathbb{Q}) \le \liminf_n c(\mathbb{Q}^n)$. The function $c(t\mathbb{Q}^0 + (1-t)\mathbb{Q})$ is convex for $t \in [0,1]$ and convexity implies upper semi continuity at the end points of the interval $[0,1]$, consequently also $c(\mathbb{Q}) \ge \limsup_n c(\mathbb{Q}^n)$ and we get $c(\mathbb{Q}) = \lim_n c(\mathbb{Q}^n)$ as desired. □

Corollary 5 *If u is relevant, then $u(\xi) = \inf\{\mathbb{E}_{\mathbb{Q}}[\xi] + c(\mathbb{Q}) \mid \mathbb{Q} \sim \mathbb{P}\}$.*

Example 24 The preceding corollary is false if the utility function is not relevant. Simply take $u(\xi) = \text{ess.inf}(\mathbf{1}_A\xi)$ where $0 < \mathbb{P}[A] < 1$. The penalty function is $c(\mathbb{Q}) = 0$ if $\mathbb{Q}[A] = 1$ and equals $+\infty$ if $\mathbb{Q}[A] < 1$. Every equivalent measure $\mathbb{Q}$ therefore satisfies $c(\mathbb{Q}) = +\infty$. Remark that this utility function is coherent.

From Bion-Nadal, [21] we recall the following definition of non-degenerate utility functions

Definition 16 *A monetary utility function u is called non-degenerate if $A \in \mathcal{F}$, $\mathbb{P}[A] > 0$ imply the existence of $\lambda > 0$ with $u(\lambda\mathbf{1}_A) > 0$.*

Proposition 21 *A non-degenerate utility function is relevant.*

Proof. Because of concavity a non-degenerate utility function u satisfies $u(\varepsilon\mathbf{1}_A) > 0$ for each A with $\mathbb{P}[A] > 0$ and each $\varepsilon > 0$. Now take $\varepsilon > 0$ and write

$$0 = u(0) \ge \frac{1}{2}u(\varepsilon\mathbf{1}_A) + \frac{1}{2}u(-\varepsilon\mathbf{1}_A).$$

Because $u(\varepsilon\mathbf{1}_A) > 0$ we must have $u(-\varepsilon\mathbf{1}_A) < 0$. □

Example 25 We take the example 19 where $c(\mathbb{Q}) = \sum_n \mathbb{Q}[A_n]^2$ where $A_n; n \ge 1$ is a given partition of Ω. We also suppose that the probability space is atomless. This defines $u(\xi) = \inf_{\mathbb{Q}}(c(\mathbb{Q}) + \mathbb{E}_{\mathbb{Q}}[\xi])$ as a utility function. We claim that this utility is relevant. But since $c(\mathbb{Q}) > 0$ for all $\mathbb{Q}$, we cannot show existence of a probability measure such that for all $\xi \ge 0, \mathbb{E}[\xi] > 0$: $c(\mathbb{Q}) - \mathbb{E}_{\mathbb{Q}}[\xi] < 0$. This shows that in the theorem we need some restriction on the size of ξ. So we needed to introduce $\mathbb{E}[\xi] \ge \delta$ for some $\delta > 0$. Let us now show that the utility function is relevant. Take $\xi \ge 0$ but not negligible, then there is n such that $\mathbb{E}[\xi\mathbf{1}_{A_n}] > 0$. Let us take a probability $\mathbb{Q}$ such that for this n: $-\eta = \mathbb{Q}[A_n]^2 - \mathbb{E}_{\mathbb{Q}}[\xi\mathbf{1}_{A_n}] < 0$. This is certainly possible since we can take $\mathbb{Q}[A_n] = \varepsilon\mathbb{P}[A_n]$ for $\varepsilon > 0$, small enough.

The mass of A_n^c will now be redistributed over the sets $A_k, k \neq n$ such that $\sum_{k\neq n} \mathbb{Q}[A_k]^2 \leq \eta/2$. This is done in the same way as in example 19. The construction yields a probability such that $c(\mathbb{Q}) - \mathbb{E}_{\mathbb{Q}}[\xi \mathbf{1}_{A_n}] \leq -\eta/2 < 0$. The one-sided derivative φ is not relevant since for each A_n, $\varphi(-\mathbf{1}_{A_n}) = 0$ whereas $\mathbb{P}[A_n] > 0$. This example can be made more general.

In [30] Cheridito et al introduced a related concept, called sensitivity to great losses.

Definition 17 *A concave monetary utility function is called sensitive to great losses if* $\mathbb{P}[\xi < 0] > 0$ *implies that* $\lim_{\lambda\to+\infty} u(\lambda\xi) = -\infty$.

Remark 27 From the definition it immediately follows that if u is coherent and sensitive to great losses, then necessarily $u(\xi) = \text{ess.inf}\,\xi$. Indeed if ξ is acceptable, i.e. $u(\xi) \geq 0$, then coherence implies $u(\lambda\xi) \geq 0$ for all $\lambda \geq 0$. Hence $\xi \geq 0$ a.s. .

Remark 28 In the definition of sensitivity to great losses, it is sufficient to require that $\mathbb{P}[\xi < 0] > 0$ implies that there is $\lambda > 0$ such that $u(\lambda\xi) < 0$. Indeed the concavity then implies $\lim_{\lambda\to+\infty} u(\lambda\xi) = -\infty$.

Proposition 22 *If u is a Fatou, concave, monetary utility function, then sensitivity to great losses is equivalent to: for each A with $\mathbb{P}[A] > 0$, we have* $\sup_{c(\mathbb{Q})<\infty} \mathbb{Q}[A] = 1$.

Proof Suppose u is sensitive to great losses. For $\mathbb{P}[A] > 0$ and $\delta > 0$ look at the function $\xi = \mathbf{1}_{A^c} - \delta\mathbf{1}_A$. There exists a $\lambda > 0$ such that $u(\lambda\xi) < 0$. Hence there is $\mathbb{Q} \in \mathbf{P}$ such that $\lambda\mathbb{E}_{\mathbb{Q}}[\xi] + c(\mathbb{Q}) < -1$. Because $\infty > c(\mathbb{Q}) \geq 0$, we must have $\mathbb{Q}[A^c] - \delta\mathbb{Q}[A] < 0$. This yields $\mathbb{Q}[A] \geq \frac{1}{1+\delta}$, proving that $\sup_{c(\mathbb{Q})<\infty} \mathbb{Q}[A] = 1$. Conversely we take ξ with $\mathbb{P}[\xi < 0] > 0$. Then there is $\delta > 0$ with $\mathbb{P}[\xi < -\delta] > 0$. Let $A = \{\xi < -\delta\}$. Obviously we have $u(\xi) \leq u(\|\xi\|\mathbf{1}_{A^c} - \delta\mathbf{1}_A$. Take now $\mathbb{Q}$ such that $c(\mathbb{Q}) < \infty$ and $\mathbb{Q}[A] \geq 1-\varepsilon$ (with $\varepsilon > 0$ to be fixed). We get for $\lambda > 0$: $u(\lambda\xi) \leq \lambda\left(\|\xi\|\mathbb{Q}[A^c] - \delta\mathbb{Q}[A]\right) + c(\mathbb{Q}) \leq \lambda\left(\varepsilon\|\xi\| - \delta(1-\varepsilon)\right) + c(\mathbb{Q})$. If ε is chosen small enough so that the first term is negative, we get that $\lim_\lambda u(\lambda\xi) = -\infty$. □

For coherent utilities u defined by the closed convex set $\mathcal{S} \subset L^1$, we get the following version (a restatement of the Halmos-Savage theorem). We do not give a proof since it is contained in Theorem 25 (use the observation that $c(\mathbb{Q}) = 0$ for $\mathbb{Q} \in \mathcal{S}$ and $= +\infty$ for $\mathbb{Q} \notin \mathcal{S}$).

Theorem 26 *For a Fatou coherent utility function, the following are equivalent:*

1. *u is relevant, i.e. for $\xi \geq 0, \mathbb{P}[\xi > 0] > 0$ we have $u(-\xi) < 0$, meaning there is $\mathbb{Q} \in \mathcal{S}$ with $\mathbb{E}_{\mathbb{Q}}[\xi] > 0$.*

2. *$A \in \mathcal{F}$, $\quad \mathbb{P}[A] > 0$, imply $u(-\mathbf{1}_A) < 0$, meaning there is $\mathbb{Q} \in \mathcal{S}$ with $\mathbb{Q}[A] > 0$.*

3. *There exist $\mathbb{Q} \in \mathcal{S}; \mathbb{Q} \sim \mathbb{P}$.*

Remark 29 We remark that relevance does not imply strict monotonicity! For instance, take an atomless space Ω and consider the set $\mathcal{S}_2 = \{\mathbb{Q} \mid \frac{d\mathbb{Q}}{d\mathbb{P}} \leq 2\}$. Then u_2 is relevant (because $\mathbb{P}$ itself belongs to $\mathcal{S}_2$) but not strictly monotone. If A is such that $0 < \mathbb{P}[A] < \frac{1}{2}$, then $u_2(\mathbf{1}_A) = \inf\{\mathbb{Q}[A] \mid \mathbb{Q} \in \mathcal{S}\} = 0$. To see this equality, simply remark that $\frac{1}{\mathbb{P}[A^c]}\mathbf{1}_{A^c}$ is an element of $\mathcal{S}$. Of course we have $u_2(\mathbf{1}_A) = u(0) = 0$, a contradiction to strict monotonicity!

Remark 30 The examples defined in Section 4.11 do not always satisfy the relevance property, even if the original utility function does. We follow the notation of Section 4.11, i.e. $u^1(\xi) = u(\xi + \eta)$ where u is coherent and Fatou, $u(\eta) = 0$. For u we take the TailVar with level $1 > \alpha > 0$. Clearly u is relevant. For η we take the indicator function of a set B with probability $1 - \alpha$. For $A \subset B$, $u^1(-\mathbf{1}_A) = 0$ since $-\mathbf{1}_A + \eta \geq 0$, hence $u^1(-\mathbf{1}_A) = u(-\mathbf{1}_A + \mathbf{1}_B) = 0$.

In the following definition we assume that $u : L^\infty \to \mathbb{R}$ is a Fatou, concave, monetary utility function. This function is extended to $u\colon L^\infty - L^0_+ \to \mathbb{R} \cup \{-\infty\}$ as explained in Section 4.9.

Definition 18 *An element $\eta \in L^\infty - L^0_+$ is called minimal if for all $\eta' \leq \eta$ with $\mathbb{P}[\eta' < \eta] > 0$ we have $u(\eta') < u(\eta)$.*

Clearly u is relevant if and only if 0 is minimal. Suppose that $\eta \in L^\infty - L^0_+$ with $u(\eta) = 0$. Define $u^1 : L^\infty \to \mathbb{R}$ as $u^1(\xi) = u(\xi + \eta)$. As easily seen u^1 is relevant if and only if η is minimal. The interest comes from the following result.

Proposition 23 *Suppose that u is relevant. If $\eta \in L^\infty - L^0_+$ and $u(\eta) = 0$ there is $\eta' \leq \eta$ with $u(\eta') = u(\eta) = 0$, $\eta' \in L^\infty - L^0_+$ and η' is minimal.*

Proof. The proof is based on Zorn's lemma. Because u is relevant we have the existence of $\mathbb{Q} \sim \mathbb{P}$ such that $\mathbb{E}_\mathbb{Q}[\eta] + c(\mathbb{Q}) \leq 2$. This implies that there is $\mathbb{Q} \sim \mathbb{P}$ with $c(\mathbb{Q}) < \infty$. Let $(\eta_i)_i$ be a completely ordered system (for the relation $\leq$ a.s.) with $u(\eta_i) = u(\eta) = 0$ and $\eta_i \leq \eta$. Because $\mathbb{E}_\mathbb{Q}[\eta_i] + c(\mathbb{Q}) \geq 0$ we get that $\mathbb{E}_\mathbb{Q}[\eta_i] \geq -c(\mathbb{Q}) > -\infty$. Clearly the order on the set $(\eta_i)_i$ is equivalent to the order given by $\mathbb{E}_\mathbb{Q}[\eta_i]$. Take a decreasing sequence η_{i_n} such that $\lim_n \mathbb{E}_\mathbb{Q}[\eta_{i_n}] = \inf_i \mathbb{E}_\mathbb{Q}[\eta_i]$. The element $\eta' = \lim_n \eta_{i_n}$ still satisfies $u(\eta') = u(\eta)$ (since u is Fatou) and $\eta' \leq \eta_i$ for all i. Furthermore η' is still in $L^\infty - L^0_+$, since $\mathbb{E}_\mathbb{Q}[\eta'] \geq -c(\mathbb{Q})$. Zorn's axiom now says that there is a minimal element η' in the set $\{\xi \in L^\infty - L^0_+ \mid u(\xi) = u(\eta), \xi \leq \eta\}$. □

Remark 31 Even if $\eta \in L^\infty$, there is no guarantee that there is minimal element $\eta' \in L^\infty$ with $\eta' \leq \eta$ and $u(\eta') = u(\eta)$.

To check whether an element is minimal is not easy, the following criterion gives a sufficient condition. To simplify notation let us say that $\mu \in \mathbf{P^{ba}}$ is equivalent to $\mathbb{P}$, we write $\mu \sim \mathbb{P}$, if $\mathbb{P}[A] > 0$ implies that $\mu(A) > 0$. Following Yosida-Hewitt, [135], we can decompose μ in its sigma-additive part, μ_a, and its purely finitely additive part, μ_s, i.e. $\mu = \mu_a + \mu_s$. For μ_s we can find a countable decomposition of Ω, $\Omega = \cup_n B_n$, where B_n is a sequence of pairwise disjoint sets such that $\mu_s(B_n) = 0$ for all n. It now follows that $\mu \sim \mathbb{P}$ if and only if $\mu_a \sim \mathbb{P}$ in the usual sense of sigma-additive measures. In one direction this is clear. If $\mu_a \sim \mathbb{P}$ then necessarily $\mathbb{P}[A] > 0$ implies $\mu(A) \geq \mu_a(A) > 0$. Conversely if μ_a is not equivalent to $\mathbb{P}$ then there is a set D such that $\mu_a(D) = 0$ whereas $\mathbb{P}[D] > 0$. Since $\mu_s(B_n \cap D) = 0$ and since $\mathbb{P}[D] = \sum_n \mathbb{P}[D \cap B_n]$, we get that there must be at least one index n such that $\mathbb{P}[B_n \cap D] > 0$. But $\mu(D \cap B_n) = \mu_a(B_n \cap D) \leq \mu_a(D) = 0$.

Lemma 7 *Suppose that $\xi \in L^\infty$ and suppose there exists $\mu \in \mathbf{P^{ba}}$ such that $\mu \sim \mathbb{P}$ and $\mu(\xi) + c(\mu) = u(\xi)$. The element ξ is then minimal.*

Proof. Take $\eta \leq \xi$ and $\mathbb{P}[\eta < \xi] > 0$. We then have $\mu(\eta - \xi) < 0$ and hence $u(\eta) \leq \mu(\eta) + c(\mu) < \mu(\xi) + c(\mu) = u(\xi)$. □

Corollary 6 *Let u be a Fatou concave utility function given by the penalty function $c\colon \mathbf{P} \to \mathbb{R}_+ \cup \{+\infty\}$. Suppose that $u(\eta) = 0$ and suppose that there is $\varepsilon > 0$ and a sequence $\mathbb{Q}^n \in \mathbf{P}$ such that $\frac{d\mathbb{Q}^n}{d\mathbb{P}} \geq \varepsilon$ and $\mathbb{E}_{\mathbb{Q}^n}[\eta] + c(\mathbb{Q}_n) \to 0$. The element η is then minimal.*

Proof. Let $\mu \in \mathbf{P^{ba}}$ be adherent to the sequence $\mathbb{Q}_n$. For all $A \in \mathcal{F}$, μ satisfies $\mu(A) \geq \varepsilon \mathbb{P}[A]$. Therefore $\mu \sim \mathbb{P}$ in the sense of the lemma. □

Remark 32 We recall that even if u is relevant (but only concave and not coherent), this does not imply that there is a probability measure $\mathbb{Q}$ such that $\mathbb{E}_{\mathbb{Q}}[\xi] \geq 0$ for all $\xi \in \mathcal{A}$. The same remark can be made here. If $u(\eta) = 0$ and η is minimal, this does not imply that there is a probability measure $\mathbb{Q}$ such that $\mathbb{E}_{\mathbb{Q}}[\eta] + c(\mathbb{Q}) = 0$. In other words, minimal elements are not necessarily support points for hyperplanes given by probability measures. One way of making such examples goes as follows. Let $\mathcal{S} = \{\mathbb{Q} \mid \frac{d\mathbb{Q}}{d\mathbb{P}} \geq 1/2\}$ and let u be the coherent measure defined by $\mathcal{S}$. The previous corollary implies that every $\eta \in L^\infty$ is minimal. So take η with $u(\eta) = 0$ and define $u^1(\xi) = u(\xi + \eta)$. Because $\mathcal{S}$ is not weakly compact, James's theorem allows us to choose η in such a way that $u(\eta) = \inf_{\mathbb{Q} \in \mathcal{S}} \mathbb{E}_{\mathbb{Q}}[\eta]$ is an infimum that is not attained by some measure $\mathbb{Q} \in \mathcal{S}$. The concave utility function u^1 is Fatou, is relevant but there is no supporting hyperplane at $\xi = 0$.

Proposition 24 *The following are equivalent*

1. *Each* $\xi \in L^\infty$ *is minimal.*
2. *If* $\mu \in \mathbf{P^{ba}}$ *and* $\xi \in L^\infty$ *are such that* $\mu(\xi) + c(\mu) = u(\xi)$*, then* $\mu \sim \mathbb{P}$.

Proof. It follows from the lemma that the second condition implies the first. Suppose now that there is $\xi \in L^\infty$, $\mu(\xi) + c(\mu) = u(\xi)$ and μ is not equivalent to $\mathbb{P}$. By definition there exists A such that $\mathbb{P}[A] > 0$ and $\mu(A) = 0$. The variable $\eta = \xi + \mathbf{1}_A$ then satisfies: $u(\eta) \leq \mu(\eta) + c(\mu) = \mu(\xi) + c(\mu) = u(\xi)$. Consequently $u(\eta) = u(\xi), \xi \leq \eta, \mathbb{P}[\xi < \eta] > 0$ and η is not minimal. □

We conclude this section with some extension to relevance for concave monetary utility functions u, that are not necessarily Fatou. The reader who believes that these results are not relevant can skip these remarks. The results could have been proved in the beginning of this section and as the reader can check — as an exercise — they imply the results for the Fatou case. However we preferred to treat the case of Fatou utility measures first since the arguments are less complicated.

The basic ingredient is a minimax theorem of Ky Fan, [98], see also König, [88] for extensions and relations to ther minimax theorems. We write the theorem in a way that we can apply it directly. The monetary utility function u is supposed to be relevant and c denotes its Fenchel-Legendre transform or penalty function. The set $\mathbf{P^{ba}}$ is equipped with the weak* topology $\sigma(\mathbf{ba}, L^\infty)$, it is then a compact set. The function c is lower semi continuous for this topology. As we did in previous sections we will suppose that the probabibility space is atomless.

Theorem 27 *Let T be a convex set of lower semi continuous convex functions, defined on $\mathbf{P^{ba}}$ and taking values in $(-\infty, +\infty]$. Then there is $\mu^0 \in \mathbf{P^{ba}}$ with*

$$\sup_{f\in T} \min_{\mu\in\mathbf{P^{ba}}} f(\mu) = \sup_{f\in T} f(\mu^0).$$

We apply this theorem in the following way. For each $\xi \in L^\infty$ we define the function

$$f_\xi(\mu) = c(\mu) - \mu(\xi) = c(\mu) + \mu(-\xi).$$

For $\delta > 0$ the set $K_\delta = \{\xi \mid 1 \geq \xi \geq 0, \mathbb{E}[\xi] \geq \delta\}$ gives us a convex set of lower semi continuous functions T_δ. The functions f_ξ are clearly convex. Moreover for each ξ we get $\min_{\mu\in\mathbf{P^{ba}}}\{f_\xi)(\mu) = u(-\xi)$ and for elements in K_δ, the outcome is bounded away from 0 as the following lemma shows.

Lemma 8 *For each $\delta > 0$ we have*

$$\sup_{\xi\in K_\delta} u(\xi) < 0.$$

Proof of the lemma Suppose that there is a sequence η_n of elements in K_δ such that $u(-\eta_n) \to 0$. We may and do suppose that the sequence η_n converges $\sigma(L^\infty, L^1)$ to $\xi \in K_\delta$. By taking good convex combinations of $\eta_n, \eta_{n+1}, \ldots$, we then get a sequence $\xi_n \in K_\delta$ such that $\xi_n \to \xi$ in probability. By Egoroff's theorem we may take a subsequence (still denoted by ξ_n) and a set A of probability $\mathbb{P}[A] \geq 1 - \delta/2$. such that $\xi_n \mathbf{1}_A$ converges to $\xi\mathbf{1}_A$ in L^∞. Of course by monotonicity we still have $u(-\xi_n\mathbf{1}_A) \to 0$. But norm convergene then implies that also $u(-\xi\mathbf{1}_A) = 0$. Since u is relevant and since obviously $\mathbb{E}[\xi\mathbf{1}_A] \geq \delta - \delta/2 > 0$ we must have that $u(-\xi\mathbf{1}_A) < 0$, a contradiction. □

The minimax theorem now gives the existence of $\mu^\delta \in \mathbf{P^{ba}}$ such that

$$\sup_{\xi\in K_\delta} \left(c(\mu^\delta) - \mu^\delta(\xi)\right) < 0.$$

This implies that $\inf_{\xi\in K_\delta} \mu^\delta(\xi) > c(\mu^\delta)$. The meaure μ^δ is now split in its absolutely continuous part μ^δ_a and its purely discontinuous part μ^δ_p. We claim that

$$\inf_{\xi\in K_\delta} \mu^\delta(\xi) = \inf_{\xi\in K_\delta} \mu^\delta_a(\xi).$$

Indeed there is partition of Ω into pairwise disjoint sets A_n such that $\mu_p^\delta(A_n) = 0$ for each n. Let us define $B_n = \cup_{1\le k\le n} A_k$. On each set B_n we have that $\mu_a^\delta = \mu^\delta$. Because the probability space is atomless there is a set A, $\mathbb{P}[A] = \delta$ such that $\inf_{\xi\in K_\delta} \mu_a^\delta(\xi) = \mu_a^\delta[A]$. For each n we define $\eta_n = \mathbf{1}_{A\cap B_n} + \frac{\delta-\mathbb{P}[B_n\cap A]}{\mathbb{P}[B_n\setminus A]}\mathbf{1}_{B_n\setminus A}$. For n big enough we have that $\eta_n \le 1$ and we also have $\mathbb{E}[\eta_n] = \delta$. Since $\frac{\delta-\mathbb{P}[B_n\cap A]}{\mathbb{P}[B_n\setminus A]}$ tends to zero we have that

$$\begin{aligned}
\inf_{\xi\in K_\delta} \mu_a^\delta(\xi) &= \mu_a^\delta[A] \\
&\ge \mu_a^\delta(B_n\cap A) \\
&= \mu_a^\delta[\eta_n] - \frac{\delta-\mathbb{P}[B_n\cap A]}{\mathbb{P}[B_n\setminus A]}\mu_a^\delta[B_n\setminus A] \\
&= \mu^\delta[\eta_n] - \frac{\delta-\mathbb{P}[B_n\cap A]}{\mathbb{P}[B_n\setminus A]}\mu^\delta[B_n\setminus A] \\
&\ge \inf_{\xi\in K_\delta} \mu^\delta[\xi] - \frac{\delta-\mathbb{P}[B_n\cap A]}{\mathbb{P}[B_n\setminus A]}\mu^\delta[B_n\setminus A] \\
&\to \inf_{\xi\in K_\delta} \mu^\delta[\xi].
\end{aligned}$$

Since $\mu_a^\delta \le \mu^\delta$, this shows that $\inf_{\xi\in K_\delta} \mu_a^\delta(\xi) = \inf_{\xi\in K_\delta} \mu^\delta(\xi) > c(\mu^\delta) \ge 0$.

For each n we now use the previous construction to get a measure $\mu^{2^{-n}}$. We then take $\mu = \sum_n 2^{-n}\mu^{2^{-n}}$. The absolutely continuous part of μ is denoted by μ_a and it equals $\sum_n \mu_a^{2^{-n}}$. The convexity of c then shows that for $1\ge\xi\ge 0$, $\mathbb{E}[\xi] > 0$ we have

$$\mu_a(\xi) \ge c(\mu) + \Delta(\mathbb{E}[\xi]),$$

where Δ is a strictly positive function defined on $(0,1)$. In particular we get $\mu_a \sim \mathbb{P}$.

4.15 Ordering on utility functions, monotone convergence

Definition 19 *For $u_1, u_2\colon L^\infty \to \mathbb{R}$ two monetary utility functions, we say that $u_1 \le u_2$ if for all $\xi \in L^\infty$: $u_1(\xi) \le u_2(\xi)$.*

Proposition 25 *Suppose that $u_1, u_2\colon L^\infty \to \mathbb{R}$ are two monetary utility functions, defined by resp. the functions $c_1, c_2\colon \mathbf{P} \to \mathbb{R}_+ \cup \{\infty\}$. Then*

$u_1 \leq u_2$ if and only if $c_1 \leq c_2$. If both are coherent and defined by $\mathcal{S}_1^{\mathbf{ba}}, \mathcal{S}_2^{\mathbf{ba}}$, this is equivalent to $\mathcal{S}_1^{\mathbf{ba}} \supset \mathcal{S}_2^{\mathbf{ba}}$.

Proof Simply observe that for a concave monetary utility function and $\mu \in \mathbf{P}^{\mathbf{ba}}$:

$$c(\mu) = \sup\{u(\xi) - \mathbb{E}_\mu[\xi] \mid \xi \in L^\infty\}, u(\xi) = \inf\{\mathbb{E}_\mu[\xi] + c(\mu) \mid \mu \in \mathbf{P}^{\mathbf{ba}}\}.$$

Proposition 26 *Suppose that $u_n \colon L^\infty \to \mathbb{R}$ is a decreasing sequence of concave monetary utility functions. Let c_n be the corresponding sequence of penalty functions and let $\mathcal{A}_n$ be the corresponding sequence of acceptance sets. The limit $u(\xi) = \lim_n u_n(\xi)$ defines a concave monetary utility function, its acceptance set is given by $\mathcal{A} = \cap_n \mathcal{A}_n$. If every u_n is Fatou, then u is Fatou. The penalty function c of u satisfies $c(\mu) \leq c_n(\mu)$ but it is not necessarily equal to $\lim_n c_n(\mu)$.*

Proof Because the sequence $u_n(\xi)$ is bounded by $\|\xi\|_\infty$, the limit exists and is finite. The function u is clearly concave and monetary. Since $u(\xi) \geq 0$ if and only if for all n, $u_n(\xi) \geq 0$, we get that $\mathcal{A} = \cap_n \mathcal{A}_n$. The previous proposition shows the statement on the penalty functions. That teh function c can be different from $\lim_n c_n(\mu)$ is seen by the example $u_n(\xi) = -\frac{1}{\alpha_n}\mathbb{E}[\exp(-\alpha_n \xi)]$ where $\alpha_n \uparrow \infty$. The penalty functions are given by $c_n(\mathbb{Q}) = \frac{1}{\alpha_n}\mathbb{E}\left[\frac{d\mathbb{Q}}{\mathbb{P}}\log\left(\frac{d\mathbb{Q}}{\mathbb{P}}\right)\right]$. We have $c_n(\mathbb{Q})$ tends to zero or it stays equal to $+\infty$, whereas $c(\mathbb{Q}) = 0$ for every $\mathbb{Q}$. □

Remark 33 If we put $\overline{c}(\mu) = \lim_n c_n(\mu)$, then clearly

$$u(\xi) = \inf_n u_n(\xi) = \inf\{\mathbb{E}_\mu[\xi] + \overline{c}(\mu) \mid \mu \in \mathbf{P}^{\mathbf{ba}}\}.$$

Hence $\overline{c}$ also defines u but there is no guarantee that this function is lower semi continuous. The function c canbe obtained from $\overline{c}$ by closing the epigraph $\{(\alpha, \mu) \mid \alpha \geq \overline{c}(\mu)\}$ in $\overline{\mathbb{R}_+} \times \mathbf{P}^{\mathbf{ba}}$. The details are left as an exercise.

Remark 34 If all the functions u_n are coherent and defined by $\mathcal{S}_n^{\mathbf{ba}}$, then u is also coherent and is defined by the closure of $\cup_n \mathcal{S}_n^{\mathbf{ba}}$. The previous remark can be applied and translates as follows: the union $\cup_n \mathcal{S}_n^{\mathbf{ba}}$ is not necessarily closed. In case all the coherent functions are Fatou, then we can replace $\mathcal{S}_n^{\mathbf{ba}}$ by $\mathcal{S}_n = \mathcal{S}_n^{\mathbf{ba}} \cap \mathbf{P}$. We get that $\mathcal{S}$ is the closure (in $\mathbf{P}$) of $\cup_n \mathcal{S}_n$

Proposition 27 *Suppose that $u_n \colon L^\infty \to \mathbb{R}$ is an increasing sequence of concave monetary utility functions. Let c_n be the corresponding sequence of* *penalty functions. The limit $u(\xi) = \lim_n u_n(\xi)$ defines a concave monetary utility function. The penalty function c of u satisfies $c(\mu) = \lim_n c_n(\mu) = \sup_n c_n(\mu)$.*

Proof It is easily seen that u is a concave monetary utility function. And we have

$$c(\mu) = \sup\{u(\xi) - \mathbb{E}_\mu[\xi] \mid \xi \in L^\infty\} = \sup_n \sup\{u_n(\xi) - \mathbb{E}_\mu[\xi] \mid \xi \in L^\infty\}$$
$$= \sup_n c_n(\mu).$$

Exercise 15 Rephrase the previous proposition for coherent measures (general as well as Fatou) and give corresponding statements for their sets of scenarios.

Example 26 If every u_n is Fatou, then u is not necessarily Fatou. Take $\Omega = \mathbb{N}, \mathcal{F} = 2^{\mathbb{N}}$ and equipped with the probability measure $\mathbb{P}[n] = 2^{-n}$. Take $\mathcal{S}_n$ to be the set of all probabilities supported by the set $\{n+1, n+2, \ldots\}$. Clearly $u_n(\xi) = \inf_{k>n} \xi(k)$ and $u(\xi) = \liminf_k \xi(k)$. The function u is not Fatou and is given by the set $\mathbb{F}$ of purely finitely additive measures. Moreover $\cap_n \mathcal{S}_n = \emptyset, \cap_n \mathcal{S}_n^{\mathbf{ba}} = \mathbb{F}$. This remark almost contains a solution to the previous exercise.

4.16 Utility functions defined on bigger spaces

Later we will show that there is no possibility to define real valued concave utility functions for all random variables. But maybe that one can define good utility functions on smaller spaces than L^0. Recently, especially in modelling operational risk, Neslehova, Embrechts and Chavez-Demoulin,[106] got interested in spaces containing random variables that are Pareto distributed, or more generally with distributions having fat tails, so that the random variables are not integrable. The following shows that also on spaces that are smaller than L^0, there is no hope of finding a reasonable finitely valued risk measure or utility function. The space on which we will prove the impossibility theorem are solid and rearrangement invariant. These spaces include spaces such as L^p with $p < 1$.

Throughout this section we will assume that E satisfies

1. E is rearrangement invariant: if $\xi \in E$, if ξ and η have the same law or distribution then also $\eta \in E$.

2. E is solid, i.e. $\xi \in E$ and $|\eta| \leq |\xi|$ imply that $\eta \in E$

3. $E \supset L^\infty$.

Rearrangement invariant spaces satisfy the following stronger property.

Lemma 9 *Let $(A_t)_{t\in J\subset]0,1[}$ be an increasing family of sets such that $\mathbb{P}[A_t] = t$. Let ξ be an element of E. There exists a random variable η having the same distribution as ξ and such that on A_t we have $\eta \leq q_t(\xi)$.*

Proof. We complete the system as in Theorem 1. We get a uniformly $]0,1[$ distributed random variable γ such that $\{\gamma \leq t\} = A_t$. We then define $\eta = q_\gamma(\xi)$. The random variable η has the same law as ξ and satisfies all the desired properties. □

The function u satisfies the following properties, in other words u is what we call a monetary utility function.

1. $u\colon E \to \mathbf{R}$, $u(0) = 0$,

2. if $\xi \in E$ and $\xi \geq 0$ then $u(\xi) \geq 0$,

3. u is monetary, i.e. for $\xi \in E$ and $a \in \mathbf{R}$ we have: $u(\xi + a) = u(\xi) + a$,

4. u is concave.

Remark 35 We do not assume any continuity property except that u is nonnegative for nonnegative random variables. In particular we do not assume that u has the Fatou property and we do not require that u is law invariant (rearrangement invariant). We only require that u is defined on a rearrangement invariant, solid space.

We first prove that on these bigger space, the utility function remains in a special sense, monotone. This is the same as in Proposition 6 but this time we cannot use the boundedness of the random varables. So we need a more general proof.

Lemma 10 *The function u is monotone in the following sense. If $\xi \leq \eta$ are elements of E, if moreover $\eta \in L^\infty$, then $u(\xi) \leq u(\eta)$.*

Proof. We may suppose that $u(\xi) = 0$. It is then sufficient to show that $u(\eta) \geq 0$. Let $1 \geq \varepsilon > 0$ and let $\alpha \geq \max(2, 2\|\eta\|_\infty)/\varepsilon$. We claim that $\alpha(\eta - \xi + \varepsilon) + \xi \geq 0$. Indeed on the set $\{\xi \geq -2\|\eta\|_\infty\}$ we have $\alpha(\eta - \xi + \varepsilon) + \xi \geq 0$ since $\alpha\varepsilon \geq 2\|\eta\|_\infty$. On the set $\{\xi \leq -2\|\eta\|_\infty\}$ we have

$\alpha(\eta - \xi) \geq \alpha(-\|\eta\|_\infty - \xi) \geq -\xi$ since $\alpha \geq 2$. Since u is nonnegative for nonnegative random variables we find that $u(\alpha(\eta - \xi + \varepsilon) + \xi) \geq 0$. Since u is concave we then get for $0 \leq \lambda = \frac{1}{\alpha} \leq 1$: $u(\eta + \varepsilon) = u(\lambda(\alpha(\eta - \xi + \varepsilon) + \xi) + (1-\lambda)\xi) \geq \lambda u(\alpha(\eta - \xi + \varepsilon) + \xi) + (1-\lambda)u(\xi) \geq 0$. Since ε was arbitrary we proved $u(\eta) \geq 0$. □

Remark 36 The proof is a little bit curious. The fact that $\eta \in L^\infty$ seems to be needed. However if u is coherent (hence superadditive), then $u(\eta) \geq u(\eta - \xi) + u(\xi)$ would give a trivial proof.

We now use the representation theorem for monetary utility functions. This produces a function $c \colon \mathbf{P}^{\mathbf{ba}} \to \mathbb{R}_+ \cup \{\infty\}$ such that for all $\xi \in L^\infty$, $u(\xi) = \inf\{\mu(\xi) + c(\mu) \mid \mu \in \mathbf{P}^{\mathbf{ba}}\}$.

Remark 37 We remark that the representation theorem is only stated for bounded random variables. We do not claim any representation for unbounded elements of E.

Theorem 28 *Suppose that $E \setminus L^\infty \neq \emptyset$, then u satisfies*

1. *For $0 \leq k < \infty$ the set $\{\mu \in \mathbf{P}^{\mathbf{ba}} \mid c(\mu) \leq k\}$ is weakly compact in L^1. Hence $c(\mu) < \infty$ implies that μ is sigma-additive and is absolutely continuous with respect to $\mathbb{P}$, i.e. $\mu \in L^1$. Furthermore $0 = \min_{\mu \in \mathbf{P}} c(\mu)$.*

2. *u is continuous from below: for non-decreasing sequences $\xi_n \uparrow \xi$, uniformly bounded in L^∞, i.e. $\sup \|\xi_n\|_\infty < \infty$, we have $\lim u(\xi_n) = u(\xi)$. This implies the weaker property that u is continuous from above, i.e. has the Fatou property.*

Proof. Take k a real number $0 \leq k < \infty$ and suppose that the set $\{\mu \in \mathcal{P}^{\mathbf{ba}} \mid c(\mu) \leq k\}$ contains a measure that is not sigma-additive. The Yosida-Hewitt decomposition theorem, [135], allows us to write $\mu = \mu^a + \mu^s$ where $\mu^a \in L^1$ and μ^s is purely finitely additive. Moreover if $\mu^s \neq 0$, there is a decreasing sequence of sets, say $(A_n)_n$ such that $\mu^s(A_n) \geq \varepsilon > 0$ and $\mathbb{P}[A_n] \downarrow 0$. Let us now take $\xi \in E \setminus L^\infty$. We may suppose that $\xi \leq 0$ (since E is a solid vector space).
Let $\beta_n = \inf\{x \mid \mathbf{P}[\xi \leq x] \geq \mathbf{P}[A_n]\}$. Because ξ is unbounded we have that $\beta_n \to -\infty$. Now by rearrangement – see below – we may suppose that $\xi \leq \beta_n$ on the set A_n. Monotonicity as in the lemma above, implies $u(\xi) \leq u(\beta_n \mathbf{1}_{A_n})$. The representation theorem then implies that the latter term is bounded by $u(\beta_n \mathbf{1}_{A_n}) \leq \mu(\beta_n \mathbf{1}_{A_n}) + c(\mu) \leq \beta_n \mu^s(A_n) + k \leq \beta_n \varepsilon + k$.

Since β_n tends to $-\infty$ this would imply that $u(\xi) \le -\infty$, a contradiction to the hypothesis that u is real-valued. So we proved that $\mu^s = 0$ and consequently the set $\{\mu \in \mathbf{P^{ba}} \mid c(\mu) \le k\}$ is a weakly compact subset of L^1. An easy compactness argument shows that the infimum in $0 = \inf_{\mu \in \mathcal{P}^{\mathbf{ba}}} c(\mu)$ is now a minimum. The continuity from below is now a consequence of weak compactness. □

Theorem 29 *With the above notation we have $E \subset L^1$.*

Proof. We may suppose that there is $\xi \in E \setminus L^\infty$ since otherwise $E = L^\infty$ and the statement becomes trivial. Take $\mathbb{Q} \in L^1$ a probability measure such that $c(\mathbb{Q}) < \infty$, we can even take $c(\mathbb{Q}) = 0$ but it does not simplify the proof. The existence of such a sigma additive probability measure is guaranteed by the weak compactness property and $\inf\{c(\mu) \mid \mu \in \mathbf{P^{ba}}\} = 0$. Of course we have $u(\eta) \le \mathbf{E}_\mathbb{Q}[\eta] + c(\mathbb{Q})$ for any $\eta \in L^\infty$. By the monotonicity and the Beppo Levi theorem this inequality extends to nonpositive elements of E. For given $\xi \in E$, ξ unbounded, we have that $|\xi| \in E$ and we may by rearrangement, suppose that there is $\beta > 0$ such that $\{|\xi| \ge \beta\} \subset \{\frac{d\mathbb{Q}}{d\mathbb{P}} \ge 1/2\}$. The change of ξ to a rearrangement does not change the problem since rearrangements have the same integral under $\mathbb{P}$! We then find

$$-\infty < u(-|\xi|) \le \mathbb{E}_\mathbb{Q}[-|\xi|] + c(\mathbb{Q}).$$

In particular we find that $\mathbb{E}_\mathbb{Q}[|\xi|] < \infty$. This implies that $\mathbf{E}_\mathbb{P}[|\xi|\mathbf{1}_{\{|\xi|\ge\beta\}}] < \infty$. Hence $\mathbb{E}_\mathbb{P}[|X|] < \infty$. □

Remark 38 The above theorem is related to the automatic continuity theorem for positive linear functionals defined on ordered spaces. From the continuity of such functionals it is easily derived that the space E cannot be too big, see [18] and [31, 32] for a discussion. The difference with our result and the Namioka-Klee theorem is that we replaced the completeness assumption by the hypothesis that the space is rearrangement invariant. Together with the assumption that the space is solid, this is a convenient substitute to construct elements of E.

Example 27 Let u be coherent and defined on a space E (rearrangement invariant and solid) containing an unbounded random variable. The theorem then says that there is a weakly compact convex set $\mathcal{S}$ of probability measures $\mathbb{Q} \in L^1$ such that for $\xi \in L^\infty$: $u(\xi) = \min_{\mathbb{Q}\in\mathcal{S}} \mathbb{E}_\mathbb{Q}[\xi]$. There is no reason to believe that the same representation holds for all elements $\xi \in E$. As the following example shows this is related to (the failure of) a density property of L^∞ in the space E. Indeed let us recall some facts from chapter 2. The following conjugate Young functions: $\Phi(x) = (x+1)\log(x+1) - x, \Psi(y) = \exp(y) - y - 1$ define the following Orlicz spaces, see [93] for more details:

$$\begin{aligned} L^\Phi &= \{\xi \in L^0 \mid \mathbb{E}_\mathbb{P}[\Phi(|\xi|)] < \infty\}, \\ L^\Psi &= \{\xi \mid \text{ there is } \alpha > 0, \mathbb{E}_\mathbb{P}[\Psi(\alpha|\xi|)] < \infty\}, \\ L^{(\Psi)} &= \{\xi \mid \text{ for all } \alpha > 0, \mathbb{E}_\mathbb{P}[\Psi(\alpha|\xi|)] < \infty\}. \end{aligned}$$

The latter space is the closure of L^∞ in L^Ψ. It is clear that $L^{(\Psi)} \neq L^\Psi$, e.g. look at a random variable ξ that is exponentially distributed with density $\exp(-x)\mathbf{1}_{\{x>0\}}$. Furthermore L^Φ is the dual of $L^{(\Psi)}$ and L^Ψ is the dual of L^Φ. Take now ξ exponentially distributed and take $\mu \in (L^\Psi)^*$ so that $\mu \geq 0$, μ is zero on $L^{(\Psi)}$ and $\mu(\xi) \neq 0$. Since $\xi + L^\Psi_+$ is at a strictly positive distance from $L^{(\Psi)}$ (prove this as an exercise), the Hahn-Banach theorem gives the existence of such an element. We now define $u(\eta) = \mathbb{E}_\mathbb{P}[\eta] + \mu(\eta)$. The functional u defined on $E = L^\Psi$ is linear, positive and monetary. When restricted to $L^{(\Psi)}$ and hence to L^∞, it coincides with the expectation operator. But on $E = L^\Psi$ it is different, since $\mu(\xi) \neq 0$. This shows that the representation theorem does not hold for all elements of E. Of course the reason is that L^∞ is not dense in the space E and approximation by bounded random variables is not possible. We also remark that the utility function u defined on L^Ψ is not the extension defined in Section 4.9. Indeed this extension would give $\mathbb{E}_\mathbb{P}[\xi]$ for $\xi \in L^\infty - L^0_+$.

Chapter 5

Law Determined Monetary Utility Functions

The present section deals with three results. One is due to Jouini, Schachermayer and Touzi [84], and says that a concave utility function that only depends on the law of the random variable is necessarily Fatou. The second result due to Kusuoka [96], characterises these law invariant coherent measures. The third result due to Frittelli and Rosazza-Gianin, [72] characterises the convex law invariant risk measures.

5.1 The Fatou property

Definition 20 *A utility function $u\colon L^\infty \to \mathbb{R}$ is called law determined if $u(\xi) = u(\eta)$ as soon as ξ and η have the same distribution (or law).*

Remark 39 Such utility functions have been called *law invariant* but many researchers don't like this expression since it suggests that u does not change if the law of the variable changes, another expression is law equivariant. The following theorem was proved by Jouini-Schachermayer-Touzi [84]. Their proof uses that the probability space is standard. The present proof — almost the same as the one from [84] — does not use this assumption, see also [128].

Theorem 30 *Let $u\colon L^\infty \to \mathbb{R}$ be a quasi-concave utility function that is law-determined. Suppose that u satisfies the semi continuity property: for each $a \in \mathbb{R}$, the set $K = \{\xi \mid u(\xi) \geq a\}$ is norm closed. Then u has the Fatou property.*

Proof. The basic ingredient is the following result

Lemma 11 *Let the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *be atomless. Let* $(\xi_n)_{n\geq 1}$ *be a uniformly bounded sequence that converges in probability to the random variable* ξ. *Then*

1. *for each* n *there is a natural number* N_n *and there are random variables* $\xi_{n,1} \dots \xi_{n,N_n}$, *each* $\xi_{n,j}$ *having the same law as* ξ_n,
2. *the sequence* $\frac{1}{N_n}(\xi_{n,1} + \dots \xi_{n,N_n})$ *tends to* ξ *in* L^∞ *norm.*

Proof of the Theorem. We will prove the lemma later. For the moment let us show how it implies the theorem. We have to show that for each $\alpha \in \mathbb{R}$, the convex set $K = \{\eta \mid u(\eta) \geq \alpha\}$ is weak* closed in L^∞. By the Krein-Smulian theorem it is sufficient to show that uniformly bounded sequences in K that converge in probability, have a limit that is still in K. So let us suppose that for each n, $\xi_n \in K$, suppose that $\sup \|\xi_n\|_\infty < \infty$ and suppose that ξ_n tends to ξ in probability. We have to show that $u(\xi) \geq \alpha$ or what is the same $\xi \in K$. By the lemma we have the existence of convex combinations $\frac{1}{N_n}(\xi_{n,1} + \dots \xi_{n,N_n})$ — each $\xi_{n,j}$ having the same law as ξ_n — that converge to ξ in L^∞ norm. Since u is law-determined and quasi-concave we have that each $\xi_{n,j} \in K$ therefore also $\frac{1}{N_n}(\xi_{n,1} + \dots \xi_{n,N_n}) \in K$. Because K is norm closed we get $\xi \in K$ as required. □

Proof of the Lemma. There is no loss in generality to suppose that $\|\xi_n\|_\infty \leq 1$. We first replace ξ by an elementary function $\eta = \sum_{i=1}^k \alpha_i \mathbf{1}_{A_i}$ in such a way that $\|\xi - \eta\|_\infty \leq \varepsilon$, $\|\eta\|_\infty \leq 1$, $\min_i \mathbb{P}[A_i] = \delta > 0$ and the sets $A_i, i = 1, \dots k$ are disjoint. For n big enough, say $n \geq n_0(\varepsilon)$, and for each i, we will have that $\mathbb{P}[|\xi_n - \eta| > 2\varepsilon \mid A_i] \leq \varepsilon$. For such n we will now construct the variables $\xi_{n,1}, \dots, \xi_{n,N_n}$. These variables will be constructed on each A_i separately. The idea is to construct $\xi_{n,1}, \dots, \xi_{n,N_n}$ in such a way that on each A_i, the conditional distribution of $\xi_{n,j}$ is the same as the conditional distribution of ξ_n. Let us now fix N_n so that $\frac{1}{N_n} < \varepsilon$. Take k_n so that $\frac{k_n - 1}{N_n} \leq \varepsilon < \frac{k_n}{N_n}$. Remark that this implies that $\frac{k_n}{N_n} \leq 2\varepsilon$. Because $\mathbb{P}$ is atomless we can divide A_i in N_n disjoint sets $A_{i,j}$, each having the probability $\frac{1}{N_n}\mathbb{P}[A_i]$ and such that $\{|\xi_n - \alpha_i| > 2\varepsilon\} \cap A_i \subset \cup_{j=1}^{k_n} A_{i,j}$. This is possible since $\mathbb{P}[|\xi_n - \eta| > 2\varepsilon \mid A_i] \leq \varepsilon \mathbb{P}[A_i]$ and by the choice of k_n. The sets $A_{i,j}$ also depend on n but we drop this index to keep the notation simple. On each of the sets $A_{i,k_n+1}, \dots, A_{i,N_n}$ we have that $|\xi_n - \eta| \leq 2\varepsilon$. Let us now put $\xi_{n,1} = \xi_n$. For each $j \geq 2$, we use the cyclic permutation that maps 1 to j, 2 to $j+1$ etc, more precisely s is mapped to $\pi_j(s) = (s + j - 1)$ if $s \leq N_n - j + 1$ and to $\pi_j(s) = (s + j - 1 - N_n)$ if $s > N_n - j + 1$. For such

j we define $\xi_{n,j}$ so that on the set $A_{i,\pi_j(s)}$, the variable $\xi_{n,j}$ has the same conditional distribution as ξ_n has on the set $A_{i,s}$. This is possible because $\mathbb{P}$ is atomless. Because all the sets $A_{i,s}$ have the *same* probability, it follows that the variables $\xi_{n,j}$ have the same conditional distribution on A_i as ξ_n. This implies that the law of $\xi_{n,j}$ is the same as the law of ξ_n. Let us now look at the average $\frac{1}{N_n}(\xi_{n,1}+\ldots\xi_{n,N_n})$ on each set A_i. The difference between $\xi_{n,j}$ and η is for $N_n - k_n$ terms bounded by 2ε and for k_n terms it is bounded by $\|\xi_{n,j}-\eta\|_\infty \le 2$. This gives a bound

$$\left\|\frac{1}{N_n}(\xi_{n,1}+\ldots\xi_{n,N_n})-\eta\right\|_\infty \le \frac{N_n-k_n}{N_n}2\varepsilon+\frac{k_n}{N_n}2 < 2\varepsilon+4\varepsilon=6\varepsilon,$$

hence

$$\left\|\frac{1}{N_n}(\xi_{n,1}+\ldots\xi_{n,N_n})-\xi\right\|_\infty \le 7\varepsilon.$$

To finish the proof we continue in a standard way, using some diagonalisation argument. Let us sketch the details. For each ε of the form $\varepsilon = 1/k$ we get a number $n_0(\varepsilon) = n_0(1/k)$. We may suppose that this sequence is strictly increasing. For $n_0(1/k) \le n < n_0(1/(k+1))$ we perform the construction above. This ends the proof of the lemma. □

Remark 40 A careful analysis of the proof shows that the numbers N_n have to be big enough. This means that we can for each n, take N_n as big as we want. For the moment we do not see how to use this extra feature. The idea of the proof is of course the law of large numbers. The symmetry needed in the law of large numbers is taken over by the cyclic permutations. We could also have used the set of all permutations of $\{1,\ldots,N_n\}$, but the combinatorics are then a little bit more complicated: more counting is required.

5.2 A Representation of probability measures as nonincreasing functions

This section deals with some results of measure theory. These results are independent of the rest of the chapter. The results are probably known but for completeness we give proofs instead of leaving them as exercises.

Lemma 12 *Let* $\eta\colon (0,1] \to \mathbb{R}_+$ *be a nonincreasing function, then* $\eta'(x) = \lim_{y\downarrow x, y>x}\eta(y)$ *defines a right continuous function on* $(0,1)$*. We have* $\eta' = \eta$*,* *a.s. . If $\eta_1 = \eta$ a.s. and if η_1 is nonincreasing, then η_1 yields the same function η'.*

Exercise 16 The preceding lemma dealt with functions on $(0,1)$. Give a definition of "nonincreasing" that is adapted to classes of random variables (instead of functions). Prove that there is always a right continuous representative of such class.

Lemma 13 *The set of nonincreasing random variables $\eta\colon (0,1) \to \mathbb{R}_+$ such that $\int_{(0,1)} \eta(x)\,dx = 1$ forms a convex closed set $\mathbf{C}$, in $L^1[0,1]$. On this set the weak convergence and the strong convergence are the same. A subset $\mathbf{H} \subset \mathbf{C}$ is relatively weakly compact if and only if it is strongly relatively compact and this property is equivalent to: for all $\varepsilon > 0$ there is $\delta > 0$ such that*

$$\sup_{\eta \in \mathbf{H}} \int_{(0,\delta)} \eta \le \varepsilon.$$

Proof Because the elements of $\mathbf{C}$ are nonincreasing, the characterisation of relatively weakly compact sets as uniformly integrable sets (the Dunford-Pettis theorem), immediately yields the last claim. We still have to show that if $\eta_n \to \eta$, weakly in $\mathbf{C}$, then the convergence is also a norm convergence. By the preceding lemma we may suppose that η, η_n are right continuous. To show norm convergence it is then sufficient to show that $\eta_n \to \eta$ a.s. . Take $x \in (0,1)$ and suppose that x is a continuity point of η, we will show that $\eta_n(x) \to \eta(x)$. Because almost every point is a continuity point, this will complete the proof. For given $\varepsilon > 0$, take $\delta > 0$ such that $\eta(x+\delta) \ge \eta(x) - \varepsilon$ and $\eta(x-\delta) \le \eta(x) + \varepsilon$. By weak convergence $\int_{(x-\delta,x)} \eta_n \to \int_{(x-\delta,x)} \eta$. By monotonicity and by the choice of δ, we then get $\limsup \eta_n(x)\delta \le (\eta(x)+\varepsilon)\,\delta$. This proves $\limsup \eta_n(x) \le \eta(x) + \varepsilon$. In the same way, by integrating over the interval $(x, x+\delta)$ we prove $\liminf \eta_n(x) \ge \eta(x) - \varepsilon$. Because ϵ was arbitrary we have shown that $\lim \eta_n(x)$ exists and is equal to $\eta(x)$. □

Lemma 14 *If $\eta \in \mathbf{C}$ there is a probability measure ν on $(0,1]$ such that almost surely*

$$\eta(x) = \int_{(0,1]} \left(\frac{1}{a} \mathbf{1}_{[0,a)}(x) \right) \nu(da).$$

Conversely this formula associates withe every probability measure on $(0,1]$ a nonincreasing function $\eta \in \mathbf{C}$.

Proof We may and do suppose that η is right continuous. Because η is non-increasing, there is a $\sigma-$finite nonnegative measure μ on $(0,1]$ such that almost surely $\eta(x) = \mu\left((x,1]\right)$. We claim that $\int x\,\mu(dx) = 1$ so that $\nu(du) = u\,\mu(du)$ is a probability measure. This follows from Fubini's theorem, also called integration by parts. Indeed

$$\begin{aligned}\int_{(0,1]} x\,\mu(dx) &= \int_{(0,1]}\int_{(0,1]} \mathbf{1}_{\{u<x\}}\,du\,\mu(dx)\\ &= \int_{(0,1]} du\left(\int_{(0,1]} \mathbf{1}_{\{u<x\}}\mu(dx)\right) = \int_{(0,1]} du\,\eta(u) = 1.\end{aligned}$$

We can now write

$$\eta(x) = \int_{(0,1]} \mathbf{1}_{x<u}\mu(du) = \int_{(0,1]} \left(\frac{1}{u}\mathbf{1}_{(0,u)}(x)\right)\,\nu(du).$$

The converse is proved in the same calculations. □

Exercise 17 If we integrate the variable η and define $\phi(x) = \int_x^1 \eta(u)\,du$, we get a representation of nonnegative convex functions $\phi : [0,1] \to \mathbb{R}_+$, $\phi(0) = 1, \phi(1) = 0$ and with some little extra effort you get a representation of nonnegative convex functions $\phi : \mathbb{R}_+ \to \mathbb{R}_+$, $\phi(0) = 1, \lim_{x\to\infty}\phi(x) = 0$. The result is well known and used in probability theory in the theory of characteristic functions. Try e.g. $\phi(x) = \exp(-|x|), \phi(x) = \frac{1}{1+x^2}$.

Remark 41 The fanatic reader can check that the above relation can be understood by looking at the operator $T\colon C[0,1] \to C[0,1]$ defined as $(Tf)(a) = \frac{1}{a}\int_0^a f(x)\,dx$ for $a \neq 0$ and $(Tf)(0) = f(0)$. This operator is continuous, is not weakly compact, and its transpose is defined by the above relation.

Lemma 15 *Let $\mathcal{M}^1_+(0,1]$ be the set of all probability measures on $(0,1]$, equipped with the weak* topology induced by the continuous functions $C[0,1]$. With each probability measure $\nu \in \mathcal{M}^1_+(0,1]$ we associate the non-increasing function $\eta_\nu(x) = \int_{(0,1]}\left(\frac{1}{a}\mathbf{1}_{[0,a)}(x)\right)\,\nu(da)$. The mapping $\mathcal{M}^1_+(0,1] \to L^1$ is a homeomorphism between the space $\mathcal{M}^1_+(0,1]$ and the space $\mathbf{C}$, equipped with the norm topology.*

Remark 42 In probability theory this convergence is usually called weak convergence (for probability measures), it is the same topology as the one induced by the continuous and bounded functions on $(0,1]$, the continuous functions on $[0,1]$ zero at 0 or the functions on $[0,1]$ that are restrictions of smooth functions on $\mathbb{R}$. And there are many more ways to define this topology, see [20] for details and equivalences.

Proof The previous lemma already showed that the mapping is a bijection between $\mathcal{M}^1_+(0,1]$ and $\mathbf{C}$. The definition of the weak* topology shows that if $\nu^n \to \nu$, then almost surely $\eta_{\nu^n} \to \eta_\nu$. Since all the elements η_ν are nonnegative and have integral equal to 1, the almost sure convergence implies the convergence in L^1 (by Scheffe's lemma). The converse is proved in the following way. Let $\eta_n(x) = \int_{(0,1]} \left(\frac{1}{a}\mathbf{1}_{[0,a)}(x)\right) \nu_n(da)$ and suppose that $\eta_n \to \eta$ in L^1, where $\eta(x) = \int_{(0,1]} \left(\frac{1}{a}\mathbf{1}_{[0,a)}(x)\right) \nu(da)$. We have to show that $\nu_n \to \nu$ in the weak* topology for measures on $(0,1]$. For this it is sufficient to show that $\int f(a)\nu_n(da) \to \int f(a)\nu(da)$ for every smooth function f. But such a function f can be represented as $f(a) = \frac{1}{a}\int_0^a g(u)\,du$ where $g(x) = xf'(x) + f(x)$. In the language of remark 41 $f = Tg$. We get

$$\begin{aligned}\int f(a)\nu_n(da) &= \int \frac{1}{a}\int_0^a g(x)\,dx\,\nu_n(da)\\ &= \int g(x)\eta_n(x)\,dx \to \int g(x)\eta(x)\,dx = \int f(a)\nu(da).\end{aligned}$$

□

Corollary 7 *With the notation of the above proposition we get that compact sets H of $\mathbf{C}$ are in one-to-one correspondence with compact sets K of $\mathcal{M}_+(0,1]$. A set $K \subset \mathcal{M}_+(0,1]$ is weak* compact if and only if for all $\varepsilon > 0$ there is $\delta > 0$ such that* $\sup\{\nu((0,\delta]) \mid \nu \in K\} \le \varepsilon$.

Remark 43 What happens for the Dirac measure concentrated at 0? Let us see what happens with the sequence ν^n where ν^n is the Dirac measure concentrated at $1/n$. The corresponding function η^n is $n\mathbf{1}_{(0,1/n)}$ and this sequence is not converging in $L^1[0,1]$. The sequence $\int \eta^n(x)q_x\,dx$ – where q is the quantile function or increasing rearrangement of a random variable ξ – converges to ess.inf ξ. But we can write the definition of η in another way. With ν we associate the integral of η_ν, i.e. the function $H_\nu(x) = \int_{(0,x]} \eta_\nu(u)\,du$. The function H_ν is then the distribution function of probability measure with density η_ν. We can then extend the function H_ν to the measures defined on $[0,1]$. The Dirac measure then gets the function $H(x) = 1$ for all $x \in [0,1]$ whereas an arbitrary probability measure ν on $[0,1]$

gets the function $H_\nu(0) = \nu(\{0\})$ and $H_\nu(t) = H_\nu(0) + \int_{(0,1]} \nu(da)\frac{1}{a}\min(a,t)$ for $t > 0$. With a little bit of liberal thinking this can be written as $H_\nu(t) = \int_{[0,1]} \nu(da)\frac{1}{a}\min(a,t)$ for all $t \in [0,1]$.

Exercise 18 Analyse the continuity properties of the mapping $\mathcal{M}^1_+[0,1] \to \mathcal{M}^1_+[0,1]$ where the image of ν is the measure with distribution function H_ν. Compare with the remark on the operator T of remark 41.

5.3 Law Determined Utilities

Theorem 31 *Suppose that the probability space is atomless. Then the concave monetary utility function is law determined if and only if the penalty function is law determined, i.e. if $\frac{d\mathbb{Q}}{d\mathbb{P}}$ and $\frac{d\mathbb{Q}'}{d\mathbb{P}}$ have the same law, then $c(\mathbb{Q}) = c(\mathbb{Q}')$.*

Proof. Suppose that u is law determined. We need to show that if $\frac{d\mathbb{Q}}{d\mathbb{P}}$ and $\frac{d\mathbb{Q}'}{d\mathbb{P}}$ have the same law, $c(\mathbb{Q}) = c(\mathbb{Q}')$. We will give the proof when both $c(\mathbb{Q})$ and $c(\mathbb{Q}')$ are finite. The extension to the general case is done using the same idea and is left as an exercise. Fix $\varepsilon > 0$ and take λ with $u(\lambda) = 0, c(\mathbb{Q}) \le \mathbb{E}_\mathbb{Q}[-\lambda] + \varepsilon$. We now use the results of Chapter 2 on non atomic spaces. There are $[0,1]$ uniformly distributed random variables υ, υ_1 such that $\frac{d\mathbb{Q}}{d\mathbb{P}} = f \circ \upsilon, \lambda = l \circ \upsilon_1$ and where f, l are nondecreasing. Clearly we have $c(\mathbb{Q}) \le \mathbb{E}_\mathbb{Q}[-\lambda] + \varepsilon = \mathbb{E}_\mathbb{Q}[-l \circ (1-\upsilon)] + \varepsilon$. But because u is law determined we have $u(l \circ (1-\upsilon) = u(\lambda) = 0$. Because $\frac{d\mathbb{Q}}{d\mathbb{P}}$ and $\frac{d\mathbb{Q}'}{d\mathbb{P}}$ have the same law, we may write $\frac{d\mathbb{Q}'}{d\mathbb{P}} = f \circ \upsilon'$ where υ' is $[0,1]$ uniformly distributed. Of course $l \circ (1-\upsilon')$ has the same law as λ and hence can be used to estimate $c(\mathbb{Q}')$. We get $c(\mathbb{Q}') \ge \mathbb{E}_{\mathbb{Q}'}[-l \circ (1-\upsilon')] = \mathbb{E}_\mathbb{Q}[-l \circ (1-\upsilon)] \ge c(\mathbb{Q}) - \varepsilon$. This shows that $c(\mathbb{Q}') \ge c(\mathbb{Q})$ and by symmetry we get equality. The converse is proved along the same ideas. □

Corollary 8 *If the coherent utility u is given by the scenario set $\mathcal{S}$, then u is law determined if and only if $\mathbb{Q} \in \mathcal{S}$, $\frac{d\mathbb{Q}}{d\mathbb{P}}$ and $\frac{d\mathbb{Q}'}{d\mathbb{P}}$ have the same law, imply $\mathbb{Q}' \in \mathcal{S}$. In other words if and only if the set $\mathcal{S}$ is rearrangement invariant.*

The idea is to represent law determined utilities with quantiles or better with the family of TailVar-quantities. We will give the representation of concave utility functions. The result includes Kusuoka's representation, [96]. as well as the generalisation due to Frittelli and Rosazza-Gianin, [18]. Because the probability space is atomless, we are able to reduce the problem to the interval $[0,1]$. We fix a random variable υ, defined on $(\Omega,\mathcal{F},\mathbb{P})$ such that υ is uniformly distributed on the interval $[0,1]$. The random variable υ then defines an imbedding $\psi\colon L^\infty[0,1]\to L^\infty(\Omega), \psi(\xi)=\xi\circ\upsilon$. If u is a law determined monetary utility function on $L^\infty(\Omega)$, then $v(\xi)=u(\psi(\xi))=u(\xi\circ\upsilon)$ defines a law determined utility function on $L^\infty[0,1]$. If q_x denotes the quantile function of a random variable $\lambda\in L^\infty(\Omega)$, then $\psi(q)$ has the same law as λ, hence $v(q)=u(\lambda)$. If we can represent v then we can also represent u. Because v is Fatou we already have that there is a convex, lsc function $c\colon \mathbf{P}[0,1]\to\overline{\mathbb{R}_+}$ such that

$$v(\xi)=\inf\left\{\int_{[0,1]}\xi(x)\eta(x)\,dx+c(\eta)\mid \eta\in L^1[0,1],\eta\geq 0,\int_{[0,1]}\eta(x)\,dx=1\right\}.$$

Suppose now that $\xi\in L^\infty[0,1]$ is increasing, then it equals its quantile function. To recover $v(\xi)$ we may suppose that ξ is increasing and because $c(\eta)=c(\eta_*)$, we get

$$v(\xi)=\inf\left\{\int_{[0,1]}\xi(x)\eta(x)\,dx+c(\eta)\mid \eta\in\mathbf{P}[0,1],\eta \text{ non-increasing}\right\}.$$

Theorem 32 *If $u\colon L^\infty(\Omega)\to\mathbb{R}$ is a law determined, concave, monetary utility function then there is a convex, lsc function*

$$c\colon \mathcal{M}^1_+(0,1]\to\overline{\mathbb{R}_+},$$

such that

1. $\inf\{c(\nu)\mid \nu\in\mathcal{M}^1_+(0,1]\}=0$
2. *for all $\xi\in L^\infty$ we have*
$$u(\xi)=\inf\left\{\int\nu(d\alpha)u_\alpha(\xi)+c(\nu)\mid \nu\in\mathcal{M}^1_+(0,1]\right\},$$
where u_α represents the TailVar utility function at level $\alpha>0$:
3. *If u is coherent then c is the indicator function of a convex set $\mathcal{S}\subset\mathcal{M}^1_+(0,1]$ and we get*
$$u(\xi)=\inf\left\{\int\nu(d\alpha)u_\alpha(\xi)\mid \nu\in\mathcal{S}\right\}.$$

Proof The utility function $v = u \circ \psi$ has a penalty function c. Let us put for $\nu \in \mathcal{M}^1_+(0,1]$, $c(\nu) = c(\eta_\nu)$. Because of the continuity, proved above, this function c is lsc and is certainly convex. We have for $\xi \in L^\infty(\Omega)$:

$$\begin{aligned}
u(\xi) &= v(q(\xi)) \\
&= \inf\left\{\int q_x(\xi)\eta(x)\,dx + c(\eta) \mid \eta \in \mathbf{P}[0,1], \eta \text{ is decreasing}\right\} \\
&= \inf\left\{\int q_x(\xi)\int_{(0,1]} \frac{1}{a}\mathbf{1}_{(0,a)}(x)\nu(da)\,dx + c(\nu) \mid \nu \in \mathcal{M}^1_+(0,1]\right\} \\
&= \inf\left\{\int_{(0,1]}\left(\int q_x(\xi)\frac{1}{a}\mathbf{1}_{(0,a)}(x)\,dx\right)\nu(da) + c(\nu) \mid \nu \in \mathcal{M}^1_+(0,1]\right\} \\
&= \inf\left\{\int_{(0,1]} u_a(\xi)\,\nu(da) + c(\nu) \mid \nu \in \mathcal{M}^1_+(0,1]\right\}
\end{aligned}$$

If u is coherent then the penalty function is an indicator of a convex weak* closed convex set $\mathcal{S} \subset \mathcal{M}^1_+(0,1]$ and hence the expression can be simplified to

$$u(\xi) = \inf\left\{\int \nu(d\alpha)u_\alpha(\xi) \mid \nu \in \mathcal{S}\right\}.$$

□

Theorem 33 *If u is a law determined concave utility function, then there is a convex lsc function $\overline{c}\colon \mathcal{M}^1_+[0,1] \to \overline{\mathbb{R}_+}$ such that*

1. $\inf\{\overline{c}(\nu) \mid \nu \in \mathcal{M}^1_+[0,1]\} = 0$
2. *for all $\xi \in L^\infty$ we have*

$$u(\xi) = \inf\left\{\int \nu(d\alpha)u_\alpha(\xi) + \overline{c}(\nu) \mid \nu \in \mathcal{M}^1_+[0,1]\right\},$$

   where u_α represents the TailVar utility function at level $\alpha > 0$ and $u_0(\xi) = \text{ess.inf}\,\xi$.

