COHERENT RISK MEASURES ON GENERAL PROBABILITY SPACES

Freddy Delbaen, Eidgenössische Technische Hochschule, Zürich

March 10, 2000

Abstract. We extend the definition of coherent risk measures, as introduced by Artzner, Delbaen, Eber and Heath, to general probability spaces and we show how to define such measures on the space of all random variables. We also give examples that relates the theory of coherent risk measures to game theory and to distorted probability measures. The mathematics are based on the characterisation of closed convex sets $\mathcal{P}_\sigma$ of probability measures that satisfy the property that every random variable is integrable for at least one probability measure in the set $\mathcal{P}_\sigma$.

Key words and phrases. capital requirement, coherent risk measure, capacity theory, convex games, insurance premium principle, measure of risk, Orlicz spaces, quantile, scenario, shortfall, subadditivity, submodular functions, value at risk.

The author acknowledges financial support from Credit Suisse for his work and from Société Générale for earlier versions of this paper. Special thanks go to Artzner, Eber and Heath for the many stimulating discussions on risk measures and other topics. I also want to thank Maaß for pointing out extra references to related work. Discussions with Kabanov were more than helpful to improve the presentation of the paper. Part of the work was done during Summer 99, while the author was visiting Tokyo Institute of Technology. The views expressed are those of the author.

1. Introduction and Notation

The concept of coherent risk measures together with its axiomatic characterization was introduced in the paper [ADEH1] and further developed in [ADEH2]. Both these papers supposed that the underlying probability space was finite. The aim of this paper is twofold. First we extend the notion of coherent risk measures to arbitrary probability spaces, second we deepen the relation between coherent risk measures and the theory of cooperative games. In many occasions we will make a bridge between different existing theories. In order to keep the paper self contained, we sometimes will have to repeat known proofs. In March 2000, the author gave a series of lectures at the Cattedra Galileiana at the Scuola Normale di Pisa. The subject of these lectures was the theory of coherent risk measures as well as applications to several problems in risk management. The interested reader can consult the lecture notes [D2]. Since the original version of this paper (1997), proofs have undergone a lot of changes. Discussions with colleagues greatly contributed to the presentation. The reader will also notice that the theory of convex games plays a special role in the theory of coherent risk measures. It was Dieter Sondermann who mentioned the theory of convex games to the author and asked about continuity properties of its core, see [D1]. It is therefore a special pleasure to be able to put this paper in the Festschrift.

Throughout the paper, we will work with a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. With $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ (or $L^\infty(\mathbb{P})$ or even L^∞ if no confusion is possible), we mean the space of all equivalence classes of bounded real valued random variables. The space $L^0(\Omega, \mathcal{F}, \mathbb{P})$ (or $L^0(\mathbb{P})$ or simply L^0) denotes the space of all equivalence classes of real valued random variables. The space L^0 is equipped with the topology of convergence in probability. The space $L^\infty(\mathbb{P})$, equipped with the usual L^∞ norm, is the dual space of the space of integrable (equivalence classes of) random variables, $L^1(\Omega, \mathcal{F}, \mathbb{P})$ (also denoted by $L^1(\mathbb{P})$ or L^1 if no confusion is possible). We will identify, through the Radon–Nikodym theorem, finite measures that are absolutely continuous with respect to $\mathbb{P}$, with their densities, i.e. with functions in L^1. This may occasionally lead to expressions like $\|\mu - f\|$ where μ is a measure and $f \in L^1$. If $\mathbb{Q}$ is a probability defined on the σ-algebra $\mathcal{F}$, we will use the notation $\mathbf{E}_\mathbb{Q}$ to denote the expected value operator defined by the probability $\mathbb{Q}$. Let us also recall, see [DS] for details, that the dual of $L^\infty(\mathbb{P})$ is the Banach space $\mathbf{ba}(\Omega, \mathcal{F}, \mathbb{P})$ of all bounded, finitely additive measures μ on $(\Omega, \mathcal{F})$ with the property that $\mathbb{P}(A) = 0$ implies $\mu(A) = 0$. In case no confusion is possible we will abbreviate the notation to $\mathbf{ba}(\mathbb{P})$. A positive element $\mu \in \mathbf{ba}(\mathbb{P})$ such that $\mu(1) = 1$ is also called a finitely additive probability, an interpretation that should be used with care. To keep notation consistent with integration theory we sometimes denote the action $\mu(f)$ of $\mu \in \mathbf{ba}(\mathbb{P})$ on the bounded function f, by $\mathbf{E}_\mu[f]$. The Yosida-Hewitt theorem, see [YH], implies for each $\mu \in \mathbf{ba}(\mathbb{P})$, the existence of a uniquely defined decomposition $\mu = \mu_a + \mu_p$, where μ_a is a σ–addtive measure, absolutely continuous with respect to $\mathbb{P}$, i.e. an element of $L^1(\mathbb{P})$, and where μ_p is a purely finitely additive measure. Furthermore the results in [YH] show that there is a countable partition $(A_n)_n$ of Ω into elements of $\mathcal{F}$, such that for each n, we have that $\mu_p(A_n) = 0$.