3. *If u is coherent then $\overline{c}$ is the indicator function of a convex set $\overline{\mathcal{S}} \subset \mathcal{M}^1_+[0,1]$ and we get*

$$u(\xi) = \inf\left\{\int \nu(d\alpha)u_\alpha(\xi) \mid \nu \in \overline{\mathcal{S}}\right\}.$$

Proof We use the preceding theorem and we "close" the function c in the right way. The topology on the set $\mathcal{M}^1_+[0,1]$ is – as usual – the weak* topology induced by the functions $C[0,1]$, see [20]. Since c is lsc and convex, the graph of c:

$$\left\{(x,\nu) \mid x \in \mathbb{R}; \nu \in \mathcal{M}^1_+(0,1]; x \geq c(\nu)\right\}$$

is convex and closed. If we take the closure of this set in the space $\mathbb{R}_+ \times \mathcal{M}^1_+[0,1]$ we get a convex set, that defines a function $\overline{c}$. The restriction of $\overline{c}$ to $\mathcal{M}^1_+(0,1]$ is precisely c. If c was the indicator of a convex set $\mathcal{S}$, then $\overline{c}$ is the indicator of the closure $\overline{\mathcal{S}}$ of $\mathcal{S}$ in the compact space $\mathcal{M}^1_+[0,1]$. Since we put $u_0(\xi) = \text{ess.inf}\,\xi$, the functions $[0,1] \to \mathbb{R}, \alpha \to u_\alpha(\xi)$ are continuous. From this it follows that

$$u(\xi) = \inf\left\{\int \nu(d\alpha)u_\alpha(\xi) + c(\nu) \mid \nu \in \mathcal{M}^1_+[0,1]\right\}.$$

The rest is trivial. □

Exercise 19 Prove that the restriction of $\overline{c}$ to $\mathcal{M}^1_+(0,1]$ is indeed c (and not smaller). The main point in this exercise is to realise that something has to be proved.

Remark 44 Let us emphasize that $\overline{\mathcal{S}}$ was obtained via the set $\mathcal{S}$. But there is no guarantee that the set $\overline{\mathcal{S}}$ has a minimal property. Indeed we have

$$\begin{aligned}
\text{ess.inf}\,\xi &= \int_{[0,1]} q_x(\xi)\,\mu(dx) \text{ where } \mu \text{ is the Dirac measure concentrated at } 0\\
&= \inf_{\nu\in\mathcal{M}^1_+(0,1]} \int_{(0,1]} q_x(\xi)\,\nu(dx)\\
&= \inf_{\nu\in\mathcal{M}^1_+[0,1]} \int_{[0,1]} q_x(\xi)\,\nu(dx).
\end{aligned}$$

This can be explained as follows. We got the set $\mathcal{S}$ out of the representation theorem for a utility function, v, on $L^\infty[0,1]$. But to get a representation for the utility function, u, on $L^\infty(\Omega)$, we only needed the increasing (non-decreasing) elements of $L^\infty[0,1]$. Of course on this subset the representation of v can be given by more sets, $\mathcal{S}$, or for concave functions by more functions c. The function c is not uniquely defined.

For probability measures $\mu, \nu \in \mathcal{M}^1_+[0,1]$ we say that $\mu \succeq \nu$ if and only if $\int f\, d\mu \geq \int f\, d\nu$ for every nondecreasing continuous function f. This is saying that μ is more concentrated to the right than ν. Integration by parts immediately yields that the property is equivalent to $\mu[x,1] \geq \nu[x,1]$ for every $x \in [0,1]$. In insurance this ordering is called stochastic dominance of order one and it is one of the many orderings that can be defined using convex cones of "test functions". This procedure is well known in Choquet theory. In the following theorem the function c or the set $\mathcal{S}$ are obtained by the procedure above. The theorem shows that c or $\mathcal{S}$ have a certain maximality property.

Theorem 34 *Let u be law determined and given by the function $\overline{c}\colon \mathcal{M}^1_+[0,1]$, then $\mu \succeq \nu$ implies $c(\mu) \leq \overline{c}(\nu)$. If u would be law determined and coherent and given by the set $\mathcal{S} \subset \mathcal{M}^1_+[0,1]$, then $\nu \in \mathcal{S}$ implies that $\mu \in \mathcal{S}$.*

Proof For a measure ν on $[0,1]$, the function $\overline{c}(\nu)$ is obtained as

$$\overline{c}(\nu) = -\inf\left\{\int_{[0,1]} u_a(\xi)\,\nu(da) \mid \xi \in L^\infty(\Omega); u(\xi) \geq 0\right\}.$$

Since $u_a(\xi)$ is nondecreasing the result immediately follows □

Exercise 20 Prove the statement about $\overline{c}(\nu)$, check what happens for $a = 0$ and complete the proof.

5.4 Weak compactness property

Theorem 35 *Let $c\colon \mathcal{M}^1_+(0,1] \to \overline{\mathbb{R}_+}$ be the penalty function of a law determined utility function u. The function u satisfies the weak compactness property of Theorem 24 if and only if: for $k > 0$. $\varepsilon > 0$, there is $\delta > 0$ such that*

$$\sup\{\nu(0,\delta) \mid c(\nu) \leq k\} \leq \varepsilon.$$

In this case the set

$$\{\eta \mid \eta \in \mathbf{P}[0,1], \eta \text{ non-increasing }, c(\eta) \leq k\}$$

is compact in L^1. In the notation of the preceding section, this is the same as

$$\overline{c}(\nu) = +\infty \text{ if } \nu(\{0\}) > 0,$$

or for coherent utility functions

$$\nu(\{0\}) > 0 \text{ implies } \nu \notin \overline{\mathcal{S}}.$$

Remark 45 We warn the reader that this does not imply that the set $\{\eta \mid \eta \in \mathbf{P}[0,1], c(\eta) \leq k\}$ is compact for the norm topology. One can only deduce that this set is weakly compact in L^1 but this of course is already known since u satisfies the weak compactness property.

Proof If for $k > 0$. $\varepsilon > 0$, there is $\delta > 0$ such that

$$\sup\{\nu(0,\delta) \mid c(\nu) \leq k\} \leq \varepsilon,$$

we can conclude that $\{\nu \mid c(\nu) \leq k\}$ is weak* compact in $\mathcal{M}^1_+(0,1]$ (see Billingsley, [20]). The image set $\{\eta_\nu \mid c(\nu) \leq k$ is therefore compact in L^1 and hence uniformly integrable. The set $\{\eta \mid c(\eta) \leq k\}$ is therefore also uniformly integrable and hence weakly compact (since convex and closed). If the condition does not hold we have the existence of $k > 0$, $\varepsilon > 0$ as well as the existence of sequences $\nu^n \in \mathcal{M}^1_+(0,1]$ such that $c(\nu^n) \leq k$ and $\nu^n(0, \frac{1}{n}) \geq \varepsilon$. This implies

$$\begin{aligned}\int_{(0,\frac{1}{n})} \eta_{\nu^n}(x)\, dx &= \int_{(0,\frac{1}{n})} \left(\int \nu^n(da) \frac{1}{a} \mathbf{1}_{(0,a)}(x) \right) dx \\ &= \int \nu^n(da) \frac{1}{a} \min(a, \frac{1}{n}) \geq \nu^n(0, \frac{1}{n}) \geq \varepsilon.\end{aligned}$$

The sequence η_{ν^n} cannot be uniformly integrable and the weak compactness property is not fulfilled. □

Theorem 36 *Let u be a coherent law determined utility function satisfying the weak compactness property. Then u has the following representation:*

$$u(\xi) = \inf \left\{ \int \nu(d\alpha) u_\alpha(\xi) \mid \nu \in \mathcal{S} \right\},$$

where the set $\mathcal{S}$ is weak compact in $\mathcal{M}^1_+(0,1]$. This statement is equivalent to $\lim_{\delta \to 0} \sup_{\nu \in \mathcal{S}} \nu(0,\delta) = 0$ or to $\mathcal{S}$ is closed in $\mathcal{M}^1_+[0,1]$.*

Proof The proof follows immediately from the previous theorem since c is the indicator function of $\mathcal{S}$. □

Chapter 6

Operations on utility functions

6.1 Minimum of two coherent utility functions.

Let u_1 and u_2 be two coherent utility functions. Just to give an interpretation, they could stand for two different measures of risk calculated as $\rho_1 = -u_1$ and $\rho_2 = -u_2$. One of the utility functions (or risk measure) could be the manager's or the supervisor's utility, the other the shareholder's. If both groups must be pleased (a new phenomenon in management behaviour), it is natural to ask for a risk measure which is more severe than each of the two, that is:

$$\rho \equiv \rho_1 \vee \rho_2$$

We leave it to the reader to check that ρ is indeed a coherent risk measure that also satisfies the Fatou property if ρ_1 and ρ_2 do. Taking the max of two risk measures is the same as taking the minimum of the coherent utility functions. For simplicity we only treat the more complicated case of utility functions that have the Fatou property.

If we call $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}$ the accceptance sets (the first describing u_1, the second u_2 and the third $u = \min(u_1, u_2)$ respectively) and we define $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S}$ to be the related families of probabilities, we have:

$$\begin{aligned} \mathcal{A} &= \mathcal{A}_1 \cap \mathcal{A}_2 \\ \mathcal{S} &= conv\,(\mathcal{S}_1, \mathcal{S}_2) \end{aligned}$$

Actually:

$$\mathcal{A} = \{\xi \mid u(\xi) \geq 0\} = \{\xi \mid u_1(\xi) \geq 0 \,\text{and}\, u_2(\xi) \geq 0\} = \mathcal{A}_1 \cap \mathcal{A}_2.$$

Since the acceptance set characterises the risk measure, we can find the corresponding set $\mathcal{S}$:

$$\begin{aligned} \xi \in \mathcal{A} &\Leftrightarrow \xi \in \mathcal{A}_1 \quad \text{and} \quad \xi \in \mathcal{A}_2 \\ &\Leftrightarrow \forall \mathbb{Q}_1 \in \mathcal{S}_1 : \ \mathbb{Q}_1[\xi] \geq 0 \text{ and } \forall \mathbb{Q}_2 \in \mathcal{S}_2 : \ \mathbb{Q}_2[\xi] \geq 0 \\ &\Leftrightarrow \forall \mathbb{Q} \in \mathcal{S} : \quad \mathbb{Q}[\xi] \geq 0 \end{aligned}$$

We still have to prove that $\mathcal{S} = conv(\mathcal{S}_1, \mathcal{S}_2)$ is closed.

Let $(\mathbb{Y}_n)_n$ be a sequence in $\mathcal{S}$ converging in L^1-norm to a certain $\mathbb{Y}$ (remember that we identify probabilities with their densities). By definition there exist $\mathbb{P}_n \in \mathcal{S}_1$ and $\mathbb{Q}_n \in \mathcal{S}_2$ and $t_n \in [0,1]$ such that $\mathbb{Y}_n = t_n\mathbb{P}_n + (1-t_n)\mathbb{Q}_n$. We may suppose that $t_n \to t \in [0,1]$ (if not, take a converging subsequence).

There are now two possible cases:
a) if t_n or $1-t_n$ tends to 0, then we have either $\mathbb{Q}_n \to \mathbb{Y}$ or $\mathbb{P}_n \to \mathbb{Y}$ and then $\mathbb{Y} \in \mathcal{S}_1$ or $\mathbb{Y} \in \mathcal{S}_2$.
b) $0 < t < 1$. By dropping a finite number of terms, we may suppose that there is a number $c \in (0,1)$ such that $c \leq t_n \leq 1-c$. Now:

$$\mathbb{P}_n[A] \leq \frac{1}{t_n}\mathbb{Y}_n[A] \leq \frac{1}{c}(\mathbb{Y}_n[A])$$

and therefore the sequence $\left(\frac{d\mathbb{P}_n}{d\mathbb{P}}\right)_{n\geq 1}$ is dominated by the strongly convergent sequence $\left(\frac{d\mathbb{Y}_n}{d\mathbb{P}}\right)_n$. It is therefore uniformly integrable and hence a relatively weakly compact sequence. We may, by selecting a subsequence, suppose that $\mathbb{P}_n \to \mathbb{P}_0$ weakly $\sigma(L^1, L^\infty)$ and since $\mathcal{S}_1$ is convex closed, we have $\mathbb{P}_0 \in \mathcal{S}_1$. Similarly we get $\mathbb{Q}_0 \in \mathcal{S}_2$. Finally $\mathbb{Y} = t\mathbb{P}_0 + (1-t)\mathbb{Q}_0$ belongs to $conv(\mathcal{S}_1, \mathcal{S}_2)$.
□

6.2 Minimum of concave utility functions

We will now do the same analysis for two concave utility functions, u_1, u_2. They define resp. the acceptance sets $\mathcal{A}_1, \mathcal{A}_2$. The new set is $\mathcal{A} = \mathcal{A}_1 \cap \mathcal{A}_2$. The set $\mathcal{A}$ is weak* closed if both sets $\mathcal{A}_1, \mathcal{A}_2$ are weak* closed. In general $\mathcal{A}$ is norm closed as the intersection of norm closed sets. But we will only do the Fatou case, the other case being similar – up to small topological technicalities which we leave (with pleasure) to the reader. The convex penalty functions are resp. c_1, c_2. So we have $\xi \in \mathcal{A}$ if and only if for all $\mathbb{Q}$: $\min(c_1(\mathbb{Q}), c_2(\mathbb{Q})) + \mathbb{E}_\mathbb{Q}[\xi] \geq 0$. However the function $\min(c_1(\mathbb{Q}), c_2(\mathbb{Q}))$ is not convex. In the previous section we also had to replace the union of the two sets $\mathcal{S}_1, \mathcal{S}_2$ by their convex hull. In this case we replace the function $\min(c_1(\mathbb{Q}), c_2(\mathbb{Q}))$ by the largest convex function, smaller than both $c_1(\mathbb{Q})$ and $c_2(\mathbb{Q})$). This function is constructed as follows. The epigraph is simply the convex hull of the union of the two epigraphs, $\{(\mathbb{Q}, \alpha) \mid \alpha \in \mathbb{R}; \alpha \geq$

$c_i(\mathbb{Q})\}$. So we look at the set

$$\mathcal{C} = \{\lambda(\mathbb{Q}^1, \alpha^1) + (1-\lambda)(\mathbb{Q}^2, \alpha^2) \mid \alpha^i \geq c_i(\mathbb{Q}^i), 0 \leq \lambda \leq 1\}.$$

We proceed as in the previous section and show that this set is closed. Let $0 \leq \lambda_n \leq 1$ and let $(\mathbb{Q}^1_n, \alpha^1_n), (\mathbb{Q}^2_n, \alpha^2_n)$ be selected in the epigraphs of resp c_1, c_2. Furthermore let $\lambda_n \mathbb{Q}^1_n + (1-\lambda_n)\mathbb{Q}^2_n \to \mathbb{Q}$ and $\lambda_n \alpha^1_n + (1-\lambda_n)\alpha^2_n \to \alpha$. We have to show that $(\mathbb{Q}, \alpha) \in \mathcal{C}$. We may – eventually we take a subsequence – suppose that $\lambda_n \to \lambda$. If $\lambda = 0$ or 1 things are easy. Say $\lambda = 1$, the other case is similar. Then

$$\begin{aligned}
\alpha = \alpha_1 &= \lim(\lambda_n \alpha^1_n + (1-\lambda_n)\alpha^2_n) \\
&\geq \liminf(\lambda_n c_1(\mathbb{Q}^1_n) + (1-\lambda_n)c_2(\mathbb{Q}^2_n)) \\
&\geq \liminf \lambda_n c_1(\mathbb{Q}^1_n) \\
&\geq c_1(\mathbb{Q}^1) = c_1(\mathbb{Q}) \text{ since } \mathbb{Q} = \mathbb{Q}^1.
\end{aligned}$$

In case $0 < \lambda < 1$ we get – as in the previous section – that the sequences $\mathbb{Q}^1_n, \mathbb{Q}^2_n$ are uniformly integrable. So we may suppose that they converge weakly to resp. $\mathbb{Q}^1, \mathbb{Q}^2$. Of course $\mathbb{Q} = \lambda\mathbb{Q}^1 + (1-\lambda)\mathbb{Q}^2$. We remark that the penalty functions, being convex, are also lower semi continuous for the weak topology $\sigma(L^1, L^\infty)$ – this is a good exercise. So we now get

$$\begin{aligned}
\alpha &= \lim \lambda_n \alpha^1_n + (1-\lambda_n)\alpha^2_n \\
&\geq \liminf(\lambda_n c_1(\mathbb{Q}^1_n) + (1-\lambda_n)c_2(\mathbb{Q}^2_n)) \\
&\geq \liminf \lambda_n c_1(\mathbb{Q}^1_n) + \liminf(1-\lambda_n)c_2(\mathbb{Q}^2_n) \\
&\geq \lambda c_1(\mathbb{Q}^1) + (1-\lambda)c_2(\mathbb{Q}^2).
\end{aligned}$$

This proves that $(\mathbb{Q}, \alpha) \in \mathcal{C}$ as desired. Analytically we define

$$\begin{aligned}
c(\mathbb{Q}) &= \inf\{\alpha \mid (\mathbb{Q}, \alpha) \in \mathcal{C}\} \\
&= \inf\{\lambda c_1(\mathbb{Q}^1) + (1-\lambda)c_2(\mathbb{Q}^2) \mid \mathbb{Q} = \lambda\mathbb{Q}^1 + (1-\lambda)\mathbb{Q}^2\}.
\end{aligned}$$

Since $\mathcal{C}$ is closed and is the epigraph of c, we have that c is lower semi continuous and convex. Of course $\inf_{\mathbb{Q}} c(\mathbb{Q}) \leq \min(\inf_{\mathbb{Q}} c_1(\mathbb{Q}), \inf_{\mathbb{Q}} c_2(\mathbb{Q})) = 0$. So it defines a utility function. To see that it defines the utility function u, we show that $c(\mathbb{Q}) = \sup\{\mathbb{E}_{\mathbb{Q}}[-\xi] \mid \xi \in \mathcal{A}\}$. This is almost the definition of c.

$$\begin{aligned}\xi \in \mathcal{A} \Leftrightarrow & \xi \in \mathcal{A}_1 \text{ and } \xi \in \mathcal{A}_2 \\ \Leftrightarrow & \text{ for all } \mathbb{Q}^1, \mathbb{Q}^2 : \mathbb{E}_{\mathbb{Q}^1}[\xi] + c_1(\mathbb{Q}^1) \geq 0 \text{ and } \mathbb{E}_{\mathbb{Q}^2}[\xi] + c_2(\mathbb{Q}^1) \geq 0 \\ \Leftrightarrow & \text{ for all } 0 \leq \lambda \leq 1 \text{ for all } \mathbb{Q}, \mathbb{Q}^1, \mathbb{Q}^2 \text{ with } \mathbb{Q} = \lambda \mathbb{Q}^1 + (1-\lambda)\mathbb{Q}^2 : \\ & \quad \mathbb{E}_{\mathbb{Q}}[\xi] + \lambda_n c_1(\mathbb{Q}^1_n) + (1-\lambda_n) c_2(\mathbb{Q}^2_n) \geq 0 \\ \Leftrightarrow & \text{ for all } \mathbb{Q} : \mathbb{E}_{\mathbb{Q}}[\xi] + c(\mathbb{Q}) \geq 0.\end{aligned}$$

Remark 46 In convex analysis, the construction above is known as the inf-convolution or convex convolution, see [118]. It is denoted $c = c_1 \Box c_2$. More on this in the next section.

6.3 Inf Convolution of coherent utility functions

With the obvious notation, if u_1 and u_2 are given coherent utility functions, both having the Fatou property with their corresponding sets: $\mathcal{A}_1$, $\mathcal{S}_1$, $\mathcal{S}_1^{\mathbf{ba}}$, $\mathcal{S}_1 = \mathcal{S}_1^{\mathbf{ba}} \cap L^1$ and $\mathcal{A}_2$, $\mathcal{S}_2$, $\mathcal{S}_1^{\mathbf{ba}}$, $\mathcal{S}_2 = \mathcal{S}_2^{\mathbf{ba}} \cap L^2$. We can construct other utility functions by taking $\mathcal{S} = \mathcal{S}_1 \cap \mathcal{S}_2$, $\mathcal{S}_0^{\mathbf{ba}} = \mathcal{S}_1^{\mathbf{ba}} \cap \mathcal{S}_2^{\mathbf{ba}}$ or by taking $\mathcal{A} = \overline{conv\,(\mathcal{A}_1, \mathcal{A}_2)}^{\sigma(L^\infty, L^1)} = \overline{\mathcal{A}_1 + \mathcal{A}_2}^{\sigma(L^\infty, L^1)}$ or even $\mathcal{A}_0 = \overline{conv\,(\mathcal{A}_1, \mathcal{A}_2)}^{\|\,\|_\infty} = \overline{\mathcal{A}_1 + \mathcal{A}_2}^{\|\,\|_\infty}$. The closure is either taken in the norm topology or in the weak* topology $\sigma(L^\infty, L^1)$. If we take the closure in the norm topology we only get a coherent utility function. If we take the closure in the weak* topology we get a coherent utility function with the Fatou property. We will study both cases and relate them to a familiar construction from convex analysis. We will also show that both constructions can be different. We first show that $\mathcal{A}$ and $\mathcal{S}$ correspond:

Proposition 28 *$\mathcal{A}$ and $\mathcal{S}$ correspond, i.e. $\mathcal{S} = \mathcal{S}_{\mathcal{A}}$, where*

$$\begin{aligned}\mathcal{S}_{\mathcal{A}} &= \{\mathbb{Q} \mid \mathbb{Q} \ll \mathbb{P} \text{ a probability such that for all } \xi \in \mathcal{A} : \mathbb{E}_{\mathbb{Q}}[\xi] \geq 0\} \\ &= \{\mathbb{Q} \mid \mathbb{Q} \ll \mathbb{P} \text{ a probability such that for all } \xi \in \mathcal{A}_1 + \mathcal{A}_2 : \mathbb{E}_{\mathbb{Q}}[\xi] \geq 0\}\end{aligned}$$

The coherent utility function constructed from $\mathcal{A}$ (or what is the same from $\mathcal{S}$) is denoted by u. The utility function u satisfies the Fatou property.

Proof. We first show that $\mathcal{S} \supset \mathcal{S}_{\mathcal{A}}$. If $\xi \notin \mathcal{A}$ then by the Hahn Banach theorem (remember that the dual space of L^∞ with the weak* topology is exactly L^1) there exists an $f \in L^1$ such that $\mathbb{E}[f\xi] < 0$ and $\mathbb{E}[f\eta] \geq 0$ for every $\eta \in \mathcal{A}$. Since $\mathcal{A}$ contains $\mathbf{1}_A$ for every $A \in \mathcal{F}$, f will be nonnegative a.s. . Now, f can be assumed to be normalized, so we have obtained a $\mathbb{Q} \in \mathcal{S}_1 \cap \mathcal{S}_2$, $d\mathbb{Q} = f\,d\mathbb{P}$, which is strictly negative on ξ. Now we show that $\mathcal{S} \subset \mathcal{S}_{\mathcal{A}}$. If $\xi \in \mathcal{A}$ we have to prove that for every $\mathbb{Q} \in \mathcal{S}$: $\mathbb{E}_{\mathbb{Q}}[\xi] \geq 0$. Let us start with $\xi \in \mathit{conv}\,(\mathcal{A}_1, \mathcal{A}_2) = \mathcal{A}_1 + \mathcal{A}_2$, where the equality holds because the $\mathcal{A}_i$ are convex cones. Then if $\mathbb{Q} \in \mathcal{S}$, $\mathbb{Q}$ belongs to both $\mathcal{S}_i$ and taking into account that ξ can be written as $\eta + \zeta$ with $\eta \in \mathcal{A}_1$ and $\zeta \in \mathcal{A}_2$, we have that $\mathbb{E}_{\mathbb{Q}}[\xi] \geq 0$. Rewritten, this means that $0 \leq \int \xi \frac{d\mathbb{Q}}{d\mathbb{P}}\,d\mathbb{P}$ for every $\xi \in \mathit{conv}\,(\mathcal{A}_1, \mathcal{A}_2)$ and for every $\mathbb{Q} \in \mathcal{S}$. By fixing $\mathbb{Q}$, the set $\{\eta \in L^\infty \mid \mathbb{E}_Q[\eta] \geq 0\}$ is weak* closed and contains $\mathit{conv}\,(\mathcal{A}_1, \mathcal{A}_2)$: therefore it contains the weak* closure of the latter set, that is, it contains $\mathcal{A}$. □

Remark 47 In case $\mathcal{S}_1 \cap \mathcal{S}_2 = \emptyset$ we get that $\mathcal{A}_1 + \mathcal{A}_2$ is dense in L^∞ for the weak* topology. We will see by an example that this does not imply that $L^\infty = \mathcal{A}_1 + \mathcal{A}_2$. Throughout this section the function u will always be constructed using the set $\mathcal{S}_1 \cap \mathcal{S}_2$. Of course it only makes sense if $\mathcal{S}_1 \cap \mathcal{S}_2$ is nonempty.

Proposition 29 *$\mathcal{A}_0$ and $\mathcal{S}_0^{\mathbf{ba}}$ correspond, i.e.*

$$\begin{aligned}\mathcal{S}_0^{\mathbf{ba}} &= \{\mu \mid \mu \in \mathbf{P}^{\mathbf{ba}} \text{ such that for all } \xi \in \mathcal{A}_0 : \mu(\xi) \geq 0\} \\ &= \{\mu \mid \mu \in \mathbf{P}^{\mathbf{ba}} \text{ such that for all } \xi \in \mathcal{A}_1 + \mathcal{A}_2 : \mu(\xi) \geq 0\}.\end{aligned}$$

The coherent utility function constructed from $\mathcal{A}_0$ (or what is the same from $\mathcal{S}_0^{\mathbf{ba}}$) is denoted by u_0.

Proof. . The proof is a copy of the proof for the Fatou case. The difference lies in the fact that the set $\mathcal{A}_0$ is only norm closed and therefore we can only work with elements of **ba**. □

Remark 48 The reader can check that if $\mathcal{A} \subset L^\infty$ is a cone such that $\mathcal{A} \supset L^\infty_+$, then the two expressions

$$\sup\{\alpha \mid \xi - \alpha \in \mathcal{A}\}$$

and

$$\sup\{\alpha \mid \xi - \alpha \in \overline{\mathcal{A}}^{\|\cdot\|_\infty}\},$$

(meaning that we take the closure of $\mathcal{A}$ for the L^∞ norm), are the same. This is easily seen because $\xi - \alpha \in \overline{\mathcal{A}}^{\|\cdot\|_\infty}$ implies that for every $\varepsilon > 0$, we have that $\xi - \alpha + \varepsilon \in \mathcal{A}$. We also remark that, even for finite Ω (starting with 3 points), the set $\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2$ need not be closed.

Proposition 30 *Suppose $\mathcal{S}_1 \cap \mathcal{S}_2 \neq \emptyset$. Let $\bar{u}$ be a coherent utility function having the Fatou property and let it be bigger than u_1 and u_2, then $\bar{u} \geq u$.*

Proof. . Let $\bar{u}$ be given by $\overline{\mathcal{S}}$. Then $\overline{\mathcal{S}} \subset \mathcal{S}_1$ and $\overline{\mathcal{S}} \subset \mathcal{S}_2$, because $\bar{u} \geq u_1$ and $\bar{u} \geq u_2$. Therefore $\overline{\mathcal{S}} \subset \mathcal{S}_1 \cap \mathcal{S}_2$ and hence $\bar{u} \geq u$. $\square$

Remark 49 Of course the proposition does not make sense if $\mathcal{S}_1 \cap \mathcal{S}_2 = \emptyset$. In this case there will be no coherent utility function that satisfies the Fatou property and that is bigger than both u_1 and u_2. This is the same as saying that $\mathcal{A}_1 + \mathcal{A}_2$ is dense in L^∞ for the weak* topology.

If we would like to define a coherent utility function (not necessarily having the Fatou property) $\tilde{u}$, with the property that it is the smallest coherent utility function such that $\tilde{u} \geq u_1 \vee u_2$, we can take a similar construction as in the previous section:

$$
\begin{aligned}
&\tilde{u}(\xi) \\
&= \sup\{t u_1(\xi_1) + (1-t) u_2(\xi_2) \mid \xi = t\xi_1 + (1-t)\xi_2;\ \xi_1, \xi_2 \in L^\infty;\ 0 \leq t \leq 1\} \\
&= \sup\{u_1(t\xi_1) + u_2((1-t)\xi_2) \mid \xi = t\xi_1 + (1-t)\xi_2;\ \xi_1, \xi_2 \in L^\infty;\ 0 \leq t \leq 1\} \\
&= \sup\{u_1(\eta) + u_2(\xi - \eta) \mid \eta \in L^\infty\}
\end{aligned}
$$

The construction can be explained by looking at the hypographs of u_1 and u_2, i.e. the sets $\{(\xi, \alpha) \mid \alpha \leq u_i(\xi)\}$. The hypograph $\{(\xi, s) \mid \tilde{u}(\xi) \geq s\}$ is constructed as the closed convex hull of the hypographs of u_1 and u_2. This utility function is usually denoted by $u_1 \Box u_2$ and it is called the convex (or should we say concave?) convolution of u_1 and u_2 In convex function theory,[118], this convolution is also referred to as the inf-convolution or infimal convolution. For concave functions it should then be called the sup-convolution but we are not so fundamentalist on these nomenclature. The convex convolution of coherent utility functions can be characterised using the duality $(L^\infty, \mathbf{ba})$. It is an easy exercise to see that the coherent utility function $\tilde{u}$ is given by $\mathcal{S}_1^{\mathbf{ba}} \cap \mathcal{S}_2^{\mathbf{ba}}$ and hence it is equal to u_0. From the previous proposition we conclude that $u_1 \Box u_2$ has the Fatou property if and only if the following holds (where the bar indicates $\sigma(\mathbf{ba}, L^\infty)$ closure):

$$
\overline{\mathcal{S}_1^{\mathbf{ba}} \cap \mathcal{S}_2^{\mathbf{ba}} \cap L^1} = \mathcal{S}_1^{\mathbf{ba}} \cap \mathcal{S}_2^{\mathbf{ba}}\,.
$$

This is equivalent to: $\overline{\mathcal{S}_1 \cap \mathcal{S}_2} = \mathcal{S}_1^{\mathbf{ba}} \cap \mathcal{S}_2^{\mathbf{ba}}$, where again, the bar indicates $\sigma(\mathbf{ba}, L^\infty)$ closure. So we get:

Proposition 31 *u and $u_0 = u_1 \Box u_2$ coincide if and only if $u_1 \Box u_2$ has the Fatou property. This is the case when for instance $\mathcal{S}_1$ (or $\mathcal{S}_2$) is weakly compact. xxxx include 2019 result with Orihuela!*

Remark 50 If we would write the coherent functions with their penalty functions we get

1. u_0 is represented by the function $c_0 \colon \mathbf{P}^{\mathbf{ba}} \to \overline{\mathbb{R}_+} : c(\mu) = c_1(\mu) + c_2(\mu)$,
2. u is represented by the function $c \colon \mathbf{P} \to \overline{\mathbb{R}_+} : c(\mathbb{Q}) = c_1(\mathbb{Q}) + c_2(\mathbb{Q})$,
3. the lsc extension of c to $\mathbf{P}^{\mathbf{ba}}$ — defined as $c(\mu) = \sup_{\xi \in \mathcal{A}} \mu(-\xi)$ — is not necessarily equal to c_0. It is the case if and only if u_0 is Fatou.

Remark 51 If $\mathcal{S}_1$ is the set of all probability measures absolutely continuous with respect to $\mathbb{P}$, then for every Fatou coherent utility function u_2 we have $u_1 \Box u_2 = u_2$ and hence satisfies the Fatou property. This is easily seen by the equalities $\mathcal{S}_1 \cap \mathcal{S}_2 = \mathcal{S}_2$ and $\mathcal{S}_1^{\mathbf{ba}} \cap \mathcal{S}_2^{\mathbf{ba}} = \mathcal{S}_2^{\mathbf{ba}}$. In this case $u_1(\xi) = \text{ess.inf}(\xi)$ and $u_1 \Box u_2 = u_2$ for every (not necessarily Fatou) coherent utility function.

Example 28 Let $(A_n)_{n \geq 1}$ be a measurable partition of Ω into sets with $\mathbb{P}[A_n] > 0$. For each n, let e_n be the measure with density $\frac{\mathbf{1}_{A_n}}{\mathbb{P}[A_n]}$. The sets $\mathcal{S}_1$ and $\mathcal{S}_2$ are defined as follows:

$$\mathcal{S}_1 = \overline{conv}\,(e_1, e_3, e_4, \ldots)$$
$$\mathcal{S}_2 = \overline{conv}\left(e_1, \left(\frac{e_2 + n e_n}{1+n}\right)_{n \geq 3}\right).$$

Clearly, $\mathcal{S}_1 \cap \mathcal{S}_2 = \{e_1\}$ and $\mathcal{S}_1^{\mathbf{ba}} \cap \mathcal{S}_2^{\mathbf{ba}}$ contains, besides the vector e_1, the adherent points in $\mathbf{ba}$ of the sequence $(e_n, n \geq 1)$. The measure $u_1 \Box u_2$ is therefore not the same as u and $u_1 \Box u_2$ does not have the Fatou property.

Example 29 We take the same sequences as in the previous example but this time we define:

$$\mathcal{S}_1 = \overline{conv}\,(e_3, e_4, \ldots)$$
$$\mathcal{S}_2 = \overline{conv}\,(\frac{e_2 + n e_n}{1+n}, n \geq 3)\,.$$

Clearly, $\mathcal{S}_1 \cap \mathcal{S}_2 = \emptyset$ and $\mathcal{A}_1 + \mathcal{A}_2$ is $\sigma(L^\infty, L^1)$ dense in L^∞. However, $\mathcal{A}_1 + \mathcal{A}_2$ is not norm dense in L^∞, since $\mathcal{S}_1^{\mathbf{ba}} \cap \mathcal{S}_2^{\mathbf{ba}} \neq \emptyset$.

Example 30 We consider a finite Ω (to avoid topological difficulties) and we suppose the regulator agreed that the positions $\eta_1, \ldots, \eta_n$ are acceptable. In this context positions are just vectors in $\mathbb{R}^\Omega$. The minimal convex cone $\mathcal{A}_i$ containing $L^\infty_+ = \{\xi \geq 0\}$ and η_i is the set $\{\xi + \lambda\eta_i \mid \lambda \geq 0; \xi \geq 0\}$: the purpose is to construct a risk measure under which each of the originally given positions $(\eta_i)_{i=1}^n$, is still acceptable. Therefore we take $\mathcal{A} = conv\,(\mathcal{A}_i; i \leq n)$ so that our utility function u will be $u_1 \Box \cdots \Box u_n$. As an exercise we write the rest of the remark in terms of risk measures. The reader should of course adapt the definition of the convex convolution!!

We have:

$$\begin{aligned}
\rho(\xi) &= \inf \left\{ \rho_1(\xi_1) + \ldots \rho_n(\xi_n) \mid \xi = \sum_{i=1}^n \xi_i \right\} \\
&= \inf \left\{ \alpha_1 + \ldots + \alpha_n \mid \exists \lambda_i \in \mathbb{R}^+, \exists f_i \in \mathbb{R}^\Omega_+ \quad \alpha_i + \xi_i = f_i + \lambda_i \eta_i, \xi = \sum_{i=1}^n \xi_i \right\} \\
&= \inf \left\{ \alpha \mid \xi + \alpha \geq \sum_{i=1}^n \lambda_i \eta_i \text{ where } \lambda_i \geq 0 \right\}
\end{aligned}$$

We notice that the specification of the values of $\rho(\eta_i)$ is not required and that the risk measure can be equal to $-\infty$ (which is the case if $\cap_{i \leq n} \mathcal{S}_i = \emptyset$). The problem of calculating $\rho(\xi)$ can be restated as a linear program:

$$\begin{cases}
\max_{\mathbb{Q}} \mathbb{E}_{\mathbb{Q}}[-\xi] \text{ subject to} \\
\sum_\omega Q(\omega) = 1 \,, \mathbb{Q}(\omega) \geq 0 \\
\mathbb{E}_{\mathbb{Q}}[\eta_i] \geq 0
\end{cases}$$

and the preceeding equality is the usual dual-primal linear program relation.

Remark 52 The relation between primal and dual program can be worked out for the case of general Ω and it yields an example of the duality gap. Let us illustrate this as follows. Consider the primal program:

$$\begin{aligned}
&\min \mathbb{Q}[\xi] \\
&\text{subject to } \mathbb{Q} \in L^1 \text{ a probability measure } , \forall \eta \in \mathcal{A}_1 + \mathcal{A}_2 : \mathbb{E}_{\mathbb{Q}}[\eta] \geq 0.
\end{aligned}$$

The dual program is

$$\begin{aligned}
&\max \alpha \\
&\text{subject to } \xi - \alpha \in \mathcal{A}_1 + \mathcal{A}_2.
\end{aligned}$$

The dual program of this, written in **ba** is then

$$\min \mu(\xi)$$
$$\text{subject to } \mu \in \mathbf{P}^{\mathbf{ba}} \text{ such that for all } \eta \in \mathcal{A}_1 + \mathcal{A}_2 : \mu(\eta) \geq 0.$$

The second and the third program yield the value $u_0(\xi) = (u_1 \Box u_2)(\xi)$ whereas the first (primal program) yields the possibly much smaller value $u(\xi)$. It can even happen that the first program is not feasible (since $\mathcal{S}_1 \cap \mathcal{S}_2 = \emptyset$ as in Example 29, whereas the second and third program remain feasible (since $\mathcal{S}_1^{\mathbf{ba}} \cap \mathcal{S}_2^{\mathbf{ba}} \neq \emptyset$).

6.4 The inf convolution of concave utility functions

The idea is the same as in the previous section but we need to work with the penalty functions. So we start with $u_1, \mathcal{A}_1, c_1$ and $u_2, \mathcal{A}_2, c_2$. We suppose that an agent has the possibility to split an element $\xi \in L^\infty$ into two parts $\eta + \zeta = \xi$. With the first one he gets a utility $u_1(\eta)$, with the second one $u_2(\zeta)$. Of course he wants to do this as good as possible. So she looks for

$$(u_1 \Box u_2)(\xi) = \sup\{u_1(\eta_1) + u_2(\eta_2) \mid \eta_1 + \eta_2 = \xi\}.$$

This can also be described as follows. The hypograph of u_i is defined as $\mathcal{G}_i = \{(\xi, \alpha) \mid \alpha \leq u_i(\xi)\}$. Because u_1 and u_2 are concave, the sets $\mathcal{G}_i$ are convex. The hypograph of $u_1 \Box u_2$ is more or less the sum of the hypographs, $\mathcal{G}_1 + \mathcal{G}_2$. This sum is convex and it defines a function with value at ξ given by $\sup\{u_1(\eta_1) + u_2(\eta_2) \mid \eta_1 + \eta_2 = \xi\}$. So it defines $u_1 \Box u_2$. Of course we should prove that this function is well defined, that it is concave etc. We will do that in the same way as for coherent functions. Before we start let us recall that if u_1, u_2 both are Fatou, this does not imply that $u_1 \Box u_2$ has the Fatou property! This explains why we do the analysis for general concave monetary utility functions. When we use the penalty functions c_1, c_2, we systematically work on the set $\mathbf{P}^{\mathbf{ba}}$.

Lemma 16 *The function $u_1 \Box u_2 \colon L^\infty \to \mathbb{R} \cup \{+\infty\}$ is concave and hence it is either identically $+\infty$ or it is finite everywhere.*

Proof. Standard in convex analysis, let us give a proof – for completeness. Clearly $u_1 \Box u_2(\xi) > -\infty$. Take $\xi, \xi', 0 < \lambda < 1$. Take $k < u_1 \Box u_2(\xi), k' <$

$u_1\Box u_2(\xi')$. We have the existence of $\eta_1, \eta_1', \eta_2, \eta_2'$ such that $u_1(\eta_1)+u_2(\eta_2) > k$ and $u_1(\eta_1') + u_2(\eta_2') > k'$. Concavity of u_1, u_2 implies

$$\begin{aligned}
&u_1(\lambda\eta_1 + (1-\lambda)\eta_1') + u_2(\lambda\eta_2 + (1-\lambda)\eta_2') \\
&\geq \lambda u_1(\eta_1) + (1-\lambda)u_1(\eta_1') + \lambda u_2(\eta_2) + (1-\lambda)u_2(\eta_2') \\
&\geq \lambda k + (1-\lambda)k'.
\end{aligned}$$

This shows that

$$(u_1\Box u_2)(\lambda\xi + (1-\lambda)\xi') \geq \lambda(u_1\Box u_2)(\xi) + (1-\lambda)(u_1\Box u_2)(\xi').$$

Concave functions $u\colon L^\infty \to \mathbb{R}\cup\{+\infty\}$ either are identically $+\infty$ or are finite everywhere. This is seen as follows. Take a point ξ where $u(\xi) < +\infty$ and a point η where $u(\eta) = +\infty$. Look at the concave function $\phi\colon t \to u(t\xi + (1-t)\eta)$. On the real line we have that this function is finite for $t = 1$ and infinite for $t = 0$. By concaviy $\phi(1) \geq \frac{1}{2}(\phi(0) + \phi(2)) = +\infty$, a contradiction. So either $u \equiv +\infty$ or $u < +\infty$ everywhere. □

Proposition 32 *Are equivalent*

1. *For all* $\xi \in L^\infty$: $u_1\Box u_2(\xi) < \infty$
2. $u_1\Box u_2(0) < \infty$
3. $\mathcal{A}_1 + \mathcal{A}_2 \neq L^\infty$
4. $\mathcal{A}_1 + \mathcal{A}_2$ *is not norm dense in* L^∞

Proof. Suppose that $\mathcal{A}_1+\mathcal{A}_2$ is norm dense in L^∞. Then for $\xi \in L^\infty$ there exists η such that $\|\xi - \eta\| \leq 1$ and $\eta \in \mathcal{A}_1 + \mathcal{A}_2$. Because $\mathcal{A}_1 + \mathcal{A}_2 + L^\infty_+ \subset \mathcal{A}_1 + \mathcal{A}_2$ this means: $\xi + 1 \in \mathcal{A}_1 + \mathcal{A}_2$. Hence we get $\mathcal{A}_1 + \mathcal{A}_2 = L^\infty$. If $\mathcal{A}_1 + \mathcal{A}_2 = L^\infty$ then for $k \geq 1$ we have $\eta_1 \in \mathcal{A}_1, \eta_2 \in \mathcal{A}_2$ with $\eta_1 + \eta_2 = -k$, hence $u_1\Box u_2(0) \geq u_1(\eta_1) + \eta_2(\eta_2 + k) \geq k$, proving that $u_1\Box u_2(0) = +\infty$. The other implications are trivial or proved in the previous lemma. □

Proposition 33 *Suppose* $u_1\Box u_2$ *is well defined* ($\mathcal{A}_1 + \mathcal{A}_2 \neq L^\infty$) *then*

1. $u_1\Box u_2 : L^\infty \to \mathbb{R}$
2. $u_1\Box u_2$ *is monetary*

3. *$u_1\Box u_2(\xi)\ge 0$ if $\xi\ge 0$ and $\xi\ge\eta$ implies $u_1\Box u_2(\xi)\ge u_1\Box u_2(\eta)$, so we also get $|u_1\Box u_2(\xi)-u_1\Box u_2(\eta)|\le\|\xi-\eta\|_\infty$*

4. *$u_1\Box u_2$ is concave*

5. *The set $\{\xi\mid u_1\Box u_2(\xi)\ge 0\}$ is the norm closure of $\mathcal{A}_1+\mathcal{A}_2$, more precisely $u_1\Box u_2(\xi)>0$ if and only of there are elements $\eta_1\in\mathcal{A}_1$, $\eta_2\in\mathcal{A}_2$, $u_1(\eta_1)>0$, $u_2(\eta_2)>0$ and $\xi=\eta_1+\eta_2$.*

6. *$u_1\Box u_2(0)=0$ if and only if the set $\{\mu\in\mathbf{P^{ba}}\mid c_1(\mu)=0\}\cap\{\mu\in\mathbf{P^{ba}}\mid c_2(\mu)=0\}\ne\emptyset$. In this case $u_1\Box u_2$ defines a concave utility function on L^∞.*

Proof. Points 1 and 4 have been proved above. Points 2, 3 can be proved using the definition, point 5 again follows from the definition. Point 6 requires more attention. If $c_1(\mu)=c_2(\mu)=0$ we have for $\xi\in\mathcal{A}_1+\mathcal{A}_2$: $\mu(\xi)\ge 0$. This shows that for $\varepsilon<0$ we must have $\varepsilon\notin\mathcal{A}_1+\mathcal{A}_2$. Hence also for all $\varepsilon<0$ we must have $\varepsilon\notin\overline{\mathcal{A}_1+\mathcal{A}_2}$. Hence $u_1\Box u_2(\varepsilon)<0$. But then $u_1\Box u_2(0)\le 0$ by continuity. Since the other equality was shown above we get $u_1\Box u_2(0)=0$. Conversely, suppose that $\{\mu\in\mathbf{P^{ba}}\mid c_1(\mu)=0\}\cap\{\mu\in\mathbf{P^{ba}}\mid c_2(\mu)=0\}=\emptyset$. Since both sets are weak* compact and convex we can strictly separate them. This gives the existence of η with

$$-\min\{\mu(-\eta)\mid c_1(\mu)=0\}=\max\{\mu(\eta)\mid c_1(\mu)=0\}<\min\{\mu(\eta)\mid c_2(\mu)=0\}.$$

The study of the one sided derivative showed that $\min\{\mu(\eta)\mid c_2(\mu)=0\}=\lim_{\varepsilon\downarrow 0}\frac{u_2(\varepsilon\eta)}{\varepsilon}$ and similarly for u_1. As a result the element η satisfies for ε small enough:

$$u_1(-\varepsilon\eta)+u_2(\varepsilon\eta)>0.$$

This implies $u_1\Box u_2(0)>0$. □

Example 31 Suppose that f satisfies $0<f<1$ a.s. with $\text{ess.inf}\, f=0, \text{ess.sup}\, f=1$, e.g. f is uniformly distributed over $[0,1]$. Let $c_1(\mathbb{Q})=\mathbb{E}_\mathbb{Q}[f]$ and $c_2(\mathbb{Q})=\mathbb{E}_\mathbb{Q}[1-f]$. Clearly $c_1+c_2=1$ and $u_1\Box u_2(0)=1$. We leave it to the reader to find an example in the case where Ω consists of two points. Hint: look at $\mathcal{A}_1=\{(x,y)\mid x\ge 0, y\ge 0 \text{ or } -1\le x\le 0, y\ge -\frac{1}{4}x\}$ and for $\mathcal{A}_2$ interchange the role of x and y.

Proposition 34 *The penalty function c of $u_1\Box u_2$, is equal to $c(\mu)=c_1(\mu)+c_2(\mu)$. Consequently $u_1\Box u_2$ is well defined if and only if $dom(c_1)\cap dom(c_2)\ne\emptyset$. The hypograph of $u_1\Box u_2$ is the norm closure of the set $\mathcal{G}_1+\mathcal{G}_2$. $u_1\Box u_2(0)=\min\{c_1(\mu)+c_2(\mu)\mid\mu\in\mathbf{P^{ba}}\}$.*

Proof. This is done by straightforward calculation, it is even valid when $\mathcal{A}_1 + \mathcal{A}_2 = L^\infty$. Let $\mathcal{A}$ be the norm closure of $\mathcal{A}_1 + \mathcal{A}_2$. Then

$$\begin{aligned} c(\mu) &= \sup_{\xi \in \mathcal{A}} \mu(-\xi) = \sup_{\xi \in \mathcal{A}_1 + \mathcal{A}_2} \mu(-\xi) \\ &= \sup_{\eta_1 \in \mathcal{A}_1} \mu(-\eta_1) + \sup_{\eta_2 \in \mathcal{A}_2} \mu(-\eta_2) = c_1(\mu) + c_2(\mu). \end{aligned}$$

Now by the Hahn-Banach theorem, $\mathcal{A} \neq L^\infty$ if and only if there is $\mu \in \mathbf{P^{ba}}$ with $c(\mu) < \infty$. The domain of a convex function is defined as $\{\mu \mid c_1(\mu) < \infty\}$ and hence $u_1 \Box u_2$ is well defined if and only if $dom(c_1) \cap dom(c_2) \neq \emptyset$. The last line should be obvious. By definition $\mathcal{G}_1 + \mathcal{G}_2 \subset \mathcal{G} = \{(\xi, \alpha) \mid \alpha \leq u_1 \Box u_2(\xi)\}$. The latter set is closed by continuity if $u_1 \Box u_2 < \infty$ and trivially if $u_1 \Box u_2 \equiv +\infty$, hence $\overline{\mathcal{G}_1 + \mathcal{G}_2} \subset \mathcal{G}$. But by construction of $u_1 \Box u_2$, we must have $\mathcal{G} \subset \overline{\mathcal{G}_1 + \mathcal{G}_2}$ so they must be equal. The value $u_1 \Box u_2(0) = \min_\mu c(\mu) = \min_\mu (c_1(\mu) + c_2(\mu))$. □

Remark 53 The function $u_1 \Box u_2$ is not equal to the smallest concave function that is greater than both u_1 and u_2. The latter would be given by the convex hull of the union of the hypographs $\mathcal{G}_1, \mathcal{G}_2$. That both can be different can be seen on the following trivial example. Take u concave and monetary with acceptance set $\mathcal{A}$. The convolution $u \Box u$ has the acceptance set $\mathcal{A} + \mathcal{A} = 2\mathcal{A}$ (since $\mathcal{A}$ contains 0 and is convex). In case $u \Box u$ would be the smallest utility function greater than u than we would have $u \Box u = u$ or $2\mathcal{A} = \mathcal{A}$. This would imply that $\mathcal{A}$ is a cone, i.e. u is coherent.

Remark 54 If u_2 is coherent and given by the scenario set $\mathcal{S}_2$, then

$$\sup_{\xi \in L^\infty} \{u_1(\xi) - \sup_{\nu \in \mathcal{S}^{\mathbf{ba}}} \nu(\xi)\} = u_1 \Box u_2(0) = \min_{\nu \in \mathcal{S}^{\mathbf{ba}}} c_1(\nu).$$

We leave the interpretation of this equality to the reader.

Remark 55 If both u_1, u_2 are Fatou, there is no guarantee that $u_1 \Box u_2$ is Fatou. See [40] for an example. But the reader can already guess where the difficulties are. For the penalty function $c = c_1 + c_2$ we should show that for every μ there is a net of elements $\mathbb{Q}^\alpha$ such that $c(\mu) = \lim c(\mathbb{Q}^\alpha)$. For each i we can find nets $\mathbb{Q}_i^\alpha$ that give $c_i(\mu) = c_i(\mathbb{Q}_i^\alpha)$ but there is no reason why we should be able to find the same net for both penalty functions.

We will see by examples that $\mathcal{A}_1 + \mathcal{A}_2$ is not always closed. In fact we can show that

Proposition 35 $\mathcal{A}_1 + \mathcal{A}_2$ *is closed if and only if the sum of the hypographs is closed. This is equivalent to: for all* ξ *there are elements* η_1, η_2 *such that* $\eta_1 + \eta_2 = \xi$ *and* $u_1 \Box u_2(\xi) = u_1(\eta_1) + u_2(\eta_2)$.

Proof. First suppose that $\mathcal{G}_1 + \mathcal{G}_2$ is closed. Then it is equal to the hypograph $\mathcal{G}$ of $u = u_1 \Box u_2$. Let $\xi \in \mathcal{A} = \{\eta \mid u(\eta \geq 0\}$ be such that $u(\xi) = 0$. Then $(\xi, 0) \in \mathcal{G}$. Since $\mathcal{G} = \mathcal{G}_1 + \mathcal{G}_2$ we can find two elements (η_1, α_1), (η_2, α_2) such that $\eta_1 + \eta_2 = \xi$, $\alpha_1 \leq u_1(\eta_1), \alpha_2 \leq u_2(\eta_2)$ and $\alpha_1 + \alpha_2 = 0$. This implies $\xi = \eta_1 - \alpha_1 + \eta_2 - \alpha_2$ with $\eta_1 - \alpha_1 \in \mathcal{A}_1$ and $\eta_2 - \alpha_2 \in \mathcal{A}_2$. But this implies $u_1(\eta_1) \geq \alpha_1$ and $u_2(\eta_2) \geq \alpha_2$ Since $u_1(\eta_1) + u_2(\eta_2) \leq u(\xi) = 0$, we must have $u_1(\eta_1) = \alpha_1$ and $u_2(\eta_2) = \alpha_2$. Conversely suppose that $\mathcal{A}_1 + \mathcal{A}_2 = \mathcal{A}$ is closed. Take $(\xi, \alpha) \in \mathcal{G}$. We have $\xi - \alpha \in \mathcal{A}$ and hence we find two elements η_1, η_2 with $u_1(\eta_1) \geq 0, u_2(\eta_2) \geq 0$ and $\xi - \alpha = \eta_1 + \eta_2$. This can be rewritten as $\xi = \eta_1 + (\eta_2 + \alpha)$. But then $(\eta_1, 0) \in \mathcal{G}_1$ and $(\eta_2 + \alpha, \alpha) \in \mathcal{G}_2$ since $u_2(\eta_2 + \alpha) = u_2(\eta_2) + \alpha \geq \alpha$. So $(\xi, \alpha) \in \mathcal{G}_1 + \mathcal{G}_2$. □

The existence of the elements η_1, η_2 such that $\eta_1 + \eta_2 = \xi$ and $u_1 \Box u_2(\xi) = u_1(\eta_1) + u_2(\eta_2)$ means that we can actually solve the optimisation problem. In case $u(\xi) = 0$, we can select the solution in such a way that $\eta_1 + \eta_2 = \xi$ and $u_1(\eta_1) = 0 = u_2(\eta_2)$. Unfortunately the set $\mathcal{A}_1 + \mathcal{A}_2$ is not always closed, not even in nice examples. This is not just an infinite dimensional feature. Even in $\mathbb{R}^2$, i.e. for Ω having two points, the sum of two closed convex sets need not be closed and one can find (as an exercise) two concave utility functions u_1, u_2 such that the sum $\mathcal{A}_1 + \mathcal{A}_2$ is not closed. However this will give $u_1 \Box u_2(0) > 0$, $\mathbb{R}^2$ is simply too small to have better examples. But in $\mathbb{R}^3$, one can find two coherent utility functions (hence $u_1 \Box u_2(0) = 0$) such that $\mathcal{A}_1 + \mathcal{A}_2$ is not closed. These are good exercises in geometry.

Exercise 21 Give examples as described in the last paragraph.

6.5 Product of coherent utility functions

Let $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ be two probability spaces. We consider the product space $(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mathbb{P}_1 \otimes \mathbb{P}_2)$ and we would like to define a coherent utility function, the most liberal one, given two utility functions u_1 and u_2, defined on Ω_1 and Ω_2 respectively.

For a probability measure $\mathbb{Q}$ on Ω, we define $\mathbb{Q}_1$ and $\mathbb{Q}_2$ to be the marginal probabilities of $\mathbb{Q}$ on Ω_1 and Ω_2 (that is, $\mathbb{Q}_1[A_1] = \mathbb{Q}[A_1 \times \Omega_2]$ and similarly for $\mathbb{Q}_2$). If as usual, $\mathcal{S}_i$ and $\mathcal{A}_i$ represent the family of probabilities and the set of acceptable positions for u_i, we define:

$$\tilde{\mathcal{S}}_1 = \{\mathbb{Q} \mid \mathbb{Q} \ll \mathbb{P}; \mathbb{Q}_1 \in \mathcal{S}_1\}$$
$$\tilde{\mathcal{S}}_2 = \{\mathbb{Q} \mid \mathbb{Q} \ll \mathbb{P}; \mathbb{Q}_2 \in \mathcal{S}_2\}.$$

We suppose for simplicity that u_1 and u_2 are relevant, the general case is left to the reader. If $f \in \mathcal{A}_1$, a "reasonable" request is that $\tilde{f}(\omega_1, \omega_2) = f(\omega_1)$ should be acceptable; the same should hold for $g \in \mathcal{A}_2$. So we put

$$\tilde{\mathcal{A}}_1 = \{f + h \mid f \in \mathcal{A}_1, h \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}), h \geq 0\}$$
$$\tilde{\mathcal{A}}_2 = \{g + h \mid g \in \mathcal{A}_2, h \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}), h \geq 0\}$$
$$\mathcal{A} = \tilde{\mathcal{A}}_1 + \tilde{\mathcal{A}}_2.$$

Clearly the set $\mathcal{A}$ is a convex cone. However, it is also $\sigma(L^\infty, L^1)$ closed. To see this – less trivial statement – we use the Krein-Smulian theorem. So let us suppose $(\phi_n)_n \subset \mathcal{A}$, $\|\phi_n\|_\infty \leq 1$ and $\phi_n \xrightarrow{\mathbb{P}} \phi$. We have to show that $\phi \in \mathcal{A}$. Each ϕ_n can be written as $\phi_n = f_n + g_n + h_n$, where $f_n \in \mathcal{A}_1$, $g_n \in \mathcal{A}_2$ and $h_n \geq 0$. Take $\mathbb{Q}_1 \in \mathcal{S}_1$, $\mathbb{Q}_2 \in \mathcal{S}_2$, $\mathbb{Q}_1 \sim \mathbb{P}_1$, $Q_2 \sim \mathbb{P}_2$ and let $\mathbb{Q} = \mathbb{Q}_1 \otimes \mathbb{Q}_2$. Of course, $\mathbb{Q} \in \tilde{\mathcal{S}}_1 \cap \tilde{\mathcal{S}}_2$. Furthermore $\mathbb{Q}_1$ and $\mathbb{Q}_2$ are the marginal probabilities of $\mathbb{Q}$, so that there is no conflict in the notation. We clearly have $1 \geq \mathbb{E}_\mathbb{Q}[f_n + g_n + h_n] \geq \mathbb{E}_{\mathbb{Q}_1}[f_n] + \mathbb{E}_{\mathbb{Q}_2}[g_n]$. Both terms are nonnegative since $f_n \in \mathcal{A}_1$ and $g_n \in \mathcal{A}_2$. Therefore, $\mathbb{E}_{\mathbb{Q}_1}[f_n]$ and $\mathbb{E}_{\mathbb{Q}_2}[g_n]$ are between 0 and 1. We may and do suppose that $\mathbb{E}_{\mathbb{Q}_1}[f_n]$ and $\mathbb{E}_{\mathbb{Q}_2}[g_n]$ converge (if not, we take a subsequence). Since $f_n + g_n + h_n \leq 1$, we also get $f_n + g_n \leq 1$ and hence $f_n + \mathbb{E}_{\mathbb{Q}_2}[g_n] \leq 1$. Indeed for $\mathbb{Q}$, f_n and g_n are independent and the inequality results by taking conditional expectation with respect to $\mathcal{F}_1 \otimes \{\emptyset, \Omega_2\}$. Since $\mathbb{E}_{\mathbb{Q}_2}[g_n] \geq 0$, we get $f_n \leq 1$. Similarly, we get $g_n \leq 1$. We now replace f_n and g_n by respectively $f_n \vee (-2)$ and $g_n \vee (-2)$. Necessarily we have $f_n \vee (-2) \geq f_n$ and therefore $f_n \vee (-2) \in \mathcal{A}_1$, also $g_n \vee (-2) \in \mathcal{A}_2$. But this requires a correction of h_n. So we get:

$$\phi_n = f_n \vee (-2) + g_n \vee (-2) + h_n - (-2 - f_n)^+ - (-2 - g_n)^+ .$$

The function $h_n - (-2 - f_n)^+ - (-2 - g_n)^+$ is still nonnegative. To see this, we essentially have the following two cases.

On the set $\{f_n < -2\} \cap \{g_n < -2\}$ we have:

$$h_n - (-2 - f_n)^+ - (-2 - g_n)^+ = (h_n + f_n + g_n) + 4 \geq -1 + 4 > 0 .$$

On the set $\{f_n \ge -2\} \cap \{g_n < -2\}$ we have:

$$\begin{aligned} h_n - (-2 - f_n)^+ - (-2 - g_n)^+ &= h_n + 2 + g_n \\ &= (h_n + f_n + g_n) + (2 - f_n) \\ &\ge -1 + 1 \ge 0. \end{aligned}$$

The other cases are either trivial or similar.

So we finally may replace the functions as indicated and we may suppose that $\phi_n = f_n + g_n + h_n$, where $-2 \le f_n \le 1$, $-2 \le g_n \le 1$, $f_n \in \mathcal{A}_1$, $g_n \in \mathcal{A}_2$, $h_n \ge 0$.

Since the sequences $(f_n)_n$ and $(g_n)_n$ are bounded, we can take convex combinations of them, (still denoted by the same symbols), that converge in probability. So finally we get $f_n \to f$, $g_n \to g$ in probability. Of course this implies $f \in \mathcal{A}_1$ and $g \in \mathcal{A}_2$. But then we necessarily have that $h_n = \phi_n - f_n - g_n$ converges in probability, say to a function h. Of course, $h \ge 0$. So finally we get $\phi = f + g + h$ with $f \in \mathcal{A}_1$, $g \in \mathcal{A}_2$ and $h \ge 0$.

The polar cone of $\mathcal{A}$ can be described by the sets $\tilde{\mathcal{S}}_1$ and $\tilde{\mathcal{S}}_2$. Indeed

$$\begin{aligned} \mathcal{S} &= \{\mathbb{Q} \mid \mathbb{Q} \text{ a probability and } \forall u \in \mathcal{A} \quad \mathbb{Q}[u] \ge 0\} \\ &= \{\mathbb{Q} \mid \mathbb{Q} \text{ a probability and } \forall f \in \mathcal{A}_1 \quad \mathbb{Q}[f] \ge 0 \text{ and } \forall g \in \mathcal{A}_2 \quad \mathbb{Q}[g] \ge 0\} \\ &= \{\mathbb{Q} \mid \mathbb{Q} \text{ a probability and } \forall f \in \mathcal{A}_1 \quad \mathbb{Q}_1[f] \ge 0 \text{ and } \forall g \in \mathcal{A}_2 \quad \mathbb{Q}_2[g] \ge 0\} \\ &= \{\mathbb{Q} \mid \mathbb{Q}_1 \in \mathcal{S}_1 \,,\ \mathbb{Q}_2 \in \mathcal{S}_2\} \\ &= \tilde{\mathcal{S}}_1 \cap \tilde{\mathcal{S}}_2. \end{aligned}$$

Moreover

$$\begin{aligned} u(\xi) &= \sup\{\alpha \mid -\alpha + \xi \in \mathcal{A}\} \\ &= \sup\{\alpha \mid \exists f \in \mathcal{A}_1 \,, \exists g \in \mathcal{A}_2 \quad \xi - \alpha \ge f + g\}. \end{aligned}$$

The previous lines also imply that the sets $\tilde{\mathcal{A}}_1$ and $\tilde{\mathcal{A}}_2$ are $\sigma(L^\infty, L^1)$ closed. Their polars are precisely given by $\tilde{\mathcal{S}}_1$ and $\tilde{\mathcal{S}}_2$ respectively. Indeed:

$$\begin{aligned} &\{\mathbb{Q} \mid \mathbb{Q} \text{ proba. and for all } u \in \tilde{\mathcal{A}}_1 : \mathbb{Q}[u] \ge 0\} \\ &\qquad = \{\mathbb{Q} \mid \mathbb{Q} \text{ proba. and for all } u \in \mathcal{A}_1 : \mathbb{Q}[u] \ge 0\} \end{aligned}$$

and the latter is equal to $\{\mathbb{Q} \mid \forall f \in \mathcal{A}_1 \quad \mathbb{Q}_1[u] \ge 0\}$, which is exactly $\tilde{\mathcal{S}}_1$. Therefore we get that: $\tilde{\mathcal{A}}_1 = \{\phi \mid \forall \mathbb{Q} \in \tilde{\mathcal{S}}_1 \quad \mathbb{Q}[\phi] \ge 0\}$.

Remark 56 Even if $\mathcal{S}_1$ and $\mathcal{S}_2$ consist of a single point, the family $\mathcal{S}$ can be very "big". For instance, let's take $\Omega_1 = \Omega_2 = \mathbb{T}$, where $\mathbb{T}$ is the one dimensional torus (that is the circle S^1). On $\mathbb{T}$ we consider the Borel σ-algebra and as reference probability we take the normalized Lebesgue measure m, while $\mathcal{S}_1$ and $\mathcal{S}_2$ will coincide with $\{m\}$. If we take the product space $\mathbb{T} \times \mathbb{T}$ and we consider the set $A_\varepsilon = \{(e^{i\theta}, e^{i\phi}) \mid |e^{i\theta} - e^{i\phi}| \leq \varepsilon\}$ then $\lim_{\varepsilon \to 0} m(A_\varepsilon) = 0$; and by taking $\mathbb{Q}_\varepsilon$ equal to the uniform distribution on A_ε we have that Q_ε belongs to $\mathcal{S}$, for each ε. But the family $(\mathbb{Q}_\varepsilon)_\varepsilon$ is not uniformly integrable: therefore $\mathcal{S}$ is not at all small, it isn't even weakly compact! It is still an unsolved problem to characterise the extreme points of the convex set of measures on $\mathbb{T} \times \mathbb{T}$ so that the marginals are m.

Chapter 7

Convex games and utility functions

The aim of this chapter is to investigate the relations between convex games and coherent risk measures. The theory of convex games was developed by Shapley, [126] and David Schmeidler [123]. There were definitions related to submodular functions and hence relations to convex games in [56] and the theory of submodular functions goes at least back to Bergmann (1925). The theory is also related to the theory of capacities, [33], [27] and see [81]. We start with a couple of definitions.

Definition 21 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A convex game on $(\Omega, \mathcal{F})$ is a function $v : \mathcal{F} \to \mathbb{R}_+$ such that:*

$$v(\emptyset) = 0$$

$$\text{for all } A, B \in \mathcal{F} : v(A) + v(B) \leq v(A \cap B) + v(A \cup B)$$

We say that v is continuous with respect to $\mathbb{P}$ if

$$\mathbb{P}(A) = 1 \text{ implies } v(A) = 1$$

We say that v is Fatou if

$$A_n \downarrow A \text{ implies } v(A_n) \downarrow v(A)\,.$$

Remark 57 In capacitytheory the convexity relation is called 2-alternating, see [33, 81]. The Fatou property is a continuity property which will enable us to use the duality (L^1, L^∞) instead of $(L^\infty, \mathbf{ba})$. It is related to the Fatou property of an associated coherent utility function.

The idea is that we need to distribute an amount of money $v(\Omega)$, over the players (the elements of Ω). Players can form coalitions described by the structure $\mathcal{F}$. Each coalition has an intrinsic value, an amount of money that they can get by playing on their own. Convex games are such that synergies are obtained when coalitions join. For properties of convex games, we refer to Schmeidler, [123], and Delbaen, [37].

Definition 22 *For a convex game v we define the core of v as*

$$\mathcal{C}(v) = \{\mu \mid \mu \in \mathbf{ba},\ \mu(\Omega) = v(\Omega) \ \text{ and for all } A \in \mathcal{F}: \quad \mu(A) \geq v(A)\}\,.$$

The $\sigma-$core is defined as:

$$\mathcal{C}^\sigma(v) = \{f \mid f \in L^1, \mathbb{E}[f] = v(\Omega) \quad \text{and for all } A \in \mathcal{F}: \quad \mathbb{E}[f\mathbf{1}_A] \geq v(A)\}\,.$$

Again the idea is to distribute an amount of money over the different players. The influence of the players is over the coalitions they can form. So in order to be fair the allocation of the money, described by a finitely additive measure μ should be such that all coalitions are happy, i.e. $\mu(A) \geq v(A)$. Otherwise there is no reason for them to joint the others. The sigma-core is the intersection of the core with the space L^1. It was studied by [108] and [39].

7.1 Non-emptiness of the core

In this section we will show that the core of a convex game is non-empty. Of course this is well known ([**?**] for finite games and[123]) but the proof allows to explain the relation with utility theory. The associated coherent utility function u will satisfy an extra property that has a nice interpretation in risk management. The basic ingredient is (again) a theorem of Schmeidler [124], for which we will also give a complete proof.

Lemma 17 *If v is a convex game then for $B \in \mathcal{F}$, $v_A(B) = v(A\cap B)$ defines a convex game on the space $(A, A \cap \mathcal{F})$.*

Lemma 18 *Let v be a convex game and let $A \in \mathcal{F}$. Define $w_A(B) = v(A \cup B) - v(A)$. Then $w_A \colon A^c \cap \mathcal{F} \to \mathbb{R}$ is again a convex game defined on the space $(A^c, A^c \cap \mathcal{F})$.*

Proof. For $B_1, B_2 \subset A^c$ we have

$$\begin{aligned} w_A(B_1) + w_A(B_2) &= v(A \cup B_1) - v(A) + v(A \cup B_2) - v(A) \\ &\leq v(B_1 \cup B_2 \cup A) + v((B_1 \cap B_2) \cup A) - 2v(A) \\ &\leq w_A(B_1 \cup B_2) + w_A(B_1 \cap B_2). \end{aligned}$$

□

Lemma 19 *Let $A \in \mathcal{F}$ and let w_A be defined as in the previous lemma. Let $\mu \in \mathcal{C}(v_A)$ and let $\nu \in \mathcal{C}(w_A)$, then $\lambda \in$ **ba** defined as $\lambda(B) = \mu(A \cap B) + \nu(A \cap A^c)$ defines an element of $\mathcal{C}(v)$.*

Proof. This is easy since obviously $\lambda(\Omega) = v(\Omega)$ and

$$\begin{aligned}\lambda(B) &= \mu(A \cap B) + \nu(B \cap A^c)\\ &\geq v(A \cap B) + v((B \cap A^c) \cup A) - v(A)\\ &\geq v(A \cap B) + v(B \cup A) - v(A) \geq v(B)\end{aligned}$$

□

Theorem 37 *If v is a convex game then $\mathcal{C}(v) \neq \emptyset$.*

Proof. We first prove the result for $\mathcal{F}$ finite and then proceed using a compactness argument. In case $\mathcal{F}$ is finite we may replace Ω by a finite set and take $\mathcal{F} = 2^\Omega$. So let us suppose that $\Omega = \{1, 2, \ldots, N\}$. We will use induction on N. Clearly there is nothing to prove when $N = 1$. So suppose that the core is non-empty for $N - 1$. Define $A = \{1\}$ and use the previous lemma. The induction hypothesis gives an element $\nu \in \mathcal{C}(w_A)$ and trivially an element $\mu \in \mathcal{C}(v_A)$. The element λ constructed in the previous lemma is then in $\mathcal{C}(v)$.

For general $\mathcal{F}$ we use a compactness argument. First we recall that the restriction mapping

$$\mathbf{ba}_+(\Omega, \mathcal{F}) \to \mathbf{ba}_+(\Omega, \mathcal{F}'),$$

is onto. This is a straightforward consequence of the Hahn-Banach theorem in its analytic form. Suppose that $\nu \in \mathbf{ba}_+(\Omega, \mathcal{F}')$ then it defines a linear form on $L^\infty(\Omega, \mathcal{F}')$ of norm $\|\nu\| = \nu(\Omega)$. The Hahn-Banch theorem allows to find an extension μ to a linear form on $L^\infty(\Omega, \mathcal{F})$ of the *same* norm. This equality in norm shows that $\mu \geq 0$ (prove this as an exercise!). For each finite subalgebra $\mathcal{F}' \subset \mathcal{F}$ we define

$$\mathcal{C}(\mathcal{F}') = \{\mu \in \mathbf{ba}(\Omega, \mathcal{F}) \mid \text{ for each } B \in \mathcal{F}' : \mu(B) \geq v(B)\}.$$

The onto character of the restriction map as well as the fact that for finite games the core is non-empty, shows that these sets are non-empty. They are weak* compact since they are weak* closed sets of the ball of radius $v(\Omega)$ in $\mathbf{ba}(\Omega, \mathcal{F})$. The collection of sets $\{\mathcal{C}(\mathcal{F}') \mid \mathcal{F}' \text{ finite}\}$ have the finite intersection property and hence their intersection is non-empty. Obviously their intersection is $\mathcal{C}(v)$. □

Theorem 38 *Let v be a convex game and let $A \in \mathcal{F}$, then there is an element $\mu \in \mathcal{C}(v)$ such that $\mu(A) = v(A)$.*

Proof. This is easy. Take $\mu \in \mathcal{C}(v_A)$ and $\nu \in \mathcal{C}(w_A)$. The previous theorem shows that both sets are non-empty. The Lemma 19 then produces an element $\lambda \in \mathcal{C}(v)$ with $\lambda(A) = v(A)$. □

A more refined application of the same lemma yields

Theorem 39 *Let v be a convex game. Let $A_1 \supset A_2 \supset \dots A_n$ be a finite chain of elements of $\mathcal{F}$. Then there is an element $\lambda \in \mathcal{C}(v)$ such that for all i: $\lambda(A_i) = v(A_i)$.*

Proof. We use induction on n. For $n = 1$ this is the previous theorem. Suppose the theorem is proved for $n - 1$. Then we can find elements $\mu \in \mathcal{C}(v_{A_2})$ with $\mu(A_i) = v(A_i)$ for $i \geq 2$. There is also an element $\nu \in \mathcal{C}(w_{A_2})$ such that $\nu(A_1 \setminus A_2) = w_{A_2}(A_1 \setminus A_2) = v(A_1) - v(A_2)$. Lemma xx produces $\lambda \in \mathcal{C}(v)$ with $\lambda(A_i) = v(A_i)$ for all $i \geq 1$. □

Theorem 40 *Let v be a convex game and let $\{A_i \mid i \in I\}$ be a totally ordered set of elements of $\mathcal{F}$. Then there is an element $\mu \in \mathcal{C}(v)$ such that for all $i \in I$: $\mu(A_i) = v(A_i)$.*

Proof. Follows from the preceding theorem by a compactness argument. □

From now on we suppose that the convex game v is normalised, i.e. $v(\Omega) = 1$. The core then consists of finitely additive probability measures. For simplicity we also suppose that the game is continuous with respect to $\mathbb{P}$. This is not really needed but it brings us immediately in the scope of coherent utility functions defined on L^∞ instead of defining them on spaces of random variables. The reader can pursue the analysis without this assumption if she wants to do so. The core $\mathcal{C}(v)$ is then a subset of $\mathbf{ba}(\Omega, \mathcal{F}, \mathbb{P})$.

Theorem 41 *Let v be a convex game and let $A_1 \supset A_2 \supset \dots A_n$ be a chain of elements of $\mathcal{F}$. Let $\alpha_1, \alpha_2, \dots, \alpha_n$ be nonnegative numbers. Let $\xi = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$. We have*

$$\begin{aligned} u_v(\xi) &= \inf\{\mu(\xi) \mid \mu \in \mathcal{C}(v)\} = \sum_i \alpha_i v(A_i) \\ &= \int_0^{+\infty} v(\xi \geq x)\, dx = \int_0^{+\infty} v(\xi > x)\, dx. \end{aligned}$$

Proof. By definition of the core we have $u_v(\xi) \geq \sum_i \alpha_i v(A_i)$. But since there is an element $\mu \in \mathcal{C}(v)$ with $\mu(A_i) = v(A_i)$ we necessarily get equality. Also

$$\int_0^{+\infty} v(\xi > x)\,dx = \int_0^{+\infty} v(\xi \geq x)\,dx = \int_0^{\alpha_1} + \int_{\alpha_1}^{\alpha_1+\alpha_2} + \ldots = \sum_i \alpha_i v(A_i).$$

□

Remark 58 These equalities, basically due to Choquet [33], can be found in Schmeidler [123] and Delbaen [37]. Just for convenience of the reader let us rephrase the equality above for indicator functions: $u_v(\mathbf{1}_A) = v(A)$. The knowledge of the utility function on the indicator functions therefore allows to calculate the utility functions for bounded random variables. The formula of the theorem can easily be extended to all bounded random variables. The positivity of ξ can be overcome by replacing ξ by $\xi + \|\xi\|_\infty$ and then using the monetary property.

Theorem 42 *Let v be a convex game, we have for nonnegative random variables ξ:*

$$u_v(\xi) = \int_0^{+\infty} v(\xi \geq x)\,dx = \int_0^{+\infty} v(\xi > x)\,dx.$$

Proof. Let us take a sequence of random variables ξ_n such that $\xi_n \downarrow \xi$, ξ_n only takes a finite number of values and $1 \geq \varepsilon_n = \|\xi_n - \xi\|_\infty \to 0$. We clearly have that $v(\xi_n \geq x) \geq v(\xi \geq x)$ This shows that

$$\int_0^{+\infty} v(\xi_n \geq x)\,dx \geq \int_0^{+\infty} v(\xi \geq x)\,dx.$$

But we also have that $\xi \geq \xi_n - \varepsilon_n$ and hence $v(\xi \geq x) \geq v(\xi_n \geq x + \varepsilon_n)$. Therefore

$$\begin{aligned}
\int_0^{+\infty} v(\xi \geq x)\,dx &= \int_0^{\|\xi\|_\infty} v(\xi \geq x)\,dx \\
&\geq \int_0^{\|\xi\|_\infty} v(\xi_n \geq x + \varepsilon_n)\,dx \\
&= \int_{\varepsilon_n}^{\|\xi\|_\infty + \varepsilon_n} v(\xi_n \geq x)\,dx \\
&\geq \int_0^{\|\xi_n\|_\infty} v(\xi_n \geq x)\,dx - \varepsilon_n = \int_0^{\infty} v(\xi_n \geq x)\,dx - \varepsilon_n
\end{aligned}$$

So we have that

$$u_v(\xi) = \lim_n u_v(\xi_n) = \lim_n \int_0^{+\infty} v(\xi_n \geq x)\, dx = \int_0^{+\infty} v(\xi \geq x)\, dx.$$

□

Theorem 43 *With the notation of the previous theorems we have that u_v is Fatou if and only if v is Fatou.*

Proof. If u_v is Fatou then we must have $v(A_n) = u_v(\mathbf{1}_{A_n})$ decreases to $u_v(\mathbf{1}_A) = v(A)$ if $A_n \downarrow A$. The converse is also true. If $\xi_n \downarrow \xi \geq 0$, then $\{\xi_n \geq x\} \downarrow \{\xi \geq x\}$ and hence for all $x \geq 0$:

$$v(\xi_n \geq x) \downarrow v(\xi \geq x).$$

From this it follows that

$$u_v(\xi_n) = \int_0^{\|\xi_1\|_\infty} v(\xi_n \geq x)\, dx \downarrow \int_0^{\|\xi_1\|_\infty} v(\xi \geq x)\, dx = u_v(\xi).$$

□

Theorem 44 *If v is a convex game then v satisfies the Fatou property if and only if the sigma core $\mathcal{C}^\sigma(v)$ is weak* dense in the core $\mathcal{C}(v)$.*

Proof. The core $\mathcal{C}(v)$ is weak* compact and hence the utility function u_v is given by this set. But as seen before in Section 4.6, u_v is Fatou if and only if $\mathcal{C}^\sigma(v) = \mathcal{C}(v) \cap L^1$ is weak* dense in $\mathcal{C}(v)$. □

In this case we can be more precise:

Theorem 45 *Suppose that v is a convex game with the Fatou property. If $A_1 \supset A_2 \ldots \supset A_n$ is a finite non-increasing family, there exists $\mathbb{Q} \in \mathcal{C}^\sigma(v)$ with $\mathbb{Q}(A_i) = v(A_i)$ for all $i \leq n$.*

Proof. . The proof of this theorem is not easy. It relies on the theorem of Bishop-Phelps. We take $\xi = \sum_{i=1}^n \mathbf{1}_{A_i}$. We now take an arbitrary $0 < \varepsilon < \frac{1}{8}$. By the Bishop-Phelps theorem there is $\eta \in L^\infty$, with $\|\xi - \eta\|_\infty < \varepsilon$ and η attains its infimum on $\mathcal{C}^\sigma(v)$. Of course we may replace η by $\eta + \varepsilon$ and hence we get $\eta \geq 0$. This means that there exists $\mathbb{Q}^0 \in \mathcal{C}^\sigma(v)$ such that:

$$\mathbb{E}_{\mathbb{Q}^0}[\eta] = \inf\{\mathbb{E}_{\mathbb{Q}}[\eta] \mid \mathbb{Q} \in \mathcal{C}^\sigma(v)\} = \int_0^\infty v(\eta > \alpha)\, d\alpha$$

This also implies $\int_0^\infty \mathbb{Q}^0(\eta > \alpha)\, d\alpha = \int_0^\infty v(\eta > \alpha)\, d\alpha$. Since $\mathbb{Q}^0 \in \mathcal{C}^\sigma(v)$ we have $\mathbb{Q}^0(\eta > \alpha) \geq v(\eta > \alpha)$ and therefore for almost every α we necessarily have $\mathbb{Q}^0(\eta > \alpha) = v(\eta > \alpha)$. Now for each $0 \leq k < n$ we take $k + \frac{1}{4} < \alpha < k + \frac{3}{4}$ where α has the above property and, since for such α we necessarily have $\{\eta > \alpha\} = A_{k+1}$, we get $\mathbb{Q}^0[A_{k+1}] = v(A_{k+1})$ for $k = 0 \dots n-1$. □

Remark 59 The conclusion of the theorem was already known for $\mu \in \mathcal{C}(v)$ (see Delbaen, [37]). The sigma core was studied by J. Parker, [108]. The results here – due to the author – extend her results, see also [39] for extra features. The next proposition, not contained in [39], is even better.

Proposition 36 *If $(A_n)_{n\geq 1}$ is a sequence in $\mathcal{F}$, if $A_n \downarrow A$ and if v is Fatou, there is an element $\mathbb{Q} \in \mathcal{C}^\sigma(v)$ such that for all n: $\mathbb{Q}[A_n] = v(A_n)$.*

Proof. Replacing the game v by w_A which is still Fatou (prove it as an exercise), we may reduce the problem to $A = \emptyset$. Afterwards we may glue such an element in $\mathcal{C}^\sigma(w_A)$ with a sigma additive measure in the core of the Fatou game v_A. So we suppose $A_n \downarrow \emptyset$ and remark that the Fatou property now implies that $v(A_n)$ tends to 0. We put $A_0 = \Omega$. For eack $k \geq 0$ we define a game v_k on the set $A_k \setminus A_{k+1}$. The game is for $B \subset A_k \setminus A_{k+1}$ given by the expression $v_k(B) = v(B \cup A_{k+1}) - v(A_{k+1})$. This game is still Fatou and therefore we may find an element $\mathbb{Y}_k \in \mathcal{C}^\sigma(v_k)$. Of course $\mathbb{Y}_k$ is supported on $A_k \setminus A_{k+1}$. Also $\|\mathbb{Y}_k\|_1 = v(A_k) - v(A_{k+1})$. Let us put $\mathbb{Q} = \sum \mathbb{Y}_k$. This sum converges in L^1 and defines a sigma additive measure of total mass equal to $\sum_{k\geq 0}(v(A_k) - v(A_{k+1})) = 1$. Clearly $\mathbb{Q}[A_k] = v(A_k)$ for all $k \geq 0$. We still have to check that $\mathbb{Q} \in \mathcal{C}(v)$. For each K we take an element $\mu_{K+1} \in \mathcal{C}(v_{A_{K+1}})$. A repeated application of Lemma 19 shows that $\sum_{j\leq k\leq K} \mathbb{Y}_k + \mu_{K+1} \in \mathcal{C}(v_{A_j})$, hence $\sum_{0\leq k\leq K} \mathbb{Y}_k + \mu_{K+1} \in \mathcal{C}(v)$. Since μ_{K+1} has total mass equal to $v_{A_{K+1}}$ the sequence tends to 0. Hence $\mathbb{Q} = \sum_{k\geq 0} \mathbb{Y}_k$ being the limit of sequence in $\mathcal{C}(v)$ is also in $\mathcal{C}(v)$. Since $\mathbb{Q}$ is already sigma-additive, the proof is complete. □

Exercise 22 The same ideas as above allow to prove a more difficult version. Suppose again that v is Fatou and convex. Let β be a *countable* ordinal and let $(A_\alpha)_{\alpha\leq\beta}$ be a non-increasing family of sets. Show that there is an element $\mathbb{Q} \in \mathcal{C}^\sigma(v)$ such that for all $\alpha \leq \beta$: $\mathbb{Q}[A_\alpha] = v(A_\alpha)$.

Remark 60 As the following theorem shows, the statement of the proposition is not always true for increasing sequences.

Theorem 46 *Let v be a convex game then $\mathcal{C}(v) \subset L^1$ if and only if $\mathcal{C}^\sigma(v)$ is weakly compact. This happens if and only if $A_n \uparrow \Omega$ implies $v(A_n) \uparrow 1$.*

Proof. Fairly easy since $\mathcal{C}^\sigma(v)$ is weakly compact if and only if it is uniformly integrable. This is the case if and only if $A_n \uparrow \Omega$ implies $v(A_n) = \inf_{\mathbb{Q}\in\mathcal{C}^\sigma(v)} \mathbb{Q}(A_n) \uparrow 1$. □

Corollary 9 *Suppose v has the weak compactness property from the previous theorem. The following holds:*

1. *If $A_n \uparrow A$ then $v(A_n) \uparrow v(A)$*
2. *If $\{A_i \mid i \in I\}$ is totally ordered then there is probability $\mathbb{Q} \in \mathcal{C}(v)$ such that for all $i \in I$: $\mathbb{Q}(A_i) = v(A_i)$.*

   Conversely each of these properties implies that v has the weak compactness property

Proof. The direct implications are rather trivial since the first follows from the uniform integrability of the set $\mathcal{C}(v) \subset L^1$ and the second is a restatement of Theorem 40 and $\mathcal{C}^\sigma(v) = \mathcal{C}(v)$. The converse is less easy. The first item implies weak compactness since this only requires the case $A = \Omega$. For the second item let $A_n \uparrow \Omega$. The family $\{A_n \mid n \geq 1\}$ is totally ordered. Therefore there is $\mathbb{Q} \in \mathcal{C}^\sigma(v)$ with $\mathbb{Q}[A_n] = v(A_n)$ for all n, hence $v(A_n) \uparrow 1$. □

Remark 61 The above corollary was already present in [37] and in [81] with a slight correction as in their subsequent paper in 1974. We remark that the conditions in [81], especially their condition (4) implies that the core of the associated game is weakly compact in the topology $\sigma(L^1, L^\infty)$. This compactness condition is stronger than the "tightness" condition as e.g. in [20]. The reader should not confuse these different notions of compactness.

Example 32 If $1 \leq \beta \leq +\infty$ then $v(A) = \mathbb{P}(A)^\beta$ defines a convex game. If $\beta = \infty$, the σ-core is the whole family of absolutely continuous probabilities, whereas if $\beta = 1$, $\mathcal{C}^\sigma$ is the singleton $\{\mathbb{P}\}$. We also have for nonnegative ξ: $u(\xi) = \int_0^\infty \mathbb{P}(\{\xi > x\})^\beta \, dx$. These utility functions, or better their related risk measures, were studied by Delbaen, [39],[48] and many others, see [26] and the references therein.

Proposition 37 *If $f : [0,1] \to [0,1]$ is a convex function such that $f(0) = 0$ and $f(1) = 1$, then $v(A) = f(\mathbb{P}(A))$ defines a convex game. The set $\mathcal{C}^\sigma$ is weakly compact iff f is continuous at 1. Conversely if $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless, then the game $v(A) = f(\mathbb{P}[A])$ is convex if and only if f is convex.*

Proof. Let us show that such a convex function f indeed defines a convex game. Let A, B be given an define $\gamma = \mathbb{P}[A \cap B]$, $\alpha = \mathbb{P}[A \setminus (A \cap B)], \beta = \mathbb{P}[B \setminus (A \cap B)]$, then $\mathbb{P}[A \cup B] = \alpha + \beta + \gamma$. We have to show that $f(\alpha + \gamma) + f(\beta + \gamma) \leq f(\alpha + \beta + \gamma) + f(\gamma)$. Let μ be the derivative of f, i.e. the nonnegative measure defined as $\mu([0, x]) = f(x)$. Because f is convex we have that $\mu]x, x + y]$ is a non-decreasing function of y. So $f(\alpha+\gamma)-f(\gamma) = \mu]\gamma, \alpha+\gamma]) \leq \mu(]\beta+\gamma, \alpha+\beta+\gamma]) = f(\alpha+\beta+\gamma)-f(\beta+\gamma)$. This is precisely what we needed to prove. The weak compactness follows immediately from Theorem 46. Conversely suppose that we work in an atomless space and suppose that $f(\mathbb{P}[A]) = v(A)$ defines a convex game. This means that $f(0) = 0$ and $f(1) = 1$. Let $\alpha_1, \alpha_2, \alpha_3$ be three nonnegative numbers such that $\alpha_1 + \alpha_2 + \alpha_3 \leq 1$. Convexity of v implies that $f(\alpha_1 + \alpha_2 + \alpha_3) + f(\alpha_3) \geq f(\alpha_1 + \alpha_3) + f(\alpha_2 + \alpha_3)$. Here we use that the space is atomless since then we can realise these numbers as probabilities of sets $A \setminus (A \cap B), B \setminus (A \cap B), A \cap B$. If we put $\alpha_3 = 0$ we get $f(\alpha_1 + \alpha_2) \geq f(\alpha_1) + f(\alpha_2)$. This already implies that f is monotone and hence it is a Borel measurable function. If we take $\alpha_1 = \alpha_2$ we get $f(2\alpha_1 + \alpha_3) + f(\alpha_3) \geq 2f(\alpha_1 + \alpha_3)$ We can rewrite this as $f(\frac{x+y}{2}) \leq \frac{1}{2}(f(x) + f(y))$ for all $0 \leq x, y \leq 1$. Because of monotonicity this is enough to prove convexity. (The latter part is left to the reader as an exercise: first prove that the convexity inequality holds for diadic numbers, then extend). $\square$

Example 33 An example of such a function is:

$$f(x) = \begin{cases} 0 \text{ for } x \leq 1 - \frac{1}{k} \\ k\left(x - \left(1 - \frac{1}{k}\right)\right) \text{ for } 1 - \frac{1}{k} \leq x \leq 1\,, \end{cases}$$

where of course $k \geq 1$. We will check that $\mathcal{C}^\sigma$ is $\mathcal{S}_k$ of Example 9. (We remark that the sets $\mathcal{S}_{p,k}$ cannot be obtained via convex games, the related risk measures are not commonotone, see [39] for a proof). The utility functions related to such “distorted” probability measures were introduced by Yaari [133] and Denneberg, see [48]. Denneberg used them as premium calculation principles. Later Denneberg extended the theory of non-linear expectations, also called Choquet integrals, see [49]. In Section 7.3, we will give a more detailed analysis of these distorted probabilities.

Proposition 38 *If $f = (x-s)^+/(1-s)$ and if the space is atomless, the core is the set of all probability measures $\mathbb{Q}$ with $\frac{d\mathbb{Q}}{d\mathbb{P}} \le \frac{1}{s}$.*

Proof. An element $\mathbb{Q}$ is in the core if and only if for $\mathbb{P}[A] \ge 1-s$ we have $\mathbb{Q}[A] \ge \frac{\mathbb{P}[A]-s}{1-s}$. This can be rewritten as $\mathbb{Q}[B] \le \frac{1}{1-s}\mathbb{P}[B]$ for all B with $\mathbb{P}[B] \le s$. Since the space is atomless, any set can be written as the disjoint union of sets of measure smaller than $1-s$, hence we get $\mathbb{Q}[B] \le \frac{1}{1-s}\mathbb{P}[B]$ for all sets B. This is equivalent to $\frac{d\mathbb{Q}}{d\mathbb{P}} \le \frac{1}{1-s}$. The converse is easier and can be done by direct calculation.

7.2 Commonotone utility functions

According to its behaviour with respect to commonotone variables we define

Definition 23 *A coherent utility function u is called commonotone if $u(\xi+\eta) = u(\xi) + u(\eta)$ for every commonotone couple (ξ, η).*

Remark 62 We could have given the definition of commonotonicity for concave utility functions. However such concave utility functions are then necessary positively homogeneous since for all ξ we would have $u(2\xi) = u(\xi+\xi) = u(\xi) + u(\xi) = 2u(\xi)$. If u is concave, this relation implies that u is positively homogeneous.

Theorem 47 *If v is a convex game then the coherent utility function u_v is commonotone.*

Proof. Let ξ, η be commonotone and let $\mu \in \mathcal{C}(v)$ be an element such that for all $x \in \mathbb{R}$: $\mu(\xi+\eta \ge x) = v(\xi+\eta \ge x)$. Because ξ, η are monotone functions of $\xi+\eta$ we have that each set of the form $\{\xi \ge y\}$ is of the form $\{\xi+\eta \ge x\}$ for some x. So we have that $\mu(\xi \ge y) = v(\xi \ge y)$ (and the same applies to η). So we have that

$$\mu(\xi) = u_v(\xi), \mu(\eta) = u_v(\eta), \mu(\xi+\eta) = u_v(\xi+\eta).$$

This implies $u_v(\xi+\eta) = u_v(\xi) + u_v(\eta)$ as desired. □

The converse is also true as shown by David Schmeidler, [124].

Theorem 48 *A coherent utility function u originates from a convex game iff u is commonotone.*

Proof. Let u be commonotone and let it be defined by the convex, weak* compact set $\mathcal{S} \subset \mathbf{P^{ba}}$. Let us put for $A \in \mathcal{F}$: $v(A) = u(\mathbf{1}_A)$. Take two sets $A, B \in \mathcal{F}$. Because $\mathbf{1}_{A\cap B}$ and $\mathbf{1}_{A\cup B}$ are commonotone and because u is coherent we have

$$u(\mathbf{1}_A) + u(\mathbf{1}_B) \le u(\mathbf{1}_A + \mathbf{1}_B) = u(\mathbf{1}_{A\cap B} + \mathbf{1}_{A\cup B}) = u(\mathbf{1}_{A\cap B}) + u(\mathbf{1}_{A\cup B}).$$

The game v is therefore convex. Let us denote its core by $\mathcal{C}(v)$. We have to show that $\mathcal{C}(v) = \mathcal{S}$ or what is the same $u = u_v$. Because both functions are continuous for the norm topology on L^∞, we only need to check this for random variables ξ taking a finite number of values. Because both are monetary we may suppose that $\xi = \sum_1^n \alpha_i \mathbf{1}_{A_i}$, where $\alpha_i > 0$ and $A_1 \supset A_2 \supset \ldots \supset A_n$. By commonotonicity of both u and u_v we have

$$u(\xi) = \sum_i \alpha_i u(\mathbf{1}_{A_i}) = \sum_i \alpha_i v(A_i) = u_v(\xi)$$

□

Corollary 10 *Suppose that the probability space is atomless. Let u be commonotone and law determined, then there is a convex function $f\colon [0,1] \to [0,1]$ with $f(0) = 0, f(1) = 1$ such that the game $f \circ \mathbb{P}$ describes u.*

Proof. This is a particular case of Kusuoka's theorem, [96], see also [128]. It can be proved as follows. Because u is commonotone, Schmeidler's theorem says that it is given by the core of a convex game v. Because it is law determined, the value $v(A)$ is given by a function $f(\mathbb{P}[A])$. Because the game is convex, the function f must be convex by Proposition 37. □

Remark 63 We refer to [48] and [49] for another proof of the preceding corollary. Older versions were due to Yaari, [133].

Proposition 39 *Suppose that the probability space contains at least three non negligible disjoint sets (this is certainly the case if it is atomless). If u is Gâteaux differentiable at nonconstant elements $\xi \in L^\infty$ and is commonotone, then u is linear, i.e. it is given by $u(\xi) = \mu[\xi]$ for some finitely additive probability measure $\mu \in \mathbf{ba}$. If u is also Fatou, then necessarily μ is $\sigma-$additive and is absolutely continuous with respect to $\mathbb{P}$.*

Proof. Since u is commonotone, we may apply Schmeidler's theorem. This shows that for $\xi \geq 0$, $u(\xi) = \int_0^\infty v(\{\xi \geq x\})\,dx$, where v is a convex game. Take a set A such that $0 < \mathbb{P}[A] < 1$. Since u is differentiable at $\mathbf{1}_A$, we have that there is a unique element $\mu \in \mathcal{C}(v)$ with $\mu(A) = v(A)$. If $B \subset A$ or $A \subset B$, we have the existence of an element $\nu \in \mathcal{C}(v)$ with both $\nu(A) = v(A)$ and $\nu(B) = v(B)$. But then we must have $\nu = \mu$. From here we deduce that for two disjoint elements B_1 and B_2 such that $\mathbb{P}[B_1 \cup B_2] < 1$ have the existence of an element $\mu \in \mathcal{C}(v)$ such that $\mu(B_1 \cup B_2) = v(B_1 \cup B_2)$ and therefore also $\mu(B_1) = v(B_1), \mu(B_2) = v(B_2)$. This shows that $v(B_1) + v(B_2) = \mu(B_1) + \mu(B_2) = \mu(B_1 \cup B_2) = v(B_1 \cup B_2)$. In case $\mathbb{P}[B_1 \cup B_2] = 1$ we can, by hypothesis, split at least one of the two sets in two strictly smaller non negligible sets, say $C_1 \cup C_2 = B_1$. Then we have the existence of a unique element $\mu \in \mathcal{C}(v)$ that must satisfy $\mu(C_1) = v(C_1)$. This element then satisfies $\mu(B_1) = v(B_1)$ and hence also $\mu(C_2) = v(C_2)$. But μ must also satisfy $v(B_2 \cup C_1) = \mu(B_2 \cup C_1)$ and hence also $v(B_2) = \mu(B_2)$. As a conclusion we get that $v(B_1) + v(B_2) = \mu(B_1) + \mu(B_2) = 1$. So v is additive, concluding the proof. □

Remark 64 The previous theorem was observed by Sebastian Maass, see [?]. The proof shows that we only used differentiability at non constant indicator functions. It shows that differentiability and commonotonicity are not very compatible. This will have an influence on the solution of the capital allocation problem. In [51], Deprez and Gerber used functions that were Gâteaux differentiable. The previous results show that the ideas in their paper must be applied with care.

7.3 Distortion

We already looked at games of the form $v(A) = f(\mathbb{P}[A])$. Such games coming from "distorted" probabilities play an important role. They characterise the commonotone law determined utility functions. We now analyse the representation from Chapter 4 with the extra information that u is commonotone. In doing so we get a relation with TailVaR. We use the notation $v = f \circ \mathbb{P}$ where $f : [0,1] \to [0,1]$ is convex, $f(0) = 0, f(1) = 1$. For $0 \leq k \leq 1$ we define $f_k : [0,1] \to [0,1]$ as $f_k(x) = 0$ for $0 \leq x < k$, $f_k(x) = (x-k)/(1-k)$ for $k \leq x < 1$. This definition yields $f_0(x) = x$, $f_1(x) = 0$ for $x < 1$. The corresponding games are denoted by v_k. So $\mathcal{C}(v_k)$ is defining TailVaR for the level $1-k$.

Proposition 40 *Every convex function $f : [0,1] \to [0,1]$, $f(0) = 0, f(1) = 1$ can be written as a mixture of the functions f_k. More precisely there is a probability measure λ on $[0,1]$ such that $f = \int_{[0,1]} f_k \, \lambda(dk)$. f is continuous at 1 if and only if $\lambda(\{1\}) = 0$.*

Proof. If f is not continuous at 1, then we define $g(x) = f(x)/f(1-)$ for $x < 1$ and $g(1) = 1$. We find that $f = f(1-)g + (1 - f(1-))f_1$. This reduces the problem to continuous functions f. Elementary properties of integration theory and the almost sure differentiability of f – in one dimension saying that the left and right derivatives exist – then yields:

$$\begin{aligned}
f(x) &= \int_0^x f'(u)\,du \qquad \text{for } 0 \le x < 1 \\
&= \int_0^x \mu[0,u]\,du \qquad \text{where } \mu[0,u] = f'(u) \text{ is the right derivative} \\
&= \int_0^x \int_{[0,u]} \mu(ds)\,du \\
&= \int_{[0,1[} \mu(ds) \int_{s \le u \le x} du \\
&= \int_{[0,1[} \mu(ds)(x-s)^+ \\
&= \int_{[0,1[} \mu(ds)(1-s) f_s(x) \\
&= \int_{[0,1[} \lambda(ds) f_s(x) \qquad \text{where } d\lambda = (1-s)\,d\mu.
\end{aligned}$$

□

Remark 65 The above formula is a special case of the more general result of Kusuoka, [96] as we already pointed out in Chapter 5.

Every distorted probability $v = f \circ \mathbb{P}$ defines a Fatou utility function u. The above shows that

$$u(\xi) = \int_{[0,1]} \lambda(ds)\, u_s(\xi) \qquad \text{where } u_s \text{ is TailVaR with level } 1-s.$$

This formula can be proved as follows. We suppose that ξ is nonnegative.

$$\begin{aligned}
u(\xi) &= \int_0^\infty v(\xi \geq x)\, dx = \int_0^\infty f(\mathbb{P}[\xi \geq x]\, dx \\
&= \int_0^\infty \int_{[0,1]} f_s(\mathbb{P}[\xi \geq x])\, d\lambda\, dx \\
&= \int_{[0,1]} \int_0^\infty f_s(\mathbb{P}[\xi \geq x])\, dx\, d\lambda \\
&= \int_{[0,1]} u_s(\xi)\, d\lambda.
\end{aligned}$$

Since the quantile functions q_α allow to write TailVaR we get, at least for functions f that are continuous at 1:

$$\begin{aligned}
u(\xi) &= \int_{[0,1]} \lambda(ds)\, u_s(\xi) \\
&= \int_{[0,1]} \lambda(ds) \frac{1}{1-s} \int_0^{1-s} q_\alpha(\xi)\, d\alpha \\
&= \int_{[0,1]} \mu(ds) \int_0^{1-s} q_\alpha(\xi)\, d\alpha \\
&= \int_{[0,1]} d\alpha\, q_\alpha(\xi) \mu[0, 1-\alpha] \\
&= \int_{[0,1]} d\alpha\, q_\alpha(\xi) f'(1-\alpha).
\end{aligned}$$

These are of the form $\int d\alpha\, q_\alpha \phi(\alpha)$ where ϕ is a non-increasing function. See Kupper et al [2] for the investigation of a wider class of such functions. We invite the reader to extend this formula to the case where f has a discontinuity at 1, more precisely when we need a Dirac measure at 1 for the derivative of f. Hint: remember that $q_0 = \text{ess.inf}$.

Proposition 41 *Let f be a continuous distortion function and let $f'(1)$ be its left derivative at 1 (this can be $+\infty$). Then for $1-k = 1/f'(1)$, we have by convexity of f, that $f \geq f_k$. It follows that the utility function associated with f is bigger than TailVaR at level $1-k$*

The proof is already contained in the statement, so we omit the details.

Remark 66 As already mentioned above, typical examples for distortion are $f(x) = x^p$, where $1 \leq p \leq \infty$, the case $p = 1$ gives the expected value, the case $p = \infty$ gives the case ess.inf. For p an integer we can write the utility function in a different way. Suppose that $\xi \geq 0$ and let $\xi_1, \ldots, \xi_n$ be independent copies of ξ. We have that $u(\xi) = \int_0^\infty f(\mathbb{P}[\xi > x])\,dx = \int_0^\infty (\mathbb{P}[\xi > x])^n\,dx = \int_0^\infty \mathbb{P}[\min(\xi_1, \ldots, x_n) > x]\,dx$. This means that $u(\xi) = \mathbb{E}[\min(\xi_1, \ldots, \xi_n)] = \mathbb{E}[\xi_{[1]}]$, where $\xi_{[1]}$ is the first order statistic (we drop the fixed value n), i.e. the smallest among n values. The previous proposition now gives $\mathbb{E}[\min(\xi_1, \ldots, \xi_n)] \geq u'(\xi)$, where u' is TailVar at level $1/n$. We leave the economic interpretation to the reader. Other examples of distortion functions are $f(x) = 1-(1-x)^s$ where $0 < s < 1$. They have a sharp increase at $x = 1$.

Remark 67 The other order statistics $(\xi_{[2]} \ldots, \xi_{[n]})$, do not define utility functions since the distortion function f defined by the relation $f_{[j]}(\mathbb{P}[\xi \geq x]) = \mathbb{P}[\xi_{[j]} \geq x]$ (for all $\xi \in L^\infty$), is not a convex function for $j \geq 2$. This can be expected since the n-th order statistic is the maximum among n values and this is certainly not a cautious value.

Exercise 23 Find $f_{[j]}$ and prove that it is not convex. You may suppose that the probability space is atomless.

From now on and until the end of this chapter, we will make the assumption that the probability space is atomless.

Proposition 42 *Let w be a convex game such that $\mathcal{C}(w)$ is a weakly compact set of L^1. Let v_1 be defined as above, i.e. $v_1(A) = 0$ if $\mathbb{P}[A] < 1$ and $v_1(A) = 1$ if $\mathbb{P}[A] = 1$. Let $0 \leq t < 1$ and let $v = t\,w + (1-t)v_1$. The game v is Fatou but $\mathcal{C}^\sigma(v)$ has no extreme points. One can write*

$$\mathcal{C}^\sigma(v) = \{t\mu + (1-t)\mathbb{Q} \mid \mu \in \mathcal{C}(w); \mathbb{Q} \ll \mathbb{P} \text{ an arbitrary probability measure}\}.$$

Proof. Clearly

$$\mathcal{C}^\sigma(v) \supset \{t\mu + (1-t)\mathbb{Q} \mid \mu \in \mathcal{C}(w); \mathbb{Q} \ll \mathbb{P} \text{ an arbitrary probability measure}\}.$$

Furthermore the right hand side is convex and closed in L^1. If we calculate for $\xi \geq 0$:

$$\inf\{\mathbb{E}_\nu[\xi] \mid \nu = t\mu + (1-t)\mathbb{Q}; \mu \in \mathcal{C}(w); \mathbb{Q} \in \mathbf{P}\},$$

we find

$$\begin{aligned} &t\,u_w(\xi) + (1-t)\,\text{ess.inf}(\xi) \\ &= \int_0^\infty t\,w(\xi > x)dx + \int_0^\infty (1-t)\,v_1(\xi > x)dx = \int_0^\infty v(\xi > x)dx = u_v(\xi). \end{aligned}$$

It is well known – and an easy exercise – that the set of all absolutely continuous probability measures on an atomless space has no extreme points. This ends the proof. □

This means that for distortions we may limit the study to functions f that are continuous at 1. In this case there are enough extreme points since the core is weakly compact. In [37] we studied the extreme points for such games but since then much better is known, see e.g. Carlier and Dana,[26], Marinacci and Montrucchio [**?**], Brüning and Denneberg, [24] and the references given in these papers. Many of the results given in these papers rest on the classic papers of Ryff,[121]. The rest of this section is devoted to the study of extreme points of the core. The presentation is different from [26] and uses a little bit more functional analysis. The reader not familiar with these concepts can skip the proofs. Along the road, we give a – maybe new – proof of Ryff's theorem.

The first step consists in giving a description of the core of the game $v = f \circ \mathbb{P}$. Since v is a mixture of TailVaR at different levels, we expect that the core is a similar mixture, in other words we are aiming for a generalisation of Proposition 40.

Theorem 49 *Let f be continuous at* 1 *and let v be described as $v(A) = \int d\lambda\, v_s(A)$ where v_s describes TailVaR at level* $(1-s)$. *We have*

$$\mathcal{C}(v) = \left\{ \int d\lambda\, \mathbb{Q}_s \mid s \to \mathbb{Q}_s \text{ is Bochner measurable and } \lambda \text{ a.s. } \mathbb{Q}_s \in \mathcal{C}(v_s) \right\}$$

In other words the operator

$$\begin{aligned} T : & L^1([0,1] \times \Omega, \lambda \times \mathbb{P}) \to L^1(\Omega, \mathbb{P}) \\ & (T(\phi))(\omega) = \int_{[0,1]} \phi(s,\omega)\, d\lambda(s) = \mathbb{E}\left[\phi \mid \{\emptyset, [0,1]\} \times \mathcal{F}\right], \end{aligned}$$

maps the set

$$\mathcal{D} = \{\phi \mid 0 \le \phi(s,\omega) \le \frac{1}{1-s}, \int_\Omega \phi(s,.)\, d\mathbb{P} = 1, \lambda \text{ a.s. } \}$$

onto $\mathcal{C}(v)$. The set $\mathcal{D}$ is weakly compact in $L^1(\lambda \times \mathbb{P})$.

Proof. A direct calculation shows $T(\mathcal{D}) \subset \mathcal{C}(v)$. The set $T(\mathcal{D})$ is clearly convex. For each ξ having a continuous law, we have the existence of an element $\mathbb{Q}_s \in \mathcal{C}(v_s)$ such that $u_{v_s}(\xi) = \mathbb{E}_{\mathbb{Q}_s}[\xi]$. The element $\mathbb{Q}_s$ is given in an explicit way, $\phi(s,.) = \frac{1}{1-s}\mathbf{1}_{\xi \le q_{1-s}(\xi)}$. This means that $s \to \mathbb{Q}_s$ is Bochner measurable. But then we have

$$u(\xi) \le u_{T(\mathcal{D})}(\xi) \le \int d\lambda\, \mathbb{Q}_s(\xi) = \int d\lambda\, u_s(\xi) = u(\xi).$$

Since random variables with continuous laws are norm dense in L^∞ we get the equality $u(\xi) = u_{T(\mathcal{D})}(\xi)$ for all elements in L^∞. This shows that $\mathcal{C}(v)$ is contained in the closure of $T(\mathcal{D})$. So it remains to show that $T(\mathcal{D})$ is norm closed in L^1. This follows from the weak compactness property of $\mathcal{D}$. It is clear that the set $\mathcal{D}$ is convex and closed in $L^1(\lambda \times \mathbb{P})$. To prove weak compactness or uniform integrability we need to find a function $K : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{x\to\infty} \frac{K(x)}{x} = +\infty$ and such that $\sup_{\phi\in\mathcal{D}} \mathbb{E}_{\lambda\times\mathbb{P}}[K(\phi)] < \infty$. We will find a function K under the form $K(x) = x\,p(x)$ where p is nondecreasing and tends to ∞ at ∞. We take $p(x)$ so that p is increasing and $\sum_n \lambda([1-2^{-n}, 1-2^{-(n+1)}))p(2^{n+1}) < \infty$. This is possible since $\sum_{n\ge 0} \lambda([1-2^{-n}, 1-2^{-(n+1)})) = 1$. The function K is then convex. Stochastic dominance or Choquet theory (or a good application of Jensen's inequality) shows that for every s we have

$$\int_\Omega K(\phi(s,.))\, d\mathbb{P} \le K\left(\frac{1}{1-s}\right)(1-s).$$

From here we get for every $\phi \in \mathcal{D}$:

$$\begin{aligned}
\mathbb{E}_{\lambda\times\mathbb{P}}[K(\phi)] &= \int_{[0,1]} \lambda(ds) \int_\Omega d\mathbb{P}\, K(\phi(s,\omega)) \\
&\le \int_{[0,1]} \lambda(ds)\, K\left(\frac{1}{1-s}\right)(1-s) \\
&\le \int_{[0,1]} \lambda(ds)\, p\left(\frac{1}{1-s}\right) \\
&\le \sum_{n\ge 0} \lambda([1-2^{-n}, 1-2^{-(n+1)}))\, p(2^{n+1}) < \infty.
\end{aligned}$$

□

Exercise 24 If λ is supported on a set $[0,u]$ with $u < 1$, then $\mathcal{D}$ consists of functions that are bounded by $\frac{1}{1-u}$. For general λ, use an approximation argument to show that elements in $\mathcal{D}$ can be truncated by $\frac{1}{1-u}$ in a uniform way. Hint: simply make us of the fact that $\lambda([u,1)) \to 0$.

Exercise 25 This exercise or remark is only for the fanatics. We used Bochner measurability (i.e. limit of a sequence of random variables taking only finitely many values in L^1), to get jointly measurable functions. However one can show that if $\mathbb{Q}_s$ is only Pettis measurable, i.e. for every $\xi \in L^\infty$, $s \to \int \xi \, d\mathbb{Q}_s$ is measurable, then for every jointly measurable function $g(s,\omega)$, the function $s \to \int d\mathbb{Q}_s \, g(s,.)$ is still measurable. So we may define $\int_{[0,1]} d\lambda(s) \int d\mathbb{Q}_s g(s,.)$. This defines a measure that is absolutely continuous with respect to $\lambda \times \mathbb{P}$ and its RN-derivative ϕ, satisfies, for almost every s: $\phi(s,.) = \frac{d\mathbb{Q}_s}{d\mathbb{P}}$. As long as we are only concerned about integrals, we may replace Pettis measurability by Bochner measurability.

Theorem 50 *We use the notation of Theorem 49. The extreme points of the set $\mathcal{D}$ are precisely the functions ϕ with the property $\phi(s,\omega)$ is either 0 or $\phi(s,\omega) = \frac{1}{1-s}$. If ϕ is not an extreme point of $\mathcal{D}$, then $T(\phi)$ is not an extreme point of $\mathcal{C}(v)$. Consequently for an extreme point $\mathbb{Q} \in \mathcal{C}(v)$, there is exactly one point $\phi \in \mathcal{D}$ such that $T(\phi) = \mathbb{Q}$. The point ϕ is an extreme point of $\mathcal{D}$ and hence is of the form $\phi(s,\omega) = \frac{1}{1-s}\mathbf{1}_E$ where the sections $E_s = \{\omega \mid (s,\omega) \in E\}$ satisfy $\mathbb{P}[E_s] = 1-s$.*

Proof Only the first part has to be proved. Take $\phi \in \mathcal{D}$ and suppose that the set $E = \{(s,\omega) \mid 0 < \phi(s,\omega) < \frac{1}{1-s}\}$ has strictly positive measure. We will show that ϕ is not extreme in $\mathcal{D}$ and that $T(\phi)$ is not extreme in $\mathcal{C}(v)$. The function $s \to \mathbb{P}[E_s]$ is not negligible and by making the set E smaller we get a nonzero set such that either $\mathbb{P}[E_s] = 0$ or $\mathbb{P}[E_s] \geq \delta > 0$ and $E \subset \{(s,\omega) \mid 0 < \varepsilon < \phi(s,\omega) < \frac{1}{1-s}(1-\varepsilon)\}$, where δ, ε are chosen small enough. Let E^* be set $\{s \mid \mathbb{P}[E_s] > 0\}$. The injection $L^1(E^*,\lambda) \to L^1(E, \lambda \times \mathbb{P}); k \to k(s)\mathbf{1}_E$ is an isomorphism into (since $\delta > 0$) and hence the image is closed. Let η be a $[0,1]$ uniformly distributed random variable defined on Ω. Clearly $\eta\mathbf{1}_E$ is not in the image of $L^1(E^*)$. By the Hahn-Banach theorem there is a non trivial element $h \in L^\infty(E, \lambda \times \mathbb{P})$, $\|h\|_\infty = 1$ such that for all $k \in L^1(E^*)$:

$$\int_E k(s)h(s,\omega)\, d\lambda \, d\mathbb{P} = 0,$$

and $\int_E \eta(\omega)h(s,\omega)\, d\lambda \, d\mathbb{P} \neq 0$. This implies that for λ almost all s, the integral $\int_{E_s} h(s,\omega)\, d\mathbb{P} = 0$ but it also implies that $\int d\lambda(s)h(s,\omega)$ is not identically zero. The functions $\phi + \varepsilon h$ and $\phi - \varepsilon h$ are both in $\mathcal{D}$ and are different from ϕ. This shows that ϕ is not extreme. The elements $\int d\lambda(s)\,(\phi(s) + \varepsilon h(s,\omega))$ and $\int d\lambda(s)\,(\phi(s) - \varepsilon h(s,\omega))$ are different, are in $\mathcal{C}(v)$ proving that $T(\phi)$ is not extreme in $\mathcal{C}(v)$. □

Definition 24 *We say that a point $\mathbb{Q} \in \mathcal{C}(v)$ is exposed if there is $\xi \in L^\infty$ such that $\mathbb{Q}[\xi] = u_v(\xi)$ but for all other elements $\mathbb{Q}' \in \mathcal{C}(v)$ we have $\mathbb{E}_\mathbb{Q}[\xi] < \mathbb{E}_{\mathbb{Q}'}[\xi]$. We say that ξ is an exposing functional.*

Since $\mathcal{C}(v)$ is weakly compact, the exposed points are enough to recover the whole set. We have that the set of exposed points is weakly dense in the set of extreme points, see [52].

In the next paragraphs we will show that the extreme points of $\mathcal{C}(v)$ are exposed and we will describe their structure. The plan is the following. We already know that an extreme point of $\mathcal{C}(v)$ necessarily is the integral of extreme points of $\mathcal{C}(v_s)$. Then we will see that for exposed points something more can be said. Since exposed points are dense in the extreme points, we can use a limit result and finally we will then show that an extreme point is exposed. This result goes back to Ryff, see [121].

So let $\mathbb{Q} = \int d\lambda\, \mathbb{Q}_s \in \mathcal{C}(v)$, where of course we suppose that $\mathbb{Q}_s \in \mathcal{C}(v_s)$. In case $\mathbb{Q}$ is extreme we must have that for λ almost all s, the element $\mathbb{Q}_s$ is extreme in $\mathcal{C}(v_s)$ and hence defined by a set B_s of probability $1-s$.

The next step is to show that for an exposed point $\mathbb{Q}$ we know more about the sets B_s. Suppose that $\xi \geq 0$ is an exposing functional for the exposed point $\mathbb{Q}$. We then have for some Bochner measurable mapping $\mathbb{Q}_s$ with $\mathbb{Q}_s \in \mathcal{C}(v_s)$:

$$u_v(\xi) = \mathbb{Q}[\xi] = \int d\lambda\, \mathbb{Q}_s[\xi] \geq \int d\lambda\, u_{v_s}(\xi) = u_v(\xi).$$

Hence for λ almost all s we have $\mathbb{Q}_s[\xi] = u_{v_s}(\xi) = \int_0^\infty v_s(\xi \geq x)\, dx$. This implies that for all $x \geq 0$: $\mathbb{Q}_s[\xi \geq x] = v_s(\xi \geq x)$. But for λ almost all s we must then have that $\mathbb{Q}_s$ is an exposed point of $\mathcal{C}(v_s)$. If it would not be the case, then the minimising probability $\mathbb{Q}_s$ is not the unique element in $\mathcal{C}(v_s)$ that allows to calculate the value $u_{v_s}(\xi)$. This implies that the quantile q_{1-s}, at level $1-s$ satisfies $\mathbb{P}[\xi < q_{1-s}(\xi)] < 1-s < \mathbb{P}[\xi \leq q_{1-s}(\xi)]$. This can only happen for a countable number of points s. Furthermore one of these points must then have a positive λ-measure. Take one of these points say t. For this t we replace $\mathbb{Q}_t$ by a different measure $\mathbb{Q}'{}_t$ still minimising in the sense $u_{v_t}(\xi) = \mathbb{E}_{\mathbb{Q}'_t}[\xi]$. From the discussion on TailVaR we know that on the set $\{\xi < q_{1-t}\}$ the measures $\mathbb{Q}_t$ and $\mathbb{Q}'_t$ have density $1/(1-t)$ and on the set $\{\xi > q_{1-t}\}$, their densities are both 0. Take now the integral

$$\mathbb{Q}' = \int_{s \neq t} d\lambda\, \mathbb{Q}_s + \lambda\{t\}\mathbb{Q}'_t$$

This measure is different from $\mathbb{Q}$ since it is different from it on the set $\{\xi \le q_{1-t}\} \setminus \{\xi < q_{1-t}\}$. The description of the core gives us that $\mathbb{Q}' \in \mathcal{C}(v)$ and $\mathbb{E}_{\mathbb{Q}}[\xi] = \mathbb{E}_{\mathbb{Q}'}[\xi]$ by construction. This is a contradiction to the fact that $\mathbb{Q}$ was exposed.
So we get that for λ almost all s, the measure $\mathbb{Q}_s$ is either supported by $\{\xi < q_{1-s}\}$ or by $\{\xi \le q_{1-s}\}$. This means that the sets $\{\frac{d\mathbb{Q}_s}{d\mathbb{P}} > 0\}$ form a decreasing family of sets. We can express this by saying that the functions (indicator functions!) $(1-s)\frac{d\mathbb{Q}_s}{d\mathbb{P}}$ are decreasing.

If $\mathbb{Q} \in \mathcal{C}(v)$ there is a sequence of convex combinations of exposed points that tends to $\mathbb{Q}$. By taking more convex combinations, we get for $\mathbb{Q}$ a representation $\mathbb{Q} = \int d\lambda\, \mathbb{Q}_s$ with $(1-s)\frac{d\mathbb{Q}_s}{d\mathbb{P}}$ decreasing. This property is valid for all elements of the core $\mathcal{C}(v)$.

In case $\mathbb{Q}$ is an extreme point this allows us to write $\mathbb{Q} = \int d\lambda\, \mathbb{Q}_s$ where $\frac{d\mathbb{Q}_s}{d\mathbb{P}} = \frac{1}{1-s}\mathbf{1}_{B_s}, \mathbb{P}[B_s] = 1-s$ and where the system B_s is now *decreasing.* Since the space $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless, we have the existence of a random variable ξ such that for λ almost all s: $\{\xi < q_{1-s}\} = B_s = \{\xi \le q_{1-s}\}$. This ξ exposes the point $\mathbb{Q}$ in $\mathcal{C}(v)$ and also shows the first lines of the next theorem.

Theorem 51 *If v is a distortion game $v = f \circ \mathbb{P}$ where f is continuous at 1, then all extreme points of $\mathcal{C}(v)$ are exposed. Exposed points are characterised as $\mathbb{Q} = \int d\lambda\, \mathbb{Q}_s$ where $\frac{d\mathbb{Q}_s}{d\mathbb{P}} = \frac{1}{1-s}\mathbf{1}_{B_s}, \mathbb{P}[B_s] = 1-s$ and where the system B_s is decreasing. A random variable $\eta \in \mathcal{C}(v)$ is an exposed point if and only if it has the same law as the function $f'\colon [0,1] \to \mathbb{R}_+$. The set of extreme points, $\partial\mathcal{C}(v)$, is a closed G_δ set in $\mathcal{C}(v)$.*

Proof. Only the last sentence has to be proved. If $\eta = \frac{d\mathbb{Q}}{d\mathbb{P}}$ is an exposed point, we can write $\eta = \int d\lambda\, \frac{1}{1-s}\mathbf{1}_{B_s}$ where B_s is decreasing and is of the form $B_s = \{\xi \le 1-s\}$ where ξ has a uniform law on $[0,1]$. Because $d\lambda = (1-s)\, d\mu$ where $\mu[0,x] = f'(x)$ we can rewrite the integral as $\int d\mu\, \mathbf{1}_{\{\xi \le 1-s\}} = \mu([0, 1-\xi]) = f'(1-\xi)$. Of course $1-\xi$ is uniformly distributed over the interval $[0,1]$. Conversely if η has the same law as f', then, because the probability space is atomless, we can find a random variable ξ uniformly distributed over $[0,1]$ and such that $\phi = f' \circ (1-\xi)$. The sets $B_s = \{\xi \le 1-s\}$ form a decreasing system and $\phi = \int d\mu\, \mathbf{1}_{\{\xi \le 1-s\}}$. This shows that η is exposed and the random variable ξ is an exposing functional. To show that $\partial\mathcal{C}(v)$ is closed we observe that

$$\begin{aligned} &\partial\mathcal{C}(v) \\ &= \left\{ \mathbb{Q} \in \mathcal{C}(v) \mid \text{ for all } x \in \mathbb{R} : \mathbb{E}_{\mathbb{P}}\left[\exp\left(ix\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] = \int_0^1 \exp\left(ixf'(s)\right)\, ds \right\}. \end{aligned}$$

To see that is a G_δ set observe that

$$\partial\mathcal{C}(v)$$
$$= \cap_{n\geq 1, q \text{ rational}} \left\{ \mathbb{Q} \mid \left| \mathbb{E}_{\mathbb{P}} \left[\exp\left(iq \frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right] - \int_0^1 \exp\left(iqf'(s)\right) \, ds \right| < \frac{1}{n} \right\}.$$

□

Remark 68 The above result does not imply that for arbitrary $\xi \in L^\infty$ we necessarily have that ξ is an exposing functional. We do have that ξ attains its minimum in an exposed point $\mathbb{Q} \in \mathcal{C}(v)$ but this does not mean that the random variable ξ is an exposing functional. In order to make sure that ξ attains its minimum in a uniquely defined point of the core, we need extra hypothesis on ξ. If ξ is an exposing functional, then u is Gâteaux differentiable at ξ. We have seen that for convex games, the differentiability of the utility function cannot be guaranteed at all indicator functions.

Ryff's paper also shows a connection with stochastic dominance.

Definition 25 *Let ξ, η be integrable random variables, not necessarily defined on the same probability space, we say that η dominates ξ, denoted $\xi \preceq \eta$ if for all convex functions $\phi \colon \mathbb{R} \to \mathbb{R}$ we have $\mathbb{E}[\phi(\xi)] \leq \mathbb{E}[\phi(\eta)]$.*

We remark that both integrals are defined since ξ^-, η^- are integrable. By taking $\phi(x) = x$ and $\phi(x) = -x$ we get that $\mathbb{E}[\xi] = \mathbb{E}[\eta]$. Stochastic dominance is important in insurance mathematics, risk theory and the theory of decisions under uncertainty. Its importance comes from the relation with Choquet theory, see Phelps [111]. Most of the theory can be obtained by cleverly applying Choquet theory. For instance the famous theorem of Cartier-Fell-Meyer, Strassen, [111] says that $\xi \preceq \eta$ if and only if there are random variables ξ', η', defined on the same probability space, ξ and ξ' have the same law, η and η' have the same law and $\mathbb{E}[\eta' \mid \xi'] = \xi'$. This statement is then used to construct martingales and it is also used in finance.

The study of stochastic dominance requires some small introduction to the theory of convex functions. If $\phi \colon \mathbb{R} \to \mathbb{R}$ is a convex function, its derivative exists except in a countable number of points. We will – as we did above – use the right derivative. The second derivative of it is a nonnegative measure ϕ''. For each $a \in \mathbb{R}$ the convex function $\phi_a(x) = \phi(a) + \phi'(a)(x-a) + \int_{(a,\infty)} \phi''(ds)(x-s)^+$ is smaller than ϕ and for $x \geq a$ coincides with it. It follows that $\phi_a \uparrow \phi$ as $a \to -\infty$. Also this family of convex functions (for $a \leq 0$) is bounded below by the affine function $\phi(0) + \phi'(0)\,x$. The idea is that every non-decreasing convex function is a positive combination of functions of the form $(x-s)^+$. This means that properties for convex functions can be shown by first showing them for these functions and then proceeding to the limit.

Proposition 43 *$\xi \preceq \eta$ if and only if $\mathbb{E}[\xi] = \mathbb{E}[\eta]$ and the increasing rearrangements $q(\xi), q(\eta)$ satisfy:*

$$\textit{for all } 0 \leq x \leq 1: \int_{[0,x]} q_u(\xi)\,du \geq \int_{[0,x]} q_u(\eta)\,du.$$

Proof. We may suppose that ξ, η are defined on $[0,1]$ and are increasing. This allows to replace the quantiles by the functions and it makes the notation easier. Suppose that $\xi \preceq \eta$, we will show $\int_0^x \xi(s)\,ds \geq \int_0^x \eta(s)\,ds$ or what is the same since $\mathbb{E}[\xi] = \mathbb{E}[\eta]$, $\int_x^1 \xi(s)\,ds \leq \int_x^1 \eta(s)\,ds$. The indicator function of the interval $[x,1]$ is not the limit of convex functions so we need something better and in fact the monotonicity of the functions will play a role. Take $x \in [0,1]$ and take the convex function $(\xi - \xi(x))^+$. We have

$$\begin{aligned}
\int_0^1 (\xi - \xi(x))^+\,ds &= \int_x^1 \xi\,ds - \xi(x)(1-x), \\
\int_0^1 (\xi - \xi(x))^-\,ds &= -\int_0^x \xi\,ds + \xi(x)(x) \\
&= -\int_0^1 \xi\,ds + \int_x^1 \xi\,ds + \xi(x)(x),
\end{aligned}$$

the same holds for η.
Suppose first that $\xi(x) \leq \eta(x)$, we get that

$$\begin{aligned}
\int_x^1 \xi\,ds - \int_x^1 \eta\,ds &= \int_0^1 (\xi - \xi(x))^+\,ds - \int_0^1 (\eta - \eta(x))^+\,ds + (\xi(x) - \eta(x))(1-x) \\
&\leq \int_0^1 (\xi - \xi(x))^+\,ds - \int_0^1 (\eta - \xi(x))^+\,ds + (\xi(x) - \eta(x))(1-x) \\
&\leq (\xi(x) - \eta(x))(1-x) \leq 0.
\end{aligned}$$

In case $\xi(x) \geq \eta(x)$ we use the other equalities:

$$\begin{aligned}\int_x^1 \xi\,ds - \int_x^1 \eta\,ds &= \int_0^1 (\xi - \xi(x))^-\,ds - \int_0^1 (\eta - \eta(x))^-\,ds \\ &\quad + \int_0^1 \xi\,ds - \int_0^1 \eta\,ds - (\xi(x) - \eta(x))(x) \\ &\leq \int_0^1 (\xi - \xi(x))^-\,ds - \int_0^1 (\eta - \xi(x))^-\,ds - (\xi(x) - \eta(x))(x) \\ &\leq -(\xi(x) - \eta(x))(x) \leq 0.\end{aligned}$$

In either case we found $\int_x^1 \xi\,ds \leq \int_x^1 \eta\,ds$.

Conversely, for each $s \in \mathbb{R}$ we have the existence of $u \in [0,1]$ such that $\xi(u-) \leq s \leq \xi(u)$ and we then get by the hypothesis on ξ and η:

$$\begin{aligned}\mathbb{E}[(\xi - s)^+] = \int_u^1 (\xi(y) - s)\,dy &\leq \int_u^1 (\eta(y) - s)\,dy \\ &\leq \int_u^1 (\eta(y) - s)^+\,dy \leq \mathbb{E}[(\eta - s)^+].\end{aligned}$$

If ϕ is a convex function then integrating with respect to ϕ'' and using that $\mathbb{E}[\xi] = \mathbb{E}[\eta]$, gives for each $a \in \mathbb{R}$:

$$\mathbb{E}[\phi_a(\xi)] \leq \mathbb{E}[\phi_a(\eta)].$$

If $a \to -\infty$ the Beppo Levi theorem gives $\mathbb{E}[\phi(\xi)] \leq \mathbb{E}[\phi(\eta)]$, as desired. □

Proposition 44 *An element $\mathbb{Q}$ is in the core of $v = f \circ \mathbb{P}$ if and only if $\frac{d\mathbb{Q}}{d\mathbb{P}} \preceq f'$.*

Proof. Suppose $\frac{d\mathbb{Q}}{d\mathbb{P}} \preceq f'$, then for each $A \in \mathcal{F}$ with probability $\mathbb{P}[A] = x$ we have

$$\mathbb{Q}[A] = \int_A \frac{d\mathbb{Q}}{d\mathbb{P}}\,d\mathbb{P} \geq \int_0^x q_u\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\,du \geq \int_0^x f'(u)\,du = f(x) = v(A).$$

Conversely if $\mathbb{Q} \in \mathcal{C}(v)$ we have that for any $x \in [0,1]$, any A with $\mathbb{P}[A] = x$ and such that $\left\{\frac{d\mathbb{Q}}{d\mathbb{P}} < q_x\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right\} \subset A \subset \left\{\frac{d\mathbb{Q}}{d\mathbb{P}} \leq q_x\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right\}$:

$$\int_0^x q_u\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\,du = \int_A \frac{d\mathbb{Q}}{d\mathbb{P}}\,d\mathbb{P} = \mathbb{Q}[A] \geq v(A) = f(x) = \int_0^x f'(u)\,du.$$

□

Corollary 11 *If $h \in L^1$ and $h \preceq f'$ then h can be written as an integral over the extreme points of $\mathcal{C}(v)$. In other words there is a probability measure μ on $\mathcal{C}(v)$, supported by the exposed points $\partial\mathcal{C}(v)$ of $\mathcal{C}(v)$ such that $h = \int_{\partial\mathcal{C}(v)} \mathbb{Q}\, d\mu$*

Remark 69 The measure μ is necessarily supported by $\partial\mathcal{C}(v)$ since this is a Baire set, even a closed G_δ. Hence $\mu(\partial\mathcal{C}(v)) = 1$, see [111]. The separability of the support follows from results in functional analysis — the so called study of Eberlein compact sets, see [8]. In this case it could be proved by hand using the fact that $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is already measurable for a separable atomless sigma algebra and hence we can restrict everything to the case of L^1 being separable. We leave the details to the reader.

Corollary 12 *If $h \in L^1$ and $h \preceq f'$ then h is the limit of convex combinations of random variables, equal to f' in law.*

Remark 70 The previous reasoning gives an alternative proof of Ryff's theorem, [121]

7.4 Strongly exposed points

For convex sets there is a stronger notion than exposed points.

Definition 26 *If C is a convex bounded closed set in a Banach space E, then we say that $x \in C$ is strongly exposed if there is a linear functional $x^*\colon E \to \mathbb{R}$ such that for every sequence $y_n; n \geq 1$ in C, the convergence of $x^*(y_n) \to x^*(x)$ implies $\|x - y_n\| \to 0$.*

Of course this implies that x is an exposed point and that x^* is an exposing functional. It is known that for a weakly compact set in a Banach space, say C, the convex closed hull of the strongly exposed points is equal to C, [**?**],[**?**]. Because of the special nature of the core of a distorted probability, we can guess that the extreme points are not only exposed, they are even strongly exposed. Indeed all the extreme points are of the same nature (up to some isomorphisms of the probability space — if the topological nature of Ω would allow it). This means that they all have the same properties. So they all should be strongly exposed. In this special case one can give a direct proof using the characterisation of the extreme points. We suppose that $v(A) = f(\mathbb{P}[A])$ is a convex game and that the distortion f is continuous at 1.

Proposition 45 *All extreme points of $\mathcal{C}(v)$ are strongly exposed.*

Proof. Let us recall that if $\mathbb{Q}$ is an extreme point, there is a random variable ξ having a uniform law $[0,1]$ and such that ξ is exposing. This means that for all other elements, $\mathbb{Q}'$ of the core $\mathcal{C}(v)$ we have $\mathbb{E}_{\mathbb{Q}}[\xi] < \mathbb{E}_{\mathbb{Q}'}[\xi]$. What we need to show is that for a sequence $\mathbb{Q}^n$ in $\mathcal{C}(v)$, $\lim_n \mathbb{E}_{\mathbb{Q}^n}[\xi] = \mathbb{E}_{\mathbb{Q}}[\xi]$ implies that $\mathbb{Q}^n \to \mathbb{Q}$ in L^1-norm. From the discussion on the structure of the extreme points, Theorem 50, we recall that $\mathbb{Q}$ can be written as $\mathbb{Q} = \int \mathbb{Q}_s \, \lambda(ds)$ where $\mathbb{Q}_s$ is given by $d\mathbb{Q}_s = \frac{1}{1-s}\mathbf{1}_{B_s}\, d\mathbb{P}$ with $B_s = \{\xi \le 1-s\}$. At the same time we can write $\mathbb{Q}^n = \int \mathbb{Q}^n_s \, \lambda(ds)$, where $\mathbb{Q}^n_s$ is in the core of the game v_s, i.e. $\frac{d\mathbb{Q}^n_s}{d\mathbb{P}} \le \frac{1}{1-s}$. The random variable ξ satisfies (at least in λ measure): $\mathbb{E}_{\mathbb{Q}^n_s}[\xi] \to \mathbb{E}_{\mathbb{Q}_s}[\xi]$. Indeed $\mathbb{E}_{\mathbb{Q}^n}[\xi] = \int \mathbb{E}_{\mathbb{Q}^n_s}[\xi]\, d\lambda \to \int \mathbb{E}_{\mathbb{Q}_s}[\xi]\, d\lambda$. But for each s we have $\mathbb{E}_{\mathbb{Q}^n_s}[\xi] \ge \mathbb{E}_{\mathbb{Q}_s}[\xi]$. Hence $\mathbb{E}_{\mathbb{Q}^n_s}[\xi] \to \mathbb{E}_{\mathbb{Q}_s}[\xi]$ for almost every s. Now $\mathbb{Q}_s$ is an exposed point of $\mathcal{C}(v_s)$ and we will show that it is strongly exposed, meaning that $\|\mathbb{Q}^n_s - \mathbb{Q}_s\|_1 \to 0$. This will imply $\|\mathbb{Q}^n - \mathbb{Q}\|_1 \le \int d\lambda \, \|\mathbb{Q}^n_s - \mathbb{Q}_s\|_1 \to 0$. In other words our representation of the core $\mathcal{C}(v)$ allows to reduce the problem to the special case of TailVar. We now go back to the calculations in Example 9. There it was shown that for $k = \frac{1}{1-s}$ and $\alpha = 1-s$:

$$\mathbb{E}_{\mathbb{Q}^n_s}[\xi] - \mathbb{E}_{\mathbb{Q}_s}[\xi] = \int_{B_s} (\xi-\alpha)\left(\frac{d\mathbb{Q}^n_s}{d\mathbb{P}} - k\right) d\mathbb{P} + \int_{B^c_s} (\xi-\alpha)\, \frac{d\mathbb{Q}^n_s}{d\mathbb{P}} d\mathbb{P}.$$

Since the left side tends to 0 and since both terms on the right are non-negative we get that each of them tends to zero. This implies that both $(\xi-\alpha)\left(\frac{d\mathbb{Q}^n_s}{d\mathbb{P}} - k\right)\mathbf{1}_{B_s}$ and $(\xi-\alpha)\,\frac{d\mathbb{Q}^n_s}{d\mathbb{P}}\mathbf{1}_{B^c_s}$ tend to zero. Since $\mathbb{P}[\xi=\alpha]=0$ we have that $\mathbf{1}_{B^c_s}\frac{d\mathbb{Q}^n_s}{d\mathbb{P}}$ and $\left(\frac{d\mathbb{Q}^n_s}{d\mathbb{P}} - k\right)\mathbf{1}_{B_s}$ tend to zero. In other words $\frac{d\mathbb{Q}^n_s}{d\mathbb{P}} \to k\mathbf{1}_{B_s}$, all convergences taking place in probability. Because $\mathbb{Q}^n_s$ and $\mathbb{Q}_s$ are probabilities, Scheffé's lemma implies $\|\mathbb{Q}^n_s - \mathbb{Q}_s\|_1 \to 0$. □

Chapter 8

Relation with VaR

In this chapter we deepen on the relation between r utility functions and VaR. We recall that if α belongs to the interval $(0,1)$ the family $\mathcal{S}_{1/\alpha} = \{\mathbb{Q} \mid \frac{d\mathbb{Q}}{d\mathbb{P}} \leq 1/\alpha\}$ is well defined; the corresponding u_α is such that if $\mathbb{P}$ is atomless and if the distribution of ξ is continuous, $u_\alpha(\xi) = \mathbb{E}[\xi| \, \xi \leq q_\alpha(\xi)]$. Recall that $q_\alpha(\xi)$ is the α-quantile of ξ, defined as $\inf\{x \mid \mathbb{P}[\xi \leq x] > \alpha\}$. We defined the Value at Risk as $-q_\alpha(\xi)$. In case the law of ξ is not necessarily continuous, i.e. in general, we have $u_\alpha(\xi) = \frac{1}{\alpha}\int_0^\alpha q_u(\xi)\, du$.

8.1 VaR and TailVaR

Let us now come back to the relation between utility functions and VaR. The utility functon u_α is maximal in the sense that it is the maximum in the class of coherent utility functions, only depending on the distribution and smaller than q_α. More precisely the following theorem holds:

Theorem 52 *Suppose that $\mathbb{P}$ is atomless; let u be a coherent utility function verifying the additional property that if ξ and η are identically distributed, then $u(\xi) = u(\eta)$. If for every $\xi \in L^\infty$, $u(\xi)$ is smaller than $q_\alpha(\xi)$, then $u \leq u_\alpha$.*

Proof. . We first observe that utility functions that only depend on the distribution of the random variables have the Fatou property, see Section 5.1. We now prove that for every ξ, $u(\xi) \leq \mathbb{E}[\xi \mid \xi \leq q_\alpha(\xi) + \varepsilon]$. Let $A = \{\omega \mid \xi(\omega) \leq q_\alpha(\xi) + \varepsilon\}$, by definition of q_α we have the strict inequality $\mathbb{P}[A] > \alpha$. Let η be the random variable equal to ξ on A^c and equal to the number $\mathbb{E}[\xi \mid A]$ on A. $q_\alpha(\eta)$ is then equal to $\mathbb{E}[\xi \mid A]$ and we deduce from $q_\alpha(\eta) \geq u(\eta)$, that $\mathbb{E}[\xi \mid A] \geq u(\eta)$. Let us call ν the distribution of ξ given A, where A is equipped with the inherited σ algebra $\{A \cap B \mid B \in \mathcal{F}\}$ and with the conditional probability $\mathbb{P}[\,.\mid A]$. The hypothesis of the absence of atoms in Ω implies in particular the absence of atoms in A. This fact guarantees the existence on A of a sequence, say Z_n of independent (for $\mathbb{P}[\,.\mid A]$), ν-distributed random variables. Let us denote by ξ_n the random variable coinciding with the n-th element Z_n on A and with ξ (and therefore with η) outside A. The $(\xi_n)_n$ have the same distribution, equal to the distribution of ξ. By the law of large numbers, $\frac{\xi_1+\cdots+\xi_n}{n}$ converges almost surely to η. Remembering that the Fatou property holds, we finally obtain:

$$u(\eta) \geq \limsup_{n\to\infty} u\left(\frac{\xi_1 + \ldots \xi_n}{n}\right) \geq \limsup_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} u(\xi_i) = u(\xi)$$

Thus we get $u(\xi) \leq u(\eta) \leq \mathbb{E}[\xi \mid \xi \leq q_\alpha(\xi) + \varepsilon]$. If ξ has a distribution function continuous at q_α, we can pass to the limit, obtaining $u(\xi) \leq u_\alpha(\xi)$. What if the distribution of ξ is not continuous? In this case, we can find an approximating sequence ξ_n as in Proposition 1 and we have that both $u(\xi_n)$ and $u_\alpha(\xi_n)$ tend to $u(\xi)$ and $u_\alpha(\xi)$ respectively (because coherent utility functions are continuous with respect to the uniform L^∞ topology). Passing to the limit in the already established inequality $u(\xi_n) \leq u_\alpha(\xi_n)$, gives $u(\xi) \leq u_\alpha(\xi)$ for all $\xi \in L^\infty$. □

Remark 71 Kusuoka could characterise the coherent risk measures that are law invariant. His characterisation gives an alternative proof of the above result, see [96]. See also Chapter 5.

8.2 VaR as an envelope of coherent utilities

As a general result, under the hypotheses of absence of atoms, there is no smallest coherent risk measure that dominates VaR. As usual we say that ρ dominates VaR if for all $\eta \in L^\infty$ we have that $\rho(\eta) \geq \mathrm{VaR}_\alpha(\eta)$ or what is the same $u(\eta) \leq q_\alpha(\eta)$.

Theorem 53 *If $\mathbb{P}$ is atomless we have, for each $0 < \alpha < 1$:*

$$q_\alpha(\xi) = \sup\left\{u(\xi) \mid u \textit{ coherent with the Fatou property and } u \leq q_\alpha\right\}.$$

Remark 72 The theorem says that if we take the supremum over **all** coherent utility functions that dominate VaR (and not only the ones depending just on distributions) we obtain VaR, wich as we already saw, is not a coherent risk measure (remember, it's not subadditive). This shows that there is no smallest convex risk measure that dominates VaR. The proof taken from [39], is quite technical, it can be omitted.

We start the proof with the lemma that characterises the utility functions that are dominated by a quantile. We remark that we always take the right (or largest) quantile. For the left quantile there are difficulties as can be seen from [39]

Lemma 20 *A coherent utility function u, defined by $\mathcal{S}^{\mathbf{ba}}$ is dominated by q_α if and only if for each set B with $\mathbb{P}[B] > \alpha$ and for each $\varepsilon > 0$, there is a measure $\mu \in \mathcal{S}^{\mathbf{ba}}$ such that $\mu(B) > 1 - \varepsilon$.*

Proof of the lemma We first prove necessity. Take $\varepsilon > 0$ and a set B such that $\mathbb{P}[B] > \alpha$. Since $q_\alpha(\xi) = -1$ for the random variable $\xi = -\mathbf{1}_B$, we conclude from the inequality $u \le q_\alpha$, that there is a measure $\mu \in \mathcal{S}^{\mathbf{ba}}$ such that $\mu(B) \ge 1 - \varepsilon$. For the sufficiency we take a random variable ξ as well as $\varepsilon > 0$ and we consider the set $B = \{\xi \le q_\alpha + \varepsilon\}$. By definition of q_α, we have $\mathbb{P}[B] > \alpha$. By assumption there exists a measure $\mu \in \mathcal{S}^{\mathbf{ba}}$ with the property $\mu(B) \ge 1 - \varepsilon$. This gives the inequality

$$u(\xi) \le \mu[\xi] \le \mu[\xi \mathbf{1}_B] + \varepsilon \|\xi\|_\infty \le (q_\alpha(\xi) + \varepsilon) + \varepsilon \|\xi\|_\infty .$$

Since the inequality holds for every $\varepsilon > 0$, we get the result $u \le q_\alpha$. □

Proof of the Theorem We only have to show that for ξ given, we can find a coherent utility dominated by q_α and with the property that $u(\xi) \ge q_\alpha(\xi)$. For each $\varepsilon > 0$, the set $C = \{\xi \le q_\alpha + \varepsilon\}$ has measure $\mathbb{P}[C] > \alpha$. But the definition of q_α implies that $\mathbb{P}[\xi < q_\alpha] \le \alpha$. It follows that the set $D = \{q_\alpha \le \xi \le q_\alpha + \varepsilon\}$ has strictly positive measure. Take now an arbitrary set B with measure $\mathbb{P}[B] > \alpha$. Either we have that $\mathbb{P}[B \cap C^c] \ne 0$, in which case we take $h_B = \frac{\mathbf{1}_{B \cap C^c}}{\mathbb{P}[B \cap C^c]}$ or we have that $B \subset C$. In this case and because $\mathbb{P}[\xi < q_\alpha] \le \alpha$ we must have that $\mathbb{P}[B \cap D] > 0$. We take $h_B = \frac{\mathbf{1}_{B \cap D}}{\mathbb{P}[B \cap D]}$. The Fatou coherent utility function is then defined as $u(\eta) = \inf_{\mathbb{P}[B] > \alpha} \mathbb{E}_{\mathbb{P}}[\eta h_B]$. By the lemma we have that u is dominated by q_α but for the variable ξ we find that $\mathbb{E}_{\mathbb{P}}[\xi h_B]$ is always bounded below by q_α, i.e. $u(\xi) \ge q_\alpha(\xi)$. It follows that $u(\xi) = q_\alpha(\xi)$. □

Remark 73 We will not continue the study of the relation between coherent utility functions and VaR. The examples on credit risk see Chapter 4, can be used as further illustrations. We leave it to the intelligent reader to draw his/her conclusions on the use of VaR as an institutional risk measure.

Chapter 9

The Capital Allocation Problem

Let, as before, $u\colon L^\infty \to \mathbb{R}$ be a coherent utility function with the Fatou property. With u we associate the coherent risk measure $\rho(\xi) = -u(\xi)$. Imagine that a firm is organised as N trading units and let their future wealth be denoted by $\xi_1, \dots, \xi_N$, all belonging to L^∞. With these individual positions we need to associate an amount of economic capital. The idea is that the economic capital of the firm – associated to $\xi_1 + \dots + \xi_N$ – has to be divided among the individual contributions ξ_i. The total capital required to face the risk is $\rho(\sum_{i=1}^N \xi_i)$ and we have to find a "fair" way to allocate $k_1, \dots, k_N$ so that $k_1 + \dots + k_N = \rho(\sum_{i=1}^N \xi_i)$. Because the risk measure is subadditive, the individual business units can benefit from the diversification. Another point of view of the allocation problem is to distribute the gain of diversification over the different business units of a financial institution. The reason why we have to solve this problem comes from problems such as the calculation of risk adjusted returns, the correct charge of the capital costs, etc.. We will present two solutions of the capital allocation problem. Both are related to a game theoretic approach.

Here is another interpretation of the capital allocation problem. Suppose that an insurance company has issued contracts for which the claims are described by the random variables $\xi_1, \dots, \xi_n$. We assume that the claims are denoted by positive numbers. The company now wants to charge a fair premium to each of this contracts. Of course these numbers will be augmented by the cost of capital, the overhead costs, coffee for the secretaries, ... The total future position before premium income is $-(\xi_1 + \dots + \xi_n)$. This requires a premium income equal to $-u(-(\xi_1 + \dots + \xi_n))$. This is precisely $\rho\,(-(\xi_1 + \dots + \xi_n))$. The solution of the capital allocation problem allows us to find a fair allocation to each of the individual contracts. Of course this means that the premium of a contract will depend on the other contracts in the portfolio. This is not a new issue. We refer to Deprez and Gerber [51] where this was discussed.

9.1 Simple game theoretic approach.

In the previous setting, we define $k_1, \dots, k_N$ to be a fair allocation if:

1. $\sum_{i=1}^N k_i = u(\sum_{i=1}^N \xi_i)$
2. $\forall J \subseteq \{1, \dots, N\}$ we have $\sum_{j\in J} k_j \geq u(\sum_{j\in J} \xi_j)$.

The existence of a fair allocation is in fact equivalent to the non-emptiness of the core of a "balanced" game. So it is no surprise that the following theorem uses the same technique as the Bondareva-Shapley theorem in game theory. For completeness and becaude it is instructive, we include a proof.

Theorem 54 *(Bondareva-Shapley theorem for risk measures) If u is coherent then there exists a fair allocation.*

Proof. . Let $m = 2^N$ and let $\phi : \mathbb{R}^N \to \mathbb{R}^m$ be the following linear map:

$$\phi((k_i)_i) = \left(\left(\sum_{j\in J} k_j \right)_{\emptyset \neq J \subseteq \{1,\dots,N\}}, \left(-\sum_{j=1}^N k_j \right) \right)$$

We have to find $k = (k_1, \dots, k_N) \in \mathbb{R}^N$ so that for each $\emptyset \neq J \subset \{1, \dots, N\}$ we have $\phi(k)_J \geq u(\sum_{j\in J} \xi_j)$ and so that $\sum_{i\leq N} k_i = u(\sum_{i\leq N} \xi_i)$.

Let $P = \left\{ ((x_J)_J, x) \mid x_J \geq u\left(\sum_{j\in J} \xi_j\right), x \geq -u\left(\sum_{i\leq N} \xi_i\right) \right\}$. The problem is reduced to showing that $\phi(\mathbb{R}^N) \cap P$ is non empty. If it were empty, by the separating hyperplane theorem, there would be $((\alpha_J)_J, \alpha)$ such that:

1. $\sum_J \alpha_J (\sum_{j\in J} k_j) - \alpha \sum_{i\leq N} k_i = 0$;
2. $\sum_J \alpha_J u(\sum_{j\in J} \xi_j) - \alpha u(\sum_{i\leq N} \xi_i) > 0$;
3. $\alpha_J \geq 0$, $\alpha \geq 0$.

Condition *1* can be written as: for each j, we have $\sum_{J\ni j} \alpha_J = \alpha$. If $\alpha = 0$, then all the α_J would be 0 but this is impossible by point *2*. Therefore we can normalize: we may suppose $\alpha = 1$. Hence we have found positive $(\alpha_J)_J$ such that $\sum_{J\ni j} \alpha_J = 1$ and verifying $\sum_J \alpha_J u(\sum_{j\in J} \xi_j) > u(\sum_{i\leq N} \xi_i)$. By coherence, it is a contradiction, since we may write:

$$\begin{aligned} u\left(\sum_{i\le N}\xi_i\right) &= u\left(\sum_j\left(\sum_{J\in j}\alpha_J\right)\xi_j\right)\\ &= u\left(\sum_J\alpha_J\left(\sum_{j\in J}\xi_j\right)\right)\\ &\ge \sum_J\alpha_J\, u\left(\sum_{j\in J}\xi_j\right). \end{aligned}$$

So there is a fair allocation. □

Remark 74 One can see that concavity alone is not sufficient to give a solution to the capital allocation problem. Indeed if we take two "players", each having the same random variable ξ, we need to find two numbers k_1, k_2 such that $k_i \ge u(\xi)$ and $k_1 + k_2 = u(2\xi)$. This is only possible if $u(2\xi) \ge 2u(\xi)$. This implies that u must be coherent.

Remark 75 There is a case where the solution of the capital allocation problem becomes trivial. Suppose that u is commonotone, i.e. given by the core of a convex game. Suppose that $\xi_1, \xi_2, \ldots, \xi_N$ are commonotone, i.e. nondecreasing functions of one random variable. Then the only fair solution is $k_i = u(\xi_i)$. Indeed we have by commonotonicity that $u(\xi_1 + \ldots + \xi_N) = \sum_i u(\xi_i)$ and hence we must have $k_i = u(\xi_i)$.

9.2 A stronger concept of fairness

The basic papers regarding this approach are Aubin, [12], Artzner-Ostroy, [9] and Billera-Heath, [19]. An allocation $k_1, \ldots, k_N$ with $k = k_1 + \ldots k_N = u(\sum_{j=1}^N \xi_j)$ is now called fair (or fair for fuzzy games) if $\forall\alpha_j$, $j = 1, \ldots, N$, $0 \le \alpha_j \le 1$ we have:

$$\sum_{j=1}^N \alpha_j k_j \ge u\left(\sum_{j=1}^N \alpha_j\xi_j\right).$$

The expression "fuzzy games" comes from the fact that we can see a vector $(\alpha_1, \ldots, \alpha_N)$ as a representation of a coalition that uses the proportion α_i

of business unit i. This requirement of being fair is therefore much stricter than the one from the previous section. It has the advantage that it is robust for "reorganisations of the firm". The cited papers as well as the paper by Deprez and Gerber, see [51], relate this problem to the existence of derivatives.

Theorem 55 *Suppose that $\xi_1, \ldots, \xi_N$ are given. Let $\xi = \xi_1 + \ldots + \xi_N$. The allocation $k_1, \ldots, k_N$ is fair if and only if there is $\mu \in \partial_\xi(u)$ with $k_i = \mu(\xi_i)$.*

Proof. Suppose that $\mu \in \partial_\xi(u)$. Define $k_i = \mu(\xi_i)$. Obviously $\sum_i \mu(\xi_i) = \mu(\xi) = u(\xi)$. But for given $0 \le \alpha_i \le 1$ we also have $\sum_i \alpha_i \mu(\xi_i) = \mu(\sum_i \alpha_i \xi_i) \ge u(\sum_i \alpha_i \xi_i)$. Conversely let $k_1, \ldots, k_N$ be fair. Set $C = \{(x_1, \ldots, x_N) \mid x_i \le k_i\} \subset \mathbb{R}^N$. Consider the mapping $\Phi : \mathcal{S}^{\mathbf{ba}} \to \mathbb{R}^N$ given by $\Phi(\mu) = (\mu(\xi_1), \ldots, \mu(\xi_N))$. The image, K, is convex and compact. Suppose that $K \cap C = \emptyset$. Then we can strictly separate the two sets. This gives a nonzero vector $(\alpha_1, \ldots, \alpha_N)$ such that

$$\sup_{x \in C} \alpha.x < \min_{\mu \in \mathcal{S}^{\mathbf{ba}}} \alpha.\Phi(\mu).$$

This implies that for all i we must have $\alpha_i \ge 0$. We can therefore divide by the maximum of α_i and get $0 \le \alpha_i \le 1$. The supremum on the left is attained for $(x_1, \ldots, x_N) = (k_1, \ldots, k_N)$. The right side gives the minimum of $\mu(\sum_i \alpha_i \xi_i)$, which is $u(\sum_i \alpha_i \xi_i)$. We get

$$\sum_i \alpha_i k_i < u\left(\sum_i \alpha_i \xi_i\right),$$

a contradiction to fairness. So we proved that $C \cap K \ne \emptyset$. In other words we found $\mu \in \mathcal{S}^{\mathbf{ba}}$ with $\mu(\xi_i) \le k_i$ for all i. If we sum we get

$$\mu\left(\sum_i \xi_i\right) = \sum_i \mu(\xi_i) \le \sum_i k_i = u\left(\sum_i \xi_i\right) \le \mu\left(\sum_i \xi_i\right).$$

But this shows that all inequalities are equalities and hence for all i: $\mu(\xi_i) = k_i$, but it also shows that $u\left(\sum_i \xi_i\right) = \mu\left(\sum_i \xi_i\right)$, proving that $\mu \in \partial_\xi(u)$ □

Corollary 13 *In case u is differentiable at ξ, i.e. $\partial_\xi(u) = \{\mu\}$, we have that*

$$k_i = \lim_{\varepsilon \to 0} \frac{u(\xi + \varepsilon \xi_i) - u(\xi)}{\varepsilon}.$$

Remark 76 The corollary also shows that we can see the capital allocated to ξ_i as a marginal contribution. In the total wealth $\xi = \xi_1 + \ldots + \xi_N$ we increase the contribution of business unit i with $\varepsilon\xi_i$ and see how the total utility changes. Then we calculate the partial derivative. This procedure was, based on heuristic arguments, introduced by [17] and it was called the Euler principle. The above proposition explains why it works and why it gives good and fair allocations. We also remark that when u is Fatou, then the derivative, if it exists, is necessarily an element of L^1, see Theorem 22.

Example 34 Another illustration of this has been given by Uwe Schmock in a paper written for Swiss Reinsurance, [125]. He proposed to use $\mathbb{E}[\xi_i \mid \xi \le q_\alpha(\xi)]$ as a capital allocation method. The previous theory shows that this is a very natural way. Indeed the risk measure corresponds to the weakly compact set $\mathcal{S}_{1/\alpha}$ of Example 9. If ξ has a continuous distribution, or more generally when $\mathbb{P}[\xi \le q_\alpha(\xi)] = \alpha$, then $\partial_\xi(u) = \{1/\alpha\, \mathbf{1}_A\}$, where $A = \{\xi \le q_\alpha(\xi)\}$. So this example fits in the above framework of differentiability. The differentiability here is on the space L^∞. If only differentiability is required on the linear span of the random variables $\xi_1, \ldots, \xi_n$, things change. For more information on this topic the reader should consult the paper by Tasche, [127].

Remark 77 In Deprez and Gerber's paper [51], the reader will find a lot of similarities with the reasoning above. Their paper is full of ideas and relations between different properties. In some sense a forerunner of "risk measures" and of the gradient principle or Euler principle. The paper is from the mathematical viewpoint not so precise and therefore does not give the same results as above. For instance there is no discussion on the existence of the derivative, neither of the uniqueness of the subgradient. One can prove that the so-called bid price in an incomplete market defines a utility function that is **nowhere** Gâteaux differentiable. In Kalkbrener's paper, [86], there is an axiomatic approach to the capital allocation problem. We should also mention the paper by Denault, [47]. The axiomatics there are a little bit different. The idea of using game theoretic ideas is present, but Denault wants to get something that is related to the Shapley value. Since the Shapley value is – for convex games – somewhere in the "middle" of the core, the solution is not related to our presentation.

Chapter 10

The extension of risk measures to L^0

10.1 L^0 and utility functions

As we already said in the introduction, L^0 is invariant under a change of probability measure and the definition of risk measures or utility functions on it deserves special attention. The following theorem shows that there is not much hope.

Theorem 56 *If $\mathbb{P}$ is atomless, there exists no functional $u\colon L^0 \to \mathbb{R}$ such that:*

1. $u(\xi + a) = \rho(\xi) + a \quad \forall a \in \mathbb{R}$ *;*
2. $u(\xi + \eta) \geq u(\xi) + u(\eta)$ *;*
3. $u(\lambda\xi) = \lambda u(\xi) \quad \forall \lambda \in \mathbb{R}_+$*;*
4. $\xi \geq 0 \to u(\xi) \geq 0$ *.*

This is a consequence of the analytic version of the Hahn-Banach theorem and of the fact that a continuous linear functional on L^0 must be necessarily null if $\mathbb{P}$ is atomless. We do not give the details. The proof is essentially the same as the proof of Nikodym's theorem, [107].

Corollary 14 *If $\mathbb{P}$ is atomless, then the quantiles q_α (defined for $0 < \alpha < 1$) cannot be superadditive. Consequently VaR is not subadditive.*

Proof. The proof is quite easy. A quantile q_α satisfies properties 1, 3 and 4. Since there is no utility function satisfying all 4 properties, q_α cannot satisfy property 2. We remark that this proof is structural and we invite the reader to find (easy) counterexamples different from the ones given in the credit risk section. We also remark that the quantiles satisfy a Fatou property. Also this is left as a (not so easy) exercise. □

10.2 Coherent functions defined on L^0

If E is a solid, rearrangement invariant vector space containing non-integrable random variables, and if we want to define a utility function on E, we need to consider a utility function u that takes infinite values. Of course we would like to have functions u such that $+\infty$ is avoided as this does not make economic sense, see [39] and [40]. A value $u(\xi) = +\infty$ is indeed meaningless, because it implies that any sum of money can be taken away without becoming unacceptable. Such a random variable would represent the dream of many risk managers or traders. It would allow them to get an enormous commission on the trade. On the contrary, $u(\xi) = -\infty$ makes sense: it represents a risky position, which no amount of money can cover. In insurance terms, the risk leading to the position ξ would not be insurable, at least not by a prudent insurance company.

So let us consider $u\colon L^\infty \to \mathbb{R}$ defined by $\mathcal{S} \subset L^1$. We first define for arbitrary random variables $\xi \in L^0$:

$$u(\xi \wedge n) = \inf_{\mathbb{Q}\in\mathcal{S}} \mathbb{E}_{\mathbb{Q}}[(\xi \wedge n)]\,.$$

We remark that the truncation is necessary to prevent the integral from being $+\infty$ (in practice, we want to avoid the influence of “too optimistic” large benefits). We then define:

$$u(\xi) = \lim_{n\to+\infty} u(\xi \wedge n)\,.$$

Of course for random variables in L^∞ this definition yields the same value, therefore there is no need to introduce a new notation. Unfortunately, $u(\xi)$ can for some $\xi \in L^0$, turn out to be $+\infty$. For instance, one could take $\xi \geq 0$ but non integrable, so that every $u(\xi \wedge n)$ is finite, while the limit is not.

We note that the following implications hold:

$$\big(\forall \xi \in L^0 : u(\xi) < +\infty\big) \Leftrightarrow \forall \xi \in L^0, \xi \geq 0 \text{ implies } \lim_{n\to+\infty} \inf_{\mathbb{Q}\in\mathcal{S}} \mathbb{E}_{\mathbb{Q}}[\xi \wedge n] < +\infty.$$

If the first inequality holds, obviously the second one is true. To prove the converse, we note that the newly defined u is monotone and that $\xi^+ = \xi + \xi^-$ then implies that $u(\xi) \leq u(\xi^+) < +\infty$.

So we have already proved the equivalence between the first two points of the following theorem:

Theorem 57 *The following conditions are equivalent:*

1. *$\forall \xi \in L^0 : u(\xi) < +\infty$;*
2. *$\forall \xi \geq 0$, $\phi(\xi) = \lim_{n\to+\infty} \inf_{\mathbb{Q}\in\mathcal{S}} \mathbb{E}_{\mathbb{Q}}[\xi \wedge n] < +\infty$;*
3. *$\exists \gamma > 0$ such that $A \in \mathcal{F}$ $\quad \mathbb{P}[A] \leq \gamma$ implies $\inf_{\mathbb{Q}\in\mathcal{S}} \mathbb{Q}[A] = 0$;*
4. *$\forall f \geq 0$ $\exists \mathbb{Q} \in \mathcal{S}$ such that $\mathbb{E}_{\mathbb{Q}}[f] < +\infty$;*
5. *$\exists \gamma > 0$ such that $\forall A \in \mathcal{F}, \mathbb{P}[A] \leq \gamma$, $\exists \mathbb{Q} \in \mathcal{S}$ with $\mathbb{Q}[A] = 0$;*
6. *$\exists \gamma > 0$, $\exists k$ such that $\forall A \in \mathcal{F}$, with $\mathbb{P}[A] \leq \gamma$ $\exists \mathbb{Q} \in \mathcal{S}$ with the properties:*
$$\begin{cases} \mathbb{Q}[A] = 0 \\ \\ \frac{d\mathbb{Q}}{d\mathbb{P}} \leq k\,. \end{cases}$$

Proof. We need to prove the equivalences from point *2* to point *6* and the scheme is:
3$\Leftrightarrow$*2* $\Rightarrow$ *6* $\Rightarrow$ *5* $\Rightarrow$ *4* $\Rightarrow$ *2*.
(*2* $\Rightarrow$ *3*)
By contradiction, if *3* is false then for every n we can find A_n with $\mathbb{P}[A_n] \leq 2^{-n}$ so that $\inf_{\mathbb{Q}\in\mathcal{S}} \mathbb{Q}[A_n] \geq \varepsilon_n > 0$. Then we define $f = \sum_{n\geq 1} \mathbf{1}_{A_n} \frac{n}{\varepsilon_n}$. By Borel-Cantelli's lemma the sum is finite almost surely. Now we can write:

$$\mathbb{E}_{\mathbb{Q}}\left[f \wedge \frac{N}{\varepsilon_N}\right] \geq \mathbb{E}_{\mathbb{Q}}\left[\left(\mathbf{1}_{A_N}\frac{N}{\varepsilon_N}\right) \wedge \frac{N}{\varepsilon_N}\right] \geq N$$

and therefore $\inf_{\mathbb{Q}\in\mathcal{S}} \mathbb{E}_{\mathbb{Q}}[f \wedge \frac{N}{\varepsilon_N}] \geq N$; letting N tend to infinity, we contradict *2*.
(*3* $\Rightarrow$ *2*)
Let's fix a positive f: since it is real valued, there exists K such that $\mathbb{P}[\{f > K\}] < \gamma$ and taking $n > K$ we get

$$\begin{aligned} \inf_{\mathbb{Q}\in\mathcal{S}} \mathbb{E}_{\mathbb{Q}}[f \wedge n] &\equiv \inf_{\mathbb{Q}\in\mathcal{S}} \mathbb{E}_{\mathbb{Q}}[(f \wedge n)\mathbf{1}_{\{f>K\}} + (f \wedge n)\mathbf{1}_{\{f\leq K\}}] \\ &\leq \inf_{\mathbb{Q}\in\mathcal{S}} (\mathbb{E}_{\mathbb{Q}}[(f \wedge n)\mathbf{1}_{\{f>K\}}] + K) \leq K \end{aligned}$$

The implications *6* $\Rightarrow$ *5* $\Rightarrow$ *4* $\Rightarrow$ *2* are easy exercises. The real challenge is proving the implication *3* $\Rightarrow$ *6*. Let $k > \frac{2}{\gamma}$ and let A, satisfying $\mathbb{P}[A] < \frac{\gamma}{2}$, be given. We will show *6* by contradiction. So let us take $H_k = \{f \mid |f| \leq k,\ f = 0 \text{ on } A\}$. If H_k and $\mathcal{S}$ were disjoint we could, by the Hahn-Banach theorem, strictly separate the closed convex set $\mathcal{S}$ and the weakly compact convex set H_k. This means that there exists an element $\xi \in L^\infty$, $\|\xi\|_\infty \leq 1$ so that

$$\sup\{\mathbb{E}[\xi f] \mid f \in H_k\} < \inf\{\mathbb{E}_\mathbb{Q}[\xi] \mid \mathbb{Q} \in \mathcal{S}\}. \tag{10.1}$$

We will show that this inequality implies that $\|\xi \mathbf{1}_{A^c}\|_1 = 0$. Indeed if not, we would have $\mathbb{P}[\{\mathbf{1}_{A^c}|\xi| > \frac{2}{\gamma}\|\xi\mathbf{1}_{A^c}\|_1\}] \leq \frac{\gamma}{2}$ and hence for each $\varepsilon > 0$ there is a $\mathbb{Q} \in \mathcal{S}$ so that $\mathbb{E}_\mathbb{Q}[A \cup \{|\xi| > \frac{2}{\gamma}\|\xi\mathbf{1}_{A^c}\|_1\}] \leq \varepsilon$. This implies that the right side of (10.1) is bounded by $\frac{2}{\gamma}\|\xi\mathbf{1}_{A^c}\|_1$. However, the left side is precisely $k\|\xi\mathbf{1}_{A^c}\|_1$. This implies $k\|\xi\mathbf{1}_{A^c}\|_1 < \frac{2}{\gamma}\|\xi\mathbf{1}_{A^c}\|_1$, a contradiction to the choice of k. Therefore $\xi = 0$ on A^c. But then property *3* implies that the right side is 0, whereas the left side is automatically equal to zero. This is a contradiction to the strict separation and the implication *3* $\Rightarrow$ *6* is therefore proved. □

Remark 78 For L^p with $p < 1$, there are utility functions $u\colon L^p \to \mathbb{R} \cup \{-\infty\}$ that do not satisfy the conditions of the above theorem. For instance we can take the distorted probability $v(A) = \mathbb{P}[A]^2$ and then show that for $p > 1/2$ the utility function is defined on L^p, i.e. it will never take the value $+\infty$. Of course it will take the value $-\infty$ for some random variables.

Exercise 26 We leave it as an exercise to show that for $\xi \in L^p$, we have $u(\xi) = -\infty$ if and only if $\xi^- \notin L^1$.

Chapter 11

Dynamic utility functions in a two period model

11.1 Notation for the two period case

We first look at the situation were we have two periods and we will restrict the discussion to utility functions having some kind of Fatou property. This means that we have the sigma-algebras $\mathcal{F}_0$ supposed to be trivial, the uncertainty modelled by $\mathcal{F}_1$ at time 1 and the final uncertainty modelled by $\mathcal{F}_2$. Many of the features of more period models and even of continuous time are already present in this case. Because of the revelation of uncertainty at the intermediate time, we need to distinguish between variables known at date 1 and variables known only at date 2. We therefore introduce the following notation. The space $L^\infty(\mathcal{F}_1)$ is the space of (classes) of bounded random variables measurable with respect to $\mathcal{F}_1$. The utility of an element $\xi \in L^\infty(\mathcal{F}_2)$ at time 0 is given by the monetary concave utility function u_0. At the intermediate time, the economic agent having the information $\mathcal{F}_1$, can have a different idea about ξ than at time 0. The knowledge that unfavourable events have happened might influence his appreciation. So at time 1 the utility is measured by an $\mathcal{F}_1$ measurable function, $u_1(\xi)$. We suppose that u_1 is monetary and concave which in this case means

1. $u_1 \colon L^\infty(\mathcal{F}_2) \to L^\infty(\mathcal{F}_1)$
2. $u(0) = 0$ and for $\xi \geq 0$ we have $u_1(\xi) \geq 0$
3. for $\eta \in L^\infty(\mathcal{F}_1)$ we have $u_1(\xi + \eta) = u_1(\xi) + \eta$
4. for $\lambda \in L^\infty(\mathcal{F}_1)$, $0 \leq \lambda \leq 1$, $\xi_1, \xi_2 \in L^\infty(\mathcal{F}_2)$ we have
   $u_1(\lambda\xi_1 + (1-\lambda)\xi_2) \geq \lambda u_1(\xi_1) + (1-\lambda)u_1(\xi_2)$
5. if $\xi_n \downarrow \xi$, all elements taken in $L^\infty(\mathcal{F}_2)$, then $u_1(\xi_n) \downarrow u_1(\xi)$.

The assumptions are clear. The monetary assumption should be taken at time 1, using the information available at time 1. The same for the concavity. Remark that we do not make any assumption about the relation between u_0 and u_1. This will be done later. Exactly as in the one period case we can prove that $\|u_1(\xi) - u_1(\eta)\|_\infty \le \|\xi - \eta\|_\infty$. And we also have that $\xi \le \eta$ implies $u_1(\xi) \le u_1(\eta)$. The set $\mathcal{A}_1 = \{\xi \mid u_1(\xi) \ge 0\}$ is a convex set, it is weak* closed because of the Fatou property and it contains the cone $L^\infty_+(\mathcal{F}_2)$ of nonnegative elements from $L^\infty(\mathcal{F}_2)$.

The convexity allows us to prove that $u_1(\xi)$ can be "localised".

Proposition 46 *If $A \in \mathcal{F}_1$ then for all ξ: $u_1(\xi \mathbf{1}_A) = \mathbf{1}_A u_1(\xi)$*

Proof. For $\xi \in L^\infty(\mathcal{F}_2)$ we have by concavity:

$$u_1(\xi\mathbf{1}_A)\mathbf{1}_A + u_1(\xi\mathbf{1}_{A^c})\mathbf{1}_{A^c} \le u_1(\xi\mathbf{1}_A\mathbf{1}_A + \xi\mathbf{1}_{A^c}\mathbf{1}_{A^c}) = u_1(\xi).$$

This implies that $u_1(\xi\mathbf{1}_A)\mathbf{1}_A \le u_1(\xi)\mathbf{1}_A$. But the cocavity also implies that $u_1(\xi)\mathbf{1}_A = u_1(\xi)\mathbf{1}_A + u_1(0)\mathbf{1}_{A^c} \le u_1(\xi\mathbf{1}_A)$. Multiplying with $\mathbf{1}_A$ gives $u_1(\xi)\mathbf{1}_A \le u_1(\xi\mathbf{1}_A)\mathbf{1}_A$, hence $u_1(\xi\mathbf{1}_A) = \mathbf{1}_A u_1(\xi)$. □

Corollary 15 *For $\xi \in L^\infty(\mathcal{F}_2)$ and $A \in \mathcal{F}_1$, we have $u_1(\xi) \ge 0$ on A if and only if $\xi\mathbf{1}_A \in \mathcal{A}_1$.*

Proposition 47 *Let $A_n; n \ge 1$ be a partition of Ω into $\mathcal{F}_1-$measurable sets, let $\xi \in L^\infty$. Then $\xi \in \mathcal{A}_1$ if and only if $\xi = \sum_n \xi^n \mathbf{1}_{A_n}$ where for each $n : \xi^n \in \mathcal{A}_1$*

Proof. If $\xi \in \mathcal{A}_1$ then for $A \in \mathcal{F}_1$: $\xi\mathbf{1}_A \in \mathcal{A}_1$ So $\xi^n = \xi\mathbf{1}_{A_n}$ defines a sequence in $\mathcal{A}_1$ with $\xi = \sum_n \xi^n\mathbf{1}_{A_n}$. Conversely, if $\xi = \sum_n \xi^n\mathbf{1}_{A_n}$ with $\xi_n \in \mathcal{F}_1$ then for all n we can write $u_1(\xi)\mathbf{1}_{A_n} = u_1(\xi\mathbf{1}_{A_n}) = u_1(\xi^n\mathbf{1}_{A_n}) = u_1(\xi^n)\mathbf{1}_{A_n} \ge 0$. So $u_1(\xi) \ge 0$ a.s. , proving that $\xi \in \mathcal{A}_1$. □

The representation theorem takes almost the same form as for the one period case. Of course we need to introduce conditional expectations. So we introduce

$$c_1(\mathbb{Q}) = \text{ess.sup}\{\mathbb{E}_\mathbb{Q}[-\xi \mid \mathcal{F}_1] \mid u_1(\xi) \ge 0\}.$$

We remark that the function $c_1(\mathbb{Q})$ is defined up to sets of $\mathbb{Q}$ measure zero. This makes it difficult to compare and compose c_1-functions for measures that are not equivalent to $\mathbb{P}$. We will avoid this problem by restricting, there where possible, to probabilities that are equivalent to $\mathbb{P}$ on the sigma algebra $\mathcal{F}_1$. This class is still bigger than the class of probability measures that are equivalent to $\mathbb{P}$ on the bigger sigma-algebra $\mathcal{F}_2$. The measures that are equivalent to $\mathbb{P}$ on $\mathcal{F}_1$ have then the same $\mathcal{F}_1-$measurable zero sets as $\mathbb{P}$ and inequalities of the form a.s. are then well defined. In calculating $c_1(\mathbb{Q})$ we take the essential supremum over an uncountable set of functions. That such a random variable exists is a standard exercise in probability theory. We can also avoid the zero-set problem as follows. For a measure $\mathbb{Q} \in \mathbf{P}$, let us write $Z_2 = \frac{d\mathbb{Q}}{d\mathbb{P}}$ and $Z_1 = \mathbb{E}_{\mathbb{P}}[Z_2 \mid \mathcal{F}_1] = \frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_1} = \frac{d(\mathbb{Q}|\mathcal{F}_1)}{d(\mathbb{P}|\mathcal{F}_1)}$. Then we can define

$$
\begin{aligned}
c_1(\mathbb{Q}) = \text{ess.inf}\{\eta\colon \Omega \to &\mathbb{R}_+ \cup \{+\infty\} \mid \eta \text{ is } \mathcal{F}_1 \text{ measurable} \\
&\text{and for all } \xi \in \mathcal{A}_1 : \mathbb{E}[-\xi Z_2 \mid \mathcal{F}_1] \le \eta Z_1\}.
\end{aligned}
$$

That this formula defines $c_1(\mathbb{Q})$ follows from Bayes' rule. It has the advantage that we only use conditional expectations with respect to $\mathbb{P}$. On the set $\{Z_1 = 0\}$ it returns the value 0 for $c_1(\mathbb{Q})$ which is not in contradiction with the previous definition since this set has $\mathbb{Q}-$measure 0. We will not always use this extension for $c_1(\mathbb{Q})$. In some case it can lead to wrong expressions. So for the moment we only see this extension as another way of defining $c_1(\mathbf{Q})$.

In the following chapters we will frequently use the following properties of the function c_1. Since each of them requires some technical changes when compared to the one-period case, we prefer to separate them in different propositions. We start with the continuity. We have the following continuity property for the function c_1.

Proposition 48 *Let $\sum_n \|\mathbb{Q}^n - \mathbb{Q}\|_1 < \infty$. With the same notation as above we then have $Z_2^n \to Z_2$ and $Z_1^n \to Z_1$, both convergences a.s. . We also have $c_1(\mathbb{Q}) \le \liminf c_1(\mathbb{Q}^n)$.*

Proof. We remark that the statement about the a.s. convergence holds. We have for every $\xi \in L^\infty$, $\lim_n \mathbb{E}[-\xi Z_2^n \mid \mathcal{F}_1] = \mathbb{E}[-\xi Z_2 \mid \mathcal{F}_1]$, a.s. . Now take $\xi \in \mathcal{A}$. For each n we have that $\mathbb{E}[-\xi Z_2^n \mid \mathcal{F}_1] \le c_1(\mathbb{Q}^n) Z_1^n$. By taking limits gives $\mathbb{E}[-\xi Z_2 \mid \mathcal{F}_1] \le \liminf_n \left(c_1(\mathbb{Q}^n) Z_1^n\right)$. The latter is equal to $\left(\liminf_n c_1(\mathbb{Q}^n)\right) Z_1$ on the set $\{Z_1 > 0\}$ whereas the former is equal to 0 on the set $\{Z_1 = 0\}$. So we get $\mathbb{E}[-\xi Z_2 \mid \mathcal{F}_1] \le \liminf c_1(\mathbb{Q}^n) Z_1$ a.s. . This shows that $c_1(\mathbb{Q}) \le \liminf c_1(\mathbb{Q}^n)$. □

Remark 79 In case the convergence of $\mathbb{Q}^n$ is slower than we required in the proposition, the result maybe wrong. We can give an example where

$\liminf_n c(\mathbb{Q}^n) = 0$ but $c_1(\mathbb{Q}) = +\infty$. This has to do with the way one calculates lim inf. This is defined pointwise and that is the reason. In case we would change the definition into something like "lim inf $-\mathbb{P}$", the proposition would hold for converging sequences of measures. The basic fact is that even when f_n is a uniformly integrable sequence tending to f a.s. , we cannot necessarily conclude that $\mathbb{E}[f_n \mid \mathcal{F}_1] \to \mathbb{E}[f \mid \mathcal{F}_1]$ a.s. . The existence of such sequences is a well known exercise in advanced probability courses. To do this operation of interchanging conditional expectations and convergence a.s. , we need a dominated convergence. We do not want to pursue this discussion.

Definition 27 or Notation *The set of probability measures that coincide with $\mathbb{P}$ on the sigma-algebra $\mathcal{F}_1$ is denoted by $\mathbf{P}_1$.*

The set $\mathbf{P}_1$ has a nice stability property that we can use to paste together several measures. We will use the construction each time we need to show that some set has a lattice property. We will use the device without mentioning it. Let A_n be a partition of Ω into $\mathcal{F}_1$ measurable sets. Let $\mathbb{Q}_n$ be a sequence in $\mathbf{P}_1$, then $\mathbb{Q}[B] = \sum_n \mathbb{Q}_n[B \cap A_n]$ defines a measure in $\mathbf{P}_1$. The convexity property of the function c_1 is proved as in the one-period case except that we need to take some precautions. So we only use the function c_1 on the set $\mathbf{P}_1$. On this set we can make convex combinations in a wider sense. Indeed if $\mathbb{Q}^1, \mathbb{Q}^2 \in \mathbf{P}_1$ and if $0 \le \lambda \le 1$ is an $\mathcal{F}_1-$measurable function, we can define the measure $\lambda \cdot \mathbb{Q}^1 + (1-\lambda) \cdot \mathbb{Q}^2$ as the measure with density $\lambda Z_2^1 + (1-\lambda) Z_2^2$. Since $\mathbb{E}[Z_2^i \mid \mathcal{F}_1] = 1$, the outcome is indeed an element of $\mathbf{P}_1$.

Proposition 49 *The function $c_1 \colon \mathbf{P}_1 \to L^0(\Omega, \mathcal{F}_1, \mathbb{P}; [0, +\infty])$ is convex in the sense that for $\mathcal{F}_1-$measurable functions λ with $0 \le \lambda \le 1$, the convexity inequality holds (we put $(+\infty).0 = 0$):*

$$c_1\left(\lambda \cdot \mathbb{Q}^1 + (1-\lambda) \cdot \mathbb{Q}^2\right) \le \lambda c_1(\mathbb{Q}^1) + (1-\lambda) c_1(\mathbb{Q}^2).$$

Proof. Let $\xi \in \mathcal{A}_1$. We have

$$\begin{aligned}\mathbb{E}[-\xi(\lambda Z_2^1 + (1-\lambda)Z_2^2) \mid \mathcal{F}_1] &= \lambda \mathbb{E}[-\xi Z_2^1 \mid \mathcal{F}_1] + (1-\lambda)\mathbb{E}[-\xi Z_2^2 \mid \mathcal{F}_1] \\ &\le \lambda c_1(\mathbb{Q}^1) + (1-\lambda) c_1(\mathbb{Q}^2).\end{aligned}$$

□

Remark 80 It is precisely the equality

$$\mathbb{E}[\xi Z_2 \mid \mathcal{F}_1] = \mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1]$$

that forces us to use the set $\mathbf{P}_1$. The conditional expectation operator $\mathbb{E}_{\mathbb{Q}}[\cdot \mid \mathcal{F}_1]$ is only affine with respect to $\mathbb{Q}$ on the set $\mathbf{P}_1$ and not on the set $\mathbf{P}$

Proposition 50 *Let $\mathbb{Q}$ be equivalent to $\mathbb{P}$ on $\mathcal{F}_1$, let $Z_2 = \frac{d\mathbb{Q}}{d\mathbb{P}}$ and $Z_1 = \mathbb{E}_{\mathbb{P}}[Z_2 \mid \mathcal{F}_1] = \frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_1} = \frac{d(\mathbb{Q}|\mathcal{F}_1)}{d(\mathbb{P}|\mathcal{F}_1)}$. Let $\mathbb{Q}'$ be the measure defined as $d\mathbb{Q}' = \frac{Z_2}{Z_1}\, d\mathbb{P}$. The measure $\mathbb{Q}'$ coincides with $\mathbb{P}$ on the sigma algebra $\mathcal{F}_1$, in our notation $\mathbb{Q}' \in \mathbf{P}_1$ and $c_1(\mathbb{Q}') = c_1(\mathbb{Q})$.*

Proof. This obvious but useful statement follows immediately from Bayes's rule: $\mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] = \mathbb{E}_{\mathbb{Q}'}[\xi \mid \mathcal{F}_1]$. □

The proposition has an extension to arbitrary elements $\mathbb{Q} \in \mathbf{P}$.

Proposition 51 *Let $\mathbb{Q} \in \mathbf{P}$ and define $d\mathbb{Q}' = \left(\frac{Z_2}{Z_1}\mathbf{1}_{\{Z_1>0\}} + \mathbf{1}_{\{Z_1=0\}}\right) d\mathbb{P}$. Then $\mathbb{Q}' \in \mathbf{P}_1$ and we have*

1. *for all $\xi \in L^\infty : \mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1]\mathbf{1}_{\{Z_1>0\}} = \mathbb{E}_{\mathbb{Q}'}[\xi \mid \mathcal{F}_1]\mathbf{1}_{\{Z_1>0\}}$, more precisely*

$$\mathbb{E}_{\mathbb{Q}'}[\xi \mid \mathcal{F}_1] = \mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1]\mathbf{1}_{\{Z_1>0\}} + \mathbb{E}_{\mathbb{P}}[\xi \mid \mathcal{F}_1]\mathbf{1}_{\{Z_1=0\}},$$

2. *$c_1(\mathbb{Q}')\mathbf{1}_{\{Z_1>0\}} = c_1(\mathbb{Q})\mathbf{1}_{\{Z_1>0\}}$ more precisely*

$$c_1(\mathbb{Q}') = c_1(\mathbb{Q})\mathbf{1}_{\{Z_1>0\}} + c_1(\mathbb{P})\mathbf{1}_{\{Z_1=0\}}.$$

Proof. Again an application of Bayes's rule. □

Proposition 52 *The set $\{\mathbb{E}_{\mathbb{Q}}[-\xi \mid \mathcal{F}_1] \mid u_1(\xi) \geq 0\}$ is a lattice. This means that if $\eta_1, \eta_2 \in \{\mathbb{E}_{\mathbb{Q}}[-\xi \mid \mathcal{F}_1] \mid u_1(\xi) \geq 0\}$ then $\max(\eta_1, \eta_2) \in \{\mathbb{E}_{\mathbb{Q}}[-\xi \mid \mathcal{F}_1] \mid u_1(\xi) \geq 0\}$. As a consequence there is an increeasing sequence $\eta_n \in \{\mathbb{E}_{\mathbb{Q}}[-\xi \mid \mathcal{F}_1] \mid u_1(\xi) \geq 0\}$ such that $\eta_n \uparrow c_1(\mathbb{Q})$, $\mathbb{Q}-a.s.$.*

Proof. If $\eta_i = \mathbb{E}_{\mathbb{Q}}[-\xi_i \mid \mathcal{F}_1]$ with $\xi_i \in \mathcal{A}_1$, let $A = \{\eta_1 > \eta_2\}$. The set $A \in \mathcal{F}_1$ and hence $\xi = \mathbf{1}_A\xi_1 + \mathbf{1}_{A^c}\xi_2 \in \mathcal{A}_1$. Clearly $\mathbb{E}_{\mathbb{Q}}[-\xi \mid \mathcal{F}_1] = \max(\eta_1, \eta_2)$. Now take a function $\phi\colon \mathbb{R} \to]-1,1[$ that is bijective and increasing. For instance we could take $\phi = \frac{2}{\pi}\arctan$. Then the set $\mathcal{B} = \{\phi(\eta) \mid \eta = \mathbb{E}_{\mathbb{Q}}[-\xi \mid \mathcal{F}_1]; \xi \in \mathcal{A}_1\}$ is still a lattice. Let $\alpha = \sup_{f\in\mathcal{B}} \mathbb{E}_{\mathbb{Q}}[f]$. Let f_n be a sequence in $\mathcal{B}$ such that $\mathbb{E}_{\mathbb{Q}}[f_n] \to \alpha$. By the lattice property we may suppose that the sequence f_n is nondecreasing. Its limit f exists $\mathbb{Q}-$a.s. and satisfies $\mathbb{E}_{\mathbb{Q}}[f] = \alpha$. It is easy to see that for all $g \in \mathcal{B}$ we must have $g \le f$, $\mathbb{Q}-$a.s. . We can now take $\eta_n = \phi^{-1}(f_n)$ and $c_1(\mathbb{Q}) = \phi^{-1}(f)$ where $\phi^{-1}(1) = +\infty$. □

Proposition 53 $\sup\{\mathbb{E}_{\mathbb{Q}}[-\xi] \mid \xi \in \mathcal{A}_1\} = \mathbb{E}_{\mathbb{Q}}[c_1(\mathbb{Q})]$.

Proof. This follows immediately from the preceding result. Take η_n as in the previous proposition. We have $\mathbb{E}_{\mathbb{Q}}[\eta_n] \uparrow \mathbb{E}_{\mathbb{Q}}[c_1(\mathbb{Q})]$. □

Remark 81 We repeat that we did not claim any relation with $c_0(\mathbb{Q}) = \sup\{\mathbb{E}_{\mathbb{Q}}[-\xi] \mid \xi \in \mathcal{A}_0\}$, where $\mathcal{A}_0 = \{\xi \mid u_0(\xi) \ge 0\}$. This will be done in the discussion on time consistency.

Theorem 58 *For an element $\xi \in L^\infty(\mathcal{F}_2)$ the following are equivalent*

1. $\xi \in \mathcal{A}_1$
2. *for every* $\mathbb{Q} \in \mathbf{P}$: $\mathbb{E}_{\mathbb{Q}}[\xi] + \mathbb{E}_{\mathbb{Q}}[c_1(\mathbb{Q})] \ge 0$.
3. *for every* $\mathbb{Q} \in \mathbf{P}$: $\mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}) \ge 0$, $\mathbb{Q}$ *a.s.* .
4. *for every* $\mathbb{Q} \in \mathbf{P}_1$: $\mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}) \ge 0$, $\mathbb{P}$ *a.s.* .

Proof The first two items are equivalent by the theory of the one period case. Item 3 clearly implies item 2. We now show that item 1 (or 2) imply item 3. Take $\mathbb{Q} \in \mathbf{P}$ and take any $\eta \in L^\infty(\mathcal{F}_1)$ with $\mathbb{E}_{\mathbb{P}}[\eta] = 1$. The measure $\mathbb{Q}$"with density

$$\frac{d\mathbb{Q}"}{d\mathbb{P}} = \eta\left(\frac{Z_2}{Z_1}\mathbf{1}_{\{Z_1>0\}} + \mathbf{1}_{Z_1=0\}}\right)$$

is a new probability and for every $\xi \in L^\infty(\mathcal{F}_2)$ we have

$$\mathbb{Q}\text{ a.s. } : \mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] = \mathbb{E}_{\mathbb{Q}"}[\xi \mid \mathcal{F}_1] \text{ and } c_1(\mathbb{Q}") = c_1(\mathbb{Q}).$$

If we take η arbitrary but supported on the set $\{Z_1 > 0\}$ and observe that

$$\mathbb{E}_{\mathbb{Q}"}[(\mathbb{E}_{\mathbb{Q}"}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}"))] = \mathbb{E}_{\mathbb{P}}[\eta\,(\mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}))] \ge 0,$$

we get that $\mathbb{Q}$ a.s. $\mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}) \ge 0$. Clearly item 3 implies item 4 and we can prove that item 4 implies 3. In case $\mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q})$ is not everywhere nonnegative we get the existence of a set $A \subset \{Z_1 > 0\}$ such that on A: $\mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}) < 0$. We now replace $\mathbb{Q}$ by the measure $\mathbb{Q}'$ (as in Proposition 51) and we get the existence of a measure $\mathbb{Q}' \in \mathbf{P}_1$ with $\mathbb{E}_{\mathbb{Q}'}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}') < 0$ on A. □

Theorem 59 *The following parametrised duality equality is valid:*

$$u_1(\xi) = \text{ess.inf}\,\{\mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}) \mid \mathbb{Q} \in \mathbf{P}_1\}\,.$$

In fact we have the slightly stronger statement: for every strictly positive $\mathcal{F}_1$ measurable function $\varepsilon \le 1$, there is $\mathbb{Q} \in \mathbf{P}_1$ such that

$$\mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}) \le u_1(\xi) + \varepsilon \quad \mathbb{Q}\ a.s.\ .$$

Proof The proof uses an exhaustion argument as well as the lattice properties. To prove the theorem we may suppose that $u_1(\xi) = 0$, otherwise we replace it by $\xi - u_1(\xi)$. Take $\varepsilon \le 1$ a strictly positive $\mathcal{F}_1$ measurable function. We will show that there is $\mathbb{Q} \in \mathbf{P}_1$ such that $\mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}) \le \varepsilon$ a.s. . This will end the proof. We start with the following lemma.

Lemma 21 *Let $\eta \in L^\infty$ be such that $u_1(\eta) < 0$ a.s. . Then for every $B \in \mathcal{F}_1$ with $\mathbb{P}[B] > 0$, there is a measure $\mathbb{Q} \in \mathbf{P}_1$ and a set $A \in \mathcal{F}_1$, $A \subset B$, $\mathbb{P}[A] > 0$ such that on the set A: $\mathbb{E}_{\mathbb{Q}}[\eta] + c_1(\mathbb{Q}) < 0$.*

Proof of the lemma This is straightforward since $\eta \mathbf{1}_B$ is not in $\mathcal{A}_1$ and hence item 4 of the previous theorem gives the desired measure and a set A with $\mathbb{E}_{\mathbb{Q}}[\eta \mathbf{1}_B \mid \mathcal{F}_1] + c_1(\mathbb{Q}) < 0$ on A. Of course we must have $A \subset B$ since $\mathbb{E}_{\mathbb{Q}}[\eta \mathbf{1}_B \mid \mathcal{F}_1] + c_1(\mathbb{Q}) \ge 0$ on B^c □

We now look at the class

$$\mathcal{C} = \{C \in \mathcal{F}_1 \mid \text{ there is } \mathbb{Q} \in \mathbf{P}_1 \text{ with } \mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}) < \varepsilon \text{ on } C\}\,.$$

This class is stable for countable unions and hence has a maximal element, say C_0. If $\mathbb{P}[C_0] < 1$, then we can apply the lemma to $\xi - \varepsilon$ and the set C_0^c to get a set A disjoint of C_0 and belonging to the class $\mathcal{C}$. This is a contradiction to the maximality of C_0. □

Corollary 16 *For every $\varepsilon \le 1$, a strictly positive $\mathcal{F}_1$ measurable function, we have the existence of $\mathbb{Q} \in \mathbf{P}_1$ with $c_1(\mathbb{Q}) \le \varepsilon$, a.s. .*

Proof Just repeat the proof of the theorem for $\xi = 0$. □

Remark 82 The previous theorem can also be proved in a different way. In the one period case it is just an application of the Hahn-Banach theorem. The situation here is different in the sense that we have a sigma-algebra $\mathcal{F}_1$. Instead of using desintegration of measures (which needs topological properties of the space Ω), we could try to use the Hahn-Banach theorem in a parametrised way. This idea was developed by Filipovic, Kupper and Vogelpoth, see [65] for details and more information on this technique.

Exactly as in the one period case we can give conditions under which we can restrict calculations to equivalent measures. We recall that if $\xi \in L^\infty_+(\mathcal{F}_2)$, then the set $\{\mathbb{E}_\mathbb{Q}[\xi \mid \mathcal{F}_1] > 0\}$ is the same for all equivalent measures $\mathbb{Q}$ and is the smallest set in $\mathcal{F}_1$ that contains $\{\xi > 0\}$. The appropriate definition of relevance is

Definition 28 *The function u_1 is called relevant if $\xi \in L^\infty_+(\mathcal{F}_2)$ implies that $u_1(-\xi) < 0$ on the set $\{\mathbb{E}[\xi \mid \mathcal{F}_1] > 0\}$.*

Exercise 27 Suppose u_1 is relevant . Show that for $\xi \in L^\infty_+(\mathcal{F}_2)$: $\{u_1(-\xi) < 0\} = \{\mathbb{E}[\xi \mid \mathcal{F}_1] > 0\}$.

Exercise 28 Show that the following two statements are equivalent

1. u_1 is relevant.
2. For every $\varepsilon > 0$, $A \in \mathcal{F}_2$, $u_1(-\varepsilon \mathbf{1}_A) < 0$ on the set $\{\mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1] > 0\}$.

Proposition 54 *Suppose that u_1 is relevant. Then for every $\mathcal{F}_1$ measurable function, $1 \geq \varepsilon > 0$, there is an equivalent measure $\mathbb{Q} \in \mathbf{P}_1$ such that $c_1(\mathbb{Q}) \leq \varepsilon$.*

Proof The proof uses exhaustion. Take $\mathbb{P}[A] > 0$, then for $\varepsilon > 0$ we have $u_1(-\varepsilon \mathbf{1}_A) < 0$ on $\{\mathbb{E}[\mathbf{1}_A \mid \mathcal{F}_1] > 0\}$. Hence there is $\mathbb{Q} \in \mathbf{P}_1$ with $\varepsilon \mathbb{E}_\mathbb{Q}[-\mathbf{1}_A \mid \mathcal{F}_1] + c_1(\mathbb{Q}) \leq u_1(-\varepsilon \mathbf{1}_A) + \varepsilon \leq \varepsilon$. This shows that $c_1(\mathbb{Q}) \leq \varepsilon + \varepsilon \mathbb{E}_\mathbb{Q}[\mathbf{1}_A \mid \mathcal{F}_1] \leq 2\varepsilon$ and $\mathbb{E}_\mathbb{Q}[\varepsilon \mathbf{1}_A] \geq \mathbb{E}_\mathbb{P}[-u_1(-\varepsilon \mathbf{1}_A)] > 0$, hence $\mathbb{Q}[A] > 0$. Using exhaustion we get a measure $\mathbb{Q} \in \mathbf{P}_1$ such that $\mathbb{Q} \sim \mathbb{P}$ and $c_1(\mathbb{Q}) \leq \varepsilon$.

Proposition 55 *Suppose that u_1 is relevant. If $\mathbb{Q} \in \mathbf{P}_1$, there is a sequence $\mathbb{Q}^n \in \mathbf{P}_1^e$ such that $c_1(\mathbb{Q}) = \lim c_1(\mathbb{Q}^n)$.*

Proof. Take $\mathbb{Q}^0 \in \mathbf{P}_1^e$ such that $c_1(\mathbb{Q}^0) < \infty$ a.s. . By the Proposition 54 above, this is possible. Define for $n \geq 1$ the measure $\mathbb{Q}^n = \frac{1}{n^2}\mathbb{Q}^0 + \frac{n^2-1}{n^2}\mathbb{Q}$. Clearly $\mathbb{Q}^n \in \mathbf{P}_1^e$ and $\sum_n \|\mathbb{Q}^n - \mathbb{Q}\| < +\infty$. We therefore have that $c_1(\mathbb{Q}) \leq \liminf c_1(\mathbb{Q}^n)$. But the convexity relation implies $c_1(\mathbb{Q}^n) \leq \frac{1}{n^2}c_1(\mathbb{Q}^0) + \frac{n^2-1}{n^2}c_1(\mathbb{Q})$. This implies that $c_1(\mathbb{Q}) \geq \limsup c_1(\mathbb{Q}^n)$. From here we get the equality $c_1(\mathbb{Q}) = \lim c_1(\mathbb{Q}^n)$. □

Theorem 60 *Suppose that u_1 is relevant, then for each element ξ in L^∞ we have*

$$\begin{aligned} u_1(\xi) &= \text{ess.inf}\{\mathbb{E}_\mathbb{Q}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}) \mid \mathbb{Q} \in \mathbf{P}^e\} \\ &= \text{ess.inf}\{\mathbb{E}_\mathbb{Q}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}) \mid \mathbb{Q} \in \mathbf{P}_1^e\} \end{aligned}$$

Proof. The proof follows from the duality relation and the previous proposition. The reader can fill in the details if she wants. □

11.2 Time Consistency

We use the same notation as in the previous section. A two period model with filtration $(\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2)$ and concave utility functions u_0, u_1. The function u_2 is simply the identity. Their penalty functions are denoted by c_0, c_1

Definition 29 *We call (u_0, u_1) time consistent (or when confusing can arise time consistent with respect to $(\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2)$) if for all pairs ξ, η, $u_1(\xi) \leq u_1(\eta)$ a.s. implies $u_0(\xi) \leq u_0(\eta)$.*

Remark 83 This definition, in a little bit different context, was introduced by Koopmans, [89], [90], [91].

We need the following notation

$$\begin{aligned} \mathcal{A}_0 &= \{\xi \mid u_0(\xi) \geq 0\} \\ \mathcal{A}_1 &= \{\xi \mid u_1(\xi) \geq 0\} \\ \mathcal{A}_{0,1} &= \{\xi \mid \xi \in L^\infty(\mathcal{F}_1); u_0(\xi) \geq 0\} = L^\infty(\mathcal{F}_1) \cap \mathcal{A}_0 \\ c_{0,1}(\mathbb{Q}) &= \sup\{\mathbb{E}_\mathbb{Q}[-\xi] \mid \xi \in \mathcal{A}_{0,1}\}. \end{aligned}$$

We can now give an equivalence theorem for time consistency.

Theorem 61 *With the notation above, are equivalent*

1. *(u_0, u_1) is time consistent*
2. *for all $\xi \in L^\infty(\mathcal{F}_2) : u_0(u_1(\xi)) = u_0(\xi)$ (recursivity)*
3. *$\mathcal{A}_0 = \mathcal{A}_{0,1} + \mathcal{A}_1$ (decomposition property)*
4. *for all $\mathbb{Q}$: $c_0(\mathbb{Q}) = c_{0,1}(\mathbb{Q}) + \mathbb{E}_\mathbb{Q}[c_1(\mathbb{Q})]$ (cocycle property).*

If (u_0, u_1) is time consistent and u_0 is relevant then u_1 is also relevant and we have $\mathcal{A}_1 = \{\xi \mid$ for all $A \in \mathcal{F}_1 : \xi \mathbf{1}_A \in \mathcal{A}_0\}$.

Proof. $1 \Rightarrow 2$. Let $\xi \in L^\infty(\mathcal{F}_2)$. Since $u_1(\xi) \le u_1(u_1(\xi))$ and $u_1(\xi) \ge u_1(u_1(\xi))$, the definition of time consistency gives $u_0(u_1(\xi)) = u_0(\xi)$. $2 \Rightarrow 1$. Let ξ, η be given and suppose that $u_1(\xi) \le u_1(\eta)$, then the monotonicity of u_0 implies $u_0(u_1(\xi)) \le u_0(u_1(\eta))$ and recursivity gives $u_0(\xi) \le u_0(\eta)$. $1, 2 \Rightarrow 3$. Take $\xi \in \mathcal{A}_0$ then we have $u_1(\xi) \in \mathcal{A}_{0,1}$. But $\xi = \xi - u_1(\xi) + u_1(\xi)$ and trivially $\xi - u_1(\xi) \in \mathcal{A}_1$ so we get $\xi \in \mathcal{A}_1 + \mathcal{A}_{0,1}$. Let now $\xi = \eta + \zeta$ with $\eta \in \mathcal{A}_1$ and $\zeta \in \mathcal{A}_{0,1}$. We will show that $\xi \in \mathcal{A}_0$. We have $u_1(\xi) = u_1(\eta) + \zeta \ge \zeta$ and hence $u_1(\xi) \in \mathcal{A}_{0,1}$. From here we see that $u_0(\xi) = u_0(u_1(\xi)) \ge 0$ and hence $\xi \in \mathcal{A}_0$. This proves that $\mathcal{A}_0 = \mathcal{A}_{0,1} + \mathcal{A}_1$. $3 \Rightarrow 1, 2$. Let us suppose that $u_0(\xi) = 0$. We can write $\xi = \eta + \zeta$ where $\eta \in \mathcal{A}_1$ and $\zeta \in \mathcal{A}_{0,1}$. We may suppose that $u_1(\eta) = 0$ since we can replace the decomposition by $\xi = \eta - u_1(\eta) + (\zeta + u_1(\eta))$. We then get $u_1(\xi) = \zeta$ and we have to show that $u_0(\zeta) = 0$. If this is not true then we have $u_0(\zeta) > 0$ ($u_0(\zeta) \ge 0$ since $\zeta \in \mathcal{A}_{0,1}$) and $\xi - u_0(\zeta) = \eta + \zeta - u_0(\zeta) \in \mathcal{A}_1 + \mathcal{A}_{0,1} \subset \mathcal{A}_0$. This gives $u_0(\xi - u_0(\zeta)) \ge 0$, of course a contradiction to $u_0(\xi) = 0$. $3 \Rightarrow 4$. Because $\mathcal{A}_0 = \mathcal{A}_1 + \mathcal{A}_{0,1}$ we have for all $\mathbb{Q}$

$$\begin{aligned}\sup_{\xi \in \mathcal{A}_0} \mathbb{E}_\mathbb{Q}[-\xi] &= \sup_{\eta \in \mathcal{A}_1} \mathbb{E}_\mathbb{Q}[-\eta] + \sup_{\zeta \in \mathcal{A}_{0,1}} \mathbb{E}_\mathbb{Q}[-\zeta] \\ &= \mathbb{E}_\mathbb{Q}[c_1(\mathbb{Q})] + c_{0,1}(\mathbb{Q}).\end{aligned}$$

$4 \Rightarrow 2$. For measures $\mathbb{Q} \in \mathbf{P}^e$, $\mathbb{Q}^a \in \mathbf{P}_1$, let us introduce two other measures, defined through the density process Z_1, Z_2 of $\mathbb{Q}$. The measure $\mathbb{Q}_1$ has density Z_1, it coincides with $\mathbb{Q}$ for elements that are $\mathcal{F}_1$−measurable. The second measure, $\mathbb{Q}_2$ has density $\frac{Z_2}{Z_1}\mathbf{1}_{\{Z_1>0\}} + \frac{d\mathbb{Q}^a}{d\mathbb{P}}\mathbf{1}_{\{Z_1=0\}}$ and on $\{Z_1 > 0\}$ it yields the same *conditional* expectation (with respect to $\mathcal{F}_1$) as the measure $\mathbb{Q}$. We observe that for every $\mathbb{Q}^a \in \mathbf{P}_1$ we have $\mathbb{E}_\mathbb{Q}[\xi] = \mathbb{E}_{\mathbb{Q}_1}[\mathbb{E}_{\mathbb{Q}_2}[\xi \mid \mathcal{F}_1]]$. If $\mathbb{Q}$ runs through the set $\mathbf{P}^e$ and $\mathbb{Q}^a$ runs through the set $\mathbf{P}_1$, then $\mathbb{Q}_2$ describes the set of all elements in $\mathbf{P}_1$. We know that $c_1(\mathbb{Q})\mathbf{1}_{\{Z_1>0\}} = c_1(\mathbb{Q}_2)\mathbf{1}_{\{Z_1>0\}}$. Conversely if $\mathbb{Q}_1$ is a measure with density $Z_1 > 0$ that is measurable with respect to $\mathcal{F}_1$, if $\mathbb{Q}_2$ is a measure in $\mathbf{P}_1$ with density $L_2 \ge 0$, then the measure defined with the density $Z_1 L_2$ is a measure in $\mathbf{P}$. This multiplicative decomposition of measures allows to write the following

$$
\begin{aligned}
u_0(\xi) &= \inf_{\mathbb{Q}\ll\mathbb{P}} (\mathbb{E}_\mathbb{Q}[\xi) + c_0(\mathbb{Q})) \\
&= \inf_{\mathbb{Q}\ll\mathbb{P},\mathbb{Q}^a\in\mathbf{P}_1} (\mathbb{E}_\mathbb{Q}[\xi] + c_{0,1}(\mathbb{Q}_1) + \mathbb{E}_{\mathbb{Q}_1}[c_1(\mathbb{Q}_2)]) \\
&= \inf_{\mathbb{Q}\ll\mathbb{P},\mathbb{Q}^a\in\mathbf{P}_1} (\mathbb{E}_{\mathbb{Q}_1}[\mathbb{E}_{\mathbb{Q}_2}[\xi \mid \mathcal{F}_1]] + \mathbb{E}_{\mathbb{Q}_1}[c_1(\mathbb{Q}_2)] + c_{0,1}(\mathbb{Q}_1)) \\
&= \inf_{\mathbb{Q}\ll\mathbb{P}} (\mathbb{E}_{\mathbb{Q}_1}[\mathbb{E}_{\mathbb{Q}_2}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}_2)] + c_{0,1}(\mathbb{Q}_1)) \\
&\text{and because the set } \{\mathbb{E}_{\mathbb{Q}_2}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}_2) \mid \mathbb{Q}_2 \in \mathbf{P}_1\} \text{ is a lattice} \\
&= \inf_{\mathbb{Q},\mathbb{Q}^a} \left(\mathbb{E}_{\mathbb{Q}_1}\left[\inf_{\mathbb{Q}_2\in\mathbf{P}_1} (\mathbb{E}_{\mathbb{Q}_2}[\xi \mid \mathcal{F}_1] + c_1(\mathbb{Q}_2))\right] + c_{0,1}(\mathbb{Q}_1)\right) \\
&= \inf_{\mathbb{Q}_1} (\mathbb{E}_{\mathbb{Q}_1}[u_1(\xi)] + c_{0,1}(\mathbb{Q}_1)) = u_0(u_1(\xi)).
\end{aligned}
$$

Suppose now that (u_0, u_1) is time consistent and that u_0 is relevant. We will show that also u_1 is relevant. Take $\xi \geq 0$ and suppose that the set $A = \{\mathbb{E}[\xi \mid \mathcal{F}_1] > 0\} \cap \{u_1(-\xi) = 0\}$ has positive probability. We may replace ξ by $\xi \mathbf{1}_A$ to get an element ξ such that $u_1(-\xi) = 0$ a.s. . Because u_0 is relevant we get

$$0 > u_0(-\xi) = u_0(u_1(\xi)) = 0,$$

a contradiction. The last line is proved as follows:

$$
\begin{aligned}
\xi \in \mathcal{A}_1 &\Leftrightarrow u_1(\xi) \geq 0 \\
&\Leftrightarrow \forall A \in \mathcal{F}_1 : u_0(u_1(\xi)\mathbf{1}_A) \geq 0 \text{ because of relevance)} \\
&\Leftrightarrow \forall A \in \mathcal{F}_1 : u_0(u_1(\xi\mathbf{1}_A)) \geq 0 \\
&\Leftrightarrow \forall A \in \mathcal{F}_1 : u_0(\xi\mathbf{1}_A) \geq 0 \\
&\Leftrightarrow \forall A \in \mathcal{F}_1 : \xi\mathbf{1}_A \in \mathcal{A}_0.
\end{aligned}
$$

□

Remark 84 The last line of the theorem is very important. It states that for a given filtration and if u_0 is relevant, there is at most one utility function u_1 such that the system becomes time consistent. So in a time consistent framework, u_0 defines the intermediate utility function. Of course not every u_0 allows for such a time consistent construction! An alternative way to develop the theory would be to start with $\mathcal{A}_0$, then define $\mathcal{A}_1$ as above, then use the decomposition property to define time consistency and afterwards define u_1. In doing so the theory gets closer to the theory of conditional expectations and it was precisely this idea that was used when the theory of $g-$expectations was introduced, see for instance Peng, [109].

Remark 85 Let us see what happens with a coherent utility function u_0 defined by $\mathcal{A}_0$ and the scenario set $\mathcal{S}$. For simplicity let us suppose that u_0 is relevant so that $\mathcal{S}^e = \mathcal{S} \cap \mathbf{P}^e$ is dense in $\mathcal{S}$ and we can work with equivalent measures. We identify the measure with the densities Z_1, Z_2. Let us define $\mathcal{A}_1 = \{\xi \mid \text{ for all } A \in \mathcal{F}_1 : \xi \mathbf{1}_A \in \mathcal{A}_0\}$. Clearly $\xi \in \mathcal{A}_0$ if and only if for all $A \in \mathcal{F}_1$ and all $\mathbb{Q} \in \mathcal{S}$: $\mathbb{E}_{\mathbb{Q}}[\xi \mathbf{1}_A] \geq 0$. This is equivalent to $\mathbb{E}_{\mathbb{P}}[Z_2 \xi \mathbf{1}_A] \geq 0$. This can be rewritten as $\mathbb{E}[Z_2 \xi \mid \mathcal{F}_1] \geq 0$ for all $\mathbb{Q} \in \mathcal{S}$. This in turn is equivalent to $\mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] \geq 0$ for all $\mathbb{Q} \in \mathcal{S}^e$ and this means that $\mathbb{E}[\frac{Z_2}{Z_1}\xi \mid \mathcal{F}_1] \geq 0$ for all $\mathbb{Q} \in \mathcal{S}^e$. This leads us to the introduction of $u_1(\xi) = \text{ess.inf}_{\mathbb{Q} \in \mathcal{S}^e} \mathbb{E}[\frac{Z_2}{Z_1}\xi \mid \mathcal{F}_1]$. To check whether such a random variable is in $\mathcal{A}_0$ (a necessary condition since $u_0(\xi) = u_0(u_1(\xi))$) we need to check that for all $\mathbb{Q}'$ we now have $\mathbb{E}[Z_1' u_1(\xi)] \geq 0$. Because of the lattice property of $\{\mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_1] \mid \mathbb{Q} \in \mathcal{S}^e\}$, this means

$$\inf_{Z_1'} \inf_{Z_1, Z_2} \mathbb{E}[Z_1' \frac{Z_2}{Z_1} \xi] \geq 0.$$

In terms of the scenario set this means $Z_1' \frac{Z_2}{Z_1}$ whenever $\mathbb{Q}', \mathbb{Q} \in \mathcal{S}^e$. This property was called rectangularity, see [63], Riedel, [114] or m-stability, [41]. So for coherent measures we need the condition that $\mathcal{S}$ is m-stable. This condition is necessary and sufficient. So for TailVar we immediately see that — in general — the utility function is NOT time consistent since $Z_1', Z_1, Z_2 \leq k$ do not imply that $Z_1' \frac{Z_2}{Z_1} \leq k$.

Chapter 12

Finite and discete Time

12.1 Time Consistency

The finite and discrete time case is almost the same as the two period case but because there are several intermediate times, it offers a couple of extra properties. The revelation of uncertainty is given by a filtration $(\mathcal{F}_t)_{0\le t\le T}$. Time only takes the values $0, 1, \ldots, T$. The sigma-algebra $\mathcal{F}_0$ is supposed to be trivial. We suppose that there is monetary concave, **relevant** utility function

$$u_0 \colon L^\infty(\mathcal{F}_T) \to \mathbb{R}.$$

Instead of introducing intermediate utility functions u_t and then discussing time consistency, we will right away construct the intermediate acceptance sets and then use Theorem 61. We therefore define for $0 \le s \le t \le T$:

$$\begin{aligned} \mathcal{A}_s &= \{\xi \in L^\infty(\mathcal{F}_T) \mid \text{ for all } A \in \mathcal{F}_s : \xi \mathbf{1}_A \in \mathcal{A}_0\} \\ \mathcal{A}_{s,t} &= \{\xi \in L^\infty(\mathcal{F}_t) \mid \text{ for all } A \in \mathcal{F}_s : \xi \mathbf{1}_A \in \mathcal{A}_0\} \\ u_s(\xi) &= \text{ess.inf}\{\eta \in L^\infty(\mathcal{F}_s) \mid \xi - \eta \in \mathcal{A}_s\} \end{aligned}$$

Definition 30 *We say that the system $(u_t)_{0\le t\le T}$ is time consistent (or simply that u_0 is time consistent) if for all $0 \le s \le v \le t \le T$*

$$\mathcal{A}_{s,t} = \mathcal{A}_{s,v} + \mathcal{A}_{v,t}.$$

Proposition 56 *In order to be time consistent it is necessary and sufficient that for all $0 \le t \le T$:*

$$\mathcal{A}_0 = \mathcal{A}_{0,t} + \mathcal{A}_t.$$

Proof. The necessity if obvious. The sufficiency must be checked. We have that $\mathcal{A}_0 = \mathcal{A}_{0,v} + \mathcal{A}_v$ and hence we can apply Theorem 61. This means that we get that for every $\eta \in \mathcal{A}_0$ automatically $u_v(\eta) \in \mathcal{A}_0$ and $\eta - u_v(\eta) \in \mathcal{A}_v$. Suppose now that $\xi \in \mathcal{A}_{s,t} \subset \mathcal{A}_0$ then $\xi - u_v(\xi) \in \mathcal{A}_v$. But then we also have $\xi - u_v(\xi) \in \mathcal{A}_{v,t}$ because ξ is $\mathcal{F}_t$ measurable. For every $A \in \mathcal{A}_s$ we also have $\mathbf{1}_A\xi \in \mathcal{A}_0$ and hence $\mathbf{1}_A u_v(\xi) = u_v(\mathbf{1}_A\xi) \in \mathcal{A}_0$. This implies that $u_v(\xi) \in \mathcal{A}_s$ and because $u_v(\xi)$ is $\mathcal{F}_v$ measursable we also have $u_v(\xi) \in \mathcal{A}_{s,v}$. The decomposition $\xi = u_v(\xi) + (\xi - u_v(\xi))$ now shows that $\mathcal{A}_{s,t} \subset \mathcal{A}_{s,v} + \mathcal{A}_{v,t}$. The converse inequality is proved in a similar way. For each $A \in \mathcal{F}_s$ we have $\mathbf{1}_A\left(\mathcal{A}_{s,v} + \mathcal{A}_{v,t}\right) = \mathbf{1}_A\mathcal{A}_{s,v} + \mathbf{1}_A\mathcal{A}_{v,t} \subset \mathcal{A}_{0,v} + \mathcal{A}_v \subset \mathcal{A}_0$ and hence $\mathcal{A}_{s,v} + \mathcal{A}_{v,t} \subset \mathcal{A}_s$. Because $\mathcal{A}_{s,v} + \mathcal{A}_{v,t} \subset L^\infty(\mathcal{F}_t)$ we then get $\mathcal{A}_{s,v} + \mathcal{A}_{v,t} \subset \mathcal{A}_{s,t}$. □

Example 35 We again use the Example 4.11. We suppose that u is time consistent and relevant. Let η be minimal with $u_0(\eta) = 0$. Define $u_0^1(\xi) = u_0(\xi + \eta)$. The acceptance set of u_0^1 is $-\eta + \mathcal{A}_0$. The function u_0^1 is time consistent and $u_t^1(\xi) = u_t(\xi + \eta - u_t(\eta))$. There are different ways to see this. One way is to leave it as an exercise. Another way is to check the decomposition property. For a stopping time σ we could define $\mathcal{A}^1_{\sigma,T}$ as $-(\eta - u_\sigma(\eta)) + \mathcal{A}_{\sigma,T}$, then check the decomposition property $\mathcal{A}_0^1 = -\eta + \mathcal{A}_0 = (-(u_\sigma(\eta) + \mathcal{A}_{0,\sigma}) + (-(\eta - u_\sigma(\eta)) + \mathcal{A}_{\sigma,T})$. The only thing to verify is that $(-(u_\sigma(\eta) + \mathcal{A}_{0,\sigma}) = (-\eta + \mathcal{A}_0) \cap L^\infty(\mathcal{F}_\sigma)$. But for an $\mathcal{F}_\sigma-$measurable element ξ we have $u_0^1(\xi) = u_0(\xi + \eta) = u_0(u_\sigma(\xi + \eta)) = u_0(\xi + u_\sigma(\eta))$ hence $\xi \in (-\eta + \mathcal{A}_0) \cap L^\infty(\mathcal{F}_\sigma)$ if and only if $\xi + u_\sigma(\eta) \in \mathcal{A}_0$ or equivalently $\xi \in (-(u_\sigma(\eta) + \mathcal{A}_{0,\sigma})$. We could also check it via the algebraic properties of u. This goes as follows (we do not give the arguments to go from one line to the next, they are left as an exercise)

$$
\begin{aligned}
\xi \in \mathcal{A}_\sigma^1 &\Leftrightarrow \text{ for all } A \in \mathcal{F}_\sigma : \mathbf{1}_A\xi \in \mathcal{A}_0^1 \\
&\Leftrightarrow \text{ for all } A \in \mathcal{F}_\sigma : u_0(\mathbf{1}_A\xi + \eta) \geq 0 \\
&\Leftrightarrow \text{ for all } A \in \mathcal{F}_\sigma : u_\sigma(\mathbf{1}_A\xi + \eta) \in \mathcal{A}_0 \\
&\Leftrightarrow \text{ for all } A \in \mathcal{F}_\sigma : \mathbf{1}_A u_\sigma(\xi + \eta) + \mathbf{1}_{A^c} u_\sigma(\eta) \in \mathcal{A}_0 \\
&\Leftrightarrow \text{ for all } A \in \mathcal{F}_\sigma : \mathbf{1}_A(\xi + \eta - u_\sigma(\eta)) + u_\sigma(\eta) \in \mathcal{A}_0 \\
&\Leftrightarrow \text{ for all } A \in \mathcal{F}_\sigma : \mathbf{1}_A u_\sigma(\xi + \eta - u_\sigma(\eta)) + \eta \in \mathcal{A}_0 \\
&\Leftrightarrow \text{ for all } A \in \mathcal{F}_\sigma : \mathbf{1}_A u_\sigma(\xi + \eta - u_\sigma(\eta)) \in \mathcal{A}_0^1 \\
&\Leftrightarrow u_\sigma(\xi + \eta - u_\sigma(\eta)) \geq 0 \\
&\Leftrightarrow \xi \in -(\eta - u_\sigma(\eta)) + \mathcal{A}_\sigma.
\end{aligned}
$$

The above characterisation of time consistent utility functions can be translated into a condition for the penalty functions c. We introduce for each

$\mathbb{Q} \sim \mathbb{P}$ and for $s \le t$:

$$c_{s,t} = \text{ess.sup}\{\mathbb{E}_{\mathbb{Q}}[-\xi \mid \mathcal{F}_s] \mid \xi \in \mathcal{A}_{s,t}\}$$
$$c_t = \text{ess.sup}\{\mathbb{E}_{\mathbb{Q}}[-\xi \mid \mathcal{F}_t] \mid \xi \in \mathcal{A}_t\}.$$

Using the two period model with times $0 \le s \le t$ we see that for $\xi \in L^\infty(\mathcal{F}_t)$

$$u_s(\xi) = \text{ess.inf}_{\mathbb{Q}\sim\mathbb{P}} \mathbb{E}_{\mathbb{Q}}[\xi \mid \mathcal{F}_s] + c_{s,t}(\mathbb{Q}).$$

It is clear that $\{\mathbb{E}_{\mathbb{Q}}[-\xi \mid \mathcal{F}_s] \mid \xi \in \mathcal{A}_{s,t}\}$ is a lattice and we can therefore permute expected values and the ess.sup operation. We get:

Proposition 57 *Let u_0 be relevant, then the following are equivalent.*

1. *u_0 is time consistent*
2. *for each $0 \le s \le v \le t \le T$ and each $\mathbb{Q} \sim \mathbb{P}$, we have the cocycle property:*
$$c_{s,t}(\mathbb{Q}) = c_{s,v}(\mathbb{Q}) + \mathbb{E}_{\mathbb{Q}}[c_{v,t}(\mathbb{Q}) \mid \mathcal{F}_s],$$
3. *for each $0 \le t \le T$ we have*
$$c_0(\mathbb{Q}) = c_{0,t}(\mathbb{Q}) + \mathbb{E}_{\mathbb{Q}}[c_t(\mathbb{Q})],$$

Remark 86 The cocycle property was introduced in Bion-Nadal, see [21] and independently by H. Föllmer and Irene Penner, see [67]. It is the generalisation to concave utility functions of m-stability or rectangularity, [41], in the coherent case. Other and earlier characterisations were given by Epstein and Schneider [63], Riedel, [114], Maccheroni, Marinacci and Rusticini [102], [103].

Proposition 58 *If u_0 is relevant and time consistent then for all $0 \le t \le T-1$: $u_t(u_{t+1}(\xi)) = u_t(\xi)$*

Proof. This can be done using the cocycle property, exactly in the same way as in Theorem 61. But we can also give a proof using the sets $\mathcal{A}$. First let us observe that if $\eta \in \mathcal{A}_0$, then using the two period model with times $0, t, T$ we get from Theorem 61 that $u_t(\eta) \in \mathcal{A}_{0,t}$ for all t. Now take $\xi \in \mathcal{F}_T$. Since for all $A \in \mathcal{F}_t$, $\mathbf{1}_A(\xi - u_t(\xi)) \in \mathcal{A}_t \subset \mathcal{A}_0$, we can apply the previous statement with $t+1$ and get $\mathbf{1}_A u_{t+1}(\xi - u_t(\xi)) \in \mathcal{A}_{0,t+1}$. This means that for all $A \in \mathcal{F}_t$: $\mathbf{1}_A\,(u_{t+1}(\xi) - u_t(\xi)) \in \mathcal{A}_{0,t+1}$. In other words we get $u_{t+1}(\xi) - u_t(\xi) \in \mathcal{A}_t$ and hence we get $u_t(u_{t+1}(\xi)) - u_t(\xi) \ge 0$, proving $u_t(u_{t+1}(\xi)) \geq u_t(\xi)$. The other inequality goes as follows. We have $\xi - u_{t+1}(\xi) \in \mathcal{A}_{t+1}$ and $u_{t+1}(\xi) - u_t(u_{t+1}(\xi)) \in \mathcal{A}_{t,t+1}$. Since $\mathcal{A}_{t,t+1} + \mathcal{A}_{t+1} = \mathcal{A}_t$, we conclude $\xi - u_t(u_{t+1}(\xi)) \in \mathcal{A}_t$. But this implies $u_t(\xi) - u_t(u_{t+1}(\xi)) \geq 0$ or $u_t(\xi) \geq u_t(u_{t+1}(\xi))$ □

We conclude with a consequence of the preceding analysis. In discrete time this is just an inductive application of Theorem 61. We leave the details to the reader.

Proposition 59 *Suppose that u_0 is relevant, then it is time consistent if and only if*

$$\mathcal{A}_0 = \mathbb{R}_+ + \mathcal{A}_{0,1} + \mathcal{A}_{1,2} + \ldots + \mathcal{A}_{T-1,T}.$$

We observe that $\mathbb{R}_+ + \mathcal{A}_{0,1} = \mathcal{A}_{0,1}$, so the first term is only present for cosmetic reasons. One possible decomposition is given by

$$\xi = u_0(\xi) + \sum_0^{T-1} (u_{t+1}(\xi) - u_t(\xi)).$$

Remark 87 The idea to have a similar decomposition in continuous time, replacing sums by integrals and time steps by *infinitesimal increments*, leads to Backward Stochastic Differential Equations or BSDE. However the analogy is not straightforward. Let us for the moment limit the analysis to the discrete time equivalent of the BSDE. We can rewrite the decomposition in another way. Let us suppose that there is $\mathbb{Q} \sim \mathbb{P}$ such that $\mathbb{E}_{\mathbb{Q}}[\xi] + c_0(\mathbb{Q}) = u_0(\xi)$. Then for all t we have

$$\mathbb{E}_{\mathbb{Q}}[u_{t+1} \mid \mathcal{F}_t] + c_{t,t+1}(\mathbb{Q}) = u_t(\xi).$$

The conditional expectation with respect to $\mathbb{Q}$ will now be replaced by a conditional expectation with respect to $\mathbb{P}$. We introduce Z_t the density process of $\mathbb{Q}$ and put $L_{t+1} = \frac{Z_{t++1}}{Z_t}$. We can then rewrite the optimality of $\mathbb{Q}$ as

$$\mathbb{E}_{\mathbb{P}}[u_{t+1}(\xi) L_{t+1} \mid \mathcal{F}_t] + c_{t,t+1}(\mathbb{Q}) = u_t(\xi)$$

or using the covariance operator

$$cov_t(\eta_1, \eta_2) = \mathbb{E}_{\mathbb{P}}[\eta_1 \eta_2 \mid \mathcal{F}_t] - \mathbb{E}_{\mathbb{P}}[\eta_1 \mid \mathcal{F}_t]\, \mathbb{E}_{\mathbb{P}}[\eta_2 \mid \mathcal{F}_t]$$
$$\mathbb{E}_{\mathbb{P}}[u_{t+1}(\xi) \mid \mathcal{F}_t] + cov_t(u_{t+1}(\xi), L_{t+1}) + c_{t,t+1}(\mathbb{Q}) = u_t(\xi).$$

Let us put $-\eta = u_{t+1}(\xi) - \mathbb{E}_{\mathbb{P}}[u_{t+1}(\xi) \mid \mathcal{F}_t]$. For convenience we introduce

$$\mathcal{Y} = \{Y \mid Y > 0; Y \in L^1(\mathcal{F}_{t+1}); \mathbb{E}_{\mathbb{P}}[Y \mid \mathcal{F}_t] = 1\},$$

and for $Y \in \mathcal{Y}$, $c_{t,t+1}(Y)$ simply denotes $c_{t,t+1}(\mathbb{Y})$ for the measure $d\mathbb{Y} = Y\, d\mathbb{P}$. The optimality of $\mathbb{Q}$ can be rewritten as

$$\begin{aligned} g(\eta) &= \text{ess.sup}\{cov_t(\eta, Y) - c_{t,t+1}(Y) \mid Y \in \mathcal{Y}\} \\ &= cov_t(\eta, L_{t+1}) - c_{t,t+1}(L_{t+1}) = cov_t(u_{t+1}(\xi), L_{t+1}) - c_{t,t+1}(L_{t+1}). \end{aligned}$$

This allows to write:

$$\mathbb{E}_{\mathbb{P}}[u_{t+1}(\xi) \mid \mathcal{F}_t] - g(\eta) = u_t(\xi),$$

or (and this is the discrete time BSDE)

$$u_{t+1}(\xi) - u_t(\xi) = -\eta + g(\eta).$$

Given $u_{t+1}(\xi)$ we can first solve the convex optimisation problem (a calculation of some kind of Fenchel-Legendre transform)

$$g(\eta) = \sup\{cov_t(u_{t+1}(\xi), Y) - c_{t,t+1}(Y) \mid Y \in \mathcal{Y}\},$$

then we can write $u_t(\xi) = \mathbb{E}_{\mathbb{P}}[u_{t+1}(\xi) \mid \mathcal{F}_t] - g(\eta)$ to get by backward recursion the next element $u_t(\xi)$.

12.2 Supermartingale property, potentials, submartingales

The analysis in the previous section will now be extended and will bring us to the introduction of potentials. Of course in finite discrete time this is rather trivial but later we will profit from the analysis.

Proposition 60 *Let u_0 be relevant and time consistent. For each $\mathbb{Q} \sim \mathbb{P}$ with $c_0(\mathbb{Q}) < \infty$, there is an increasing process $(\alpha_t(\mathbb{Q}))_{0 \le t \le T}$ such that*

1. *$\alpha_0(\mathbb{Q}) = 0$ and $\alpha(\mathbb{Q})$ is predictable, i.e. $\alpha_t(\mathbb{Q})$ is $\mathcal{F}_{t-1}$ measurable,*
2. *$c_t(\mathbb{Q})$ defines a nonnegative $\mathbb{Q}-$supermartingale with $c_T(\mathbb{Q}) = 0$, a so-called potential,*
3. *$c_t(\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}}[\alpha_T(\mathbb{Q}) - \alpha_t(\mathbb{Q}) \mid \mathcal{F}_t]$,*
4. *for each ξ, the process $(u_t(\xi) + \alpha_t(\mathbb{Q}))_t$ is a $\mathbb{Q}-$submartingale.*

Proof. That $c_T(\mathbb{Q}) = 0$ is obvious since $\mathcal{A}_T = L^\infty_+(\mathcal{F}_T)$ (the reader should check it because it uses that u_0 is relevant). The process $c(\mathbb{Q})$ is a $\mathbb{Q}-$supermartingale. Indeed, from the cocycle property it follows that all random variables are integrable and for $s \le t$, $A \in \mathcal{F}_s$ we have

$$\int_A c_s(\mathbb{Q})\, d\mathbb{Q} = \int_A (c_{s,t}(\mathbb{Q}) + c_t(\mathbb{Q}))\; d\mathbb{Q} \ge \int_A c_t(\mathbb{Q})\; d\mathbb{Q}.$$

The existence of the process α is precisely the representation of the potential c. We define inductively, $\alpha_0(\mathbb{Q}) = 0$ and

$$\begin{aligned}\alpha_t(\mathbb{Q}) &= \alpha_{t-1}(\mathbb{Q}) + \mathbb{E}_\mathbb{Q}[c_{t-1}(\mathbb{Q}) - c_t(\mathbb{Q}) \mid \mathcal{F}_{t-1}] \\ &= \alpha_{t-1}(\mathbb{Q}) + c_{t-1,t}(\mathbb{Q}) \text{ by the cocycle property.}\end{aligned}$$

The process $\alpha(\mathbb{Q})$ is clearly predictable. It satisfies:

$$\alpha_T(\mathbb{Q}) = \sum_1^T \mathbb{E}_\mathbb{Q}[c_{t-1}(\mathbb{Q}) - c_t(\mathbb{Q}) \mid \mathcal{F}_{t-1}]$$

and

$$\begin{aligned}\mathbb{E}_\mathbb{Q}[\alpha_T(\mathbb{Q}) - \alpha_t(\mathbb{Q}) \mid \mathcal{F}_t] &= \sum_{t+1}^T \mathbb{E}_\mathbb{Q}[c_{s-1}(\mathbb{Q}) - c_s(\mathbb{Q}) \mid \mathcal{F}_t] \\ &= \mathbb{E}_\mathbb{Q}[c_t(\mathbb{Q}) - c_T(\mathbb{Q}) \mid \mathcal{F}_t] = c_t(\mathbb{Q}).\end{aligned}$$

Let us now show the submartingale property. We have that

$$\begin{aligned}&\mathbb{E}_\mathbb{Q}[u_{t+1}(\xi) + \alpha_{t+1}(\mathbb{Q}) \mid \mathcal{F}_t] \\ &\quad = \mathbb{E}_\mathbb{Q}[u_{t+1}(\xi) \mid \mathcal{F}_t] + \alpha_{t+1}(\mathbb{Q}) \\ &\quad = \mathbb{E}_\mathbb{Q}[u_{t+1}(\xi) \mid \mathcal{F}_t] + c_{t,t+1}(\mathbb{Q}) + \alpha_{t+1}(\mathbb{Q}) - c_{t,t+1}(\mathbb{Q}) \\ &\quad \ge u_t(u_{t+1}(\xi)) + \alpha_{t+1}(\mathbb{Q}) - c_{t,t+1}(\mathbb{Q}) \\ &\quad = u_t(\xi) + \alpha_t(\mathbb{Q}).\end{aligned}$$

□

Proposition 61 *In case there is* $\mathbb{Q} \sim \mathbb{P}$ *with* $u_0(\xi) = \mathbb{E}_\mathbb{Q}[\xi] + c_0(\mathbb{Q})$ *the process* $(u_t(\xi) + \alpha_t(\mathbb{Q}))_{0 \le t \le T}$ *is a* $\mathbb{Q}-$*martingale.*

Proof. In the preceding proof we see that all the inequalities patch together in $u_0(\xi) \le \mathbb{E}_\mathbb{Q}[\xi + \alpha_T(\mathbb{Q})]$. But $c_0(\mathbb{Q}) = \mathbb{E}_\mathbb{Q}[\alpha_T(\mathbb{Q})]$ so that we get $u_0(\xi) \le$

$\mathbb{E}_{\mathbb{Q}}[\xi] + c_0(\mathbb{Q})$. But this is an equality by the hypothesis on $\mathbb{Q}$. Hence all the inequalities in the preceding proof must be equalities, resulting in

$$\mathbb{E}_{\mathbb{Q}}[u_{t+1}(\xi) + \alpha_{t+1}(\mathbb{Q}) \mid \mathcal{F}_t] = u_t(\xi) + \alpha_t(\mathbb{Q}).$$

□

Remark 88 The Bishop-Phelps theorem or Ekeland's variational principle shows that for a dense set $\xi \in L^\infty$ there is $\mathbb{Q}$ with $u_0(\xi) = \mathbb{E}_{\mathbb{Q}}[\xi] + c_0(\mathbb{Q})$. However there is no guarantee that $\mathbb{Q} \sim \mathbb{P}$. In fact one can show that for the bid price in an incomplete *continuous* market, either ξ is marketable or the minimising element is not equivalent. We do not pursue this theory since it requires a big portion of stochastic analysis. This is beyond the scope of these lectures.

12.3 Refinement for the case $\mathbb{Q} \ll \mathbb{P}$.

For measures $\mathbb{Q} \ll \mathbb{P}$ we can still define the process $c_t(\mathbb{Q})$. However this is only defined up to sets of $\mathbb{Q}-$measure 0. The same can be said for the representation $c_t(\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}}[\alpha_T(\mathbb{Q}) - \alpha_t(\mathbb{Q}) \mid \mathcal{F}_t]$. Also the submartingale and martingale properties remain valid. But from the Proposition 61, we cannot draw any conclusion regarding the behaviour of $u_t(\xi)$ under the measure $\mathbb{P}$. The best we can do is the following. We introduce the density process $Z_t = \mathbb{E}_{\mathbb{P}}[\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_t]$ and we put $\sigma = \inf\{t \mid Z_t = 0\}$. Strictly before time σ, i.e. on the set $\{t < \sigma\}$, we have that $\mathbb{Q} \sim \mathbb{P}$. In case $u_0(\xi) = \mathbb{E}_{\mathbb{Q}}[\xi] + c_0(\mathbb{Q})$, the calculation of $\mathbb{E}_{\mathbb{Q}}[\xi + \alpha_T(\mathbb{Q}) \mid \mathcal{F}_t]$ allows to find $u_t(\xi)$ but only for times $t < \sigma$.

Chapter 13

Stochastic notation

13.1 General Results and Notation

To discuss dynamic utility functions we will need a set-up where the revelation of uncertainty is part of the model. This requires concepts of the general theory of stochastic processes. We will work with a filtered probability space, denoted as $\left(\Omega, \mathcal{F}_\infty, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$. The timeintervalis chosento be $[0, +\infty)$. The reader can check that this is the most general case. By using suitable imbeddings it covers the case of discrete, finite as well as infinite time sets. Nevertheless in many cases we will explicitly restrict it to a finite interval and in at least one chapter we will even use a discrete time set. The filtration $\mathcal{F}$ is supposed to satisfy the usual assumptions, iėthe filtration is right continuous and $\mathcal{F}_0$ contains all the null sets of the complete $\sigma-$algebra $\mathcal{F}_\infty$. For notions from the general theory of stochastic processes, we refer the reader to [45]. If X is a stochastic process and T is a stopping time, the process X^T is defined through $X_t^T = X_{T\wedge t}$. It is called the process X stopped at T. There is also a process that starts at T and it is defined as ${}^T X_t = 0$ for $t \leq T$ and ${}^T X_t = X_t - X_T$ for $t \geq T$.

Since stochastic intervals play a special role, let us recall from [45] some of these notions. If $T \leq S$ are two stopping times, then the stochastic intervals are defined as follows

$$[\![T, S]\!] = \{(t,\omega) \mid t \in \mathbb{R}_+ \text{ and } T(\omega) \leq t \leq S(\omega)\}.$$

The other intervals are defined in a similar way. In case $T = S$ we simply write $[\![T, S]\!] = [\![T]\!] = \{(t,\omega) \mid T(\omega) < \infty\}$. If T is a stopping time and if $A \in \mathcal{F}_T$, then T_A denotes the stopping time defined as $T_A = T$ on the set A and $T_A = \infty$ on the set $A^c = \Omega \setminus A$. In particular for $t \in \mathbb{R}_+$ and $A \in \mathcal{F}_t$ we have $[\![t_A]\!] = \{t\} \times A$. With the given filtration we will construct the $\sigma-$algebras of predictable and optional sets. The predictable $\sigma-$algebra, denoted by $\mathcal{P}$, is the smallest $\sigma-$algebra on $\mathbb{R}_+ \times \Omega$ that contains sets of the form $[\![0_A]\!] = \{0\} \times A$ with $A \in \mathcal{F}_0$, as well as for each stopping time T, the stochastic interval

$$[\![0, T]\!] = \{(t, \omega) \mid t \leq T(\omega) \text{ and } t < \infty\}.$$

The optional $\sigma-$algebra, denoted by $\mathcal{O}$, is the smallest $\sigma-$algebra on $\mathbb{R}_+ \times \Omega$ that contains sets of the form $\{0\} \times A$ with $A \in \mathcal{F}_0$, as well as for each stopping time T, the stochastic interval

$$[\![0, T[\![= \{(t, \omega) \mid t < T(\omega)\}.$$

We remark that the indicator functions of elements of the generating set of $\mathcal{P}$ are left continuous adapted processes and that the indicator functions of elements of the generating sets of $\mathcal{O}$ are right continuous adapted processes. It can easily be checked that $\mathcal{P} \subset \mathcal{O}$. Let us recall that the class of predictable

sets

$$\{[\![0_A]\!] \mid A \in \mathcal{F}_0\} \cup \{]\!]T, S]\!] \mid T \leq S \text{ stopping times}\},$$

forms a semi–algebra that generates $\mathcal{S}$. The Boolean algebra generated by this class is simply

$$\mathcal{A} = \Big\{[\![0_A]\!] \cup]\!]T_0, T_1]\!] \cup]\!]T_1, T_2]\!] \ldots \cup]\!]T_{n-1}, T_n]\!] \\ \mid n \geq 1; A \in \mathcal{F}_0 \text{ and } 0 \leq T_0 \leq T_1 \leq \ldots T_n \leq +\infty \text{ are all stopping times}\Big\}.$$

The importance of this class lies in the following density result from general measure theory. The proof of the lemma is included in the proof of the Carathéodory extension theorem.

Lemma 22 *Let μ be a nonnegative finite $\sigma-$additive measure on $\mathcal{P}$, then for each $\varepsilon > 0$ and for each set $B \in \mathcal{P}$, there is a set $A \in \mathcal{A}$ such that $\mu(A \Delta B) \leq \varepsilon$.*

Since the filtration satisfies the usual assumptions, we will suppose that all the (sub–, super–, semi–) martingales are càdlàg, meaning they are right continuous and have left limits. When we deal with the construction of the Snell envelope, we will pay attention to this continuity property and the reader will notice similar difficulties as in the work of Mertens see [105] and [45], appendix, see also [58]. Although we treat the case of supermartingales with respect to a family of measures, there is no essential difference with the case of a fixed probability measure. As before we will identify, through the Radon–Nikodym theorem, finite measures ν on $\mathcal{F}_\infty$, that are absolutely continuous with respect to $\mathbb{P}$, with their densities $\frac{d\nu}{d\mathbb{P}}$, i.e. with functions in L^1. Furthermore we will sometimes identify this measure with the càdlàg martingale $Z_t = \mathbb{E}_\mathbb{P}\left[\frac{d\nu}{d\mathbb{P}} \mid \mathcal{F}_t\right]$. We hope that these identifications will not cause too many problems. If $\mathbb{Q} \sim \mathbb{P}$ then the martingale $Z_t = \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_t\right]$ has the property $\inf_{t\in\mathbb{R}_+} Z_t > 0$, $\mathbb{P}$ a.s. (see [45] page 85). If $\mathbb{Q} \sim \mathbb{P}$, Bayes' rule implies that $\mathbb{E}_\mathbb{Q}[f \mid \mathcal{F}_T] = \mathbb{E}_\mathbb{P}[f\, \frac{Z_\infty}{Z_T} \mid \mathcal{F}_T]$, here T denotes a stopping time.

If X is a (for us automatically càdlàg) semi-martingale then the square bracket of X is defined as

$$[X,X]_t = \lim_{\varepsilon\to 0}\left(X_0^2 + \sum_{i=0}^n \left(X_{T_{i+1}} - X_{T_i}\right)^2\right),$$

where the sequence of stopping times T_i satisfies $0 = T_0 \le T_1 \le \dots T_n = t$ and $\max_i (T_{i+1} - T_i) \le \varepsilon$. One can show that this limit exists and that it does not depend on the way we divide the interval $[0,t]$. The process $[X,X]$ is nondecreasing and it is right continuous. The jumps are related to the jumps of X, $\Delta[X,X] = (\Delta(X))^2$. In case X is continuous, the process $[X,X]$ is also continuous.

If M is a continuous local martingale, then the process $[M,M] = \langle M,M\rangle = \langle M\rangle$ is the unique predictable process such that $\langle M\rangle_0 = M_0^2$ and $M^2 - \langle M\rangle$ is again a local martingale. In case $M_0 = 0$, the process $\exp\left(M - \frac{1}{2}\langle M\rangle\right)$ is also a local martingale, called the stochastic exponential of M. Even if M is a uniformly integrable martingale $M_t = \mathbb{E}[M_\infty \mid \mathcal{F}_t]$, the exponential does not have to be uniformly integrable. Conversely if the stochastic exponential is uniformly integrable, the martingale M does not have to be uniformly integrable. The stochastic exponential is denoted by $\mathcal{E}(M)$ and it is the solution of the stochastic differential equation or SDE: $dX_t = X_t\, dM_t$ with initial condition $X_0 = 1$. If X is a strictly positive continuous local martingale, $X_0 = 1$, then we can write it as an exponential, namely $X = \mathcal{E}(M)$, where $dM_t = \frac{1}{X_t}\, dX_t$ defines the stochastic logarithm, sometimes denoted by $\mathcal{L}(X)$. In case $\mathbb{Q} \sim \mathbb{P}$ and the density process Z starts at $Z_0 = 1$, being the case if $\mathcal{F}_0$ is trivial, we can write Z as a stochastic exponential $Z = \mathcal{E}(M)$.

In case $\mathbb{Q}^1$ and $\mathbb{Q}^2$ are two probabilities equivalent to $\mathbb{P}$ with continuous densities Z^1, Z^2, we can write them as stcochastic exponentials wrt M^1, M^2. For a stopping time T we can then define the concatenation at time T. The density is defined as

$$Z_t = Z_t^1 \text{ for } t \le T$$

$$Z_t = Z_T^1 \frac{Z_t^2}{Z_T^2} \text{ for } t \ge T.$$

We can write $Z = \mathcal{E}(M)$ where $M = (M^1)^T + {}^T M^2$. In exponential writing the concatenation is just the sum of a stopped process and a process that starts at the time of concatenation. The interpretation is that we have two models $\mathbb{Q}^1, \mathbb{Q}^2$ and that until time T we use the first model whereas after time T we use the second model. The reader can easily see that Z is a martingale defining an equivalent measure $\mathbb{Q} \sim \mathbb{P}$. The theory of dynamic utility functions will make use of this concatenation.

13.2 The Case of a Brownian Motion

In many cases the filtration will come from a $d-$dimensional Brownian Motion W. In this case we have the representation property for local martingales, see [122]. For the density (Z) of an equivalent probability measure $\mathbb{Q}$, this leads to the existence of a $d-$dimensional predictable process q such that $Z = \mathcal{E}(q \cdot W)$. Here we use the standard notation for the stochastic integral

$$(q \cdot W)_t = \int_0^t q_u \, dW_u.$$

The density process Z is the the solution of the SDE $dZ_t = Z_t q_t \, dW_t$. This is some liberal use of vector calculus notation since we used the scalar product between q and dW. The concatenation of two measures defined by the processes q^1, q^2 is then given by the new process $q = q^1 \mathbf{1}_{[\![0,T]\!]} + q^2 \mathbf{1}_{]\!]T,\infty[\![}$. The convex combination of two measures $\mathbb{Q}^1, \mathbb{Q}^2$ can also be written as a stochastic exponential. But this time the outcome is more difficult. Let us see what happens with $\mathbb{Q} = \frac{\mathbb{Q}^1 + \mathbb{Q}^2}{2}$, its density process is given by $Z_t = \frac{Z^1 + Z^2}{2}$ but the process q defined as $Z = \mathcal{E}(q \cdot W)$, comes from:

$$\begin{aligned} Z_t q_t \, dW_t = dZ_t &= \frac{1}{2}(dZ_t^1 + dZ_t^2) \\ &= \frac{1}{2}\left(Z_t^1 q_t^1 \, dW_t + Z_t^2 q_t^2 \, dW_t\right) = Z_t \left(\frac{Z_t^1 q_t^1 + Z_t^2 q_t^2}{2 Z_t}\right) dW_t. \end{aligned}$$

This gives

$$q_t = \frac{Z_t^1 q_t^1 + Z_t^2 q_t^2}{Z_t^1 + Z_t^2},$$

for later use we note that for every t and every $\omega \in \Omega$, $q_t(\omega)$ is a convex combination of $q_t^1(\omega)$ and $q_t^2(\omega)$.

The representation property for measures $\mathbb{Q} \ll \mathbb{P}$ is a little bit more tricky. One can show that there is a predictable process such that the density process Z_t looks like

$$Z_t = \mathcal{E}(q \cdot W)_t \text{ on the set } \{t < \sigma\}.$$

Here $\sigma = \inf\{t \mid Z_t = 0\}$ is the first time the density becomes 0. The set $\{Z_\infty = 0\}$ is given by $\int_0^\sigma |q_t|^2 \, dt = +\infty$. The process q is not unique since after time σ, we can continue it with any predictable process.

13.3 Some BMO results

We now present some results on the decomposition of semimartingales. Again we suppose that the filtration comes from a Brownian Motion. A more general presentation is possible but is beyond the scope of this book. Let us introduce the following space

$$\mathbf{S}^{BMO} = \left\{ X \;\middle|\; \begin{array}{l} X \text{ is a continuous semi-martingale}, X_0 = 0 \\ X = A + M, A_0 = 0, \text{ is the Doob-Meyer decomposition} \\ A \text{ is of finite total variation} \\ M \text{ is a continuous } BMO\text{−martingale} \\ \text{there is } C \text{ such that for all } t : \mathbb{E}\left[\int_t^\infty |dA_u| \mid \mathcal{F}_t\right] \le C < \infty \end{array} \right\}$$

The space $\mathbf{S}$ can be given a norm $\|X\| = C + \|M\|_{BMO_2}$. The space is a Banach space and using the martingale convergence theorem, one can show that both A and M converge when $t \to \infty$. An equivalent norm is

$$\sup\{\|\mathbb{E}[\,|(\theta \cdot X)_\infty - (\theta \cdot X)_t| \mid \mathcal{F}_t]\|_\infty\},$$

where the sup is taken over all t and all predictable θ with $|\theta| \le 1$. This is not trivial but can be proved as follows. For each $\varepsilon > 0$ there is a predictable process $\theta, |\theta| \le 1$ such that $|\theta \cdot A| \le \varepsilon$. This allows to give an estimate for $\|M\|_{BMO_1}$ norm. Once you know that M is BMO, one can choose θ_u such that $\theta \, dA_u = |dA_u|$. Then you get the bound for $\mathbb{E}\left[\int_t^\infty |dA_u| \mid \mathcal{F}_t\right]$. From here the rest is trivial. The mapping

$$\mathbf{S}^{BMO} \to BMO; X \to M,$$

is of course continuous. For applications in utility theory we need an estimate based on the supremum of the process. But even for deterministic processes there is a big difference between the quantity $\|X\|$ and expressions such as $\sup_t |X_t|$. Nevertheless the martingale part can be estimated using the supremum of the process.

Lemma 23 *Let X be a continuous bounded submartingale bounded by a constant c. Suppose that X has the continuous Doob-Meyer decomposition $X = A + M$, then the martingale part M is in BMO, more precisely $\|M\|_{BMO_2} \leq 2c$ and $\|X\|_{\mathcal{S}^{BMO}} \leq 4c$. Consequently A_∞ has exponential moments of some order.*

Proof. The process A satisfies $\mathbb{E}[\int_t^\infty |dA_u| \mid \mathcal{F}_t] = \mathbb{E}[X_\infty - X_t \mid \mathcal{F}_t] \leq 2c$. This already shows that $X \in \mathbf{S}^{BMO}$. However this only yields a bound for the BMO_1 norm and the bound is not the best. We prefer to give a direct proof. First observe that the process $Y = (c-X)^2$ is given by the differential equation

$$dY_t = 2(c - X_t)dA_t + d\langle M, M\rangle_t + 2(c - X_t)\, dM_t.$$

Since the coefficient $c - X_t$ is nonnegative, the process Y is also a bounded submartingale and taking conditional expectations gives

$$\mathbb{E}\left[\int_t^\infty d\langle M, M\rangle_u \mid \mathcal{F}_t\right] \leq \mathbb{E}[Y_\infty - Y_t \mid \mathcal{F}_t] \leq \mathbb{E}[Y_\infty \mid \mathcal{F}_t].$$

Consequently

$$\mathbb{E}\left[\int_t^\infty d\langle M, M\rangle_u \mid \mathcal{F}_t\right] \leq 4c^2.$$

Hence $\|M\|_{BMO} \leq 2c$. Because M is BMO it has – by the John-Nirenberg inequality, see [**?**] xxx – exponential moments. Because X is bounded also A_∞ must have exponential moments. □

Lemma 24 *For $X \in \mathbf{S}^{BMO}$ and X bounded, we have*

$$\|M\|^2_{BMO_2} \leq 8\|\sup_t |X_t|\|^2_\infty + 8\|\sup_t |X_t|\,\|_\infty \|X\|$$

Proof. Let $c = \|\sup_t |X_t|\|_\infty$. The following inequalities are now straightforward. If needed one can use a localisation argument by stopping the processes when they reach a level.

$$\begin{aligned}\mathbb{E}[(M_\infty - M_t)^2 \mid \mathcal{F}_t] &\le 2\mathbb{E}[(X_\infty - X_t)^2 \mid \mathcal{F}_t] + 2\mathbb{E}[(A_\infty - A_t)^2 \mid \mathcal{F}_t] \\ &\le 8c^2 + 4\mathbb{E}\left[\int_t^\infty (A_u - A_t) dA_u \mid \mathcal{F}_t\right] \\ &\le 8c^2 + 4\mathbb{E}\left[\int_t^\infty (A_u - A_t) dX_u \mid \mathcal{F}_t\right] \\ &\le 8c^2 + 4\mathbb{E}\left[\int_t^\infty (X_\infty - X_u) dA_u \mid \mathcal{F}_t\right] \\ &\le 8c^2 + 8c\mathbb{E}\left[\int_t^\infty |dA_u| \mid \mathcal{F}_t\right].\end{aligned}$$

This shows the bound on $\|M\|_{BMO_2}$. □

Corollary 17 *If X^n is a sequence in $\mathbf{S}^{BMO}$, $\sup_n \|X^n\| < \infty$, if $\|\sup_t |X_t^n - X_t|\,\|_\infty \to 0$, then $X \in \mathbf{S}^{BMO}$ and the martingale parts of the Doob-Meyer decompositions satisfy $M^n - M \to 0$ in BMO.*

Proof. Only the statement that $X \in \mathbf{S}^{BMO}$ needs to be shown. By the inequality we get that $\lim_{n,m\to\infty}(M^n - M^m) \to 0$ in BMO. So the Cauchy sequence converges to $M \in BMO$. The finite variation part then also converges to a process A in the sense that $\sup_t |A_t^n - A_t|$ tends to zero at least in probability. That $\mathbb{E}\left[\int_t^\infty |dA_u| \mid \mathcal{F}_t\right] \le \sup_n \|X^n\|$ is straightforward. So $X \in \mathbf{S}^{BMO}$, $\|X\| \le \sup_n \|X^n\|$. □

Remark 89 Of course we did not make a statement on norm convergene in the space $\mathbf{S}^{BMO}$. As already observed above, even for deterministic bounded variation processes, the convergence in sup-norm is not the same as the convergence in variation norm.

Corollary 18 *If X^n is a uniformly bounded sequence of submartingales $|X_t^n| \le c$, if $\|\sup_t |X_t - X_t^n|\|_\infty$ tends to 0, then the Doob-Meyer decompositions $X^n = A^n + M^n, X = A + M$ satisfy M^n tends to M in BMO.*

Proof. The previous corollary shows that $\|X^n\| \le C$ (where C only depends on c). The convergence of the martingale parts now follows. □

Chapter 14

Continuous time dynamic utility functions

14.1 The model

We will develop the theory for a finite time horizon, $[0,T]$. Some of the properties remain valid for the time interval $[0,\infty)$ as well (the reader can find out where the difficulties are when working with an open end interval). The probability space is $(\Omega, \mathcal{F}_T, \mathbb{P})$ is supposed to be atomless. The revelation of uncertainty is given by the filtration $\mathcal{F} = (\mathcal{F}_t)_{0\le t\le T}$. We make the usual assumptions:

1. $A \subset B \in \mathcal{F}_T, \mathbb{P}[B] = 0$ implies $A \in \mathcal{F}_0$.
2. The filtration is right continuous: $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$ whenever $t < T$.
3. The sigma algebra $\mathcal{F}_0$ is up to sets of measure zero, trivial. This is not part of the usual "usual assumptions" but we will nevertheless include it.

We will assume that there is a **relevant** monetary concave utility function defined on $L^\infty(\mathcal{F}_T)$. This function u_0 is supposed to have the Fatou property. The acceptance sets are introduced in the same way as in the discrete case. So $\mathcal{A}_0 = \{\xi \mid u_0(\xi) \ge 0\}$ and

for a stopping time $\sigma : \mathcal{A}_\sigma = \{\xi \in L^\infty(\mathcal{F}_T) \mid \text{ for all } A \in \mathcal{F}_\sigma : \mathbf{1}_A\xi \in \mathcal{A}_0\}$
if $\sigma \le \tau$ are two stopping times: $\mathcal{A}_{\sigma,\tau} = \mathcal{A}_\sigma \cap L^\infty(\mathcal{F}_\tau)$
if $\mathbb{Q} \sim \mathbb{P}$: $c_{\sigma,\tau}(\mathbb{Q}) = \text{ess.sup}\{\mathbb{E}_\mathbb{Q}[-\xi \mid \mathcal{F}_\sigma] \mid \xi \in \mathcal{A}_{\sigma,\tau}\}$
$c_{\sigma,T}(\mathbb{Q})$ is simply denoted by $c_\sigma(\mathbb{Q})$
$u_\sigma(\xi) = \text{ess.sup}\{\eta \in L^\infty(\mathcal{F}_\sigma) \mid \xi - \eta \in \mathcal{A}_\sigma\}$

From the theory in discrete time we copy the following.

Definition 31 *The utility function is called time consistent if for all pairs of stopping times $\sigma \le \tau$ and all pairs of bounded random variables ξ, η, $u_\tau(\xi) \le u_\tau(\eta)$ implies $u_\sigma(\xi) \le u_\sigma(\eta)$.*

As in the discrete time case, this time applied to the (stopping) times $0 \le \sigma \le \tau \le \nu \le T$ we get that time consistency is equivalent to some algebraic properties for the sets $\mathcal{A}$.

Proposition 62 *For a relevant concave monetary utility function u_0 the following are equivalent*

1. *u_0 is time consistent*
2. *for stopping times $\sigma \le \tau \le \nu$: $\mathcal{A}_{\sigma,\nu} = \mathcal{A}_{\sigma,\tau} + \mathcal{A}_{\tau,\nu}$*
3. *for all $\mathbb{Q} \sim \mathbb{P}$: $c_{\sigma,\nu}(\mathbb{Q}) = c_{\sigma,\tau}(\mathbb{Q}) + \mathbb{E}_\mathbb{Q}[c_{\tau,\nu}(\mathbb{Q}) \mid \mathcal{F}_\sigma]$*
4. *for stopping times σ: $\mathcal{A}_0 = \mathcal{A}_{0,\sigma} + \mathcal{A}_\sigma$*
5. *for stopping times σ and $\mathbb{Q} \sim \mathbb{P}$: $c_0(\mathbb{Q}) = c_{0,\sigma}(\mathbb{Q}) + \mathbb{E}_\mathbb{Q}[c_\sigma(\mathbb{Q})]$*
6. *for stopping times $\sigma \le \tau$: $u_\sigma(u_\tau(\xi)) = u_\sigma(\xi)$.*

Let $\sigma \le \tau$ be two stopping times. If $\xi \in L^\infty(\mathcal{F}_\tau)$ then

$$u_\sigma(\xi) = \text{ess.inf}_{\mathbb{Q}\sim\mathbb{P}}\{\mathbb{E}_\mathbb{Q}[\xi + c_{\sigma,\tau}(\mathbb{Q}) \mid \mathcal{F}_\sigma]\}$$

For the moment, the random variables $u_\sigma(\xi)$ and $c_\sigma(\mathbb{Q})$ are not glued together to form a nicely behaved stochastic process. They are just a family of random variables or better of classes modulo equality a.s. . Of course we have some continuity properties such as $\|u_\sigma(\xi) - u_\sigma(\eta)\|_\infty \le \|\xi - \eta\|_\infty$ but these are not with respect to time.

14.2 Regularity for the processes u and c

We start this section by showing that there is a càdlàg version for the "system" $c_\sigma(\mathbb{Q})$. The next step is the representation for the potential $c_t(\mathbb{Q})$. Once this is accomplished we can prove that also the system $u_\sigma(\xi)$ has a càdlàg version. The presentation is different from the classical results obtained by [**?**, **?**, **?**] xxxx. The only standing hypothesis we need is that u_0 is relevant. We will frequently use the following construction, called concatenation at a stopping time. Let $\mathbb{Q}^1, \mathbb{Q}^2$ be two probability measures equivalent to $\mathbb{P}$. Let Z^1, Z^2 be their density processes (we take the càdlàg versions).

For a stopping time σ we define a new density as follows: for $t < \sigma$ we take $Z_t = Z_t^1$, for $t \geq \sigma$ we take $Z_t = Z_\sigma^1 \frac{Z_t^2}{Z_\sigma^2}$. As already observed before, it is easily seen that Z is the density process of a probability measure equivalent to $\mathbb{P}$. This measure will be denoted by $Q^1 \star^\sigma \mathbb{Q}^2$. Obviously the measures $\mathbb{Q}^1$ and $Q^1 \star^\sigma \mathbb{Q}^2$ coincide on the sigma algebra $\mathcal{F}_\sigma$.

Lemma 25 *Let $\mathbb{Q} \sim \mathbb{P}$ with $c_0(\mathbb{Q}) < \infty$. Let τ_n be a sequence of stopping times such that $\mathbb{P}[\tau_n < T] \to 0$. Then $c_{\tau_n}(\mathbb{Q}) \to 0$ in $L^1(\mathbb{Q})$.*

Proof. From proposition xxx in the two period model $0, \tau_n, T$ we deduce that for each n we can find a measure $\mathbb{Q}^n \sim \mathbb{P}$ such that $c_{\tau_n}(\mathbb{Q}^n) \leq n^{-1}$ a.s. . We now define a new measure, $\mathbb{Y}^n = \mathbb{Q} \star^{\tau_n} \mathbb{Q}^n$ which is the concatenation of $\mathbb{Q}$ and $\mathbb{Q}^n$ at time τ_n. Because $\mathbb{P}[\tau_n < T] \to 0$, we have that $\mathbb{Y}^n \to \mathbb{Q}$ in L^1. By the cocycle property we have that $c_0(\mathbb{Y}^n) = c_{0,\tau_n}(\mathbb{Q}) + \mathbb{E}_\mathbb{Q}[c_{\tau_n}(\mathbb{Q}^n)] \leq c_{0,\tau_n}(\mathbb{Q}) + n^{-1}$. But we also have that $\mathbb{Y}^n \to \mathbb{Q}$ and hence $\liminf c_0(\mathbb{Y}^n) \geq c_0(\mathbb{Q})$. So we already deduce that $c_0(\mathbb{Q}) \leq \liminf c_{0,\tau_n}(\mathbb{Q})$. The cocycle property also implies $c_0(\mathbb{Q}) = c_{0,\tau_n}(\mathbb{Q}) + \mathbb{E}_\mathbb{Q}[c_{\tau_n}(\mathbb{Q})]$. These inequalities imply that $\liminf \mathbb{E}_\mathbb{Q}[c_{\tau_n}(\mathbb{Q})] = 0$. But this is valid for every subsequence, so we get $\lim \mathbb{E}_\mathbb{Q}[c_{\tau_n}(\mathbb{Q})] = 0$. □

Corollary 19 *Under the hypothesis $\mathbb{Q} \sim \mathbb{P}$ with $c_0(\mathbb{Q}) < \infty$, we have: for all $\varepsilon > 0$ there is $\delta > 0$ such that for a stopping time σ, $\mathbb{P}[\sigma < T] \leq \delta$ implies $\mathbb{E}_\mathbb{Q}[c_\sigma(\mathbb{Q})] \leq \varepsilon$.*

Proposition 63 *Suppose $c_0(\mathbb{Q}) < \infty$. The family*

$$\{c_\sigma(\mathbb{Q}) \mid \sigma \text{ is a stopping time } 0 \leq \sigma \leq T\}$$

is $\mathbb{Q}-$uniformly integrable. It satisfies the supermartingale inequality: for stopping times $\sigma \leq \tau$:

$$c_\sigma(\mathbb{Q}) \geq \mathbb{E}_\mathbb{Q}[c_\tau(\mathbb{Q}) \mid \mathcal{F}_\sigma].$$

Proof. The supermartingale property follows from the discrete time analysis. The uniform integrability has to be shown. We will make a uniform estimate (i.e. not depending on σ) of $\mathbb{E}_\mathbb{Q}\left[c_\sigma \mathbf{1}_{\{c_\sigma > n\}}\right]$. We have $n\mathbb{Q}[c_\sigma(\mathbb{Q}) > n] \leq \mathbb{E}_\mathbb{Q}[c_\sigma(\mathbb{Q})] \leq c_0(\mathbb{Q})$ and hence $\mathbb{Q}[c_\sigma(\mathbb{Q}) > n] \leq c_0(\mathbb{Q})/n$. Given $\varepsilon > 0$ we can find by the preceding corrolary, a $\delta > 0$ such that $\mathbb{P}[\tau < T] \leq \delta$ implies $\mathbb{E}_\mathbb{Q}[c_\tau(\mathbb{Q})] \leq \varepsilon$. Since $\mathbb{Q} \sim \mathbb{P}$, we can find n big enough so that $\mathbb{Q}[A] \leq c_0(\mathbb{Q})/n$ implies $\mathbb{P}[A] \leq \delta$. Put now $\tau = \sigma$ if $c_\sigma(\mathbb{Q}) > n$ and $\tau = T$ if

$c_\sigma(\mathbb{Q}) \le n$. τ is a stopping time since $c_\sigma(\mathbb{Q})$ is $\mathcal{F}_\sigma-$measurable. With this notation: $\mathbb{P}[\tau < T] \le \delta$ and hence $\mathbb{E}_\mathbb{Q}\left[c_\sigma \mathbf{1}_{\{c_\sigma > n\}}\right] = \mathbb{E}_\mathbb{Q}[c_\tau(\mathbb{Q})] \le \varepsilon$. □

Lemma 26 *Let $\mathbb{Q} \sim \mathbb{P}$ and $c_0(\mathbb{Q}) < \infty$. Let $\sigma \le \tau \le \nu$ be three stopping times. Let $\varepsilon > 0$ and let $\xi \in \mathcal{A}_{\sigma,\nu}$ be chosen so that $c_{\sigma,\nu}(\mathbb{Q}) \le \mathbb{E}_\mathbb{Q}[-\xi \mid \mathcal{F}_\sigma]+\varepsilon$, which is the same as $c_{\sigma,\nu}(\mathbb{Q})+\mathbb{E}_\mathbb{Q}[\xi \mid \mathcal{F}_\sigma] \le \varepsilon$. Then also $c_{\sigma,\tau}(\mathbb{Q})+\mathbb{E}_\mathbb{Q}[u_\tau(\xi) \mid \mathcal{F}_\sigma] \le \varepsilon$.*

Proof We decompose ξ as follows $\xi = u_\tau(\xi) + (\xi - u_\tau(\xi))$. The second term is in $\mathcal{A}_{\tau,\nu}$ and the first is in $\mathcal{A}_{\sigma,\tau}$. This allows to write

$$\begin{aligned}\varepsilon &\ge c_{\sigma,\nu}(\mathbb{Q}) + \mathbb{E}_\mathbb{Q}[\xi \mid \mathcal{F}_\sigma] \\ &= (c_{\sigma,\tau}(\mathbb{Q}) + \mathbb{E}_\mathbb{Q}[u_\tau(\xi) \mid \mathcal{F}_\sigma]) \\ &\qquad + \mathbb{E}_\mathbb{Q}[c_{\tau,\nu}(\mathbb{Q}) + (\xi - u_\tau(\xi)) \mid \mathcal{F}_\sigma] \\ &= (c_{\sigma,\tau}(\mathbb{Q}) + \mathbb{E}_\mathbb{Q}[u_\tau(\xi) \mid \mathcal{F}_\sigma]) \\ &\qquad + \mathbb{E}_\mathbb{Q}[\mathbb{E}_\mathbb{Q}[c_{\tau,\nu}(\mathbb{Q}) + (\xi - u_\tau(\xi)) \mid \mathcal{F}_\tau] \mid \mathcal{F}_\sigma].\end{aligned}$$

The inner conditional expectation in the second term is nonnegative, meaning that the second term is nonnegative. But also the first term is nonnegative. So both terms must be smaller than ε. □

Proposition 64 *Let $\mathbb{Q} \sim \mathbb{P}$ and $c_0(\mathbb{Q}) < \infty$. If σ_k is a sequence of stopping times decreasing to σ, then $c_{\sigma,\sigma_k}(\mathbb{Q}) \downarrow 0$.*

Proof. That the sequence $c_{\sigma,\sigma_k}(\mathbb{Q})$ is decreasing follows from the definition and $\mathcal{A}_{\sigma,\sigma_k} \subset \mathcal{A}_{\sigma,\sigma_{k+1}}$ for all k. Let us now take $\varepsilon > 0$ and let $\xi \in \mathcal{A}_{\sigma,\sigma_1}$ be chosen to satisfy $c_{\sigma,\sigma_1}(\mathbb{Q}) + \mathbb{E}_\mathbb{Q}[\xi \mid \mathcal{F}_\sigma] \le \varepsilon$. Then by the preceding lemma, we have for all k: $c_{\sigma,\sigma_k}(\mathbb{Q}) + \mathbb{E}_\mathbb{Q}[u_{\sigma_k}(\xi) \mid \mathcal{F}_\sigma] \le \varepsilon$. Let us now take convex combinations $\eta_k \in conv\{u_{\sigma_l}(\xi) \mid l \ge k\}$ that converge almost surely to an element η. This is possible since the sequence $u_{\sigma_k}(\xi)$ is uniformly bounded by $\|\xi\|_\infty$. The same convex combinations will be taken on the decreasing sequence $c_{\sigma,\sigma_k}(\mathbb{Q})$. Since $\eta_k \in \mathcal{A}_{\sigma,\sigma_k}$ we have that $u_\sigma(\eta_k) \ge 0$ and hence also $u_\sigma(\eta) \ge 0$. But the limit η belongs to $L^\infty(\mathcal{F}_\sigma)$ and hence $u_\sigma(\eta) = \eta$ showing that $\eta \ge 0$. This results in $\lim_k c_{\sigma,\sigma_k}(\mathbb{Q})+\eta \le \varepsilon$ and since $\eta \ge 0$ we must have $\lim_k c_{\sigma,\sigma_k}(\mathbb{Q}) \le \varepsilon$. Because $\varepsilon > 0$ was arbitrary, we have $\lim_k c_{\sigma,\sigma_k}(\mathbb{Q}) = 0$. □

Proposition 65 *Let $\mathbb{Q} \sim \mathbb{P}$ and $c_0(\mathbb{Q}) < \infty$. If σ_k is a sequence of stopping times decreasing to σ, then $c_{\sigma_k}(\mathbb{Q}) \to c_\sigma(\mathbb{Q})$ in L^1 and a.s. .*

Proof. The system $(c_{\sigma_k}(\mathbb{Q}), \mathcal{F}_{\sigma_k}; k \geq 1)$ forms a uniformly integrable supermartingale and hence $c_{\sigma_k}(\mathbb{Q})$ converges a.s. and in L^1. The limit $\eta = \lim c_{\sigma_k}(\mathbb{Q})$ is $\mathcal{F}_\sigma-$measurable and by the supermartingale property we also have that $\eta \leq c_\sigma(\mathbb{Q})$. But by the cocycle property $c_\sigma(\mathbb{Q}) = \mathbb{E}_\mathbb{Q}[c_{\sigma_k}(\mathbb{Q}) \mid \mathcal{F}_\sigma] + c_{\sigma,\sigma_k}(\mathbb{Q})$ and as shown above the latter term tends to 0, so we have a.s. : $\lim \mathbb{E}_\mathbb{Q}[c_{\sigma_k}(\mathbb{Q}) \mid \mathcal{F}_\sigma] = c_\sigma(\mathbb{Q})$. So we have $\mathbb{E}_\mathbb{Q}[\eta] = \lim \mathbb{E}_\mathbb{Q}[\mathbb{E}_\mathbb{Q}[c_{\sigma_k}(\mathbb{Q}) \mid \mathcal{F}_\sigma]] = \mathbb{E}_\mathbb{Q}[c_\sigma(\mathbb{Q})]$, resulting in $\eta = c_\sigma(\mathbb{Q})$. □

Theorem 62 *Let $\mathbb{Q} \sim \mathbb{P}$ and $c_0(\mathbb{Q}) < \infty$. There is a càdlàg process V such that for all stopping times $V_\sigma = c_\sigma(\mathbb{Q})$. The process V is a $\mathbb{Q}-$supermartingale and is a potential of class D. In the future it will be denoted by $c_{\cdot}(\mathbb{Q})$. There a predictable càdlàg process $\alpha(\mathbb{Q})$ so that $c_t(\mathbb{Q}) = \mathbb{E}_\mathbb{Q}[\alpha_T(\mathbb{Q}) - \alpha_t(\mathbb{Q}) \mid \mathcal{F}_t]$, if we normalise $\alpha(\mathbb{Q})$ with $\alpha_0(\mathbb{Q}) = 0$, the process $\alpha(\mathbb{Q})$ is uniquely defined.*

Proof. The family $c_t(\mathbb{Q})$ is a supermartingale with $t \to \mathbb{E}_\mathbb{Q}[c_t(\mathbb{Q})]$ being right continuous. By the modification theorem for supermartingales see [45], there is a càdlàg version, called V. For every t, the process V satisfies $V_t = c_t(\mathbb{Q})$ a.s. . We still have to check the property for stopping times. For σ having rational values, we deduce from the the previous equalities that $V_\sigma = c_\sigma(\mathbb{Q})$. If the stopping σ takes arbitrary values, then we choose a decreasing sequence of stopping times $(\sigma_k)_k$, only taking rational values with $\sigma_k \downarrow \sigma$. For each k we have $V_{\sigma_k} = c_{\sigma_k}(\mathbb{Q})$ a.s. . Since $V_{\sigma_k} \to V_\sigma$ (a.s.) by right-continuity and $c_{\sigma_k}(\mathbb{Q}) \to c_\sigma(\mathbb{Q})$ (in L^1), by proposition xxx, we find that $V_\sigma = c_\sigma(\mathbb{Q})$ a.s. . By proposition xxx the process V is a potential of class D. According to Rao's theorem, see [45], there is a uniquely defined predictable càdlàg process $\alpha(\mathbb{Q})$ so that $V_t = \mathbb{E}_\mathbb{Q}[\alpha_T(\mathbb{Q}) - \alpha_t(\mathbb{Q}) \mid \mathcal{F}_t]$. □

We can now start the proof of the modification of $u_{\cdot}(\mathbb{Q})$. The next lemmata form the basic ingredient of the proof.

Lemma 27 *For two stopping times $\sigma \leq \tau$ and for measures $\mathbb{Q} \sim \mathbb{P}$ with $c_0(\mathbb{Q}) < \infty$ the utilities satisfy the submartingale inequality*

$$\mathbb{E}_\mathbb{Q}[u_\tau(\xi) + \alpha_\tau(\mathbb{Q}) \mid \mathcal{F}_\sigma] \geq u_\sigma(\xi) + \alpha_t(\mathbb{Q}).$$

Proof. By definition $\mathbb{E}_\mathbb{Q}[\xi \mid \mathcal{F}_\sigma] + c_{\sigma,\tau} \geq u_\sigma(\xi)$. But the cocycle property says that $c_\sigma(\mathbb{Q}) = c_{\sigma,\tau} + \mathbb{E}_\mathbb{Q}[c_\tau(\mathbb{Q}) \mid \mathcal{F}_\sigma]$. With the representation of the process $c_{\cdot}(\mathbb{Q})$ we can transform this into

$$\begin{aligned} c_{\sigma,\tau}(\mathbb{Q}) &= c_\sigma(\mathbb{Q}) - \mathbb{E}_\mathbb{Q}[c_\tau(\mathbb{Q}) \mid \mathcal{F}_\sigma] \\ &= \mathbb{E}_\mathbb{Q}[\alpha_T(\mathbb{Q}) - \alpha_\sigma(\mathbb{Q}) \mid \mathcal{F}_\sigma] - \mathbb{E}_\mathbb{Q}\left[\mathbb{E}_\mathbb{Q}[\alpha_T(\mathbb{Q}) - \alpha_\tau(\mathbb{Q}) \mid \mathcal{F}_\tau] \mid \mathcal{F}_\sigma\right] \\ &= \mathbb{E}_\mathbb{Q}[\alpha_\tau(\mathbb{Q}) - \alpha_\sigma(\mathbb{Q}) \mid \mathcal{F}_\sigma]. \end{aligned}$$

The above can now be used in the inequality $\mathbb{E}_\mathbb{Q}[u_\tau(\xi) \mid \mathcal{F}_\sigma] + c_{\sigma,\tau}(\mathbb{Q}) \geq u_\sigma(\xi)$. This yields

$$\mathbb{E}_\mathbb{Q}[u_\tau(\xi) + \alpha_\tau(\mathbb{Q}) \mid \mathcal{F}_\sigma] \geq u_\sigma(\xi) + \alpha_t(\mathbb{Q}).$$

□

Lemma 28 *Let u_0 be relevant. Let σ_n be a decreasing sequence of stopping times such that $\sigma_n \downarrow \sigma$. For every $\xi \in L^\infty(\mathcal{F}_T)$, the utilities satisfy $u_{\sigma_n}(\xi) \to u_\sigma(\xi)$, a.s. .*

Proof. Fix $\varepsilon > 0$ and take a measure $\mathbb{Q} \sim \mathbb{P}$ such that

$$\mathbb{E}_\mathbb{Q}[u_{\sigma_1}(\xi) \mid \mathcal{F}_\sigma] + c_{\sigma,\sigma_n}(\mathbb{Q}) \leq u_\sigma(\xi) + \varepsilon.$$

Concatenating at $\mathbb{Q}$ with a measure $\mathbb{Q}'$ such that $c_0(\mathbb{Q}') < \infty$ and $\mathbb{Q}' \sim \mathbb{P}$ allows to assume that $c_0(\mathbb{Q}) < \infty$. The system $(u_{\sigma_n}(\xi) + \alpha_{\sigma_n(\mathbb{Q})}, \mathcal{F}_{\sigma_n})$ forms an inverse $\mathbb{Q}-$submartingale (moreover it is uniformly integrable) and hence it converges a.s. by the martingale convergence theorem. Since $\alpha_{\sigma_n}(\mathbb{Q}) \downarrow \alpha_\sigma(\mathbb{Q})$ (a.s.) this implies that also $u_{\sigma_n}(\xi) \to \eta$ (a.s.), where $\eta \in L^\infty(\mathcal{F}_\sigma)$. The submartingale property implies that $\eta \geq u_\sigma(\xi)$. Taking conditional expectations with respect to $\mathcal{F}_\sigma$ then gives $\mathbb{E}_\mathbb{Q}[u_{\sigma_n}(\xi) \mid \mathcal{F}_\sigma] \to \eta$ (a.s.). Since $u_\sigma(u_{\sigma_n}(\xi)) = u_\sigma(\xi)$ we get by lemma xxx and by the choice of $\mathbb{Q}$:

$$u_\sigma(\xi) \leq \mathbb{E}_\mathbb{Q}[u_{\sigma_n}(\xi) \mid \mathcal{F}_\sigma] + c_{\sigma,\sigma_n}(\mathbb{Q}) \leq u_\sigma(\xi) + \varepsilon.$$

Taking limits as $n \to \infty$ and using $c_{\sigma,\sigma_n}(\mathbb{Q})$, then gives: $u_\sigma(\xi) \leq \eta \leq u_\sigma(\xi)+\varepsilon$ (a.s.). Since this is true for all $\varepsilon > 0$, we conclude $\eta = u_\sigma(\xi)$, a.s. . So we conclude that $u_{\sigma_n}(\xi) \to u_\sigma(\xi)$ a.s. . Before we continue we emphasize that the a.s. convergence of the submartingale(s) was taken for $\mathbb{Q}$ but that since $\mathbb{Q} \sim \mathbb{P}$, it is a.s. for the measure $\mathbb{P}$ as well. □

Theorem 63 *Let u_0 be relevant. For $\xi \in L^\infty(\mathcal{F}_T)$, there is a càdlàg process V such that for every stopping time σ: $V_\sigma = u_\sigma(\xi)$. In the sequel we will denote this process by $u_{\cdot}(\xi)$. If $c_0(\mathbb{Q}) < \infty$, the process $(V_t + \alpha_t(\mathbb{Q}))_{0\leq t\leq T}$ is a $\mathbb{Q}-$submartingale.*

Proof. The proof is now a standard application of the modification theorem. Take $\mathbb{Q} \sim \mathbb{P}$ with $c_0(\mathbb{Q}) < \infty$. The mapping $t \to \mathbb{E}_{\mathbb{Q}}[u_t(\xi) + \alpha_t(\mathbb{Q})]$ is right continuous (by lemmaxxx above) and forms a $\mathbb{Q}-$submartingale (by lemma xxx above). By the modification theorem, see [45], there is a right continuous version. Because $\alpha(\mathbb{Q})$ is càdlàg , we get a right continuous version V, for $(u_t(\xi))_{0\le t\le T}$. As with the process c we only have to check that the a.s. equality for each t extends to an a.s. equality for each stopping time σ. This is done in the same way as for the process $c_{.}(\mathbb{Q})$ but we use lemma xxx above. □

Remark 90 The existence of a càdlàg modification was shown in [**?**] for coherent measures under the hypothesis that there is an equivalent measure $\mathbb{Q} \sim \mathbb{P}$ in the scenario set $\mathcal{S}$. In Bion-Nadal this was extended to the case of concave utility functions under the assumption that there is $\mathbb{Q} \sim \mathbb{P}$ with $c_0(\mathbb{Q}) = 0$. In this presentation, where we first investigate the process c and then the process u, we do not need this assumption. It is replaced by the weaker property that u_0 is relevant. The existence of a càdlàg modification for the processes $c_{.}(\mathbb{Q})$ was first shown in Delbaen-Peng-Rosazza-Gianin, citeDPRxxx but again under the assumption that there is $\mathbb{Q} \sim \mathbb{P}$ with $c_0(\mathbb{Q}) = 0$. In fact in that paper the standing assumption that $c_0(\mathbb{P}) = 0$ was used. The proof was different from the one above and the existence of a càdlàg modification for $u_{.}(\xi)$ (i.e. Bion-Nadal's result) was used. Apparently we only need that u_0 is relevant. The reader familiar with the construction of Snell's envelope certainly saw a big analogy. For more details we refer to [45] for an explanation on Snell's envelope and the technical problems involved.

14.3 Construction of time consistent utilities

We will work with a Brownian filtration generated by a $d-$dimensional Brownian Motion denoted by W. For a measure $\mathbb{Q} \sim \mathbb{P}$ we write the density process as a stochastic exponential, $\mathcal{E}(q \cdot W)$. This uniquely defines the predictable process q. For each $(t, \omega) \in [0, T] \times \Omega$ we give a convex function, denoted by f_t (remark that – according to bad stochastic practice – we do not write the ω):

$$f_t \colon \mathbb{R}_d \to \overline{\mathbb{R}_+}.$$

We require the following properties

1. $\inf_{x\in\mathbb{R}^d} f_t(x) = 0$

2. f_t is lower semi continuous, i.e. for each k, the set $\{x \in \mathbb{R}^d \mid f_t(x) \le k\}$ is convex and closed.

3. for each $\varepsilon > 0$, we require the existence of a "selection" q that is predictable, that satisfies $\mathcal{E}(q \cdot W)$ is uniformly integrable and $f_t(q_t) \le \varepsilon$. Moreover we require that the probability measure $d\mathbb{Q} = \mathcal{E}(q \cdot W)_T \, d\mathbb{P}$ is equivalent to $\mathbb{P}$ or what is the same $\int_0^T |q_t|^2 \, dt < \infty$ a.s. .

Now we define:

$$c_0(\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}}\left[\int_0^T f_t(q_t)\, dt\right] \text{ and for } s \le u \le T$$
$$c_{s,u}(\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}}\left[\int_s^u f_t(q_t)\, dt \mid \mathcal{F}_s\right].$$

We will show that this defines a time consistent utility function and in a later chapter we will show that in a BM framework and under an extra assumption, every time consistent utility function is of this form. Let us verify the basic properties of the function c_0.

Proposition 66 *The function c_0 is convex.*

Proposition 67 *The function c_0 satisfies $\inf_{\mathbb{Q}\sim\mathbb{P}} c_0(\mathbb{Q}) = 0$.*

Proposition 68 *The function c_0 is lower semi-continuous.*

Proposition 69 *The function c_0 can be extended in a natural way to the set of all probabilities $\mathbb{Q} \ll \mathbb{P}$.*

The set $\mathcal{A}_\sigma$ can now be characterised as follows:

$$\mathcal{A}_\sigma = \left\{\xi \mid \mathbb{E}_{\mathbb{Q}}\left[\xi + \int_\sigma^T f_t(q_t)\, dt \mid \mathcal{F}_\sigma\right] \ge 0 \text{ for all } \mathbb{Q} \sim \mathbb{P}\right\}.$$

14.4 The time independent case

Here we study in greater detail the description of the previous section. We suppose that the function f only depends on the "control" variable q, i.e. $f\colon \mathbb{R}^d \to \overline{\mathbb{R}_+}$. In this case we can be more precise.

Lemma 29 *Suppose $q^n; n \ge 1$ is a sequence of elements in $\mathbb{R}^d$ such that $\lim_n |q^n| = +\infty$ and $\lim_n \frac{f(q^n)}{|q^n|^2} = 0$. Suppose that $\varepsilon_n \to 0$ satisfies*

1. $\varepsilon_n f(q^n) \to 0$

2. $t_n = \varepsilon_n |q^n|^2 \to +\infty$

Suppose that $]\!]\sigma_n, \tau_n]\!]$ *are stochastic intervals of length* ε_n *and define the measures* $\mathbb{Q}^n$ *as*

$$\frac{d\mathbb{Q}^n}{d\mathbb{P}} = \mathcal{E}\left(\left(q^n \mathbf{1}_{]\!]\sigma_n,\tau_n]\!]}\right) \cdot W\right)_T .$$

The sequence $\mathbb{Q}^n$ *satisfies*

1. $\frac{d\mathbb{Q}^n}{d\mathbb{P}} \to 0$ *in probability.*

2. $c_0(\mathbb{Q}^n) = \mathbb{E}_{\mathbb{Q}^n}\left[\int_0^T f(q_u^n \mathbf{1}_{\sigma_n < u \le \tau_n})\, du\right] \to 0.$

Proof. Because $\int_0^T q_u^n \mathbf{1}_{\sigma_n < u \le \tau_n}\, du$ is deterministic, the stochastic exponentials define probability measures $\mathbb{Q}^n$. Since Brownian Motion starts afresh from σ_n, the densities $\frac{d\mathbb{Q}^n}{d\mathbb{P}}$ have the same law as $\exp(\sqrt{t_n}N - t_n/2)$ (where N is a standard normal variable) and hence tends to 0 in probability. Since $\int_0^T f(q_u^n \mathbf{1}_{\sigma_n < u \le \tau_n})\, du = \varepsilon_n f(q^n)$ is deterministic and tends to 0, we have $c_0(\mathbb{Q}^n) \to 0$. The existence of a sequence $\varepsilon_n > 0$ with the properties needed, is a trivial exercise. □

Proposition 70 *The following are equivalent*

1. $\liminf_{|q|\to\infty} \frac{f(q)}{|q|^2} > 0$

2. *for each* $k < \infty$*, the set* $\{\mathbb{Q} \mid c_0(\mathbb{Q}) \le k\}$ *is weakly compact.*

Proof. In case $\liminf_{|q|\to\infty} \frac{f(q)}{|q|^2} = 0$, we can find a sequence q^n as in the previous lemma. The measures $\mathbb{Q}^n$ defined there cannot form a uniformly integrable sequence, but $c_0(\mathbb{Q}^n) \to 0$. In case $\liminf_{|q|\to\infty} \frac{f(q)}{|q|^2} > 0$, there is a constant $c > 0$ such that for all $q \in \mathbb{R}^d$: $f(q) \ge -c + c|q|^2$ and hence we get

$$\mathbb{E}_{\mathbb{Q}}\left[\int_0^T f(q_u)\, du\right] \ge -cT + c\,\mathbb{E}_{\mathbb{Q}}\left[\int_0^T |q_u|^2\, du\right].$$

An easy application of the Girsanov–Maruyama formula shows:

$$\mathbb{E}_{\mathbb{Q}}\left[\int_0^T \frac{1}{2}|q_u|^2\, du\right] = \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right],$$

hence for given $k < \infty$:

$$\sup_{c_0(\mathbb{Q})\le k} \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] < \infty.$$

A simple application of the criterion of de la Vallée-Poussin shows that the set $\{\mathbb{Q} \mid c_0(\mathbb{Q}) \le k\}$ is then weakly compact (remember it is already closed and convex). □

14.5 The coherent case

From the example in section xxx, we can find the structure of the coherent time consistent utility functions. Of course this is a little premature because we did not yet prove that the example exhausts all time consistent utility functions. This will be done in the next chapter. So let $c_0(\mathbb{Q}) = \mathbb{E}_\mathbb{Q}[\int_0^T f_t(q_t)\,dt]$ with $f_t(0) = 0$ to make sure that $c_0(\mathbb{P}) = 0$. f_t is convex and takes values in $\overline{\mathbb{R}_+}$. To be coherent, the penalty function c_0 only takes values $0, +\infty$ and hence an application of the measurable selection theorem shows that also f_t only takes the values $0, +\infty$. Translated in sets this means that we have a set $\mathcal{C} \subset [0,T] \times \Omega \times \mathbb{R}^d$ such that for all $(t,\omega) : (t,\omega,0) \in \mathcal{C}$. The set $\mathcal{C}$ is in the product sigma algebra given by the predicatble sets on $[0,T] \times \Omega$ and the Borel sigma algebra on $\mathbb{R}^d$. The sets $\{q \mid (t,\omega,q) \in \mathcal{C}\}$ are convex and closed and the set of scenarios is given by the closure of

$$\mathcal{S}^e = \left\{\mathbb{Q} \mid \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_T} = \exp\left(\int_0^T q_u\,dW_u - \frac{1}{2}\int_0^T |q_u|^2\,du\right);\ (u,\omega,q_u(\omega)) \in \mathcal{C}\right\}.$$

Of course we only take selections q that are predictable and that produce uniformly integrable exponential martingales. We do NOT claim that all selections produce uniformly integrable martingales! This kind of structure was introduced and studied in [**?**] m-stable.

Example 36 Here we will give an example of an m-stable set such that $\mathcal{S}$ is not weakly compact but every element in it satisfies $\frac{d\mathbb{Q}}{d\mathbb{P}} > 0$ a.s. . We work on the interval $[0,\infty]$ but using a time transformation, this can be transformed in the interval $[0,1]$. We suppose that the probability space carries a one-dimensional Brownian Motion W and that the filtration is the natural filtration of W. For more information on Bessel processes and local martingales, we refer to [122]. The process R defined by $R_0 = 1$ and $dR_t = \frac{1}{R_t} dt + dW_t$ is a Bes^3 process. It will never reach 0 and $\lim_{t\to\infty} R_t = +\infty$ a.s. . Let us now take $L_t = \frac{1}{R_t}$. Of course we have $dL_t = -L_t^2\, dW_t$ and we know that L is a strict local martingale with $\mathbb{E}[L_t] < 1$ for all $t > 0$. The process R satisfies $dR_t = L_t\, dt + dW_t$, trivial but useful as we shall see later. Let $\tau = \inf\{t \mid L_t = 1/2\}$. Since $\lim_{t\to\infty} L_t = 0$ we conclude that $\tau < \infty$ a.s. . The set of strategies is defined as

$$\mathcal{S} = \left\{ \mathcal{E}(q \cdot W)_\tau \;\middle|\; \begin{array}{l} \mathcal{E}(q \cdot W) \text{ is uniformly integrable and} \\ |q_t| \le L_t \text{ for all } t \le \tau \\ \text{ and } q_t = 0 \text{ for } t > \tau \end{array} \right\}.$$

By construction $\mathcal{S}$ is m-stable, convex, closed and hence defines a coherent time consistent utility function u. All elements in $\mathcal{S}$ are equivalent to $\mathbb{P}$. Indeed if $\mathcal{E}(q \cdot W) \in \mathcal{S}$, then $\int_0^\tau |q_u|^2\, du \le \int_0^\tau L_u^2\, du < \infty$ a.s. and hence $\mathcal{E}(q \cdot W)_\tau > 0$ a.s. . The element $q_u = -L_u \mathbf{1}_{]\!]0,\tau]\!]}$ does not define a probability measure. Indeed $\mathcal{E}(q \cdot W) = L^\tau$ is not uniformly integrable since $\mathcal{E}(q \cdot W)_\tau = 1/2$. If we define $\tau_n = \inf\{t \mid L_t \ge n\} \wedge \tau$, then $L_{\tau_n} \in \mathcal{S}$, $L_{\tau_n} \to L_\tau$ and this shows that $\mathcal{S}$ is not weakly compact.

By James's theorem, see xxx, we know that there is an element η such that $u(\eta) = 0$ but $\mathbb{E}_\mathbb{Q}[\eta] > 0$ for all $\mathbb{Q} \in \mathcal{S}$. The idea to construct such an element is the following. We try to find η in such a way that the candidate measure is given by $q_u = -L_u \mathbf{1}_{]\!]0,\tau]\!]}$.

xxxxxx use the BMO results

Lemma 30 *Let $\eta \in L^\infty(\mathcal{F}_\tau)$ and suppose that the utility process is $u_t(\eta) = \int_0^t L_u |Z_u|\, du - \int_0^t Z_u\, dW_u$. If there is $\mathbb{Q} \in \mathcal{S}$ such that $\mathbb{E}_\mathbb{Q}[\eta] = 0$, then $L_u |Z_u| = q_u Z_u$, i.e. $q_u = L_u\, \text{sign}(Z_u)$ on $\{Z \neq 0\}$.*

Proof. For a measure $\mathbb{Q} \in \mathcal{S}$ given by the process q we have

$$\mathbb{E}_\mathbb{Q}[\eta] = \mathbb{E}_\mathbb{Q}\left[\int_0^\tau (L_u |Z_u| - q_u Z_u)\, du\right].$$

This needs to be verified. The integrand is nonnegative and therefore we can use a stopping time argument using

$$\sigma_k = \inf\left\{ t \mid \mathcal{E}(q \cdot W)_t \ge k \text{ or } \int_0^t L_u |Z_u|\, du \ge k \right\} \wedge \tau.$$

By the preceding result we have that $\int_0^\tau L_u|Z_u|\,du < \infty$ a.s. , therefore $\mathbb{P}[\sigma_k = \tau] \to 1$. The boundedness of all processes involved now allows to affirm that

$$\mathbb{E}_{\mathbb{Q}}[u_{\sigma_k}(\eta)] = \mathbb{E}_{\mathbb{Q}}\left[\int_0^{\sigma_k} (L_u|Z_u| - q_u Z_u)\,du\right].$$

Since $|u_{\sigma_k}(\eta)| \le \|\eta\|_\infty$, the left side converges to $\mathbb{E}_{\mathbb{Q}}[\eta]$ and the right side converges to $\mathbb{E}_{\mathbb{Q}}\left[\int_0^\tau (L_u|Z_u| - q_u Z_u)\,du\right]$ by Beppo Levi's theorem. Passing to the limit therefore completes the verification. The rest of the proof is now easy. For all $\mathbb{Q} \in \mathcal{S}$ the expression is nonnegative and if $\mathbb{E}_{\mathbb{Q}}[\int_0^\tau (L_u|Z_u| - q_u Z_u)\,du] = 0$, then – because $\mathbb{Q} \sim \mathbb{P}$ – necessarily $(L_u|Z_u| - q_u Z_u) = 0$ on $]\!]0,\tau]\!]$. □

We could now try to use $Z_u = -1$. This is not good since this results in the outcome $\eta = R_\tau - R_0 = 1$, the process $1 = u_t(1)$ is not given by $\int_0^t L_u|Z_u|du - \int_0^t Z_u\,dW_u$. So we must be more careful.

Lemma 31 *Suppose that $\eta = \int_0^\tau L_u|Z_u|\,du - \int_0^\tau Z_u\,dW_u \in L^\infty$, suppose that $L_u|Z_u| = q_u^0 Z_u$ on $]\!]0,\tau]\!]$ and suppose that $\mathcal{E}(q^0 \cdot W)$ is uniformly integrable. Then*

$$u_t(\eta) = \int_0^t L_u|Z_u|\,du - \int_0^t Z_u\,dW_u.$$

Proof. The proof is almost contained in the proof of the previous lemma. As in the previous lemma we get for $\mathbb{Q} \in \mathcal{S}$

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[\eta \mid \mathcal{F}_t] &= \int_0^t L_u|Z_u|\,du - \int_0^t Z_u\,dW_u + \mathbb{E}_{\mathbb{Q}}\left[\int_t^\tau (L_u|Z_u| - q_u Z_u)\,du \mid \mathcal{F}_t\right]\\ &\ge \int_0^t L_u|Z_u|\,du - \int_0^t Z_u\,dW_u.\end{aligned}$$

For the measure with density $\mathcal{E}(q^0 \cdot W)$ we get equality. □

Corollary 20 *Let $\tau_n = \inf\{t \mid L_t \ge n\} \wedge \tau$ and let $\eta_n = R_{\tau_{n+1}} - R_{\tau_n}$. Then*

$$u_t(\eta_n) = \int_{\tau_n\wedge t}^{\tau_{n+1}\wedge t} L_u\,du + \int_{\tau_n\wedge t}^{\tau_{n+1}\wedge t} dW_u.$$

Proposition 71 *Let $Z = \sum_n -2^{-n}\mathbf{1}_{]\!]\tau_n,\tau_{n+1}]\!]}$, then $\eta = \sum 2^{-n}\eta_n$ satisfies:*

$$u_t(\eta) = \int_0^t L_u|Z_u|\,du - \int_0^t Z_u\,dW_u.$$

But there is no measure $\mathbb{Q} \in \mathcal{S}$ with $\mathbb{E}_{\mathbb{Q}}[\eta] = 0$.

Proof. As the previous lemma shows: for each n we have

$$u_t\left(\sum_{k=1}^{k=n} 2^{-k}\eta_k\right) = \int_0^{t\wedge\tau_{n+1}} L_u|Z_u|\,du - \int_0^{t\wedge\tau_{n+1}} Z_u\,dW_u.$$

Since $\sum_{k=1}^{k=n} 2^{-k}\eta_k$ tends to η in the L^∞ norm we get the desired result. From lemma xxx we know that the only candidate to have $\mathbb{E}_\mathbb{Q}[\eta] = 0$ is given by the stochastic exponential L^τ. However this does not define a probability measure. □

We are now ready to prove that there are relevant time consistent utility functions with $c_0(\mathbb{Q}) > 0$ for all $\mathbb{Q}$.

Theorem 64 *The element η of the preceding proposition is minimal, $u_0(\eta) = 0$ and consequently $u^1(\xi) = u(\xi+\eta)$ defines a relevant utility function. The utility function u^1 is time consistent. The penalty function satisfies: $c_0^1(\mathbb{Q}) = \mathbb{E}_\mathbb{Q}[\eta] > 0$ for all $\mathbb{Q} \in \mathcal{S}$ and $c_0^1(\mathbb{Q}) = +\infty$ for $\mathbb{Q} \notin \mathcal{S}$.*

Proof. The preceding propositions show that $u_0(\eta) = 0$. Since the sequence L^{τ_n} is in $\mathcal{S}$ and since $L^{\tau_n} \geq 1/2$, the lemma xxx can be applied and we conclude that η is minimal. The time consistency follows from general arguments as seen before in section xxx. □

Example 37 The previous example can be made more general. We work on a filtered probability space $(\Omega, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$. The filtration is generated by a Brownian Motion W, which can be more dimensional. Suppose that L is a continuous local martingale defined on the finite time interval $[0,T]$. Suppose that for $t < T$: $L_t > 0$, $L_0 = 1$ and $L_T = 0$. Let $R = 1/L$ and let $\tau = \inf\{t \mid L_t = 1/2\}$. Clearly $\tau < T$ a.s. and $L^\tau \geq 1/2$. The process L is the solution of the stochastic differential equation: $L_0 = 1$ and $dL_t = q_t^0 L_t dW_t$ (with obvious notational changes in case W is more dimensional). The process R satisfies (for $t < \tau$):

$$dR_t = -q_t^0 R_t dW_t + (q_t^0)^2 R_t\,dt = -q_t^0 R_t dW_t + q_t^0(q_t^0 R_t)\,dt$$

The set $\mathcal{S}$ of test probabilities is defined as

$$\mathcal{S} = \left\{\mathbb{Q} \mid \frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}(q\cdot W),\ \text{with } |q| \leq |q^0|\mathbf{1}_{]\!]0,\tau]\!]}\right\}.$$

All measures in $\mathcal{S}$ are equivalent to $\mathbb{P}$. The process L^τ itself does not define a probability measure.

The random variable $R_\tau - R_0$ is nothing else than the constant 1 but R_t is not equal to the utility function $u_t(R_\tau - R_0)$ defined by $\mathcal{S}$. As in the previous example we introduce stopping times τ_k and we can carry out the same analysis. We leave the details to the reader.

Chapter 15

The Set of Local Martingale Measures

15.1 Risk Neutral Measures

In this chapter we will prove that for locally bounded processes, the set of martingale measures forms an m–stable set. This allows us to apply our previous results to situations occurring in finance. We will also see which m–stable sets can occur as sets of martingale measures for finite dimensional processes. The latter characterisation is not fully complete since it will only be done in the context of continuous filtrations. Throughout this section we will use the following notation, see [43] for more information.

On the filtered probability space $\left(\Omega, \mathcal{F}_\infty, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$, let $S\colon \mathbb{R}_+ \times \Omega \to \mathbb{R}^d$ be an adapted càdlàg process that takes values in the $d-$dimensional space $\mathbb{R}^d$. We suppose that the process is locally bounded and that the original measure is a local martingale measure for the process S. This is a simplification when compared to the situation in finance, but it simplifies notation without destroying its generality. Since the process S is locally bounded, the set

$$\mathcal{S} = \{\mathbb{Q} \ll \mathbb{P} \mid \text{ the process } S \text{ is a local martingale for } \mathbb{Q}\}$$

is a closed convex set. As the following shows, it is also m–stable. The associated time consistent utility function will be investigated later.

Proposition 72 *The set $\mathcal{S}$ is m–stable.*

Proof. We can suppose that the process S is bounded (in the same way as in [43]). That the set S is convex and closed is then obvious. The m–stability is also quite obvious. Let us take $\mathbb{Q}^1, \mathbb{Q}^2 \in \mathcal{S}^e$. Let Z^1, Z^2 be the associated density processes. If σ is a stopping time, we have to show that the density process defined as $Z_t = Z^1_t$ for $t \geq \sigma$ and $Z_t = Z^1_\sigma \frac{Z^2_t}{Z^2_\sigma}$, is still in $\mathcal{S}$. To show this, it is sufficient to show that the process ZS is a $\mathbb{P}-$martingale. This is easy. Indeed first observe that the process $Z^1 S$ is a $\mathbb{P}-$martingale (since $\mathbb{Q}^1 \in \mathcal{S}$). The same applies to Z^2 and hence the process $\mathbf{1}_{t\geq\sigma}\left(Z_t^2 S_t - Z_\sigma^2 S_\sigma\right)$ is also a $\mathbb{P}-$martingale. It follows that the process:

$$Z_t S_t = Z^1_{t\wedge\sigma} S_{t\wedge\sigma} + \frac{Z^1_\sigma}{Z^2_\sigma}\mathbf{1}_{t\geq\sigma}\left(Z_t^2 S_t - Z_\sigma^2 S_\sigma\right)$$

is also a $\mathbb{P}-$martingale. □

To avoid complicated notation we first introduce some extra notions. We restrict ourselves to the case of a continuous price process S. As above we may and do suppose that S is bounded. If X is a local martingale then there is a decomposition of X with respect to S. This decomposition, called the Kunita-Watanabe-Galtchouk decomposition, allows to write X as a sum of two local martingales. One is a stochastic integral with respect to S, the other part M is strongly orthogonal to S. So let us write $X = H \cdot S + M$. Saying that X is strongly orthogonal to S means that $H \cdot S$ is strongly orthogonal to S. This means that the vector H is orthogonal to the predictable range of S. See the appendix for a definition and for the definition of the projection P. In other words it means that the measure $H'\ d\langle S, S\rangle H = 0$ and this implies that $H \cdot S = 0$. This can only happen when the price process has some redundancy.

Theorem 65 *With the notation of the preceding sections and under the assumption that S is continuous we have that*

$$\mathcal{S}^e = \left\{ \mathcal{E}(X) \,\middle|\, \begin{array}{l} \mathcal{E}(X)_\infty > 0, \\ X \textit{ is strongly orthogonal to } S,\ \mathcal{E}(X) \textit{ is unif. integrable} \end{array} \right\}.$$

Proof. The proof is very easy. If $\mathcal{E}(X)$ is a uniformly integrable, nonnegative martingale, where $X = H \cdot S + M$ is the Kunita-Watanabe-Galtchouk decomposition, then $\mathcal{E}(X)S$ is a martingale if and only if X is strongly orthogonal to S. This is equivalent to $H \cdot S$ being strongly orthogonal to S. The latter is equivalent to the fact that every coordinate of S is strongly orthogonal to $H \cdot S$ and hence to the fact that $H'\ d\langle S, S\rangle H = 0$. This in turn is equivalent to the property $PH = H$. □

There is also a converse to this theorem. The interpretation of such a converse theorem is the following. Given a convex closed set of probabilities, when does there exist a finite dimensional process, say S, such that the given set is the set of absolutely continuous martingale measures for the process S?

A necessary condition is certainly that the set is m–stable. In the continuous case the answer is given by the following theorem.

Theorem 66 *Let $\mathcal{S}$ be a stable set of probability measures. Let the filtration be so that every local martingale is the stochastic integral with respect to the $d-$dimensional local martingale M. Let $\mathcal{S}$ be given by the closure of*

$$\mathcal{S}^e = \{\mathcal{E}(q \cdot M) \mid q \in \Phi \text{ and } \mathbb{E}\left[(\mathcal{E}(q \cdot M))_\infty\right] = 1\},$$

where the set-valued predictable process Φ is convex and closed valued. Then the set $\mathcal{S}$ is a set of equivalent local martingale measures for a price process if and only if each $\Phi(t,\omega)$ is a subspace. If the predictable projection-valued process P is the orthogonal projection on the space $\Phi(t,\omega)$, then the price process S can be chosen as $S = (Id_{\mathbb{R}^d} - P) \cdot M$.

Proof. The proof is a reformulation of the above theorem and the results of the appendix. The details are left to the reader. □

Remark 91 The situation can be generalised to the setting of theorem xxx, in the sense that we may suppose that M only generates the continuous local martingales. This means that every local martingale is given by a decomposition of the form $H \cdot M + N$, where N is purely discontinuous. In that case we get the following theorem

Theorem 67 *With the above notation we have that the closure $\mathcal{S}$ of the set*

$$\mathcal{S}^e = \left\{ \mathcal{E}(q \cdot M + N) \;\middle|\; \begin{array}{l} q \in \Phi \\ \mathcal{E}(q \cdot M + N) \text{ uniformly integrable and strictly positive} \\ N \text{ is purely discontinuous} \end{array} \right\},$$

is a set of risk neutral measures if and only if each $\Phi(t,\omega)$ is a subspace. If the predictable projection valued process P is the orthogonal projection on the space $\Phi(t,\omega)$, then the price process S can be chosen as $S = (Id_{\mathbb{R}^d} - P) \cdot M$.

15.2 Appendix on the predictable range.

If M is a $d-$dimensional martingale then it may happen that on some time intervals — or on some predictable sets — coordinates are linearly dependent. To avoid difficulties coming from this redundancy we will introduce the predictable range of M. We will only need the concept for continuous martingales. We will not give full details of the proofs, most of them being straightforward. Readers familiar with the theory of stochastic processes can skip this section that indeed does not contain anything new. Only the presentation is (maybe) of some interest.

Lemma 32 *The set*

$$\mathcal{K} = \{q \mid q \text{ predictable } d-\text{dimensional and } q \cdot [M,M] = 0\}$$

is a vector space of predictable processes satisfying

$$\text{if } q \in \mathcal{K}, \text{ if } h \text{ is predictable and real-valued, then } h\,q \in \mathcal{K}$$

Lemma 33 *The space*

$$\mathcal{K}^\perp = \{q \mid q \text{ predictable } d-\text{dimensional and for all } k \in \mathcal{K} : q.k = 0\}$$

is a vector space of predictable processes satisfying

$$\text{if } q \in \mathcal{K}^\perp, \text{ if } h \text{ is predictable and real-valued, then } h\,q \in \mathcal{K}^\perp$$

We will introduce a measure that is related to the bracket $[M,M]$. The measure is defined as follows

$$\mu(A) = \mathbb{E}\left[\int_0^\infty e^{-Trace[M,M]_t}(\mathbf{1}_A)_t \, d(Trace[M,M])_t\right].$$

It is easily seen to be finite. We use the almost standard notation of operator theory. If $x, y \in \mathbb{R}^d$, then $x \otimes y$ denotes the rank-one operator $z \to (z.x)y$. If $x = y$ and $\|x\| = 1$, this operator is the orthogonal projection on the line generated by x.

Lemma 34 (and Definition) *There exist predictable process* $e_j : \mathbb{R}_+ \times \Omega \longrightarrow \mathbb{R}^d$, $j = 1, \ldots, d$ *such that* μ *almost everywhere*

1. *for* $j \le d-1$ *we have* $\{e_{j+1} \neq 0\} \subset \{e_j \neq 0\}$.
2. *Either* $\|e_j(\omega)\| = 1$ *or* $e_j(\omega) = 0$.
3. *For* $j \neq k$ *we have* $e_j.e_k = 0$.
4. $q \in \mathcal{K}^\perp$ *if and only if there are real-valued predictable processes* $h_1, \ldots, h_d$ *such that* $q = \sum_{j \le d} h_j\, e_j$.

5. *The orthogonal projection (depending on t and ω), $P = \sum_j e_j \otimes e_j$ satisfies $q \in \mathcal{K}^\perp$ if and only if on $\mathbb{R}_+ \times \Omega$ we have $Pq = q$.*

6. *We call the range of P the predictable range for the process M.*

7. *P satisfies: $q \cdot [M, M] = 0$ if and only if $Pq = 0$.*

Proof We will not prove all the statements, the reader can easily fill in the details. The only tricky point is how to get the predictable processes e_k. The measure μ clearly satisfies: if $q = 0$, μ almost everywhere, we have $q \cdot M = 0$. This will allow us to replace predictable processes q by processes that are equal to q, μ a.e.. We now put $\mathcal{K}_1 = \mathcal{K}^\perp$ and we look at the class

$$\mathcal{C} = \{\{q \neq 0\} \mid q \in \mathcal{K}_1\}.$$

It is easily seen that this class of predictable sets is stable for countable unions. Indeed let q^n be a sequence in $\mathcal{K}_1$. Without loss of generality we may suppose that each q^n has a norm equal to either 0 or 3^{-n}, eventually we replace q^n by $\frac{q^n}{3^n \|q^n\|} \mathbf{1}_{\{q^n \neq 0\}}$. We can now verify that $q = \sum_n q^n$ satisfies $\{q \neq 0\} = \cup_n \{q^n \neq 0\}$. Since the class $\mathcal{C}$ is stable for countable unions, it has up to $\mu-$negligable sets a biggest element, coming from say an element q. Of course we may and do suppose that $\|q\|$ is either 0 or 1. Let us put $e_1 = q$. Now we look at the class

$$\mathcal{K}_2 = \{q \in \mathcal{K}_1 \mid q.e_1 = 0\},$$

and we continue with $\mathcal{K}_1$ replaced by $\mathcal{K}_2$. We again find an element e_2 with maximal support. Of course the maximality of the support of e_1 implies that $\{e_2 \neq 0\} \subset \{e_1 \neq e_2\}$. At least μ a.e., but it is easy to adapt the processes in such away that the inclusion holds as sets. We continue by induction and observe that the procedure stops after d steps, i.e. $\mathcal{K}_{d+1} = \{0\}$. We now prove item 4. Let the space obtained using the procedure of item 4 be $\mathcal{L}$. We claim that $\mathcal{L} = \mathcal{K}^\perp$, up to equality μ a.e.. If not then we take an element $q \in \mathcal{L} \setminus \mathcal{K}^\perp$, q not equal to zero μ a.e.. Replacing q by $q - Pq$ then gives an element $q \in \mathcal{K}^\perp$ such that $q.e_j = 0$ for all $j \leq d$. This, by induction, implies that $q \in \mathcal{K}_j$ for each $j \leq d$. Consequently we must have $\{q \neq 0\} \subset \{e_j \neq 0\}$ for each j. In the points where $q(t, \omega)$ is not zero this means that the vectors $e_j(t, \omega), j = 1 \ldots d$ are all nonzero and orthogonal. But then $q(t, \omega)$ is perpendicular to a basis of $\mathbb{R}^d$, a contradiction to q not equal to zero μ a.e.. □

Chapter 16

Characterisation of the Penalty Function c

16.1 Notation and the Main Theorem

We will work with the filtration $\mathcal{F}$ of a $d-$dimensional Brownian Motion W. The time interval is bounded and is supposed to be $[0,T]$.

The notation will remain fixed and u_0 is a time consistent utility function satisfying the Fatou property. We will suppose that $c_0(\mathbb{P}) = 0$, a property equivalent to $\mathbb{E}_{\mathbb{P}}[\xi] \geq 0$ for each $\xi \in \mathcal{A}_0 = \{\eta \mid u_0(\eta) \geq 0\}$. This of course implies that u_0 is relevant and $c_t(\mathbb{P}) = 0$ for all t. The utility process is the càdlàg process u_t and the penalty process (also càdlàg) is c_t. The acceptance sets at intermediate times are denoted by $\mathcal{A}_\sigma = \{\xi \mid u_\sigma(\xi) \geq 0\}$. We know from section xxx that $u_t(\xi)$ is a $\mathbb{P}-$submartingale and its Doob-Meyer decomposition will be written as $u_t(\xi) = u_0(\xi) + A_t - \int_0^t Z_u \, dW_u$. In case we need more variables we will write A^ξ, Z^ξ.

We will show the following representation theorem:

Theorem 68 *Suppose that u_0 is Fatou and time consistent. Suppose that the filtration $\mathcal{F}$ is given by a $d-$dimensional Brownian Motion W, defined on the bounded time interval $[0,T]$. Suppose that $c_0(\mathbb{P}) = 0$. Under these assumptions there is a function*

$$f \colon \mathbb{R}^d \times [0,T] \times \Omega \to \overline{\mathbb{R}_+},$$

such that

1. *for each $(t,\omega) \in [0,T] \times \Omega$, the function $f(.,t,\omega)$ is convex on $\mathbb{R}^d$,*
2. *for each $(t,\omega) \in [0,T] \times \Omega$, $f(0,t,\omega) = 0$,*
3. *for each $x \in \mathbb{R}^d$, the function $f(x,.,.)$ is predictable,*

4. *the function f is measurable for $\mathcal{B}\times\mathcal{P}$, where $\mathcal{B}$ is the Borel $\sigma-$algebra on $\mathbb{R}^d$ and $\mathcal{P}$ is the predictable $\sigma-$algebra on $[0,T]\times\Omega$,*

5. *for each $\mathbb{Q}\ll\mathbb{P}$ we have*

$$c_0(\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}}\left[\int_0^T f(q_t(.),t,.)\,dt\right].$$

16.2 Some measure theory

From section xxx we recall that for $\mathbb{Q}\ll\mathbb{P}$, $c_0(\mathbb{Q})<\infty$, the process $c_t(\mathbb{Q})$ is a $\mathbb{Q}-$supermartingale of class D. We also proved that there is an increasing process $\alpha_t(\mathbb{Q})$ such that $c_t(\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}}[\alpha_T(\mathbb{Q})-\alpha_t(\mathbb{Q})\mid\mathcal{F}_t]$, $\mathbb{Q}$ a.s. . In case $\mathbb{Q}\ll\mathbb{P}$ the density process will be denoted by L (we will drop the index $\mathbb{Q}$). The density process can be written as a stochastic exponential $L=\mathcal{E}(q\cdot W)$. Of course q is only defined up to the stopping time $\tau=\inf\{t\mid L_t=0\}$. In case we need more measures $\mathbb{Q}^1,\mathbb{Q}^2,\ldots$, the exponentials will be denoted $q^1,q^2,\ldots$. In most of the lemmas below we will use the assumption $\mathbb{Q}\sim\mathbb{P}$. This is not always needed but is done to simplify the presentation. At the end we will then use a localisation procedure to prove the final result for all measures $\mathbb{Q}\ll\mathbb{P}$.

In case β is an increasing (in better English but worse mathematics, non-decreasing) predictable càdlàg process, $\beta_0=0,\beta_T\in L^1$, we can associate a measure on the $\sigma-$algebra $\mathcal{P}$ of the predictable sets. The measure ν is defined as $\nu(H)=\mathbb{E}\left[\int_0^T \mathbf{1}_H\,d\beta\right]$. On sets of the form $B\times]s,t]$, with $B\in\mathcal{F}_s$ this gives $\mathbb{E}[\mathbf{1}_B(\beta_t-\beta_s)]$. These sets form a semi-algebra that generates the $\sigma-$algebra $\mathcal{P}$. Because we work with Brownian motion, the Lebesgue measure m on $[0,T]$ and the product measure $dm\times d\mathbb{P}$ play a fundamental role. We can decompose the measure $d\beta$ into two parts, the part, $d\beta^{ac}$ that is absolutely continuous with respect to dm and the singular part $d\beta^s$. This decomposition can be done for each $\omega\in\Omega$. One can show that the decompositions can be"glued" together. But one can also decompose the measure $d\nu$ into a part that is absolutely continuous with respect to $dm\times d\mathbb{P}$ and a part singular to it. This yields the same decomposition. So there exists a predictable set H such that $d\beta^{ac}=\mathbf{1}_H d\beta$ and $d\beta^s=\mathbf{1}_{H^c}d\beta$. In case $d\beta\ll dm$, the derivatives $d\beta/dm$ can be glued together so that they form a predictable process and in fact they build $d\beta/(dm\times d\mathbb{P})$. All this is part of the general theory of stochastic processes, see [45].

The following result from measure theory will be used.

Lemma 35 *Let $\nu^i, i \in I$ be a family of measures on $\mathcal{P}$. Suppose that the family contains the zero measure and that there is a measure μ such that for all $i \in I : \nu^i \leq \mu$. Suppose that for every $\varepsilon > 0$ and for every partition of $[0,T] \times \Omega$ in sets of the form $B \times]s,t]$, with $B \in \mathcal{F}_s$, there is a $j \in I$ such that for every element in the given partition: $(\sup_i \nu^i)(B \times]s,t]) \leq \nu^j(B \times]s,t]) + \varepsilon$. Then for each predictable set H we have $(\sup_i \nu^i)(H) = \sup_i \nu^i(H)$.*

The element $(\sup_i \nu^i)$ is calculated as the supremum of a family of measures where the outcome is a measure on $\mathcal{P}$. This supremum exists if the family is bounded above by a fixed measure. The expression $(\sup_i \nu^i)(H) = \sup_i \nu^i(H)$ is calculated for every set H separately and in general it does not define a measure. We will not prove this straightforward lemma.

Lemma 36 *For $\mathbb{Q} \sim \mathbb{P}$, $c_0(\mathbb{Q}) < \infty$, the increasing process $\alpha(\mathbb{Q})$ is the smallest measure defined on $\mathcal{P}$ such that for all $\xi \in L^\infty$:*

$$d\alpha_t(\mathbb{Q}) \geq -dA_t^\xi + q_t Z_t^\xi \, dt.$$

Proof. Since the measure $\mathbb{Q}$ is fixed, we will drop the indices $\mathbb{Q}$. Let us define $d\nu^\xi = -dA_t^\xi + q_t Z_t^\xi \, dt$. We need to show that the measure $d\alpha = \sup_\xi d\nu^\xi$, where the sup is taken in the space of measures defined on $\mathcal{P}$. We first show that $d\alpha \geq \sup_\xi d\nu^\xi$. For this it is sufficient to show that for each ξ, $B \in \mathcal{F}_s$ and $s < t \leq T$ we have

$$\mathbb{E}_\mathbb{Q}\left[(\alpha_t - \alpha_s)\mathbf{1}_B\right] \geq \mathbb{E}_\mathbb{Q}\left[\int_s^t -dA_u^\xi + q_u Z_u^\xi \, du\right].$$

This is easy since $\alpha + u(\xi)$ is a $\mathbb{Q}-$submartingale and hence

$$\mathbb{E}_\mathbb{Q}\left[\alpha_t + A_t^\xi - (Z^\xi \cdot W)_t \mid \mathcal{F}_s\right] \geq \alpha_s + A_s^\xi - (Z^\xi \cdot W)_s.$$

This can be rewritten as

$$\begin{aligned}\mathbb{E}_\mathbb{Q}[\alpha_t - \alpha_s \mid \mathcal{F}_s] &\geq \mathbb{E}_\mathbb{Q}\left[-A_t^\xi + A_s^\xi + \int_s^t Z_u^\xi \, dW_u \mid \mathcal{F}_s\right] \\ &\geq \mathbb{E}_\mathbb{Q}\left[-A_t^\xi + A_s^\xi + \int_s^t q_u Z_u^\xi \, du \mid \mathcal{F}_s\right],\end{aligned}$$

which is what we need. We now show the converse inequality. Here we will use the measure theoretic lemma stated above. The family of measures $-dA_u^\xi + q_u Z_u^\xi \, du$ is order bounded by the measure defined by $\alpha(\mathbb{Q})$. Let us give a partition in sets of the form $B\times]s,t]$. Of course refining the partition allows to find a finite sequence of times $0=t_0<t_1<\ldots<t_n=T$ and for each k a partition of Ω into $\mathcal{F}_{t_k}-$measurable sets $B_1,\ldots,B_{N_k}$. From section xxx (same xi lemma) we recall that for every $\varepsilon>0$ there is an element $\xi\in\mathcal{A}_0$ such that for every $k<n$: $\mathbb{E}_\mathbb{Q}[c_{t_k}(\mathbb{Q})]\le\mathbb{E}_\mathbb{Q}[\mathbb{E}_\mathbb{Q}[-\xi\mid\mathcal{F}_{t_k}]]+\varepsilon$. This means that for every $\varepsilon>0$ there is ξ such that for all k and $j\le N_k$: $\mathbb{E}_\mathbb{Q}[\int_{t_k}^{t_{k+1}}\mathbf{1}_{B_j}\,d\alpha_u]\le\mathbb{E}_\mathbb{Q}[\int_{t_k}^{t_{k+1}}\mathbf{1}_{B_j}\left(-dA_u^\xi+q_uZ_u^\xi\,du\right)]+\varepsilon$. The assumption of the lemma is therefore satisfied. This means that for all $B\in\mathcal{F}_s$ and all $s<t$, the supremum over all ξ of $\mathbb{E}_\mathbb{Q}\left[\int_{B\times]s,t]}\left(-dA_u^\xi+q_uZ_u^\xi\,du\right)\right]$ is exactly equal to $\mathbb{E}_\mathbb{Q}\left[\int_s^t\mathbf{1}_B\,d\alpha_u(\mathbb{Q})\right]$. □

The lemma can be rephrased as follows

Theorem 69 *Suppose $\mathbb{Q}\sim\mathbb{P},c_0(\mathbb{Q})<\infty$. The measure $d\alpha(\mathbb{Q})$ is the smallest measure such that for all ξ, the process $u_.(\xi)+\alpha(\mathbb{Q})$ is a $\mathbb{Q}-$submartingale.*

16.3 Convexity and stochastic integrals

Proposition 73 *Suppose that $\mathbb{Q}^1,\mathbb{Q}^2\sim\mathbb{P}$ and $c_0(\mathbb{Q}^1)+c_0(\mathbb{Q}^2)<\infty$. Suppose that $0<\lambda<1$ is a predictable process and suppose that $q=\lambda q^1+(1-\lambda)q^2$ defines a probability measure $\mathbb{Q}$. We have the following convexity inequality*

$$d\alpha(\mathbb{Q})\le\lambda\,d\alpha(\mathbb{Q}^1)+(1-\lambda)\,d\alpha(\mathbb{Q}^2),$$

and hence $c_0(\mathbb{Q})<\infty$.

Proof. We only have to show that for each ξ: $q_tZ_t^\xi\,dt-dA_t^\xi\le\lambda_t\,d\alpha_t(\mathbb{Q}^1)+(1-\lambda_t)\,d\alpha_t(\mathbb{Q}^2)$. This is done by a straightforward calculation.

$$\begin{aligned}q_tZ_t^\xi\,dt-dA_t^\xi&=\lambda_t\left(q_t^1Z^\xi-dA_t^\xi\right)+(1-\lambda_t)\left(q_t^2Z^\xi-dA_t^\xi\right)\\&\le\lambda_t\,d\alpha_t(\mathbb{Q}^1)+(1-\lambda_t)\,d\alpha_t(\mathbb{Q}^2).\end{aligned}$$

□

Corollary 21 *Let $\mathbb{Q}\sim\mathbb{P}$, $c_0(\mathbb{Q})<\infty$. Suppose that H is predictable and that $\mathcal{E}((q\mathbf{1}_H)\cdot W)$ and $\mathcal{E}((q\mathbf{1}_{H^c})\cdot W)$ define probability measures $\mathbb{Q}^H$ and $\mathbb{Q}^{H^c}$. Then $d\alpha(\mathbb{Q}^H)=\mathbf{1}_H\,d\alpha(\mathbb{Q})$ and $d\alpha(\mathbb{Q})=d\alpha(\mathbb{Q}^H)+d\alpha(Q^{H^c})$.*

Proof. We apply the convexity result with $\lambda = \mathbf{1}_H$, $q^1 = q$ and $q^2 = 0$. We get $d\alpha(\mathbb{Q}^H) \le \mathbf{1}_H d\alpha(\mathbb{Q})$. Consequently $d\alpha(\mathbb{Q}^H)$ is supported by H. Next we apply the convexity result with $q^1 = q\mathbf{1}_H$ and $q^2 = q\mathbf{1}_{H^c}$. We get $d\alpha(\mathbb{Q}) \le \mathbf{1}_H d\alpha(\mathbb{Q}^H) + \mathbf{1}_{H^c} d\alpha(\mathbb{Q}^{H^c})$, We now multiply with $\mathbf{1}_H$ and get $\mathbf{1}_H d\alpha(\mathbb{Q}) \le \mathbf{1}_H d\alpha(\mathbb{Q}^H) = d\alpha(\mathbb{Q}^H)$. This gives $d\alpha(\mathbb{Q}^H) = \mathbf{1}_H d\alpha(\mathbb{Q})$. □

Remark 92 Using a localisation argument together with some approximation, one can show that to get $d\alpha(\mathbb{Q}^H) = \mathbf{1}_H\, d\alpha(\mathbb{Q})$, we only need that $\mathcal{E}((q\mathbf{1}_H)\cdot W)$ defines a probability measure. We leave this as an exercise for the reader.

Lemma 37 *Let $u_0(\xi) = 0$ and suppose that $du_t(\xi) = dA_t - Z_t\, dW_t$ is the Doob-Meyer decomposition of the process $Y = u(\xi)$. Suppose that H is predictable and that the stochastic integral $\mathbf{1}_H \cdot Y$ is bounded. The element $(\mathbf{1}_H \cdot Y)_T$ belongs to $\mathcal{A}_0$.*

Proof. If H is an elementary set, i.e. there is an increasing finite sequence of stopping times $0 = \sigma_0 \le \sigma_1 \le \ldots \le \sigma_n = T$ as well sets $B_k \in \mathcal{F}_{\sigma_k}$ such that $H = \cup_{k=0}^{n-1} B_k \times]\!]\sigma_k, \sigma_{k+1}]\!]$, the statement $(\mathbf{1}_H \cdot Y)_T \in \mathcal{A}_0$ follows from the decomposition property. For general predictable sets H we proceed by approximation. There is a sequence of elementary sets H^n such that for the measure $dA_t + dm$, $\mathbf{1}_{H^n} \to \mathbf{1}_H$ a.s. . Of course the elements $\mathbf{1}_{H^n} \cdot Y$ are bounded since each H^n is elementary, but there is no guarantee that they are uniformly bounded. So we will localise these processes. Let c be such that $\sup_t |(\mathbf{1}_H \cdot Y)_t| \le c$ and $c \ge \|\xi\|_\infty$. We may suppose that $c > 0$ since otherwise there is nothing to prove. We know that $\sup_t |(\mathbf{1}_{H^n} \cdot Y)_t - (\mathbf{1}_H \cdot Y)_t| \to 0$ in probability. Hence $\mathbb{P}[\sup_t |(\mathbf{1}_{H^n} \cdot Y)_t| > 2c] \to 0$. Let $\sigma_n = \inf\{t \mid |(\mathbf{1}_{H^n} \cdot Y)_t| > 2c\}$. The jump at time σ_n is part of the jump of Y and hence bounded by $2c$. Hence $\sup_{t \le \sigma_n} |(\mathbf{1}_{H^n} \cdot Y)_t| \le 4c$. Let us now replace H^n by $H^n \cap [\![0, \sigma_n]\!]$ (we keep the notation H^n). These sets are still elementary and converge to H. Since $(\mathbf{1}_{H^n} \cdot Y)_T \in \mathcal{A}_0$ and since this sequence tends to $(\mathbf{1}_H \cdot Y)_T \in \mathcal{A}_0$, we get $(\mathbf{1}_H \cdot Y)_T \in \mathcal{A}_0$. □

Remark 93 The hypothesis that $\mathbf{1}_H \cdot Y$ is bounded is needed. The decomposition theorem cannot be generalised to the situation where we decompose $\mathcal{A}_0$ into two parts: $\mathcal{A}^H$ and $\mathcal{A}^{H^c}$, where $\mathcal{A}^H$ is obtained by the procedure of the previous lemma.

Lemma 38 *For $\xi \in L^\infty$, $u_0(\xi) = 0$ with Doob-Meyer decomposition $u(\xi) = A - (Z \cdot W)$, let H be predictable with $\mathbb{E}_\mathbb{P}[\int_{[0,T]} \mathbf{1}_{H^c}\, dt] = 0$ and such that*

$\mathbf{1}_H dA = dA^{ac}$ *where* dA^{ac} *is the absolutely continuous part of* dA. *Let* σ_n *be defined as* $\sigma_n = \inf\{t \mid |A^{ac}_t - (Z \cdot W)_t| \geq n\}$. *The element* $\xi^n = A^{ac}_{\sigma_n} - (Z \cdot W)_{\sigma_n}$ *is in* $\mathcal{A}_0$.

Proof. This follows from the previous lemma applied to sets of the form $H \cap]\!]0, \sigma_n]\!]$. Since $\mathbb{E}_{\mathbb{P}}[\int_{[0,T]} \mathbf{1}_{H^c}\, dt] = 0$, the stochastic integral $(Z \cdot W)$ is the same as $(\mathbf{1}_H Z \cdot W)$. We remark that the existence of the set H is given by the Lebesgue decomposition of the measure dA with respect to $dm \times d\mathbb{P}$. The pointwise decomposition (for almost every ω separately) is the same as the decomposition on the space $[0, T] \times \Omega$. □

Lemma 39 *Let* $\mathbb{Q} \sim \mathbb{P}$ *and* $c_0(\mathbb{Q}) < \infty$. *Then*

$$d\alpha_t = \sup_{dA^\xi \ll dm} (-dA^\xi_t + q_t Z^\xi_t\, dt).$$

In other words, to calculate the measure $d\alpha$ *we can restrict to the set of elements* ξ *such that the Doob-Meyer decomposition has a finite variation component that is absolutely continuous with respect to the Lebesgue measure.*

Proof. Clearly $d\alpha_t \geq \sup_{dA^\xi \ll dm}(-dA^\xi_t + q_t Z^\xi_t\, dt)$. But the previous lemma shows that for each ξ, $-dA^\xi + q_t Z^\xi dt \leq \sup_n(-dA^{\xi^n,ac} + q_t Z^{\xi^n} dt)$ where ξ^n is defined as in the proof of the previous lemma. □

Corollary 22 $c_0(\mathbb{Q}) = \sup\{\mathbb{E}_{\mathbb{Q}}[-\xi] \mid u_0(\xi) = 0$ *and* $dA^\xi \ll dm\}$.

Theorem 70 *Let* $\mathbb{Q} \sim \mathbb{P}$ *and* $c_0(\mathbb{Q}) < \infty$. *The measure* $d\alpha(\mathbb{Q})$ *is absolutely continuous with respect to Lebesgue measure. Consequently the process* $c_.(\mathbb{Q})$ *is continuous.*

Proof. There is not much to prove since the supremum of measures absolutely continuous with respect to the measure $dm \times d\mathbb{P}$ is still absolutely continuous with respect to $dm \times d\mathbb{P}$. □

In case $\mathbb{Q} \sim \mathbb{P}$, $c_0(\mathbb{Q}) < \infty$, we can write $d\alpha(\mathbb{Q}) = \phi\, dm \times d\mathbb{P}$, where ϕ is a predictable process.

16.4 Proof of the Main Theorem

For each measure $\mathbb{Q} \sim \mathbb{P}$, $c_0(\mathbb{Q}) < \infty$ we get a derivative ϕ. We adapt the notation as follows. If q is predictable and defines a measure $\mathbb{Q} \sim \mathbb{P}$, then we write $\phi(q,t,\omega)$ as the value of the derivative $d\alpha(\mathbb{Q})/(dm \times d\mathbb{P})$. For the moment the value of $\phi(q,t,\omega)$, depends on the measure $\mathbb{Q}$ and not just on the value of q at time t. So we cannot yet write $\phi(q_t(\omega),t,\omega)$. But this function has a special structure. Because of the localisation shown in lemma xxx above, the value of the function ϕ at the point (t,ω) is only depending on $q_t(\omega)$. This is not so easy and it is the topic of a series of lemmata. The first step is to exploit the absolute continuity of the processes $\alpha(\mathbb{Q})$.

Lemma 40 *There is a sequence ξ^n of elements in L^∞ such that*

1. $u_0(\xi^n) = 0$, $u(\xi^n) = A^n - Z^n \cdot W$, $dA^n \ll dm$
2. $(\frac{dA^n}{dm}, Z^n)$ *is dense – for the $L^1(dm \times d\mathbb{P})$ norm – in the set* $\{(\frac{dA^\xi}{dm}, Z^\xi) \mid \xi \in L^\infty, dA^\xi \ll dm\}$.

Proof. The proof is easy and follows directly from the separability of the space $L^1([0,T] \times \Omega, \mathcal{P}, dm \times d\mathbb{P})$. □

For each n we take a version of the random variables dA^n/dm and Z^n. We will abusively denote these by $\frac{dA^n}{dm}(t,\omega), Z^n(t,\omega)$ to indicate that we work with functions on $[0,T] \times \Omega$ and not just with classes modulo $dm \times d\mathbb{P}$. We may suppose that these functions are predictable.

Lemma 41 *Suppose $\mathbb{Q} \sim \mathbb{P}$ and $c_0(\mathbb{Q}) < \infty$. Then with the notation of the previous paragraph*

$$\phi(q,t,\omega) = \sup_n \left(-\frac{dA^n}{dm}(t,\omega) + q(t,\omega) Z^n(t,\omega) \right)$$

Proof. This follows from lemma xxx and the density property for the sequence ξ^n. □

Since the right hand side is a pointwise supremum we get that $\phi(q,t,\omega)$ only depends on $q(t,\omega)$. We now define

$$\mathcal{C} = \left\{ (x,t,\omega) \mid x \in \mathbb{R}^d ; \sup_n \left(-\frac{dA^n}{dm}(t,\omega) + x Z^n(t,\omega) \right) < \infty \right\}.$$

Clearly $\mathcal{C} \in \mathcal{B} \times \mathcal{P}$ where $\mathcal{B}$ is the Borel $\sigma-$algebra on $\mathbb{R}^d$. If $\mathbb{Q} \sim \mathbb{P}, c_0(\mathbb{Q}) < \infty$, then q is $dm \times d\mathbb{P}$ almost everywhere a selection of $\mathcal{C}$. We can now define the function

$$f(x,t,\omega) = \sup_n \left(-\frac{dA^n}{dm}(t,\omega) + x Z^n(t,\omega) \right).$$

Lemma 42 *$f\colon \mathbb{R}^d \times [0,T] \times \Omega \to \overline{\mathbb{R}_+}$. It is convex on $\mathbb{R}^d$ and is lsc. $f(x,t,\omega) < \infty$ if and only if $x \in \mathcal{C}_{(t,\omega)}$. f is measurable for $\mathcal{B} \times \mathcal{P}$.*

Lemma 43 *For each $\mathbb{Q} \ll \mathbb{P}$ we have*

$$c_0(\mathbb{Q}) \leq \mathbb{E}_{\mathbb{Q}}\left[\int_0^T f_u(q_u)\,du\right]$$

Proof. By definition $c_0(\mathbb{Q}) = \sup\{\mathbb{E}_{\mathbb{Q}}[-\xi] \mid u_0(\xi) = 0\}$. If we write $u(\xi) = A - Z \cdot W$, then an application of the Girsanov formula, in its version due to Lenglart, [99] we get $c_0(\mathbb{Q}) = \sup\{\mathbb{E}_{\mathbb{Q}}[-A^\xi_T + \int_0^T q_u Z^\xi_u\,du] \mid u_0(\xi) = 0\}$. Exactly as in lemma xxx above we can write

$$\begin{aligned}
c_0(\mathbb{Q}) &= \sup\{\mathbb{E}_{\mathbb{Q}}[-A^\xi_T + \int_0^T q_u Z^\xi_u\,du] \mid u_0(\xi) = 0, dA^\xi \ll dm\} \\
&= \sup\{\mathbb{E}_{\mathbb{Q}}\left[\int_0^T \left(-\frac{dA^\xi}{dm} + q_u Z^\xi_u\,du\right)\right] \mid u_0(\xi) = 0, dA^\xi \ll dm\} \\
&\leq \mathbb{E}_{\mathbb{Q}}\left[\int_0^T f_u(q_u)\,du\right].
\end{aligned}$$

Theorem 71 *For each $\mathbb{Q} \ll \mathbb{P}$ we have*

$$c_0(\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}}\left[\int_0^T f(q_t(.), t, .)\,dt\right].$$

Proof. If $\mathbb{Q} \sim \mathbb{P}, c_0(\mathbb{Q}) < \infty$, then the equality is part of the construction of the function f. For other measures we have to work a little bit. Suppose first that $c_0(\mathbb{Q}) < \infty$. Let L be the continuous density process of $\mathbb{Q}$. Define

$$\sigma^n = \inf\{t \mid L_t \leq 1/n\}.$$

Of course $\sigma^n \uparrow \sigma = \inf\{t \mid L_t = 0\}$. Let $d\mathbb{Q}^n = L_{\sigma^n}\,d\mathbb{P}$. The cocycle property shows that $c_0(\mathbb{Q}^n)$ is increasing and hence $c_0(\mathbb{Q}^n) \uparrow c_0(\mathbb{Q})$. We have

$$\begin{aligned}
c_0(\mathbb{Q}) &= \lim_n c_0(\mathbb{Q}^n) \\
&= \lim_n \mathbb{E}\left[\int_0^{\sigma^n} L_u f_u(q_u)\,du\right] \\
&= \mathbb{E}\left[\int_0^{\sigma} L_u f_u(q_u)\,du\right] = \mathbb{E}_{\mathbb{Q}}\left[\int_0^T f_u(q_u)\,du\right].
\end{aligned}$$

The case $c_0(\mathbb{Q}) = +\infty$ is easy since this is treated in lemma xxx above. $\square$

Chapter 17

law invariant case, result of Kupper, Schachermayer

In this chapter we will study the result of Kupper and Schachermayer. They proved, [94], that a time consistent, law determined utility function is necessary of entropic type. Their basic assumptions are very weak. The only non-trivial assumption is that the filtration should allow for martingales that "sufficiently move".

Chapter 18

Relation with BSDE and the use of weak compactness.

subquadratic case, superquadratic case no uniqueness, no existence, linear equation with Bessel-inverse coefficient ==¿ BSDE always solution, not always minimizer, allows a relevant utility with $c_0(\mathbb{Q}) > 0$ all measures. g subquadratic with path-dependent coefficient ==¿ always solution not unique.

18.1 Relation with BSDE

As shown in chapter xxxx, a time consistent utility function u_0 in the filtration of a $d-$dimensional Brownian Motion, is given by the expression:

$$u_0(\xi) = \inf \mathbb{E}_{\mathbb{Q}} \left[\xi + \int_0^T f_u(q_u)\, du \right].$$

Of course the expression only characterises utility functions that have the Fatou property are relevant and for which the penalty function satisfies $c_0(\mathbb{P}) = 0$. In this setting the function f has the following properties

1. $f\colon [0, T \times \Omega \times \mathbb{R}^d \to \overline{\mathbb{R}_+}$
2. for each $q \in \mathbb{R}^d$, the partial function on $[0, T] \times \Omega$ is predictable
3. for each (u, ω) the partial function is lsc and convex on $\mathbb{R}^d$
4. the function f is measurable for $\mathcal{B} \times \mathcal{P}$, where $\mathcal{B}$ is the Borel $\sigma-$algebra on $\mathbb{R}^d$ and $\mathcal{P}$ is the predictable $\sigma-$algebra on $[0, T] \times \Omega$,
5. $f(u, \omega, 0) = 0$ for each (u, ω).

The Legendre transform of $f(t, \omega, .)$ is denoted by $g(t, \omega, z)$ and it has similar properties as f. We suppose that g takes finite values. A statement equivalent to: for each (t,ω): $\lim_{|q|\to+\infty} \frac{f(t,\omega,q)}{|q|} = \infty$. This means that f has sufficient growth. It rules out the case that f is e.g. 0 on subspaces. We leave it as an exercise for the reader to show that also g is predictable in (t,ω). The following theorem gives an inequality between the solutions of BSDE and the utility process.

Theorem 72 *Suppose that the process Y is bounded and that it satisfies*

$$\begin{cases} dY_t = g_t(Z_t)\,dt - Z_t\,dW_t \\ Y_T = \xi \end{cases}.$$

We then have $Y \leq u(\xi)$

Proof Let $\mathbb{Q} \sim \mathbb{P}$ be a measure such that $c_0(\mathbb{Q}) < \infty$, then

$$\begin{aligned} &\mathbb{E}_{\mathbb{Q}}\left[\xi + \int_t^T f_u(q_u)\,du \mid \mathcal{F}_t\right] \\ &\quad = Y_t + \mathbb{E}_{\mathbb{Q}}\left[\int_t^T (g_u(Z_u) + f_u(q_u) - q_u Z_u)\,du \mid \mathcal{F}_t\right] \\ &\quad \geq Y_t \end{aligned}$$

Taking the infimum over all such measures $\mathbb{Q}$ gives $u_t(\xi) \geq Y_t$. □

The key to applications is the following "verification" theorem.

Theorem 73 *Suppose $\xi \in L^\infty$ and suppose that there exists an equivalent measure $\mathbb{Q} \sim \mathbb{P}$ such that*

$$u_0(\xi) = \mathbb{E}_{\mathbb{Q}}\left[\xi + \int_0^T f_u(q_u)\,du\right].$$

Then there exists a process Z such that for all $t \leq T$:

$$\xi = u_t(\xi) + \int_t^T g_u(Z_u)\,du - \int_t^T Z_u\,dW_u.$$

The process defined by $Y_0 = u_0(\xi)$ and $dY_t = g_t(Z_t)\,dt - Z_t\,dW_t$ is bounded and $Y_T = \xi$. The backward stochastic differential equation

$$\begin{cases} dY_t = g_t(Z_t)\,dt - Z_t\,dW_t \\ Y_T = \xi \end{cases}$$

has a bounded solution.

Proof Let $\mathbb{Q}$ be such that $u_0(\xi) = \mathbb{E}_{\mathbb{Q}}\left[\xi + \int_0^T f_u(q_u)\,du\right]$. The density process of $\mathbb{Q}$ is given by the stochastic exponential $\mathcal{E}(q \cdot W)$. From xxx it follows that $u_t(\xi) + \int_0^t f_u(q_u)\,du$ is a $\mathbb{Q}-$martingale. The Doob-Meyer decomposition – under $\mathbb{P}$ – of $u(\xi)$ can be written as $du_t(\xi) = dA_t - Z_t dW_t$, where the process $Z \cdot W$ is a BMO-martingale according to xxxx. The decomposition under $\mathbb{Q}$ therefore becomes $du_t(\xi) = dA_t - q_t Z_t dt + f_t(q_t)dt - Z_t dW_t^{\mathbb{Q}}$. The martingale property shows that $dA_t = (q_t Z_t - f_t(q_t))dt$. But $\mathbb{Q}$ is a minimising measure and hence for all other density processes $\mathcal{E}(k \cdot W)$ we must have $dA_t \geq (k_t Z_t - f_t(k_t))dt$, which shows that $dA_t = g_t(Z_t)\,dt$. Combining these results yields $du_t(\xi) = g_t(Z_t)dt - Z_t dW_t$ as required. The other statements are trivial consequences. □

Remark 94 We do not claim uniqueness of the solution. In fact we will see later that uniqueness is not always valid.

Theorem 74 *Suppose that u satisfies the weak compactness property, then for every $\xi \in L^\infty$, the BSDE has a unique bounded solution.*

Proof The existence follows from the preceding theorem and section xxx of chapter xxx. We now prove uniqueness. Suppose that Y is a bounded solution that is different from $u(\xi)$. Because $Y \leq u(\xi)$ there is a stopping time σ such that $\mathbb{P}[u_\sigma(\xi) > Y_\sigma] > 0$. Let $\mathbb{Q}$ be a minimising measure for ξ, then $\mathbb{Q} \sim \mathbb{P}$ by xxxx. But $\mathbb{Q}$ is then also a minimising measure for $u_\sigma(\xi)$ and hence

$$u_0 \quad \textbf{xxxx}$$

Remark 95 By the Bishop-Phelps theorem, the set

$$\left\{\xi \mid \text{ there is } \mathbb{Q} \ll \mathbb{P} \text{ such that } u_0(\xi) = \mathbb{E}_{\mathbb{Q}}\left[\xi + \int_0^T f_u(q_u)\,du\right]\right\}$$

is norm dense in L^∞. But there is no guarantee that a minimising measure is equivalent to $\mathbb{P}$. This is the topic of the next theorem.

Theorem 75 *Suppose that for every $\omega \in \Omega$ there is a constant $K < \infty$ depending on ω such that for all t: $g_t(z) \leq K(1 + |z|^2)$. Suppose that $\mathbb{Q}$ is a minimising measure for ξ, then necessarily $\mathbb{Q} \sim \mathbb{P}$.*

We need the following lemma (xxx rephrase in a nicer way)

Lemma 44 *If $\phi\colon \mathbb{R}^d \to \mathbb{R}_+$ is a convex function such that $\phi(z) \le K(1+|z|^2)$ for all $z \in \mathbb{R}^d$, $|z| \ge 1$ and $q \in \partial_z\phi$ (q an element of the subgradient of ϕ in the point z) satisfies $|q| \le K(1+4|z|)$. As a consequence we have the existence of a constant C (independent of $z \in \mathbb{R}^d$, e.g. $C = 5K$ will do) such that for all z and $q \in \partial_z\phi$: $|q| \le C(1+|z|)$*

Proof of the lemma Take $z \in \mathbb{R}^d$ and $q \in \partial_z\phi$. Because of convexity we have for all $w \in \mathbb{R}^d$, with $|w| = |z|$ that $g(z+w) - g(z) \ge q \cdot w$ and hence $K(1+4|z|^2) \ge |q|\,|z|$. For $|z| \ge 1$, this shows $|q| \le K(1+4|z|)$. For $|z| \le 1$ we take $|w| = 1$ and we get $5K \ge |q|$. The two inequalities give the desired result. □

Proof of the theorem We will need the Lebesgue measure m on $[0,T]$. Let $du_t(\xi) = dA_t - Z_t dW_t$ be the Doob-Meyer decomposition of the process $u(\xi)$. Let the density process L of $\mathbb{Q}$ be given by the stochastic exponential $\mathcal{E}(q \cdot W)$. Let $\tau = \inf\{t \mid L_t = 0\}$. Since $0 < \tau$ is predictable we have that there is a sequence $\tau_n < \tau$ such that $\tau_n \uparrow \tau$. Of course $\mathbb{Q} \sim \mathbb{P}$ on $\mathcal{F}_{\tau_n}$. Under $\mathbb{Q}$ the process $u_t(\xi) + \int_0^t f_t(q_t)$ is a martingale and hence $\mathbb{Q} \times m$ a.s. : $g_t(Z_t) + f_t(q_t) - q_t Z_t = 0$ on $[\![0,\tau[\![$. Hence on $[\![0,\tau[\![$, we have that q_t is a an element of the subgradient of g_t at Z_t. This is true for the measure $\mathbb{Q}\times m$ and hence also for $\mathbb{P} \times m$ (because on $\mathcal{F}_{\tau_n}$ both measures are equivalent). Hence we have $|q_t| \le C(1+|Z_t|)$. But this shows that $\int_0^\tau |q_t|^2 dt \le C + C\int_0^\tau |Z_t|^2\,dt$ for some constant C. So we have $\mathbb{P}$ a.s. that $\int_0^\tau |Z_t|^2\,dt \le \int_0^T |Z_t|^2\,dt < \infty$ and hence also $\int_0^\tau |q_t|^2 dt < \infty$ $\mathbb{P}$ a.s. . Since $\{L_T = 0\} = \{\int_0^\tau |q_t|^2 dt = \infty\}$, we have $L_T > 0$ a.s. , meaning $\mathbb{Q} \sim \mathbb{P}$. □

Theorem 76 *Suppose that for every $\omega \in \Omega$ there is a constant K depending on ω such that for all t: $g_t(z) \le K(1+|z|^2)$. For every $\xi \in L^\infty$ the process $u(\xi)$ is a bounded solution of the BSDE:*

$$\begin{cases} dY_t = g_t(Z_t)\,dt - Z_t\,dW_t \\ Y_T = \xi \end{cases}$$

Proof By the Bishop-Phelps theorem there is a sequence $\xi_n \to \xi$ (in L^∞ norm) and such that for ξ_n there is a minimising measure $\mathbb{Q}_n$. By the preceding theorem this measure is equivalent to $\mathbb{P}$. By the "verification" theorem we can write $du_t(\xi_n) = g_t(Z_t^n)\,dt - Z_t^n dW_t$. Since $\sup_t |u_t(\xi_n) - u_t(\xi)| \le \|\xi_n - \xi\|$ we can apply xxx and hence we get that $Z^n \cdot W \to Z \cdot W$ in BMO. Hence also in L^2 and this implies $\int_0^T |Z_t^n - Z_t|^2\,dt \to 0$ in probability. The hypothesis on g then implies that in probability: $\int_0^s g_t(Z_t^n)\,dt \to \int_0^s g_t(Z_t)\,dt$ for every $s \le T$. We get $u_s(\xi) = u_0(\xi) + \int_0^s g_t(Z_t)\,dt - \int_0^s Z_t\,dW_t$. □

The general case can now be analysed further. For $\xi \in L^\infty$ we know that the process $u(\xi)$ is a $\mathbb{P}-$submartingale and hence has a Doob-Meyer decomposition $du_t(\xi) = dA_t - Z_t\,dW_t$. Because for $\mathbb{Q} \sim \mathbb{P}$, $c_0(\mathbb{Q}) < \infty$, the process $u_t(\xi) + \int_)^t f_u(q_u)\,du$ is a $\mathbb{Q}-$submartingale we must have that $dA_t \ge g_t(Z_t)\,dt$. This is proved as in theorem xxx. So we can write $du_t(\xi) = g_t(Z_t)\,dt + dC_t - Z_t\,dW_t$, where C is a nondecreasing process, with $C_0 = 0$. The discontinuity points of C must be the same as the discontinuity points of $u(\xi)$. So we get that the jumps of C must be bounded by the jumps of $u(\xi)$ and hence smaller than $2\|\xi\|$. It follows that the process C is locally bounded. We know that the existence of a minimising measure $\mathbb{Q} \sim \mathbb{P}$ implies that $C_T = 0$. There is also a relation with minimal elements (as defined in xxx).

Proposition 74 *If ξ is minimal then for each stopping time σ, the element $u_\sigma(\xi)$ is minimal.*

Proof Take $\eta \le u_\sigma(\xi)$ such that $u_0(\eta) = u_0(u_\sigma(\xi))$ and $\mathbb{P}[\eta < u_\sigma(\xi)] > 0$. Obviously we have $u_\sigma(\eta) \le u_\sigma(\xi)$. We first show that $u_\sigma(\eta) < u_\sigma(\xi)$ on a set of positive measure. Indeed if, $u_\sigma(\eta) = u_\sigma(\xi)$, then

$$\begin{aligned} u_\sigma(\xi) &\ge \mathbb{E}_\mathbb{P}[\eta \mid \mathcal{F}_\sigma] \text{ since } \eta \le u_\sigma(\xi) \\ &\ge u_\sigma(\eta) = u_\sigma(\xi), \end{aligned}$$

hence $u_\sigma(\xi) = \mathbb{E}_\mathbb{P}[\eta \mid \mathcal{F}_\sigma]$. Integrating with respect to $\mathbb{P}$ gives $\int u_\sigma(\xi)\,d\mathbb{P} = \int \eta\,d\mathbb{P}$, a contradiction to $\mathbb{P}[\eta < u_\sigma(\xi)] > 0$. This shows $\mathbb{P}[u_\sigma(\eta) < u_\sigma(\xi)] > 0$. We then get

$$\begin{aligned} u_0(\xi - (u_\sigma(\xi) - u_\sigma(\eta))) &= u_0(u_\sigma(\xi - (u_\sigma(\xi) - u_\sigma(\eta)))) \\ &= u_0(u_\sigma(\eta) = u_0(\eta) = u_0(u_\sigma(\xi)) \\ &= u_0(\xi), \end{aligned}$$

showing that ξ was not minimal, a contradiction. □

Theorem 77 *If ξ is minimal, then $C_T = 0$ and hence the process $u(\xi)$ is a bounded solution of the BSDE:*

$$\begin{cases} dY_t = g_t(Z_t)\,dt - Z_t\,dW_t \\ Y_T = \xi \end{cases}$$

Proof If $C_T \neq 0$ there is a stopping time σ, such that C_σ is bounded and $\mathbb{P}[C_\sigma > 0] > 0$. Indeed take $\varepsilon > 0$ such that $\mathbb{P}[C_T > \varepsilon] > 0$. Take now $\sigma = \inf\{t \mid C_t \geq \varepsilon\}$. Since the jumps of C are bounded we have that C_σ is bounded. We can now write

$$u_\sigma(\xi) - C_\sigma = u_0(\xi) + \int_0^\sigma g_u(Z_u)\,du - \int_0^\sigma Z_u\,dW_u.$$

This shows that we have a bounded solution of the BSDE with endpoint $u_\sigma(\xi) - C_\sigma$ and starting point $u_0(\xi)$. The proposition then gives $u_0(u_\sigma(\xi) - C_\sigma) \geq u_0(\xi)$. But we certainly have $u_0(u_\sigma(\xi) - C_\sigma) \leq u_0(u_\sigma(\xi)) = u_0(\xi)$. This shows $u_0(u_\sigma(\xi) - C_\sigma)) = u_0(\xi)$ and hence $u_\sigma(\xi)$ cannot be minimal, a contradiction to proposition xxx. We conclude that $C_T = 0$. □

Remark 96 The converse is not true. We will show that there is a utility function u and a random variable ξ such that $u(\xi)$ satisfies the associated BSDE but ξ is not minimal. This shows that the characterisation of those ξ for which the BSDE has a solution is not an easy problem. But we can show the following

Theorem 78 *Suppose that for every $\omega \in \Omega$ there is a constant $K < \infty$ depending on ω such that for all t: $g_t(z) \leq K(1 + |z|^2)$. Every $\xi \in L^\infty$ is then minimal.*

Proof The proof uses that the process $u(\xi)$ is a solution of the BSDE. Suppose that $u_0(\xi) = 0$ and $du_t(\xi) = g_t(Z_t)\,dt - Z_t\,dW_t$. For each $\omega \in \Omega$ there is a constant $K(\omega)$ such that $g_t(z) \leq K(1 + |z|^2)$. We can of course suppose that K is measurable for $\mathcal{F}_T$. To find a minimising measure we can try a selection q of $\partial_Z g$ but of course there is no guarantee that the stochastic exponential $\mathcal{E}(q \cdot W)$ is uniformly integrable. Nevertheless this is the idea behind the proof. To show that ξ is minimal we take $A \in \mathcal{F}_T, \delta > 0$ and we must show that $u_0(\xi - \delta \mathbf{1}_A) < 0$. Take N big enough so that $\mathbb{P}[A \cap \{K \leq N\}] > 0$. Define the stopping time σ as

$$\sigma = \inf\{t \mid |q_t| > 5N(1 + |Z_t|\} \wedge T.$$

By lemma xxx we have that $\{\sigma < T\} \subset \{K > N\}$ and hence $\mathbb{P}[A \cap \{\sigma = T\}] > 0$. We now show that the stochastic exponential $\mathcal{E}(q \cdot W)$ stopped at σ is uniformly integrable. Since $u_\sigma(\xi) = \int_0^\sigma g_t(Z_t)\,dt - \int_0^\sigma Z_t\,dW_t$, we have $Z \cdot W$ is in BMO. But

$$|q|\mathbf{1}_{[\![0,\sigma[\![} \leq 5N(1 + |Z|)$$

and hence $q \cdot W$ is also in BMO. Therefore $\mathcal{E}(q\mathbf{1}_{[\![0,\sigma[\![} \cdot W) = \mathcal{E}(q \cdot W)^\sigma$ is uniformly integrable. The measure $d\mathbb{Q} = \mathcal{E}(q \cdot W)_\sigma \, d\mathbb{P}$ satisfies $0 = u_0(u_\sigma(\xi)) = \mathbb{E}_\mathbb{Q}\left[u_\sigma(\xi) + \int_0^\sigma f_t(q_t)\, dt\right]$. We also have that $\mathbb{Q} \sim \mathbb{P}$. The following inequalities are now obvious

$$\begin{aligned}
u_0(\xi - \delta\mathbf{1}_A) &= u_0\left(u_\sigma(\xi - \delta\mathbf{1}_A)\right) \\
&\le u_0\left(u_\sigma(\xi - \delta\mathbf{1}_{A\cap\{\sigma=T\}})\right) \\
&\le u_0\left(u_\sigma(\xi) - \delta\mathbf{1}_{A\cap\{\sigma=T\}}\right) \\
&\le \mathbb{E}_\mathbb{Q}\left[u_\sigma(\xi) + \int_0^\sigma f_t(q_t)\, dt\right] - \delta\mathbb{Q}[A \cap \{\sigma = T\}] \\
&< 0.
\end{aligned}$$

18.2 Deterministic Drivers

Our main result is the following. xxxx print out and correct the layout

Theorem 79 *Let u be the dynamic utility function defined by xxxx . Then the following are equivalent:*

1. $\underline{\lim}_{|x|\to\infty} \frac{f(x)}{|x|^2} > 0$.
2. $\overline{\lim}_{|z|\to\infty} \frac{g(z)}{|z|^2} < \infty$.
3. *For all $k > 0$, the set $\{\mathbb{Q} \mid c_0(\mathbb{Q}) \le k\}$ is weakly compact.*
4. *For all $\xi \in L^\infty(\mathcal{F}_T)$, there exists a measure $\mathbb{Q} \ll \mathbb{P}$ such that $u_0(\xi) = \mathbb{E}_\mathbb{Q}\left[\xi + \int_0^T f(q_u)\, du\right]$.*
5. *For all $\xi \in L^\infty(\mathcal{F}_T)$, there exists a measure $\mathbb{Q} \sim \mathbb{P}$ such that $u_0(\xi) = \mathbb{E}_\mathbb{Q}\left[\xi + \int_0^T f(q_u)\, du\right]$.*
6. *For all $\xi \in L^\infty(\mathcal{F}_T)$, the BSDE $dY_t = g(Z_t)\, dt - Z_t\, dW_t$ has a unique bounded solution with $Y_T = \xi$.*
7. *u_0 is strictly monotone.*

Proof $1 \Leftrightarrow 2$: item 1 implies that there exist positive constants $a, b \in R_+$ such that $f(x) \ge a|x|^2 - b$. We then get

$$g(z) = \sup_{x\in R^d} (zx - f(x)) \le \sup_{x\in R^d} (zx - a|x|^2 + b) \le \frac{1}{4a}|z|^2 + b$$

which shows that $\overline{\lim}_{z\to\infty}\frac{g(z)}{|z|^2}<\infty$. The proof of the implication $2\Rightarrow 1$ is similar.
$1\Rightarrow 3$: It suffices to verify that for any $k>0$,

$$\left\{\frac{d\mathbb{Q}}{d\mathbb{P}}\;\Big|\; c_0(Q)=\mathbb{E}_{\mathbb{Q}}\left[\int_0^T f(q_u)\,du\right]\le k\right\}$$

is uniformly integrable. The Dunford-Pettis theorem then shows that the set is weakly compact. Since $f(x)\ge a|x|^2-b$, we get

$$k\ge \mathbb{E}_{\mathbb{Q}}\left[\int_0^T f(q_u)\,du\right]\ge a\mathbb{E}_{\mathbb{Q}}\left[\int_0^T |q_u|^2\,du\right]-b.$$

Therefore,

$$\frac{1}{2}\mathbb{E}_{\mathbb{Q}}\left[\int_0^T |q_u|^2\,du\right]\le\alpha,$$

where $\alpha=\frac{k+b}{2a}$ is a positive constant independent of $\mathbb{Q}$. It follows from

$$\begin{aligned}\frac{1}{2}\mathbb{E}_{\mathbb{Q}}\left[\int_0^T |q_u|^2\,du\right]&=\mathbb{E}_{\mathbb{Q}}\left[\int_0^T q_u\,dW_u^Q+\frac{1}{2}\int_0^T |q_u|^2\,du\right]\\&=\mathbb{E}_{\mathbb{Q}}\left[\int_0^T q_u\,dW_u-\frac{1}{2}\int_0^T |q_u|^2\,du\right]\\&=\mathbb{E}_{\mathbb{Q}}\left[\log\frac{d\mathbb{Q}}{d\mathbb{P}}\right]\end{aligned}$$

that for any $k>0$,

$$\left\{\frac{d\mathbb{Q}}{d\mathbb{P}}\;\Big|\;\mathbb{E}_{\mathbb{Q}}\left[\int_0^T f(q_u)\,du\right]\le k\right\}\subset\left\{\frac{d\mathbb{Q}}{d\mathbb{P}}\;\Big|\;\mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\log\frac{d\mathbb{Q}}{d\mathbb{P}}\right]\le\alpha\right\}.$$

From the de la Vallée Poussin theorem, we conclude that

$$\left\{\frac{d\mathbb{Q}}{d\mathbb{P}}\;\Big|\;\mathbb{E}_{\mathbb{Q}}\left[\int_0^T f(q_u)\,du\right]\le k\right\}\text{ is uniformly integrable.}$$

$3\Rightarrow 1$ This is proved by contradiction. Suppose $\liminf_{|x|\to\infty}\frac{f(x)}{|x|^2}=0$, then there exists a sequence $(x_n)_{n\ge1}$ such that $\lim_{n\to\infty}|x_n|=\infty$ and $\lim_{n\to\infty}\frac{f(x_n)}{|x_n|^2}=$

0. Put $q_n = x_n 1_{[0,\delta_n \wedge T]}$ where $\delta_n = 1\Big/\left(\sqrt{\frac{f(x_n)}{|x_n|^2}}|x_n|^2\right)$. Define the measure $\mathbb{Q}_n$ with the density process $\mathcal{E}(q_n \cdot W)$. It follows from

$$c_0(\mathbb{Q}_n) = \mathbb{E}_{\mathbb{Q}_n}\left[\int_0^T f(q_n(u))\,du\right] \le \sqrt{\frac{f(x_n)}{|x_n|^2}} \to 0,$$

that for all $k > 0$, there exists $N > 0$ such that the sequence $\{\frac{d\mathbb{Q}_n}{d\mathbb{P}}\}_{n\ge N} \subset \{\frac{d\mathbb{Q}}{d\mathbb{P}} \mid c_0(\mathbb{Q}) \le k\}$. Furthermore, we have

$$\int_0^T |q_n|^2(u)\,du = \left(1\Big/\sqrt{\frac{f(x_n)}{|x_n|^2}}\right) \wedge \left(x_n^2 T\right) \to \infty,$$

which shows that $\frac{dQ_n}{dP} = \mathcal{E}(q_n \cdot W)_T \to 0$, a.s. as $n \to \infty$. Thus $\{\frac{d\mathbb{Q}_n}{d\mathbb{P}}\}_{n\ge 1}$ is not uniformly integrable.
$3 \Leftrightarrow 4$: It is a conclusion induced by James's theorem as shown by Jouini-Schachermayer-Touzi's work [84], see also chapter xxx.
$4 \Leftrightarrow 5$: It is obvious that item 5 implies item 4. For the proof of the inverse implication, we use that item 4 is equivalent to item 2.

xxxx rewrite the next lines

In this case, by convexity, there exists a positive constant c such that $|g'(z)| \le c(|z|+1)$. For any $\xi \in L^\infty(\mathcal{F}_T)$, there is a measure $Q \ll P$ such that $U_0(\xi) = E_Q[\xi + \int_0^T f(q_u)\,du]$, then, by Proposition **??**, $U_t(\xi) + \int_0^{\tau\wedge t} f(q_u)\,du$ is a Q-martingale where $\tau = \inf\{t \in [0,T] \mid \mathcal{E}(q \cdot B)_t = 0\} \wedge T$. It follows from (**??**) that

$$dA_t = (Z_t q_t - f(q_t))\,dt \quad m \otimes Q\ a.s. \text{ on } [0,\tau],$$

where m is the Lebesgue measure on $[0,T]$. Since $dA_t \ge g(Z_t)\,dt$, $m \otimes Q$ a.s., we get

$$g(Z_t) = Z_t q_t - f(q_t) \quad m \otimes Q\ a.s.,$$

which implies $q_t = g'(Z_t)$ on $[0,\tau]$. We then have

$$\begin{aligned} \int_0^\tau |q_u|^2\,du = & \int_0^\tau (g'(Z_u))^2\,du \\ \le & \ c^2 \int_0^\tau (1 + |Z_u|)^2\,du < \infty, \end{aligned}$$

which means $P\left\{\frac{dQ}{dP} = 0\right\} = P\left\{\int_0^\tau |q_u|^2\,du = \infty\right\} = 0$. Hence $Q \sim P$.

5 $\Rightarrow$ 6: For a given $\xi \in L^\infty(\mathcal{F}_T)$, if there exists a measure $Q \sim P$ such that $U_0(\xi) = E_Q\left[\xi + \int_0^T f(q_u)\,du\right]$, it follows from Lemma xxxx that $\{U_t, Z_t\}_{0\le t\le T}$ is a solution of the following BSDE:

$$\begin{cases} dY_t = g(z_t)\,dt - Z_t dB_t; 0 \le t \le T; \\ Y_T = \xi, \xi \in L^\infty(\mathcal{F}_T); \\ Y \text{ is bounded} \end{cases}$$

where $E\left[\int_0^T |z_t|^2\,dt\right] < \infty$ and $E\left[\int_0^T g(z_t)\,dt\right] < \infty$. Since, as we have proved above, condition 5 implies $\overline{\lim}_{z\to\infty} \frac{g(z)}{|z|^2} < \infty$, the BSDE has a unique bounded solution according to Kobylanski [**?**].

6 $\Rightarrow$ 2 We will prove this in the next section. See Theorem **??**.

5 $\Rightarrow$ 7 For any $\xi \in L^\infty(\mathcal{F}_T)$, there exists an equivalent measure $Q \sim P$ such that $U_0(\xi) = E_Q[\xi + \int_0^T f(q_u)\,du]$ with $\frac{dQ}{dP} = \mathcal{E}(q \cdot B)$.

Suppose that $U_0(\eta) = U_0(\xi)$ for some $\eta \in L^\infty(\mathcal{F}_T)$ with $\eta \le \xi$, P a.s. Since

$$U_0(\eta) \le E_Q\left[\eta + \int_0^T f(q_u)\,du\right] \le E_Q\left[\xi + \int_0^T f(q_u)\,du\right] = U_0(\xi),$$

we have $E_Q[\xi - \eta] = 0$, hence $\xi = \eta$, Q a.s. Thus $\xi = \eta$, P a.s. and U_0 is strictly monotone.

7 $\Rightarrow$ 2 See Remark xxx, Remark xxx or Example xxxx . $\square$

We have proved that in the case when the generator g is at most quadratic, the dynamic utility function U is the solution of BSDE. In general, however, we have the following inequality.

Proposition 75 *For any $\xi \in L^\infty(\mathcal{F}_T)$, if BSDE (**??**) has a bounded solution Y, then we have $U(\xi) \ge Y$.*

Proof Since Y is bounded, the following calculation is justified:

$$\begin{aligned} E_Q\left[\xi + \int_t^T f(q_u)\,du \Big| \mathcal{F}_t\right] &= Y_t + E_Q\left[\int_t^T g(Z_u)\,du - \int_t^T Z_u dB_u + \int_t^T f(q_u)\,du \Big| \mathcal{F}_t\right] \\ &= Y_t + E_Q\left[\int_t^T [g(Z_u) - Z_u q_u + f(q_u)]\,du \Big| \mathcal{F}_t\right] \\ &\ge Y_t, \text{ for any } Q \sim P \text{ with } E_Q\left[\int_0^T f(q_u)\,du\right] < \infty. \end{aligned}$$

□

18.3 Backward SDEs with superquadratic growth.

Chapter 19

Applications to mathematical finance

19.1 The relation with superhedging

We follow the notation of Delbaen and Schachermayer, 1994, [43]. So let $(\Omega, (\mathcal{F}_t)_{0\leq t}, \mathbb{P})$ be a filtered probability space, satisfying the usual assumptions, and let $S : \mathbb{R}_+ \times \Omega \to \mathbb{R}^d$ be a càdlàg, locally bounded, adapted process. We suppose that the set

$$\mathcal{M}^e = \{\mathbb{Q} \mid \mathbb{Q} \text{ probability } \mathbb{Q} \sim \mathbb{P}\,,\ S \text{ is a } \mathbb{Q} \text{ local-martingale}\}$$

is non-empty. This is equivalent to the condition “NFLVR” or “no free lunch with vanishing risk”. Since S is locally bounded, the closure of the set $\mathcal{M}^e$ is the closed convex set

$$\mathcal{M}^a = \{\mathbb{Q} \mid \mathbb{Q} \text{ probability } \mathbb{Q} \ll \mathbb{P}\,,\ S \text{ is a } \mathbb{Q} \text{ local-martingale}\}.$$

Let W be the space

$$W = \{(H \cdot S)_\infty \mid H \cdot S \text{ bounded}\}.$$

It can be shown, see [43], that W is a weak* closed subspace of L^∞ and of course $W^\perp = \{f \mid \mathbb{E}[f\,g] = 0 \quad \text{for all } g \in W\}$. Clearly $\mathcal{M}^a$ is the intersection of $W^\perp$ with the set of probability measures. Let the acceptance cone be defined using the set $\mathcal{M}^a$, more precisely

$$\mathcal{A} = \{f \mid f \in L^\infty \text{ for all } \mathbb{Q} \in \mathcal{M}^a \text{ we have } \mathbb{E}_\mathbb{Q}[f] \geq 0\}\,.$$

This means that $\mathcal{A}$ is the acceptance cone of the utility function, m, constructed with $\mathbb{M}^a$. This means that

$$m(\xi) = \inf\{\mathbb{E}_\mathbb{Q}[\xi] \mid \mathbb{Q} \in \mathbb{M}^a\}.$$

The analysis in [43] shows that the set $\mathcal{A}$ can be described as

$$\mathcal{A} = \{f \in L^\infty \mid \text{ there is } H \text{ admissible such that } f \geq -(H \cdot S)_\infty\}\,.$$

The difficulty in this description is that we cannot suppose that the admissible strategy H is such that $H \cdot S$ remains bounded, i.e. defines an element in W. In case the process S is continuous, things become easier. Indeed by stopping the process $H \cdot S$ when it hits the level $\|f\|_\infty$ allows to replace the already admissible startegy H by a strategy such that the value process $H \cdot S$ is also bounded above. In that case we therefore have:

$$\mathcal{A} = \{f + h \mid f \in W, h \geq 0\}.$$

In [43] we gave a counter-example when the process S is only locally bounded. In this case the set $W + L^\infty_+$ was not even norm closed. Furthermore the norm closure of $W + L^\infty_+$ was different from $\mathcal{A}$.

From now on we suppose that the price process S is continuous.

In mathematical finance, $m(\xi)$ is the minimum price that has to be charged for a contingent claim ξ. Our theory also shows that

$$\begin{aligned} m(\xi) &= \sup\{\alpha \mid \xi - \alpha \in \mathcal{A}\} \\ &= \sup\{\alpha \mid \text{ there is } f \in W, h \geq 0, \text{ such that } \xi - \alpha = f + h\} \\ &= \sup\{\alpha \mid \text{ there is } f \in W \text{ such that } \xi \geq \alpha + f\} \end{aligned}$$

Now suppose that an economic agent has sold the contingent claim ξ. This means that his position is described by the random variable $-\xi$. The agent is now interested in the smallest amount β so that $-\xi+\beta$ is acceptable. In other words he/she is looking for the number $\beta = -m(-\xi)$. In economic terms this is not necessarily the price of ξ that should be charged. It may happen that the economic agent charges a price smaller that the amount $-m(-\xi)$ and that the rest is covered with own capital. In a real world situation, it is also possible that the regulator will use another risk meaure (or utility function) than the one we are presently investigating. The number $-m(-\xi)$ is denoted by $p(\xi)$ and can be described as

$$\begin{aligned} p(\xi) &= -\sup\{\alpha \mid -\xi - \alpha \in \mathcal{A}\} \\ &= \inf\{-\alpha \mid -\xi - \alpha \in \mathcal{A}\} \\ &= \inf\{\beta \mid -\xi + \beta \in \mathcal{A}\} \\ &= \inf\{\beta \mid \text{ there is } f \in W, h \geq 0, \text{ such that } -\xi + \beta = f + h\} \\ &= \inf\{\beta \mid \text{ there is } f' \in W \text{ such that } \xi \leq \beta + f'\} \end{aligned}$$

The number $p(\xi)$ is called the superhedging price of ξ. If an investor would have $p(\xi)$ at his disposal, he would be able to find a strategy H so that $H \cdot S$

is bounded and so that $p(\xi) + (H \cdot S)_\infty \geq \xi$. This means that after having sold ξ for the price $p(\xi)$ he could, by cleverly trading, hedge out the risky position $-\xi$. The final result would then be $p(\xi)+(H \cdot S)_\infty - \xi \geq 0$. Because $p(\xi) = \sup\{\mathbb{E}_\mathbb{Q}[\xi] \mid \mathbb{Q} \in \mathbb{M}^e\}$, the quantity $p(\xi)$ is also the maximum (or better the supremum) price that can be charged for ξ. The minimum price is simply $m(\xi) = \inf_{\mathbb{Q}\in M^a} \mathbb{E}_\mathbb{Q}[\xi]$. No agent would be willing to sell ξ for less than $m(\xi)$ and no agent would be willing to buy ξ for more than $p(\xi)$.

The mapping $m\colon L^\infty \to \mathbb{R}$ is an example of a time consistent Fatou coherent utility function. Its scenario set is $\mathbb{M}^a$.

xxxxx

From [38] we recall the following theorem.

Theorem 80 *Suppose that the filtration $\mathcal{F}$ is continuous in the sense that all the martingales are continuous. Then the following alternatives hold*

1. *Either $\mathbb{M}^a = \mathbb{M}^e$ and it is a singleton, i.e. the market is complete*
2. *or when the market is incomplete, the set $\mathbb{M}^a$ does not have extreme points, hence it cannot be weakly compact.*

This theorem has the following consequence

Theorem 81 *Under the assumption that the filtration is continuous and the market is incomplete we get that the coherent utility function m is nowhere Gâteaux differentiable.*

Proof. Differentiability at a point $\xi \in L^\infty$ means, see section xxx, that the minimising measure $\mu \in \mathbf{ba}$ is unique. However the Fatou property then implies, see xxx, that $\mu \in \mathbb{M}^a$ and that it is an exposed point. Hence μ would be an extreme point in $\mathbb{M}^a$. Since there are no extreme points in $\mathbb{M}^a$, m cannot be Gâteaux differentiable at ξ. □

19.2 Relation with other utility functions

We keep the notation of the preceding section. This also includes that we suppose the price process S to be continuous. We introduce a new coherent utility function via its defining set of probability measures. We will suppose that the new utility function, u, has the Fatou property. So let $\mathcal{S}$ be a closed convex set defining the coherent utility fnction u. The corresponding risk measure is denoted by ρ, i.e. $\rho(\xi) = -u(\xi)$. We will deal with the case $\mathcal{S}$

is weakly compact. The more general case is much more difficult. We now look at two special cases:

(a) We suppose that for all ξ we have $u(\xi) \geq m(\xi)$. This is equivalent to $\rho(\eta) \leq p(\eta)$ for all $\eta \in L^\infty$. This requirement is, by the Hahn-Banach theorem, equivalent to $\mathcal{S} \subset \mathbb{M}^a$. The condition $\mathcal{S} \subset \mathbb{M}^a$ is equivalent to $\mathcal{S} \subset W^\perp$. Therefore $u(\xi) = 0$ for all $\xi \in W$. This means that something that can be replicated does not require extra capital. In terms of risk measures it means that the risk measure ρ is less severe than the superhedging requirement.

(b) If $\mathcal{S} \cap \mathbb{M}^a = \emptyset$ then, by weak compactness of $\mathcal{S}$, the Hahn Banach theorem gives us an element $\xi \in L^\infty$ so that:

$$\inf_{\mathbb{Q} \in \mathcal{S}} \mathbb{E}_{\mathbb{Q}}[\xi] > \sup_{\mathbb{Q} \in \mathbb{M}^a} \mathbb{E}_{\mathbb{Q}}[\xi] \,.$$

This means that having sold a contingent claim, the position ξ is appreciated by much more than $p(\xi)$, but this number is the maximum price that the position ξ is worth on the market. This seems to be an overestimation. We leave it to the reader to find the interpretation of an element η such that

$$\sup_{\mathbb{Q} \in \mathcal{S}} \mathbb{E}_{\mathbb{Q}}[\eta] < \inf_{\mathbb{Q} \in \mathbb{M}^a} \mathbb{E}_{\mathbb{Q}}[\eta] \,.$$

We can push this analysis a little bit further. The hypothesis $\mathbb{M}^a \cap \mathcal{S} = \emptyset$ is equivalent to $W^\perp \cap \mathcal{S} = \emptyset$. By the Hahn-Banach theorem there exists $\xi \in L^\infty$ so that $\mathbb{E}[\xi\, f] = 0$ for all $f \in W^\perp$ and $\inf_{\xi \in \mathcal{S}} \mathbb{E}_{\mathbb{Q}}[\xi] > 0$. But of course this means $\xi \in W^{\perp\perp} = W$ and hence there is a strategy H so that $(H \cdot S)$ is bounded, $(H \cdot S)_\infty = \xi$ and $u(\xi) > 0$. This would mean that a position ξ can be completely hedged (by the strategy H) at no cost and the controlling agent or supervisor allows to reduce the capital. From this we deduce the following theorem.

Theorem 82 *If $\mathcal{S}$ is a weakly compact convex set of probability measures defining the coherent utility function u, then $\mathcal{S} \cap \mathbb{M}^a \neq \emptyset$ if and only if for all $\xi \in W$ we have $u(\xi) \leq 0$.*

We leave it to the reader to rephrase this condition for the examples $\mathcal{S}_k$, $\mathcal{S}_{p,k}$ of section xxx. It leads to necessary and sufficient conditions for the existence of local martingale measures with bounded densities (or p-integrable densities).

19.3 The mixture of a coherent utility function with a financial market

In this section we will use two coherent utility functions. One is defined through a convex closed set of probability measures $\mathcal{S}$ and is denoted by u. The other one is defined by the set $\mathbb{M}^a$ of absolutely continuous risk neutral measures of a *continuous* d-dimensional price process S. The economic agent is confronted with the following situation. He has a future wealth described by the bounded random variable ξ. The associated utility is then $u(\xi)$. Since he is able to make financial transactions he can improve his utility by adding to ξ a random variable that is attainable at zero cost. If, as in the previous sections, W denotes

$$W = \{(H \cdot S)_\infty \mid H \cdot S \text{ bounded}\},$$

the economic agent is interested in the quantity

$$\tilde{u}(\xi) = \sup\{u(\xi + \eta) \mid \eta \in W\}.$$

A little algebra allows us to change this expression into the convex convolution of u and m. Indeed, because for every $\eta \in L^\infty$ we have that $\eta - m(\eta) = Z_\eta + h_\eta$ where $Z_\eta \in W$ and $h_\eta \geq 0$, we can write

$$\begin{aligned}(u\Box m)(\xi) &= \sup\{u(\xi - \eta) + m(\eta) \mid \eta \in L^\infty\}\\ &= \sup\{u(\xi - \eta + m(\eta)) \mid \eta \in L^\infty\}\\ &= \sup\{u(\xi - Z_\eta - h_\eta) \mid \eta \in L^\infty\}\\ &= \sup\{u(\xi - Z) \mid Z \in W\}\\ &= \tilde{u}(\xi).\end{aligned}$$

It follows that $\tilde{u}$ has the Fatou property as soon as $\mathcal{S}$ is weakly compact. See the next section for a counter-example when $\mathcal{S}$ is not weakly compact.

19.4 A counterexample

We showed that $\tilde{u}$ has the Fatou property when $\mathcal{S}$ is weakly compact. This section will give a counterexample for the general case. The counterexample has its own interest since it is related to correlation trading. The idea is to hedge positions coming from one market with positions coming from another correlated market. We will not work out the interpretation of the example.

We invite the reader to make his/her philosophy about it. The example uses some stochastic integration theory as well as some facts from Brownian motion theory. The reader familiar with this theory can easily complete the details. The reader not familiar with stochastic analysis should believe the author.

There are two independent Brownian motions describing the source of uncertainty. In other words the filtration is the natural filtration coming from $B = (B^1, B^2)$, where B is a standard 2-dimensional Brownian motion. The time interval is restricted to $[0, 1]$. We suppose that there are two markets. The first market trades the financial asset S^1, the second market trades the financial asset S^2. We suppose that the measure $\mathbb{P}$ is risk neutral (this to simplify notation). The dynamics of $S = (S^1, S^2)$ is given by

$$\begin{aligned} dS_t^1 &= dB_t^1 \\ dS_t^2 &= dB_t^1 + \varepsilon_t \, dB_t^2. \end{aligned}$$

here ε is a deterministic function, rapidly decreasing to zero as $t \to 1$. We can take $\varepsilon_t = \exp(-\frac{1}{1-t})$. Of course the price processes S^i are not positive, but since we are mainly interested in stochastic integrals the processes S^i can easily be replaced by their stochastic exponentials. This only complicates the notation and obscures the idea of the example. We denote by $\mathbb{M}_1^a$ and $\mathbb{M}_2^a$ the absolutely continuous probability measures that turn resp. S^1 and S^2 into a local martingale. The utility functions are denoted by resp. m_1 and m_2. Both have the Fatou property. If $\mathcal{E}$ denotes the stochastic exponential function, the sets $\mathbb{M}_1^a$ and $\mathbb{M}_2^a$ can be decribed as the closures of:

$$\begin{aligned} \mathbb{M}_1^e = \{ &\mathcal{E}\left(H \cdot B^2\right)_1 \mid \int_0^1 H_u^2 \, du < +\infty \ a.s., \ H \text{ predictable and} \\ &\mathbb{E}\left[\mathcal{E}(H \cdot B^2)_1\right] = 1\} \\ \mathbb{M}_2^e = \{ &\mathcal{E}\left((H^1, H^2) \cdot (B^1, B^2)\right)_1 \mid \int_0^1 H_u^2 \, du < +\infty \ a.s., \ H \text{ predictable,} \\ &\mathbb{E}\left[\mathcal{E}\left((H^1, H^2) \cdot (B^1, B^2)\right)_1\right] = 1 \text{ and } H_t^1 + \varepsilon_t H_t^2 = 0\}. \end{aligned}$$

These sets are not relatively weakly compact (as we will see later). The closures of these sets in the dual of L^∞ are denoted by resp. $\mathcal{S}_1^{\mathbf{ba}}$ and $\mathcal{S}_2^{\mathbf{ba}}$. We will show that $\mathbb{M}_1^a \cap \mathbb{M}_2^a = \{\mathbb{P}\}$ but that $\mathcal{S}_1^{\mathbf{ba}} \cap \mathcal{S}_2^{\mathbf{ba}} \neq \{\mathbb{P}\}$. According to the results in section xxx this means that $m_1 \square m_2$ does not have the Fatou property.

That

$$\mathbb{M}_1^e \cap \mathbb{M}_2^e = \{\mathbb{P}\}$$

is fairly obvious. It follows from the fact that the couple (S^1, S^2) defines a complete market. It can also be verified directly. If $\mathbb{Q} \in \mathbb{M}_1^a \cap \mathbb{M}_2^a$, then any non-trivial convex combination of $\mathbb{Q}$ and $\mathbb{P}$ is in $\mathbb{M}_1^e \cap \mathbb{M}_2^e$. If $\mathcal{E}\left((H^1, H^2) \cdot (B^1, B^2)\right)_1 \in \mathbb{M}_1^e \cap \mathbb{M}_2^e$, then necessarily we must have (because it is in $\mathbb{M}_1^e$) that $H^1 = 0$. But the requirement to be in $\mathbb{M}_2^e$ and $\varepsilon_t > 0 \quad (t < 1)$ then gives $H^2 = 0$. Consequently we have that $\mathcal{E}\left((H^1, H^2) \cdot (B^1, B^2)\right)_1 = 1$. Now it is fairly trivial to see that $\mathbb{M}_1^e \cap \mathbb{M}_2^e = \{\mathbb{P}\}$ implies that also

$$\mathbb{M}_1^a \cap \mathbb{M}_2^a = \{\mathbb{P}\}.$$

To see that

$$\mathcal{S}_1^{\mathbf{ba}} \cap \mathcal{S}_2^{\mathbf{ba}} \neq \{\mathbb{P}\}$$

is less easy. Let us take the following sequence of stochastic exponentials.

$$L^n = \mathcal{E}\left(H^n \cdot B^2\right),$$

where $H^n = -(5/2)^n \mathbf{1}_{]1-2^{-n}, 1-2^{-(n+1)}]}$. xxxx maybe 5/2 is not good enough – check. This sequence is in $\mathbb{M}_1^a$. The sequence $(L_1^n)_{n \geq 1}$ is equivalent to the standard basis in ℓ^1 (see below). Therefore its adherent points all lie in $\mathbf{ba} \setminus L^1$ (see below for a sketch of this result from functional analysis). Now look at the sequence of densities $\mathcal{E}\left((-\varepsilon\, H^n, H^n) \cdot (B^1, B^2)\right)_1$. This is a sequence in $\mathbb{M}_2^a$. This sequence has the same adherent points as the sequence $\mathcal{E}\left(H^n \cdot B^2\right)_1$. This is proved by calculating the L^1-norm of their difference

$$\|\mathcal{E}\left(H^n \cdot B^2\right)_1 - \mathcal{E}\left((-\varepsilon\, H^n, H^n) \cdot (B^1, B^2)\right)_1 \|_1.$$

If we denote by $\mathbb{Q}^n$ the measure (by the way in $\mathbb{M}_1^a$) defined as $d\mathbb{Q}^n/d\mathbb{P} = \mathcal{E}\left(H^n \cdot B^2\right)_1$, the above expression is simply

$$\mathbb{E}_{\mathbb{Q}^n}\left[\left|\mathcal{E}\left(-\varepsilon\, H^n \cdot B^1\right)_1 - 1\right|\right].$$

Since under $\mathbb{Q}^n$ the process B^1 is still a Brownian motion and since all the integrands are deterministic, we get that the expression is the same as

$$\mathbb{E}_{\mathbb{P}}\left[\left|\mathcal{E}\left(-\varepsilon\, H^n \cdot B^1\right)_1 - 1\right|\right].$$

Since εH^n tends to zero uniformly on $[0, 1] \times \Omega$, we get that the above expression tends to zero. So the adherent points in $\mathbf{ba}$ of the sequences $\mathcal{E}\left((-\varepsilon\, H^n, H^n) \cdot (B^1, B^2)\right)_1$ and $\mathcal{E}\left(H^n \cdot B^2\right)_1$ are the same and consequently the sets $\mathcal{S}_1^{\mathbf{ba}}$ and $\mathcal{S}_2^{\mathbf{ba}}$ have an intersection that is much bigger than $\{\mathbb{P}\}$.

Remark 97 Let B be a Brownian motion. Let q be predictable with $|q| \leq a$. We then have $\|\mathcal{E}(q \cdot B)_1 - 1\|_1 \leq \|\mathcal{E}(q \cdot B)_1 - 1\|_2 \leq \sqrt{e^{a^2} - 1}$. Hint: it is easier to calculate the L^2 norm. We calculate

$$\begin{aligned}
\mathbb{E}\left[(\mathcal{E}(q \cdot B))_1^2\right] &= \mathbb{E}\left[\exp\left(2(q \cdot B)_1\right)\exp(-\int_0^1 q_u^2\, du)\right] \\
&= \mathbb{E}\left[\mathcal{E}((2q) \cdot B)_1 \exp\left(\int_0^1 q_u^2\, du\right)\right] \\
&\leq \exp(a^2)\mathbb{E}\left[\mathcal{E}((2q) \cdot B)_1\right] = e^{a^2}.
\end{aligned}$$

From this deduce that $\|\mathcal{E}(q \cdot B)_1 - 1\|_2^2 \leq e^{a^2} - 1$. This estimate can be used to fill in the details above.

Remark 98 We still have to show that the sequence $(L_1^n)_n$ is equivalent to the unit vector basis of ℓ^1. In other words we have to show the existence of $\delta > 0$ so that $\|\sum_n \gamma_n L_1^n\|_1 \geq \delta \sum_n |\gamma_n|$. We will only give a sketch, the details are left as exercises. The first step is to show an inequality for exponentials.

Lemma 45 *If N is a standard normal random variable then for $x > 0$ we have $\mathbb{P}[N \geq x] \leq \frac{1}{x\sqrt{2\pi}} \exp(-\frac{1}{2}x^2)$.*

Proof. This inequality is standard and easily proved:

$$\begin{aligned}
\mathbb{P}[N \geq x] &= \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} \exp\left(-\frac{1}{2}u^2\right) du \\
&\leq \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} \frac{u}{x} \exp\left(-\frac{1}{2}u^2\right) du \\
&= \frac{1}{x\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right).
\end{aligned}$$

□

Lemma 46 *Let $\alpha > 0$ and let N be a standard normal variable. Let $\xi = \exp(\alpha N - \frac{1}{2}\alpha^2)$ and let $B = \{\xi \geq 1\}$. Then $\mathbb{E}[\xi] = 1$, $\mathbb{E}[\mathbf{1}_B\, \xi] = 1 - \mathbb{P}[B]$, therefore $\mathbb{E}[\mathbf{1}_{B^c}\xi] = \mathbb{P}[B]$. We also have that $\mathbb{P}[B] \leq \frac{2}{\alpha\sqrt{2\pi}} \exp(-\frac{1}{8}\alpha^2)$.*

Proof. That $\mathbb{E}[\xi] = 1$ is obvious. If we define the measure $\mathbb{Q}$ as $d\mathbb{Q} = \xi\, d\mathbb{P}$, then the random variable N, seen under the measure $\mathbb{Q}$, is a normal random variable with standard deviation 1 but with mean $\mathbb{E}_{\mathbb{Q}}[N] = \alpha$. (This can be seen by calculating $\mathbb{E}_{\mathbb{Q}}[\exp(i\,t\,N)] = \mathbb{E}_{\mathbb{P}}[\exp(i\,t\,N)\xi]$.) In other words $(\mathbb{Q}\cdot N^{-1})$ is equal to $(\mathbb{P}\cdot(N+\alpha)^{-1})$. We now have

$$\mathbb{E}_{\mathbb{P}}[\mathbf{1}_B\,\xi] = \mathbb{Q}[B] = \mathbb{Q}\left[\alpha N - \frac{1}{2}\alpha^2 \geq 0\right] = \mathbb{Q}\left[N \geq \frac{1}{2}\alpha\right]$$
$$= \mathbb{P}\left[N + \alpha \geq \frac{1}{2}\alpha\right] = \mathbb{P}\left[N \geq -\frac{1}{2}\alpha\right] = 1 - \mathbb{P}\left[N \geq \frac{1}{2}\alpha\right] = 1 - \mathbb{P}[B].$$

The inequality in the last line is proved as follows:

$$\mathbb{P}[B] = \mathbb{P}\left[\alpha N - \frac{1}{2}\alpha^2 \geq 0\right] = \mathbb{P}\left[N \geq \frac{1}{2}\alpha\right] \leq \frac{2}{\alpha\sqrt{2\pi}}\exp\left(-\frac{1}{8}\alpha^2\right).$$

□

Each of the random variables is of the form $L_1^n = \exp(\alpha_n N_n - \frac{1}{2}\alpha_n^2)$ with $\alpha_n = \sqrt{(\frac{5}{2})^n 2^{-(n+1)}}$ and where $(N_n)_n$ is an *independent* sequence of standard normal random variables. Let $B_n = \{L_1^n \geq 1\}$ and let $A_n = B_1^c \cap \ldots B_{n-1}^c \cap B_n$. The set A_n "almost" supports the random variable L_1^n whereas $\int_{A_k} L_1^n \, d\mathbb{P}$ is "small" for $k \neq n$. We leave the details to the reader to check that this implies that the sequence $(L_1^n)_n$ is equivalent to the unit vector base of ℓ^1. One can also use the following lemma.

Lemma 47 *If $(f_n)_{n\geq 1}$ is a bounded sequence in L^1 that is not uniformly integrable, then there is a subsequence $(f_{n_k})_{k\geq 1}$ that is equivalent to the unit vector base of ℓ^1.*

Proof. This is standard. Here is a sketch. Since the sequence $(f_n)_{n\geq 1}$ is not uniformly integrable, we can find $\varepsilon > 0$, a subsequence (to be completed xxx) □

Lemma 48 *Let E be a Banach space and let $F \subset E$ be a closed subspace of E, isomorphic to the space ℓ^1. Let $(y_n)_{n\geq 1}$ be a sequence in F that corresponds to the unit vector base of ℓ^1. Let $e^{**} \in E^{**}$ be a weakly adherent point of the sequence $(y_n)_{n\geq 1}$. Then $e^{**} \in E^{**} \setminus E$.*

Proof. Let T be an isomorphism between F and ℓ^1, so that $T(y_n)$ is the unit vector base of ℓ^1. In case $e^{**} \in E$, we would have that for every k the point e^{**} would be in the weak closure, i.e. $\sigma(E, E^*)$, of the convex hull of the sequence $(y_n)_{n\geq k}$. By the Hahn-Banach theorem it is therefore in the norm closure of these convex hulls. This would mean that there are convex combinations z_k of $(y_n)_{n\geq k}$ that would converge in norm to e^{**}. In particular the sequence z_k is a Cauchy sequence. This can be translated by T and it would give us a sequence $T(z_k)$ that would be a Cauchy sequence in ℓ^1. But this sequence – or at least a subsequence – is supported by disjoint sets of coordinates. Such a situation is impossible. Therefore $e^{**} \in E^{**} \setminus E$. □

19.5 Another example

The reader might ask what happens if we take two independent Brownian motions and look at the example of the previous paragraph. More precisely let the dynamics of $S = (S^1, S^2)$ be given by

$$\begin{aligned} dS_t^1 &= dB_t^1 \\ dS_t^2 &= \qquad dB_t^2. \end{aligned}$$

Again the price processes S^i are not positive, but since we are mainly interested in stochastic integrals the processes S^i can easily be replaced by their stochastic exponentials. For simplicity of the notation we assume that the time interval is $[0, +\infty)$. We again denote by $\mathbb{M}_1^a$ and $\mathbb{M}_2^a$ the absolutely continuous probability measures that turn resp. S^1 and S^2 into a local martingale. The utility functions are denoted by resp. m_1 and m_2. Both have the Fatou property. If $\mathcal{E}$ denotes the stochastic exponential function, the sets $\mathbb{M}_1^a$ and $\mathbb{M}_2^a$ can be decribed as the closures of:

$$\begin{aligned} \mathbb{M}_1^e = \{\mathcal{E}\left(H\cdot B^2\right)_\infty \mid &\int_0^\infty H_u^2\,du < +\infty\ a.s.,\ H \text{ predictable and} \\ &\mathbb{E}\left[\mathcal{E}(H\cdot B^2)_\infty\right] = 1\} \\ \mathbb{M}_2^e = \{\mathcal{E}\left(H\cdot B^1\right)_\infty \mid &\int_0^\infty H_u^2\,du < +\infty\ a.s.,\ H \text{ predictable and} \\ &\mathbb{E}\left[\mathcal{E}\left(H\cdot B^1\right)_\infty\right] = 1\}. \end{aligned}$$

These sets are not relatively weakly compact. The closures of these sets in the dual of L^∞ are denoted by resp. $\mathcal{S}_1^{\mathbf{ba}}$ and $\mathcal{S}_2^{\mathbf{ba}}$. The question is whether $m_1 \square m_2$ has the Fatou property. That

$$\mathbb{M}_1^e \cap \mathbb{M}_2^e = \{\mathbb{P}\}$$

is fairly obvious. But what happens with $\mathcal{S}_1^{\mathbf{ba}} \cap \mathcal{S}_2^{\mathbf{ba}}$? The problem is equivalent to an approximation property in L^∞ as we shall explain now. Let us introduce

$$W_1 = \{(H \cdot B^1)_\infty \mid H \text{ predictable and } (H \cdot B^1) \text{ bounded}\}$$
$$W_2 = \{(H \cdot B^2)_\infty \mid H \text{ predictable and } (H \cdot B^2) \text{ bounded}\}.$$

Furthermore let us introduce the cones:

$$\mathcal{A}_1 = \{f + g \mid f \in W_1 \text{ and } 0 \le g \in L^\infty\} = \{h \mid \mathbb{E}_\mathbb{Q}[h] \ge 0 \text{ for all } \mathbb{Q} \in \mathbb{M}_1^e\}$$
$$\mathcal{A}_2 = \{f + g \mid f \in W_2 \text{ and } 0 \le g \in L^\infty\} = \{h \mid \mathbb{E}_\mathbb{Q}[h] \ge 0 \text{ for all } \mathbb{Q} \in \mathbb{M}_2^e\}.$$

We will prove, in a series of lemma's and remarks, that

1. The set $W_1 + W_2$ is not norm-closed. The cone $\mathcal{A}_1 + \mathcal{A}_2$ is not norm closed.
2. The norm closure of $W_1 + W_2$ is strictly contained in the set
$$L_0^\infty = \{f \in L^\infty \mid \mathbb{E}_\mathbb{P}[f] = 0\}.$$
Of course the representation theorem for martingales shows that the set $W_1 + W_2$ is weak* dense in L_0^∞.
3. The norm closure of $\mathcal{A}_1 + \mathcal{A}_2 = W_1 + W_2 + L_+^\infty$ contains the set L_0^∞.
4. The preceding can be reformulated as follows. For $\xi \in L_0^\infty$ and $\varepsilon > 0$ we have $\xi + \varepsilon \in W_1 + W_2 + L_0^\infty$.
5. $\mathcal{S}_1^{\mathbf{ba}} \cap \mathcal{S}_2^{\mathbf{ba}} = \{\mathbb{P}\}$.
6. $m_1 \Box m_2 = \mathbb{P}$.
7. A bounded hedging property: for $\xi \in L^\infty$ we have
$$\mathbb{E}_\mathbb{P}[\xi] = \inf\{\alpha \mid \text{ there are } f \in W_1, g \in W_2 \text{ with } \xi \le \alpha + f + g\}.$$

The proof of these will be divided over a series of separate results. The proofs of the first two statements are independent of the proof of the other statements. To prove 1 and 2 we need some extra information on BMO martingales.

Proposition 76 *Let $T = \inf\{t \mid |B_t^1 + B_t^2| = 1\}$. The stopping time T has a Laplace transform given by $\mathbb{E}[\exp(-\lambda^2 T)] = 1/\cosh(\lambda)$. For $s \in \mathbb{C}, |\Re(s)| \le \pi/2$ we have $\mathbb{E}\left[\exp(s^2 T)\right] = 1/\cos(s)$.*

Proof. That B^1_T and B^2_T are unbounded is easily seen, we leave the proof to the reader. The Laplace transform of T is found by standard methods. Let us look at the martingale $M_t = \exp\left(\lambda(B^1_t + B^2_t) - \lambda^2 t\right)$. Then by symmetry the variable $a = B^1_T + B^2_T$ has a distribution given by $\mathbb{P}[a = 1] = \mathbb{P}[a = -1] = 1/2$. Moreover by symmetry, the variable a is independent of T. Since for $t \le T$, $M_t \le \exp(|\lambda|)$ we can apply the stopping time theorem and we get $\mathbb{E}\left[\exp(\lambda a - \lambda^2 T)\right] = 1$. This immediately implies $\mathbb{E}\left[\exp(-\lambda^2 T)\right] = 1/\cosh(\lambda)$. Since $\cosh(z)$ is analytic around the origin and different from 0 for $|z| < \pi/2$ we get for z complex and for $|\Re(z)| < \pi/2$ that $\mathbb{E}\left[\exp(z^2 T)\right] = 1/\cosh(iz)$. For $s \in \mathbb{R}, |s| < \pi/2$ we then get $\mathbb{E}\left[\exp(s^2 T)\right] = 1/\cos(s)$. Moreover we get that $\mathbb{E}\left[\exp\left((\pi/2)^2 T\right)\right] = +\infty$ as an application of the monotone convergence theorem for $s \to \pi/2$. If $|\Re(s)| = \pi/2$ but s is not real, we can proceed by a limit argument. □

Proposition 77 *The variable B^1_T satisfies $\mathbb{E}[\exp(\alpha B^1_T)] = +\infty$ for $|\alpha| \ge \pi$. The random variable B^1_T is in BMO but not in the closure of L^∞ in BMO.*

Proof. The two processes $B^1 + B^2$ and $B^1 - B^2$ are independent processes. Furthermore the stopping time T is defined through $B^1 + B^2$ and hence independent of $B^1 - B^2$. Let us denote by $\mathcal{G}$ the $\sigma-$algebra generated by the process $B^1 + B^2$. We then get, at least for λ small enough:

$$\begin{aligned}\mathbb{E}[\exp(2\lambda B^1_T)] &= \mathbb{E}\left[\mathbb{E}\left[\exp(\lambda(B^1 - B^2)_T) \mid \mathcal{G}\right] \exp(\lambda(B^1 + B^2)_T)\right] \\ &= \mathbb{E}\left[\exp(\lambda^2 T)\exp(\lambda a)\right].\end{aligned}$$

Since a and T are independent this gives

$$\cosh(\lambda)\mathbb{E}[\exp(\lambda^2 T)] = \cosh(\lambda)\frac{1}{\cos(\lambda)}.$$

We get that $\mathbb{E}[\exp(2\lambda B^1_T)] = \frac{\cosh(\lambda)}{\cos(\lambda)}$ for $|\lambda| < \pi/2$. For $|\lambda| = \pi/2$ we find as an application of Beppo Levi's theorem that $\mathbb{E}[\exp(\pi B^1_T)] = +\infty$. And of course this implies the same inequality for $|\alpha| \le \pi$.

The statement about BMO follows from BMO-theory where it is shown that elements in the closure of L^∞ in BMO necessarily have exponential moments of all order. We do not give details since this is beyond the scope of this book. □

Proposition 78 *The variable a defined above cannot be in the $BMO-$closure of $W_1 + W_2$. Consequently the variable a cannot be in the $L^\infty-$norm-closure of $W_1 + W_2$.*

Proof. Suppose that a would be in the closure of $W_1 + W_2$ for the BMO topology. This means that there are $f^n \in W_1$ and $g^n \in W_2$ so that $f^n + g^n \to a$. It can be shown that this implies that $f^n \to B^1_T$ in BMO. Since B^1_T is not in the BMO closure of L^∞ this is a contradiction. □

Proposition 79 *The set $W_1 + W_2$ is not norm closed.*

Proof. Since $W_1 \cap W_2 = \{0\}$ the closedness of $W_1 + W_2$ would imply – by the closed graph theorem – that the projections $W_1 + W_2 \to W_i$ would be continuous. Let us define $T_n = \inf\{t \mid |B^1_t| \geq n\} \wedge T$ where T is defined in the previous proposition. Set $a^n = \mathbb{E}[a \mid \mathcal{F}_{T_n}] = B^1_{T_n} + B^2_{T_n}$. Then clearly $\|a^n\|_\infty \leq 1$ but as easily seen $\|B^1_{T_n}\|_\infty = n$. This implies that the projections $W_1 + W_2 \to W_i$ cannot be continuous. Therefore $W_1 + W_2$ is not closed. □

Remark 99 This also means that there are sequences $f_n \in W_1$ and $g_n \in W_2$ so that $\|f_n\|_\infty = \|g_n\|_\infty = 1$ and such that $\|f_n - g_n\|_\infty \to 0$. For instance we can take $f_n = \frac{B^1_{T_n}}{n}$ and $g_n = \frac{-B^2_{T_n}}{n}$.

Proposition 80 *The set $\mathcal{A}_1 + \mathcal{A}_2$ is not norm closed.*

Proof. If the set $\mathcal{A}_1 + \mathcal{A}_2$ would be norm closed then the set $W_1 + W_2 = \{b \mid b \in \mathcal{A}_1 + \mathcal{A}_2 \text{ and } \mathbb{E}[b] = 0\}$ would also be norm closed. □

We now start the proof that L^∞_0 is in the norm closure of the set $\mathcal{A}_1 + \mathcal{A}_2$. We will use the duality between $L^1 \times L^1$ and $L^\infty \times L^\infty$. This allows us to use the weak$^*-$closedness of the sets $\mathcal{A}_1$ and $\mathcal{A}_2$, in other words the weak$^*-$closedness of the convex set $\mathcal{A}_1 \times \mathcal{A}_2$ in $L^\infty \times L^\infty$.

Theorem 83 *Let L^n and Z^n be uniformly integrable martingales. Suppose that $L^n = \mathcal{E}(H^n \cdot B^2)$ and $Z^n = \mathcal{E}(K^n \cdot B^1)$ for predictable processes H^n and K^n. Then $L^n \in \mathbb{M}_1$ and $Z^n \in \mathbb{M}_1$. If $\|L^n_\infty - Z^n_\infty\|_1 \to 0$, then the sequences $(L^n_\infty)_n$ and $(Z^n_\infty)_n$ converge in L^1-norm to 1.*

Proof. Replacing L by $\frac{L+1}{2}$ and replacing Z by $\frac{Z+1}{2}$ does not change the theorem (since $\mathbb{P} \in \mathbb{M}_1 \cap \mathbb{M}_2$) and allows us to suppose that $L^n_\infty \geq 1/2$, $Z^n_\infty \geq 1/2$. Furthermore we only have to show that there is a subsequence that fulfils the conclusion. So let us suppose that $\|L^n_\infty - Z^n_\infty\|_1 \leq 4^{-n}$. We will show that the sequence L^n_∞ converges to 1 and that it is uniformly integrable. This will then prove the theorem. Let us first define the stopping times

$$T_n = \inf\left\{t \mid L^n_t \geq 2^n \text{ or } Z^n_t \geq 2^n \text{ or } |L^n_t - Z^n_t| \geq 2^{-n}\right\}.$$

Clearly $\mathbb{P}[\sup_t L_t^n \geq 2^n] \leq 2^{-n}$ and $\mathbb{P}[\sup_t Z_t^n \geq 2^n] \leq 2^{-n}$. Also we have that $\mathbb{P}[\sup_t |L_t^n - Z_t^n| \geq 2^{-n}] \leq 2^n 4^{-n} = 2^{-n}$. These inequalities follow from Doob's maximum inequality. As a consequence we have that $\mathbb{P}[T_n < \infty] \leq 2^{-n+3}$. The Borel-Cantelli lemma now implies that for almost all $\omega \in \Omega : T_n(\omega) = \infty$ for n big enough.

The central trick in the proof is that the processes $L^n Z^n$ are local martingales. This is so because the bracket $\langle L^n, Z^n \rangle = 0$. Therefore we have that $(L^n Z^n)^{T_n}$ are bounded martingales. The following inequalities are trivial

$$1 = \mathbb{E}\left[L_{T_n}^n Z_{T_n}^n\right] \leq \mathbb{E}\left[L_{T_n}^n \left(L_{T_n}^n + 2^{-n}\right)\right] \leq \mathbb{E}\left[\left(L_{T_n}^n\right)^2\right] + 2^{-n},$$

and

$$1 = \mathbb{E}\left[L_{T_n}^n Z_{T_n}^n\right] \geq \mathbb{E}\left[L_{T_n}^n \left(L_{T_n}^n - 2^{-n}\right)\right] \geq \mathbb{E}\left[\left(L_{T_n}^n\right)^2\right] - 2^{-n}.$$

It follows that $\mathbb{E}\left[\left(L_{T_n}^n - 1\right)^2\right] = \mathbb{E}\left[\left(L_{T_n}^n\right)^2\right] - 1 \to 0$. We therefore have that the sequence $(L_{T_n}^n)_n$ is uniformly integrable. But we also have that $\sup_{t \leq T_n} |L_t^n - 1| \to 0$ in probability. Since T_n becomes eventually equal to ∞, this implies that $\sup_{t \geq 0} |L_t^n - 1| \to 0$ in probability. We now turn to the sequence $\left(\frac{L_\infty^n}{L_{T_n}^n}\right)_n$. We have that these elements all have expected value equal to 1 and since $L_\infty^n \geq 1/2$, the stationary converge of T_n to ∞, implies that $\frac{L_\infty^n}{L_{T_n}^n} \to 1$. Scheffé's lemma then shows that $\frac{L_\infty^n}{L_{T_n}^n} \to 1$ in L^1 norm. This implies that the sequence $\left(\frac{L_\infty^n}{L_{T_n}^n}\right)_n$ is uniformly integrable. We can now show that the sequence L_∞^n is also uniformly integrable. This will complete the proof of the theorem. So let $\varepsilon > 0$ and let $\delta > 0$ be chosen so that $\|g\|_\infty \leq 1$ and $\|g\|_1 \leq \delta$ imply that for all n: $\|L_{T_n}^n g\|_1 \leq \varepsilon$. Let $\eta > 0$ be so that $\|g\|_\infty \leq 1$ and $\|g\|_1 \leq \eta$ imply that for all n: $\|\frac{L_\infty^n}{L_{T_n}^n} g\|_1 \leq \delta$. If $A \in \mathcal{F}_\infty$ is such that $\mathbb{P}[A] \leq \eta$ we then have that $g = \mathbb{E}[\frac{L_\infty^n}{L_{T_n}^n} \mathbf{1}_A \mid \mathcal{F}_{T_n}]$ satisfies $\|g\|_\infty \leq 1$ and $\|g\|_1 \leq \delta$. Consequently for all n we have:

$$\mathbb{E}[L_\infty^n \mathbf{1}_A] = \mathbb{E}\left[L_{T_n}^n \frac{L_\infty^n}{L_{T_n}^n} \mathbf{1}_A\right] = \mathbb{E}[L_{T_n}^n g] \leq \varepsilon.$$

This shows that $(L_\infty^n)_n$ is uniformly integrable. □

Remark 100 The previous result has a quantitative equivalent. The reader can check that the theorem also follows from the following result in integration theory. The proof of this result is much more technical than the proof above and it can be skipped. An earlier version of this technical result gave an estimate or order 1/6. It did not require the hypothesis that also $\int_\Omega g\, d\mathbb{P} = 1$ but for practical applications this is only a minor generalisation. The present proof with the better exponent 1/2 was found by Jaixxx of Fudan University Shanghai. He found this nice proof with the constant $(2+\sqrt{2})$ which is up to a factor 1/2 the constant obtained in the proof below.

Lemma 49 *Let f, g be nonnegative functions defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose that $\int_\Omega f\, d\mathbb{P} = \int_\Omega g\, d\mathbb{P} = 1$ and $\int_\Omega fg\, d\mathbb{P} \le 1$. Then the following estimate holds:*

$$\int_\Omega |f-1|\, d\mathbb{P} \le \frac{\sqrt{2}+2}{2}\|f-g\|_1^{1/2}.$$

Proof. For simplicity of the notation let us write $F = f-1, G = g-1$. Clearly $\int F = \int G = 0 \ge \int FG$. Now

$$\int_{\{f\le g\}} F^2 = \int_{\{f\le g\}} FG + \int_{\{f\le g\}} F(F-G).$$

Because $F \ge -1$ and because on the set $\{f \le g\}$ we have $f-g = F-G \le 0$, this also implies

$$\int_{\{f\le g\}} F^2 \le \int_{\{f\le g\}} FG + \int_{\{f\le g\}} |F-G| = \int_{\{f\le g\}} FG + \frac{1}{2}\|f-g\|_1.$$

A similar inequality holds for F and G interchanged:

$$\int_{\{f> g\}} G^2 \le \int_{\{f> g\}} FG + \int_{\{f> g\}} |F-G| = \int_{\{f> g\}} FG + \frac{1}{2}\|f-g\|_1.$$

If we sum these two inequalities we get

$$\int_{\{f\le g\}} F^2 + \int_{\{f>g\}} G^2 \le \|f-g\|_1,$$

where we used that $\int FG \le 0$. This implies

$$\int_{\{f\le g\}} |F| + \int_{\{f>g\}} |G| \le \left(\int \mathbf{1}_{\{f\le g\}} F^2 + \mathbf{1}_{\{f>g\}} G^2\right)^{1/2} \le \|f-g\|_1^{1/2}.$$

Since obviously $\int_{\{f>g\}} |F| \leq \int_{\{f>g\}} |G| + \int_{\{f>g\}} |F-G|$ we get that

$$\int_\Omega |F| \leq \|f-g\|_1^{1/2} + \frac{1}{2}\|f-g\|_1 \leq \left(1+\sqrt{\frac{1}{2}}\right) \|f-g\|_1^{1/2}.$$

This ends the proof of the lemma. □

Theorem 84 *If $b \in L_0^\infty$ then b is in the norm closure of $\mathcal{A}_1 + \mathcal{A}_2$. Consequently the norm closure of $\mathcal{A}_1 + \mathcal{A}_2$ is equal to $\{f \mid \mathbb{E}[f] \geq 0\}$.*

Proof. As already observed above, the set

$$D = \mathcal{A}_1 \times \mathcal{A}_2.$$

is a convex, weak*−closed subset of $L^\infty \times L^\infty$. Let $\varepsilon > 0$. For each n let us define the weak*−compact convex set

$$C_n = \{(b+\varepsilon-h, h) \mid \|h\|_\infty \leq n\}.$$

We will show that for some n, necessarily $C_n \cap D \neq \emptyset$. Suppose on the contrary that for all n we have $C_n \cap D = \emptyset$. Since the set C_n is weak*−compact this implies the existence of an element $(\phi_1^n, \phi_2^n) \in L^1 \times L^1$ so that

$$\sup\{(\phi_1^n,\phi_2^n)(b+\varepsilon-h,h) \mid \|h\|_\infty \leq n\} < \inf\{(\phi_1^n,\phi_2^n)(f+k,g+k') \mid f \in W_1, g \in W_2, k, k' \geq 0\}.$$

This necessarily implies that ϕ_1^n and ϕ_2^n are nonnegative and that $\phi_1^n(f) = 0$ for all $f \in W_1$. Similarly $\phi_2^n(g) = 0$ for all $g \in W_2$. Therefore a standard normalisation allows us to suppose that $\phi_1^n \in \mathbb{M}_1$ and $\phi_2^n \in \mathbb{M}_2$. The above inequality can then be rewritten as

$$\mathbb{E}[\phi_1^n(b+\varepsilon)] + \sup_{\|h\|_\infty \leq n} \mathbb{E}[(\phi_2^n - \phi_1^n)h] < 0.$$

This implies that $(1-\varepsilon)/n \geq \|\phi_1^n - \phi_2^n\|_1 \to 0$. The previous theorem then shows that $\phi_1^n \to 1$ in L^1. The above inequality finally reduces to

$$\mathbb{E}[b+\varepsilon] \leq 0.$$

This is clearly a contradiction and hence $C_n \cap D \neq \emptyset$ for n big enough. The definition of the sets C_n allows us to write this statement as follows. There is $h \in L^\infty$ so that $b+\varepsilon-h \in \mathcal{A}_1$ and $h \in \mathcal{A}_2$. This implies $b+\varepsilon \in \mathcal{A}_1 + \mathcal{A}_2$ as desired. Since this is true for all $\varepsilon > 0$ we find that L_0^∞ is a subset of the norm closure of $\mathcal{A}_1 + \mathcal{A}_2$. The last line of the theorem is an obvious consequence. □

Remark 101 Using the technical lemma above it is possible to give a quantitative estimate in the sense that we can show that for given $\varepsilon > 0, \|b\|_\infty \leq 1, \mathbb{E}[b] = 0$, there are functions $f \in \mathcal{A}_1, g \in \mathcal{A}_2$, $\|f\|_\infty, \|g\|_\infty \leq M$ such that $\|b - (f + g)\|_\infty \leq \varepsilon$. We leave it as an exercise in accounting to find the explicit dependence between ε and $M \approx 6/\varepsilon^2$.

Proposition 81 *For every $\xi \in L^\infty$ we have $m_1 \Box m_2(\xi) = \mathbb{E}[\xi]$.*

Proof. Since the acceptance cone of $m_1 \Box m_2$ is equal to the *norm* closure of the sum of both acceptance cones, i.e. the closure of $\mathcal{A}_1 + \mathcal{A}_2$, we get that the acceptance cone of $m_1 \Box m_2$ equals the acceptance cone of the functional $\mathbb{E}[.]$. Consequently $m_1 \Box m_2(\xi) = \mathbb{E}[\xi]$. □

Remark 102 Since $W_1 + W_2$ is not norm dense in the set L_0^∞, the Hahn-Banach theorem allows us to find a nonzero element $\mu \in \mathbf{ba}, \mu \neq \mathbb{P}$ such that $\mu(f) = 0$ for all $f \in W_1 + W_2$. However such elements will be *signed measures*. Indeed the proposition above shows that the only finitely additive *probability* measure $\mu \in \mathbf{ba}$ annihilating the set $W_1 + W_2$, is necessarily equal to $\mathbb{P}$.

Chapter 20

Utility Functions of a Process

Bellman's principle, problem with open end

20.1 Definition of the Utility Process

Until now we worked with utility functions that are defined on a space of random variables. In applications to finance, it is needed to define the utility process of positions that are better described by stochastic processes. A basic example is for instance that we look at a situation where the economic agent has written an American option. Suppose that the payout process is given by the stochastic process X. Since there is a possibility of early exercise and since the exercise time cannot be controlled by the economic agent, he should look at all possible future values of the form X_τ, where τ is a stopping time. At time 0, it is therefore important to look at the quantity $\inf\{u_0(X_\tau) \mid \tau \text{ is a stopping time}\}$. At a stopping time σ the economic agent will then look at $\text{ess.inf}\{u_\sigma(X_\tau) \mid \sigma \leq \tau \text{ and } \tau \text{ is a stopping time}\}$. Of course this is very informal and we need certain assumptions on the utility function and on the stochastic processes involved. So let us fix the following notations

1. The time interval will be a closed interval $[0, T]$ and the filtration $\mathcal{F}$ satisfies the usual assumption. We will see later that if the time interval is an open end interval there are "surprises".

2. $u_0 \colon L^\infty(\Omega, \mathcal{F}_T, \mathbb{P}) \to \mathbb{R}$ is a fixed concave, monetary utility function. In case u is time consistent, this gives a family of utility "operators" u_σ.

3. Using u_0, for the moment only defined for random variables, we will define the utility process for bounded càdlàg processes X, adapted to the filtration.

Theorem 85 *If u_0 is time consistent with associated utilities u_σ (σ stopping time), if X is a bounded, adapted, càdlàg process, then for each stopping time σ, the random variable*

$$\Psi_\sigma(X) = \text{ess.inf}\left\{u_\sigma(X_\tau) \mid \sigma \le \tau \text{ and } \tau \text{ is a stopping time}\right\},$$

is well defined. $\Psi\sigma(X)$ is $\mathcal{F}_\sigma-$measurable and bounded. There is a bounded, càdlàg , adapted process V such that $V_\sigma = \Psi_\sigma(X)$.

Proof Because X is right continuous we also have

$$\Psi_\sigma(X) = \text{ess.inf}\left\{u_\sigma(X_\tau) \mid \sigma < \tau \text{ on } \{\sigma < T\} \text{ and } \tau \text{ is a stopping time}\right\}.$$

This is easily seen. Fix $\varepsilon > 0$ and let $\sigma' = \inf\{t \ge \sigma \mid |X_t - X_\sigma| \ge \varepsilon\}$. For any stopping time $\tau \ge \sigma$, we now have $|X_\tau - X_{\tau\vee\sigma'}| \le \varepsilon$. So replacing τ by $\max(\tau, \sigma')$ only changes the outcome by at most ε. Since $\varepsilon > 0$ was arbitrary and since $\sigma' > \sigma$ on $\{\sigma < T\}$, the statement follows.

The calculation of the essential infimum is done over a set of random variables that forms a lattice. For each $\mathbb{Q} \sim \mathbb{P}$ with $c_0(\mathbb{Q}) < \infty$ and for stopping times $\sigma \le \nu$ we therefore have that

$$\Psi_\sigma(X) \le \mathbb{E}_\mathbb{Q}[\Psi_\nu(X) + \alpha_\nu(\mathbb{Q}) - \alpha_\sigma(\mathbb{Q}) \mid \mathcal{F}_\sigma].$$

If $\sigma_n \downarrow \sigma$ is a decreasing sequence of stopping times, the above shows that $\Psi_{\sigma_n}(X) + \alpha_{\sigma_n}(\mathbb{Q}) - \alpha_\sigma(\mathbb{Q})$ is an inverse submartingale. Since $\alpha(\mathbb{Q})$ is right continuous (see xxx), $\alpha_{\sigma_n}(\mathbb{Q}) - \alpha_\sigma(\mathbb{Q})$ tends to zero. It follows that $\Psi_{\sigma_n}(X)$ converges almost surely to a random variable η, which is $\mathcal{F}_\sigma-$measurable. But the submartingale inequality also shows that $\eta \ge \Psi_\sigma(X)$. We will now show that $\eta \le \Psi_\sigma(X)$. Given $\varepsilon > 0$ there exists a stopping time $\tau > \sigma$ such that $u_\sigma(X_\tau) \le \Psi_\sigma(X) + \varepsilon$. In the same way as in chapter xxx, there then exists a measure $\mathbb{Q} \sim \mathbb{P}$ such that $u_\sigma(X_\tau) + \varepsilon \ge \mathbb{E}_\mathbb{Q}[X_\tau \mid \mathcal{F}_\sigma] + c_{\sigma,\tau}(\mathbb{Q})$. We find that $\Psi_\sigma(X) + 2\varepsilon \ge \mathbb{E}_\mathbb{Q}[X_\tau \mid \mathcal{F}_\sigma] + c_{\sigma,\tau}(\mathbb{Q})$. For each n we have $\sigma \le \sigma_n \wedge \tau \le \tau$ and the cocycle property then implies

$$\begin{aligned}
\Psi_\sigma(X) + 2\varepsilon &\ge \mathbb{E}_\mathbb{Q}[X_\tau \mid \mathcal{F}_\sigma] + c_{\sigma,\tau}(\mathbb{Q}) \\
&\ge \mathbb{E}_\mathbb{Q}\left[\mathbb{E}_\mathbb{Q}[X_\tau \mid \mathcal{F}_{\sigma_n\wedge\tau}] \mid \mathcal{F}_\sigma\right] + c_{\sigma,\sigma_n\wedge\tau}(\mathbb{Q}) + \mathbb{E}_\mathbb{Q}\left[c_{\sigma_n\wedge\tau,\tau}(\mathbb{Q}) \mid \mathcal{F}_\sigma\right] \\
&\ge \mathbb{E}_\mathbb{Q}\left[\mathbb{E}_\mathbb{Q}[X_\tau \mid \mathcal{F}_{\sigma_n\wedge\tau}] + c_{\sigma_n\wedge\tau,\tau}(\mathbb{Q}) \mid \mathcal{F}_\sigma\right] + c_{\sigma,\sigma_n\wedge\tau}(\mathbb{Q}) \\
&\ge \mathbb{E}_\mathbb{Q}[u_{\sigma_n\wedge\tau}(X_\tau) \mid \mathcal{F}_\sigma] + c_{\sigma,\sigma_n\wedge\tau}(\mathbb{Q}) \\
&\ge \mathbb{E}_\mathbb{Q}[u_{\sigma_n}(X_\tau)\mathbf{1}_{\tau\ge\sigma_n} \mid \mathcal{F}_\sigma] + \mathbb{E}_\mathbb{Q}[X_\tau\mathbf{1}_{\tau<\sigma_n} \mid \mathcal{F}_\sigma] + c_{\sigma,\sigma_n\wedge\tau}(\mathbb{Q}) \\
&\ge \mathbb{E}_\mathbb{Q}[\Psi_{\sigma_n}(X)\mathbf{1}_{\tau\ge\sigma_n} \mid \mathcal{F}_\sigma] + \mathbb{E}_\mathbb{Q}[X_\tau\mathbf{1}_{\tau<\sigma_n} \mid \mathcal{F}_\sigma] + c_{\sigma,\sigma_n\wedge\tau}(\mathbb{Q}).
\end{aligned}$$

Since $\mathbf{1}_{\tau\geq\sigma_n} \to 1$ (because $\tau > \sigma$ on $\{\sigma < T\}$), the dominated convergence theorem for conditional expectations shows that the first term converges to η. The second and the third converge to 0. Since $\varepsilon > 0$ was arbitrary, this yields $\Psi_\sigma(X) \geq \eta$.

We have shown that for decreasing sequences of stopping times we always have $\Psi_{\sigma_n}(X) \to \Psi_\sigma(X)$. As in chapter xxx, this is enough to show that there is a càdlàg version for the "process" $\Psi(X)$. □

20.2 The Relation with Bellman's Principle

In this section we prove that time consistency is equivalent to the validity of Bellman's principle. The proof is almost the same as in [5] or in [41]. Especially in the case of Markov processes such a result can be of great importance. We also suppose that $\mathcal{F}_0$ is trivial. For a bounded process X and a stopping time σ we defined

$$\Psi_\sigma(X) = \text{ess.inf}\{u_\sigma(X_\tau) \mid \tau \geq \sigma\}.$$

Of course $\sigma, \tau, \ldots$ denote stopping times. We also recall that if σ is a stopping time, the process ${}^\sigma X$ is defined as ${}^\sigma X_s = 0$ if $s \leq \sigma$ and ${}^\sigma X_s = X_s - X_\sigma$ if $s \geq \sigma$. The process X^σ is defined as $X^\sigma_s = X_s$ if $s \leq \sigma$ and $X^\sigma_s = X_\sigma$ if $s \geq \sigma$.

Theorem 86 *Suppose that u_0 is a relevant monetary Fatou utility function defined on L^∞. In case the time interval is closed from the right, say $[0, T]$, with $0 \leq T < +\infty$, the following two properties are equivalent*

1. *u_0 is time consistent,*

2. *(Bellman's principle) For every bounded càdlàg adapted process X and every finite stopping time $\tau \leq T$, we have that*

$$\Psi_0(X) = \Psi_0(X^\tau + \Psi_\tau({}^\tau X)\mathbf{1}_{[\![\tau,T]\!]}).$$

Proof. We first show that Bellman's principle implies time consistency. For $\xi \in L^\infty(\mathcal{F}_T)$ we introduce the process X defined as $X_u = \|f\|_\infty$ for $u < T$ and $X_u = \xi$ for $u \geq T$. The value $\Psi_\tau(X)$ then coincides with the value $u_\tau(\xi)$ and the Bellman principle gives the recursivity for u.

Conversely let us show the Bellman principle. For a given stopping time τ we have:

$$\begin{aligned}\Psi_0(X) &= \inf_\sigma u_0(X_\sigma)\\ &= \inf_\sigma u_0(X_\sigma \mathbf{1}_{\sigma\geq\tau} + X_\sigma \mathbf{1}_{\sigma<\tau})\\ &= \inf_\sigma u_0\left(u_\tau(X_\sigma)\mathbf{1}_{\sigma\geq\tau} + X_\sigma \mathbf{1}_{\sigma<\tau}\right)\\ &= \inf_\sigma \inf_{\nu\geq\tau} u_0\left(u_\tau(X_\nu)\mathbf{1}_{\sigma\geq\tau} + X_\sigma \mathbf{1}_{\sigma<\tau}\right)\\ &= \inf_\sigma u_0\left(\Psi_\tau(X)\mathbf{1}_{\sigma\geq\tau} + X_\sigma \mathbf{1}_{\sigma<\tau}\right)\\ &= \inf_\sigma u_0\left(\left(\Psi({}^\tau X) + X_\tau\right)\mathbf{1}_{\sigma\geq\tau} + X_\sigma \mathbf{1}_{\sigma<\tau}\right)\\ &= \Psi_0(\Psi_\tau({}^\tau X)\mathbf{1}_{[\![\tau,T]\!]} + X^\tau)\end{aligned}$$

We have used that $\{u_\tau(X_\nu) \mid \nu \geq \tau\}$ form a lattice and that u_0 is Fatou. The other steps are left as an exercise for the reader. □

Remark 103 The time consistency always implies the Bellman principle, even if the time interval is not closed. The trick to replace a random variable by this special process is strange and one has the feeling that using more "intelligent" methods one should be able to avoid this gimmick. The following section shows that this is not the case

20.3 Counter-example

In this section we gave an example where the Bellman principle is valid, the measure is coherent but the utility function is not time consistent. In the case of coherent utility functions, time consistency is equivalent to the stability of the scenario set $\mathcal{S}$. Whether the Bellman principle implies the stability property is a much more delicate problem. We will give two answers. In case the set $\mathcal{S}$ is weakly compact in L^1, the answer is yes. Afterwards we will give a counter-example in the case where $\mathcal{S}$ is not weakly compact.

Proposition 82 *Suppose that the time interval is $\mathbb{R}_+$, suppose that the Bellman principle holds and suppose that the set $\mathcal{S}$ is weakly compact in L^1, then the set $\mathcal{S}$ is m–stable.*

Proof. We will adapt the proof of theorem 86 above. The idea is to show that $u_0(\xi) = u_0(u_\tau(\xi))$ for every finite stopping time τ and for every bounded function ξ that is $\mathcal{F}_\infty-$measurable. Since $\mathcal{S}$ is weakly compact the set

$$\{Z_\sigma \mid \sigma \text{ a finite stopping time, } Z \in \mathcal{S}\}$$

is still relatively weakly compact. If we replace ξ by the sequence $\xi_n = \mathbb{E}_\mathbb{P}[\xi \mid \mathcal{F}_n]$ then weak-compactness implies that *uniformly* for $\mathbb{Q} \in \mathcal{S}$, ξ_n approximates ξ in $L^1(\mathbb{Q})$. It follows that $u_0(\xi_n), u_\tau(\xi_n), u_0(u_\tau(\xi_n))$ tend to $u_0(\xi)$,$u_\tau(\xi)$, $u_0(u_\tau(\xi))$. It is therefore sufficient to prove the statement for functions that are $\mathcal{F}_n-$measurable. This is done exactly in the same way as in the proof of the theorem. □

It is clear that a counter-example will have to use the fact that the set $\mathcal{S}$ is big. For notational ease we will work on the time interval $[0,1[$. This is equivalent to the time interval $\mathbb{R}_+$ (simply use a time transform such as $u = t/(t+1)$). The use of the time interval $[0,1[$ allows us to use a Brownian Motion W defined for all times $t < \infty$ even if we only need the part before time 1. Finite stopping times will now be replaced by stopping times $\nu < 1$. The filtration we will use is the usual filtration coming from the process W. The set $\mathcal{S}$ is defined as

$$\{Z_1 \mid \mathbb{E}_\mathbb{P}[Z_1] = 1, Z_1 \geq 0, \mathbb{E}_\mathbb{P}[Z_1 \text{ sign}(W_1)] = 0\}\,.$$

It is clear that this set is not m–stable. This can be seen using the definition of m–stability but it will also follow from the results below. We first give the sequence of lemma's used to prove the Bellman principle and then we will give the details of the proofs of these lemma's. Since the Bellman principle will be valid, Ψ_0 will in fact be equivalent to the risk adjusted value

$$\Psi_0(X) = \text{ess.inf}\{\inf_{0\leq t<1} X_t\}.$$

Hence we cannot have m-stability. Indeed $u_0(\xi) = 0$ for $\xi = \text{sign}(W_1)$.

Lemma 50 *Let $\nu < 1$ be a stopping time. The set*

$$\{Z_\nu \mid Z \in \mathcal{S}\}$$

is dense in the set of all $\mathcal{F}_\nu$ measurable densities of probabilities absolutely continuous with respect to $\mathbb{P}$.

Lemma 51 *Bellman's principle is valid.*

Lemma 52 *Let $\mathcal{Q}$ be the set of all density processes Z such that*

1. $Z_1 = \mathcal{E}(q \cdot W)_1 > 0$, $\mathbb{E}_{\mathbb{P}}[Z_1] = 1$
2. $\int_0^1 q_u \, du = 0$ *a.s.* .

We then have that $\mathcal{Q} \subset \mathcal{S}$. *For* $\tau < 1$ *a stopping time, the set*

$$\{Z_\tau \mid Z \in \mathcal{Q}\}$$

is dense in the set of all probability densities on the $\sigma-$*algebra* $\mathcal{F}_\tau$.

Lemma 53 *Let* $\nu < 1$ *be a stopping time and let* q *be a predictable process, defined on* $[0,1] \times \Omega$ *so that*

1. $q_u = 0$ *for* $u \leq \nu$
2. q *is measurable for the* $\sigma-$*algebra* $\mathcal{R} \times \mathcal{F}_\nu$ *where* $\mathcal{R}$ *is the Borel* $\sigma-$*algebra on* $[0,1]$,
3. *a.s.* $\int_\nu^1 q_u^2 \, du < \infty$,

then $\mathbb{E}_{\mathbb{P}}[\mathcal{E}(q \cdot W)_1] = 1$ *and therefore* $\mathcal{E}(q \cdot W)_1$ *is the density of a probability measure, equivalent to* $\mathbb{P}$. *Moreover we have*

$$\mathbb{E}_{\mathbb{P}}[\mathcal{E}(q \cdot W)_1 \mid \mathcal{F}_\nu] = 1.$$

Proof of Lemma 53 This is almost trivial. Seen from time ν the process q is deterministic. Here are the details. For each n we put

$$A_n = \left\{ \int_\nu^1 q_u^2 \, du \leq n \right\}.$$

Clearly $A_n \in \mathcal{F}_\nu$ and the stochastic exponential $\mathcal{E}(\mathbf{1}_{A_n} q \cdot W)$ satisfies Novikov's condition. Therefore we have

$$\mathbb{E}_{\mathbb{P}}[\mathbf{1}_{A_n} \mathcal{E}(q \cdot W)_1] = \mathbb{E}_{\mathbb{P}}[\mathbf{1}_{A_n} \mathcal{E}(\mathbf{1}_{A_n} q \cdot W)_1] = \mathbb{P}[A_n].$$

We now apply Beppo Levi's theorem to conclude that $\mathbb{E}_{\mathbb{P}}[\mathcal{E}(q \cdot W)_1] = 1$ as desired. The statement on the conditional expectation follows from the fact that since $\mathbb{E}_{\mathbb{P}}[\mathcal{E}(q \cdot W)_1] = 1$, $\mathcal{E}(q \cdot W)$ must be a uniformly integrable martingale. □

Proof of Lemma 52 and 50 Let Z_τ be the density of a probability measure equivalent to $\mathbb{P}$ on $\mathcal{F}_\tau$. The process Z is supposed to be defined up to time τ. We will now extend it in such a way that it defines an element $Z \in \mathcal{Q}$. The process Z is a stochastic exponential and therefore Z_τ can be written as $Z_\tau = \mathcal{E}(q \cdot W)_\tau$. The predictable process q is defined up to time τ. Since $Z_\tau > 0$ we must have that $\int_0^\tau q_u^2\, du < \infty$ and therefore we also have that $r = \int_0^\tau q_u\, du$ is defined. If we now put for $u > \tau$

$$q_u = \frac{-r}{1-\tau}$$

we have that $q\mathbf{1}_{]\!]\tau,1[\![}$ satisfies the assumptions of lemma 53. We therefore have that

$$\mathbb{E}_{\mathbb{P}}\left[\mathcal{E}(q \cdot W)_1\right] = \mathbb{E}_{\mathbb{P}}\left[\mathbb{E}_{\mathbb{P}}\left[\mathcal{E}(q \cdot W)_1 \mid \mathcal{F}_\tau\right]\right] = \mathbb{E}_{\mathbb{P}}[1] = 1.$$

Moreover $\int_0^1 q_u\, du = \int_0^\tau q_u\, du + \int_\tau^1 q_u\, du = r + (-r) = 0$. Also, we have that $\int_0^1 q_u^2\, du = \int_0^\tau q_u^2\, du + \int_\tau^1 q_u^2\, du = \int_0^\tau q_u^2\, du + r^2/(1-\tau) < \infty$. Therefore $Z_1 > 0$ and $Z \in \mathcal{Q}$. This proves the density part of the lemma. We now prove that $\mathcal{Q} \subset \mathcal{S}$. For an element $\mathbb{Q} \in \mathcal{Q}$ we have that W is a Brownian motion with drift $q_u\, du$. Therefore the variable W_t is, under the measure $\mathbb{Q}$, equal to a gaussian random variable $+ \int_0^t q_u\, du$. For $t = 1$ this simply means that under $\mathbb{Q}$, the random variable W_1 is still a symmetric gaussian random variable with $L^2(\mathbb{Q})$ norm 1. In particular we have that $\mathbb{E}_{\mathbb{Q}}\left[\text{sign}(W_1)\right] = 0$, i.e. $\mathbb{Q} \in \mathcal{S}$. Lemma 50 immediately follows from Lemma 52. □

Proof of lemma 51 Let us suppose that X is càdlàg , bounded adapted. Furthermore let us fix a stopping time $\nu < 1$. It is clear that

$$\Psi_\nu(X) = X_\nu + \Psi_\nu({}^\nu X).$$

So we have to calculate $\Psi_\nu({}^\nu X)$. By definition we have

$$\Psi_\nu({}^\nu X) = \text{ess.inf}_{\nu \le \sigma < 1}\, \text{ess.inf}_{\mathbb{Q} \in \mathcal{S}^e} \left\{\mathbb{E}_{\mathbb{Q}}[{}^\nu X_\sigma \mid \mathcal{F}_\nu]\right\}.$$

Because of lemma 8.3 this can also be written as

$$\Psi_\nu({}^\nu X) = \text{ess.inf}_{\nu \le \sigma < 1}\, \text{ess.inf}_{\mathbb{Q} \sim \mathbb{P}} \left\{\mathbb{E}_{\mathbb{Q}}[{}^\nu X_\sigma \mid \mathcal{F}_\nu]\right\}.$$

Indeed the set

$$\{Z_\sigma \mid Z \in \mathcal{S}^e\}$$

is dense in the set

$$\{Z_\sigma \mid Z \text{ a nonnegative uniformly integrable martingale with } \mathbb{E}_{\mathbb{P}}[Z_1] = 1\}.$$

This means that the $\Psi-$operator is the same when calculated with the set $\mathcal{S}$ as with the set of all probability measures that are absolutely continuous with respect to $\mathbb{P}$. The latter set is stable and therefore the $\Psi-$operator satisfies Bellman's inequality. □

We end this analysis with the following

Corollary 23 *The m–stable hull of the set $\mathcal{Q}$ is the set of all probability measures that are absolutely continuous with respect to $\mathbb{P}$.*

Remark 104 That the set $\mathcal{S}$ is not m–stable can also be seen from the criteria in section xxx. The calculations are of course similar than the ones above but it might be of pedagogical interest to give the details. Let us have a look at the variable $\xi = \mathrm{sign}(W_1)$. Because of the definition of $\mathcal{S}$, we have that $\xi \in \mathcal{A}$. Let τ be a stopping time $0 \le \tau < 1$. We will show that $u_\tau(\xi) \notin \mathcal{A}_\tau$. According to theorem xxx, this is a contradiction to the m–stability of $\mathcal{S}$. The set $\mathcal{A}_\tau$ is the set of all $\mathcal{F}_\tau-$measurable elements η such that for all $Z \in \mathcal{S}$ we have $\mathbb{E}_\mathbb{P}[\eta Z] = \mathbb{E}_\mathbb{P}[\eta Z_\tau] \ge 0$. But lemma 50 then implies that necessarily $\eta \ge 0$. Hence $\mathcal{A}_\tau = L^\infty_+(\mathcal{F}_\tau)$. Now let us calculate $u_\tau(\xi)$. To do so, let us define the function $p : \mathbb{R} \to [-1, +1]$ by the relation $p(W_\tau) = \mathbb{E}_\mathbb{P}[\xi \mid \mathcal{F}_\tau]$. Take now $\frac{d\mathbb{Q}}{d\mathbb{P}} = Z_1 = \mathcal{E}(q \cdot W)_1 \in \mathcal{S}$, implying that W_1 has the same distribution under $\mathbb{P}$ as under $\mathbb{Q}$. Clearly we have that $W_t - \int_0^t q_u\, du$ is a $\mathbb{Q}-$Brownian motion and therefore $\mathbb{E}_\mathbb{Q}[\xi \mid \mathcal{F}_\tau] = p\left(W_\tau - \int_0^\tau q_u\, du\right)$. Therefore $u_\tau(\xi) \le \inf_n p\left(W_\tau - \int_0^\tau n\, du\right) = p\left(W_\tau - n\tau\right) = -1$. This is sufficient to guarantee that $u_\tau(\xi) \notin \mathcal{A}_\tau$.

Bibliography

[1] Credit Exposure, Diversification Risk and Coherent VaR, *Working Paper, University of Toronto, Department of Mathematics*, (1997)

[2] Angelsberg, G., Delbaen, F., Kaelin, I., Kupper, M. and Nf, J.: On a class of law invariant convex risk measures, *Fin. and Stoch.*, **XV**, 343–363,(2011)

[3] Artzner, Ph., F. Delbaen, J.-M. Eber, and D. Heath: Thinking Coherently, *RISK*, **November 97**, 68–71, (1997)

[4] Artzner, Ph., F. Delbaen, J.-M. Eber, and D. Heath: Characterisation of Coherent Risk Measures, *Mathematical Finance* **9**, 145–175, (1999)

[5] Artzner, P; Delbaen, F; Eber, J-M; Heath, D and Ku, H: Coherent Multiperiod Risk Adjusted Values and Bellman's Principle, *Ann. Oper. Res.*, **152**, 5–22, (2007),
preliminary version: http://www.math.ethz.ch/˜delbaen, (2003)

[6] Artzner, P; Delbaen, F; Koch-Medina, P.: xxxxx

[7] Amarante, M., Maccheroni, F., Marinacci, M., Montrucchio, L.: Cores of non-atomic market games, *Internat. J. Game Theory*, **34**, no. 3, 399424. (2006)

[8] Amir, D., Lindenstrauss, J.: The structure of weakly compact sets in Banach spaces, *Ann. of Math.* **88**, 35–46, (1968)

[9] Artzner, P. and Ostroy: Gradients, subgradients and economic equilibria, *Adv. in Appl. Math.*, **4**, 1984, pp. 245–259

[10] Artstein, Z.: Set-valued measures, *Trans. Amer. Math. Soc.* **165**, 103–125, (1972)

[11] Acerbi, C., Tasche, D.:Expected Shortfall: a natural coherent alternative to Value at Risk. *Wilmott (to appear)*, (2001) http://www-m4.mathematik.tu-muenchen.de/m4/pers/tasche/ (for the abstract)

[12] Aubin, J.P.: Cœur et équilibres des jeux flous à paiements latéraux, *C.R.Acad.Sci Paris Sér A*, 279, 891–894, (1974)

[13] Aumann, R.J. Integrals of Set Valued Functions, *Journ. Math. Anal. and Appl.*, **12** 1–12, (1965)

[14] Aumann, R.J., Serrano, R.: An Economic Index of Riskiness, *Journal of Political Economy*, 810–836, (2008)

[15] Banach, S.: Théorie des Opérations Linéaires, Monografje Matematyczne, Warzawa, Tom 1, (1932).

[16] Barrieu, P., El Karoui, N.: Inf convolution and optimal risk transfer *Finance Stoch*, **9**, 269–298, (2005)

[17] Bernegger, S., Patrick, G and Rüegg, S.: (1999), Internal Note of Swiss Re.

[18] Biagini, S., Frittelli, M.: On the extension of the Namioka-Klee theorem and the Fatou property, published in *Optimality and Risk, Papers published in honour of Kabanov's 60th birthday, editors Delbaen, F., Rasonyi, M and Stricker, C.*, Springer, Berlin, Heidelberg, New York (2009)

[19] Billera, L.J. and Heath, D.C.: Allocation of costs: a set of axioms yielding a unique procedure, *Mathematics of Operations Research* **1**, 1982, pp. 32–39

[20] Billingsley, P.: Probability and Measure, 1995, 3rd edition, *John Wiley and Sons*, New York

[21] Bion-Nadal, J.: Time Consistent Dynamic Risk Processes, *Stochastic Processes and their Applications* **119**, 633–654, (2009)

[22] Bonti, G., Kalkbrener, M., Lotz, C., Stahl, G.: Credit risk concentrations under stress, *Journal of Credit Risk*, **2** pp. 115–136, (2006)

[23] Briand, P., Coquet, F., Hu, Y., Mémin J., Peng, S.: A converse comparison theorem for BSDE's and related properties of $g-$expectation, *Elect. Comm. in Probab.*, **5**, 101–117,(2000)

[24] Brning, M., Denneberg, D.: The extreme points of the set of belief measures, *International Journal of Approximate Reasoning*, **48**, 670-675, (2008)

[25] Bhlmann, H.: Mathematical Methods in Risk Theory, Springer-Verlag, Berlin-Heidelberg. (1970)

[26] Carlier, G., Dana, R. A.,: Core of convex distortions of a probability, *J. Econom. Theory*, **113**, no. 2, 199222, (2003)

[27] Chateauneuf, A.: On the Use of Capacities in Modeling Uncertainty Aversion and Risk Aversion, *Journ. Math.Econ*, **20**, 343–369, (1991)

[28] Chateauneuf, A., Wakker, P.: An axiomatization of cumulative prospect theory for decision under risk, *Journal of Risk and Uncertainty* **18**, 137-145, (1999)

[29] Cheridito, P., Delbaen, F., Kupper, M.: Coherent and convex monetary risk measures for càdlàg processes, *Stoch. Proc. Appl*, **112**, 1–22, (2004)

[30] Cheridito, P., Horst, U., Kupper, M., Pirvu, T.: Equilibrium Pricing in Incomplete Markets under Translation Invariant Preferences, *preprint Humboldt University Berlin* (2009)

[31] Cheridito, P., Li,T.: Dual characterisation of properties of risk measures on Orlicz Hearts, *Math. Financ. Econ*, **2**, 29–55, (2008)

[32] Cheridito, P., Li,T.: Risk measures on Orlicz Hearts, *Math. Financ.*, **19**, 189–214, (2009)

[33] Choquet, G.: Theory of capacities, *Ann. Inst. Fourier*, **5**,(1953), pp. 131–295.

[34] Coquet, F., Hu, Y., Mémin J., Peng, S. Filtration consistent nonlinear expectations and related $g-$expectations, *Probability Theory and Related Fields*, **123**, 1–27, (2002)

[35] Christensen, J.P.R.: Topology and Borel structure:descriptive topology and set theory with applications to functional analysis and measure theory,1974, *North Holland*, Amsterdam

[36] Debreu, G. and Schmeidler, D.: The Radon-Nikodým derivative of a correspondence. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability Vol. II: Probability theory*, 41–56, (1972)

[37] Delbaen, F.: Convex Games and Extreme Points, *Journ. Math. Anal. Appli.*, **45**,(1974), pp. 210–233.

[38] Delbaen,F.: Representing Martingale Measures when Asset Prices are Continuous and Bounded. *Mathematical Finance*,**2**, 107–130, (1992)

[39] Delbaen, F.: Coherent Risk Measures on General Probability Spaces *in Advances in Finance and Stochastics*, pp. 137, Springer, Berlin (2002)

[40] Delbaen, F: Coherent Risk Measures, *Lectures given at the Cattedra Galileiana at the Scuola Normale di Pisa, March 2000*, Published by the *Scuola Normale di Pisa*, (2002)

[41] Delbaen, F: The structure of m-stable sets and in particular of the set of riskneutral measures. *In: Yor, M., Emery, M. (eds.): In Memoriam Paul-Andre Meyer - Seminaire de Probabilites XXXIX.* Berlin Heidelberg New York: Springer, 215258. (2006).

[42] Delbaen, F.: Hedging bounded claims with bounded outcomes *Adv. in Math. Econ.* 2006, **6**, 75–86

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

[44] Delbaen, F. and Schachermayer, W.: The Mathematics of Arbitrage, 2006, *Springer*, Berlin

[45] Dellacherie, C., Meyer, P.-A.: Probabilités et potentiel, Chapitres I à IV, Chapitres V à VIII, 1980, *Hermann, Paris*

[46] Detlefsen, K., Scandalo, G.: Conditional and Dynamic Convex Risk Measures, *Finance and Stochastics*, **9**, 539–561, (2005)

[47] Denault, M.: Coherent Allocation of Risk Capital, *Journal of Risk*, (1999)

[48] Denneberg, D.: Verzerrte Wahrscheinlichkeiten in der Versicherungsmathematik, quantilabhängige Prämienprinzipen.*Mathematik-Arbeitspapiere*, **34**, 1989, Universität Bremen

[49] Denneberg, D.: Non-Additive Measure and Integral, (1997), *Kluwer*, Dordrecht

[50] Diestel, J., Ruess, W., Schachermayer, W.: On weak compactness in $L^1(X)$. *Proc. Am. Math. Soc.*, **118**, 447–453. (1993)

[51] Deprez, O. and Gerber, H.U: On convex principles of premium calculation *Insurance: Mathematics and Economics*, 1985, pp. 179–189

[52] Diestel, J.: Geometry of Banach Spaces-Selected Topics, Lecture Notes in Mathematics, **485**, (1975), , *Springer Verlag*, Berlin Heidelberg New York

[53] Stochastic Differential Utility, *Econometrica*, **60**, 353–394, (1992)

[54] Continuous-Time Security Pricing: A Utility Gradient Approach, *Graduate School of Business, Research Paper 1096*, (1990)

[55] Dunford, N. and Schwartz, J.: Linear Operators, Vol I, (1958), New York, Interscience.

[56] Edmonds, J. Submodular functions, matroids, and certain polyhedra. *Combinatorial Structures and their Applications (Proc. Calgary Internat. Conf., Calgary, Alta., 1969)* 6987, Gordon and Breach, New York, (1970)

[57] Ekeland, I.: Nonconvex minimisation problem. *Bull. Amer. Math. Soc.*,**1**, 443–474, (1979)

[58] El Karoui, N.: Les aspects probabilistes du contrôle stochastique *Ecole d'Eté de Probabilités de Saint-Flour IX, 1979, Lecture Notes in Mathematics 876, Springer, Berlin, 1981*, 73–238

[59] El Karoui, N.: Backward Stochastic Differential Equations: a general introduction in *N. El Karoui and L. Mazliak eds Pitman Res. Notes Math. Ser.*, **364**, 7–26, (1997)

[60] El Karoui, N., Quenez, M.C.: Nonlinear Pricing Theory and backward stochastic differential equations *Financial Mathematics , Lecture Notes in Mathematics, 1656*, Springer Verlag, 191–246, (1996)

[61] El Karoui, N., Peng, S., Quenez, M.C.: Backward Stochastic Differential Equations in Finance, *Mathematical Finance*, **7**, 1–71, (1997)

[62] The Global Stability of Efficient Intertemporal Allocations, *Econometrica*, **55**, 329–358, (1987)

[63] Epstein, L., Schneider, M.: Recursive Multiple Priors, *Journ. Econ. Theory*, **113**, 1–13, (2003)

[64] Epstein, L., Zin, S.: Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, *Econometrica*, **57**, 937–969, (1989)

[65] Filipovic, D., Kupper, M., Vogelpoth, N.: Separation and Duality in Locally L0-Convex Modules, *Journal of Functional Analysis*, **256**, 3996–4029, (2009)

[66] Fischer, T. Coherent Risk Measures depending on Higher Moments, Ruprecht-Karls-Universität Heidelberg, Institut für International Vergleichende Wirstschafts- und Sozialstatistik (preprint), (2001).

[67] Convex Risk Measures and the Dynamics of their Penalty Functions. *Statistics and Decisions*, **24**, 61–96, (2006)

[68] Föllmer, H. and Schied, A.: Convex Measures of Risk and Trading Constraints, *Finance Stoch.* **6** 429447, (2002)

[69] Föllmer, H.; Schied, A.: Robust preferences and convex measures of risk. *Advances in finance and stochastics*, 3956, Springer, Berlin, (2002).

[70] Föllmer, H.; Kramkov, D. O.: Optional decompositions under constraints, *Probab. Theory Related Fields*, **109**, 1–25, (1997)

[71] Frittelli, M,;Rosazza-Gianin, Em.: Dynamic Cinvex Risk Measures *Risk Measures for the 21st Century*, 227–248, Wiley and Sons (2004)

[72] Frittelli, M,;Rosazza-Gianin, Em.: Law invariant convex risk measures *Adv.Math.Econ*,**7**, 33–46 (2005)

[73] Gerber, H.U.: An Introdution to Mathematical Risk Theory. Huebner Foundation for Insurance Education, Wharton School, University of Pennsylvania, Philadelphia.(1979)

[74] Grothendieck, A.: Topological Vector Spaces, 1973, New York, *Gordon and Breach*. Translation of *Espaces vectoriels topologiques, 1954, Sociedade de Matemática de S. Paulo*

[75] Gilboa, I., Schmeidler, D.: Maxmin expected utility with non-unique prior. *J. Math. Econom.*, **18**, 141–153 (1989).

[76] Gilboa, I., Schmeidler, D.: Canonical Representation of Set Functions. *Mathematics of Operations Research* **20**, 197-212 (1995)

[77] Gusso, R., Schmock, U., Seiler, D.: Credit risk modelling with Dirichlet mixture models, *Working Paper, not yet published*

[78] Hardy, G., Littlewood, J.E., Polya, G.: Inequalities. 2nd ed., Cambridge University press, Cambridge (1952).

[79] Halmos, P. and Savage, L.J.: Application of the Radon-Nikodym theorem to the theory of sufficient statistics, 1949, *Ann. Math. Stat.* **20**, 225–241

[80] Huber, P.: Robust Statistics, Wiley, London, (2004)

[81] Huber, P. and Strassen, V.: Minimax Test and the Neyman-Pearson lemma for Capacities. *Ann. Statis.*, **1**, 251–263 (1973); Correction in *Ann. Statist.* **2**, 223–224. (1974)

[82] Jaschke, S.: Quantile-VaR is the Wrong Measure to Quantify Risk for Regulatory Purposes, (2001),*check ref* http://wws.mathematik.hu-berlin.de/ jaschke/papers/ (preliminary version).

[83] Jaschke, S and Küchler, U.: Coherent Risk Measures, Valuation Bounds, and (μ, σ) Portfolio Optimisation, *check ref*, (1999), http://wws.mathematik.hu-berlin.de/ jaschke/papers/ (preliminary version).

[84] Jouini, E., Schachermayer, W., Touzi, N.: Law invariant risk measures have the Fatou property. *Advances in Mathematical Economics*,**9**, 4971, Springer, Tokyo.(2006) [

[85] Kabanov, Y: On an existence of the optimal solution in a control problem for a counting process, *Mat. Sbornik*, **119**, 431–445, (1982)

[86] Kalkbrener, M. An Axiomatic Approach to Capital Allocation, *Mathematical Finance*, **15**, 425–437, (2005)

[87] Kalkbrener, M., Lotter, H., Overbeck, L.: Sensible and efficient capital allocation for credit portfolios. *RISK*, **Jan. 2004**, 19–24.

[88] König, H.:Über das von Neumannsche Minimax Theorem. *Acta. Math. Hungar.*, **63**, 371–374 (1994)

[89] Koopmans, T. C.: Stationary ordinal utility and impatience. *Econometrica*, **28**, 287309 (1960)

[90] Koopmans, T. C., Diamond, P. A., Williamson, R. E.: Stationary utility and time perspective. *Econometrica*,**32**, 82100, (1964)

[91] Koopmans, T.C.: Concepts of optimality and their uses. *Nobel Memorial Lecture 11 December 1975* in *Math. Programming*,**11**, no. 1, 212228 (1976)

[92] Kramkov, D. O.: Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets, *Probab. Theory Related Fields*, **105**, 459–479, (1996)

[93] Krasnoselskii, M.,A. and Ruticki, J., Convex functionals and Orlicz spaces Noordhoff, Groningen, (1961)

[94] Kupper, M.: Schachermayer, W.: Representation Results for Law Invariant Time Consistent Functions. *Finance and Stochastics* (2009)

[95] Kupper, M.: Ph.D. Thesis, ETH-Zürich, April 2005, *check for website version.*

[96] Kusuoka, S.: On Law Invariant Coherent Risk Measures, *Advances in Mathematical Economics*, **3**, 83–95, (2001)

[97] Kusuoka, S. and Morimoto, J.: On Dynamic Law Invariant Coherent Risk Values, *Working Paper, Tokyo University*, (2004)

[98] Ky Fan: Minimax Theorems *Proc. Nat.Acad. Sci. U.S.A.*, **39**, 42–47, (1953)

[99] Lenglart, E.: Transformation des martingales locales par changement absolument continu de probabilits. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **39** , no. 1, 65–70, (1977)

[100] Lindenstrauss, J.: A short proof of Liapounoff's convexity theorem. *J. Math. Mech.* **15**, 971972, (1966)

[101] Leippold, M., Trojani, F. and Vanini, P.: Equilibrium Impact of Value-at-Risk regulation, *Journal of Economic Dynamics and Control*, **30**, 1277–1313, (2006)

[102] Maccheroni, F., Marinacci, M, Rusticini, A.: "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences,", *Econometrica*, **74**, 1447–1498, (2006)

[103] Dynamic Variational Preferences, *Journal of Economic Theory*, **128**, 4–44, (2006).

[104] Machina, M.J., Schmeidler, D.: A More Robust Definition of Subjective Probability, *Econometrica*, **60**, 745–780, (1992)

[105] Mertens, J.-F.: Processus stochastiques généraux et surmartingales, *Zeitschr für Wahrscheinlichkeitsrechnung und verwandte Gebiete*, **22**, 45–68, (1972)

[106] Neslehova, J., Embrechts, P., Chavez-Demoulin, V.: Infinite mean models and the LDA for operational risk *Journal of Operational Risk*, **1**, 3–25, (2006)

[107] Nikodym, O.: Contribution à la théorie des fonctionnelles linéaires en connection avec la théorie de la mesure des ensembles abstraits, *Mathematica*, 1931, **5**, 130–141

[108] Parker, J.: The Sigma-Core of a Cooperative Game *Manus. Math.* **70**, 247–253, (1991)

[109] Peng, S.: Backward SDE and related $g-$expectations, *in N. El Karoui and L. Mazliak eds Pitman Res. Notes Math. Ser.*, **364**, 141–159, (1997)

[110] Peng, S.: xxxxxx Backward SDE and related $g-$expectations, LNMA notes

[111] Phelps, R. R.: Lectures on Choquet's theorem. *D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London* (1966), *Second edition: Lecture Notes in Mathematics,* **1757***, Springer-Verlag, Berlin* (2001)

[112] Phelps, R.,R.: Convex Functions, Monotone Operators and Differentiability, 2nd Edition, Lecture Notes in Mathematics,**1364**, (1993), Berlin Heidelberg New York, Springer Verlag.

[113] Protter, P.: Stochastic Integration and Differential Equations: a new approach, Springer-Verlag, Berlin, 1990

[114] Riedel, F.: Dynamic Coherent Risk Measures, *Stoch. Proc. Appl.*, **112**, 185-200, (2004)

[115] Roorda, B., Schumacher, J.M., Engwerda, J.: Coherent Acceptability Measures in Multiperiod Models, *Math. Fin.*, **15**, 589–612, (2005)

[116] Rosazza Gianin, E.: Risk measures via g-expectations, *Insurance: Mathematics and Economics*, **39**, 19-34, 2006

[117] Rosenmüller, J.: Some properties of convex set functions. *Arch. Math.* **22**, 420430, (1971).

[118] Rockafellar, R.: Convex Analysis, *Princeton Mathematical Series*,**28**, (1970), Princeton, New Jersey, Princeton University Press.

[119] Rockafellar, R.T., Uryasev, S.: Optimization of Conditional Value-At-Risk, *The Journal of Risk*, **2**, 21-41, (2000)

[120] Rudin, W.: Functional Analysis, 1973, New York, *Mc Graw Hill*

[121] Ryff, J. V.: Extreme points of some convex subsets of $L^1(0,1)$. *Proc. Amer. Math. Soc.*, **18**, 10261034, (1967)

[122] Revuz, D. and Yor. M: Continuous Martingales and Brownian Motion, second edition, Springer-Verlag, Berlin, 1994

[123] Schmeidler, D.: Cores of Convex Games, *J. Math. Anal. Appli.*,**40**, (1972), pp. 214–225.

[124] Schmeidler, D.: Integral Representation without Additivity, *Proc. Amer. Math. Soc.*,**97**, (1986), pp. 255–261.

[125] Straumann, D., Schmock, U. (1999), Euler principle in connection with Tailvar. *Slides available under* http://www.fam.tuwien.ac.at/ schmock/slides/AllocationSlidesOct1999.pdf

[126] Shapley, L.: Notes on N-person games, VII: Cores of convex games', *Rand Memorandum, RM 4571 PR, Rand Corp., Santa Monica, Calif.* (1965)

[127] Tasche, D.: Risk contributions and performance measurement (1999), **http://www.math.ethz.ch/˜tasche/** or **http://www-m1.mathematik.tu-muenchen.de/m4/Papers/Tasche/**.*xxxx check ref*

[128] Tsukahara, H.: Some Properties of distortion risk measures, *Adv. Math. Econ.*, **12**, 153–166, (2009)

[129] Wakker, P.: Continious Subjective Expected Utility with Non-additive Probabilities, *Journ. Math. Econ.*, **18**, 1–27, (1989)

[130] Wang, S.: Premium Calculation by Transforming the Layer Premium density reference?? xxxx

[131] Wang, S.S., V.R. Young, and H.H. Panjer: Axiomatic Characterization Of Insurance Prices, *Insurance: Mathematics and Economics* 21:173–183 (1997)

[132] Wang, T . A Class of Multi-Prior Preferences, *working paper*, (2002), http://www.ccfr.org.cn/english/N&E&R/research.php

[133] Yaari, M. E.: The dual theory of choice under risk. *Econometrica*, 1987, **55**, 95115

[134] Walley

[135] Yosida, K., Hewitt, E.: Finitely Additive Measures, *Transactions of the American Mathematical Society*, **72**, 46–66, (1952)

[136] Zitkovic, G A filtered version of the Bipolar Theorem of Brannath and Schachermayer, *Journal of Theoretical Probability*, **15**, 41–61, (2002)