The paper is organised as follows. In section 2 we repeat the definition of coherent risk measure and relate this definition to submodular and supermodular functionals. We will show that using bounded finitely additive measures, we get the same results as in [ADEH2]. This section is a standard application of the duality theory between L^∞ and its dual space **ba**. The main purpose of this section is to introduce the notation. In section 3 we relate several continuity properties of coherent risk measures to properties of a defining set of probability measures. This section relies heavily on the duality theory of the spaces L^1 and L^∞. Examples of coherent risk measures are given in section 4. By carefully selecting the defining set of probability measures, we give examples that are related to higher moments of the random variable. Section 5 studies the extension of a coherent risk measure, defined on the space L^∞ to the space L^0 of all random variables. This extension to L^0 poses a problem since a coherent risk measure defined on L^0 is a convex function defined on L^0. Nikodym's result on L^0, then implies that, at least for an atomless probability $\mathbb{P}$, there are no coherent risk measures that only take finite values. The solution given, is to extend the risk measures in such a way that it can take the value $+\infty$ but it cannot take the value $-\infty$. The former ($+\infty$) means that the risk is very bad and is unacceptable for the economic agent (something like a risk that cannot be insured). The latter ($-\infty$) would mean that the position is so safe that an arbitrary amount of capital could be withdrawn without endangering the company. Clearly such a situation cannot occur in any reasonable model. The main mathematical results of this section are summarised in the following theorem

Theorem. *If $\mathcal{P}_\sigma$ is a norm closed, convex set of probability measures, all absolutely continuous with respect to $\mathbb{P}$, then the following properties are equivalent:*

(1) *For each $f \in L^0_+$ we have that*

$$\lim_n \inf_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_\mathbb{Q}[f \wedge n] < +\infty$$

(2) *There is a $\gamma > 0$ such that for each A with $\mathbb{P}[A] \le \gamma$ we have*

$$\inf_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbb{Q}[A] = 0.$$

(3) *For every $f \in L^0_+$ there is $\mathbb{Q} \in \mathcal{P}_\sigma$ such that $\mathbf{E}_\mathbb{Q}[f] < \infty$.*
(4) *There is a $\delta > 0$ such that for every set A with $\mathbb{P}[A] < \delta$ we can find an element $\mathbb{Q} \in \mathcal{P}_\sigma$ such that $\mathbb{Q}[A] = 0$.*
(5) *There is a $\delta > 0$, as well as a number K such that for every set A with $\mathbb{P}[A] < \delta$ we can find an element $\mathbb{Q} \in \mathcal{P}_\sigma$ such that $\mathbb{Q}[A] = 0$ and $\|\frac{d\mathbb{Q}}{d\mathbb{P}}\|_\infty \le K$.*

In the same section 5, we also give extra examples showing that, even when the defining set of probability measures is weakly compact, the Beppo Levi type theorems do not hold for coherent risk measures. Some of the examples rely on the theory of non–reflexive Orlicz spaces. In section 6 we discuss, along the same lines as in [ADEH2], the relation with the popular concept, called Value at Risk and denoted by VaR. Section 7 is devoted to the relation between convex games, coherent risk measures and non–additive integration. We extend known results on the sigma–core of a game to cooperative games that are defined on abstract measure spaces and that do not necessarily fulfill topological regularity assumptions. This work is based on previous work of Parker, [Pa] and of the author [D1]. In section 8 we give some explicit examples that show how different risk measures can be.

2. The General Case

In this section we show that the main theorems of the papers [ADEH1] and [ADEH2] can easily be generalised to the case of general probability spaces. The only difficulty consists in replacing the finite dimensional space $\mathbb{R}^\Omega$ by the space of bounded measurable functions, $L^\infty(\mathbb{P})$. In this setting the definition of a coherent risk measure as given in [ADEH1] can be written as:

Definition 2.1. *A mapping $\rho : L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ is called a coherent risk measure if the following properties hold*

1. *If $X \geq 0$ then $\rho(X) \leq 0$.*
2. *Subadditivity: $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.*
3. *Positive homogeneity: for $\lambda \geq 0$ we have $\rho(\lambda X) = \lambda\rho(X)$.*
4. *For every constant function a we have that $\rho(a + X) = \rho(X) - a$.*

Remark. We refer to [ADEH1] and [ADEH2] for an interpretation and discussion of the above properties. Here we only remark that we are working in a model without interest rate, the general case can "easily" be reduced to this case by "discounting". Although the properties listed in the definition of a coherent risk measure have a direct interpretation in mathematical finance, it is mathematically more convenient to work with the related submodular function, ψ, or with the associated supermodular function, ϕ. The definitions we give below differ slightly from the usual ones. The changes are minor and only consist in the part related to positivity, i.e. to part one of the definitions.

Definition 2.2. *A mapping $\psi: L^\infty \to \mathbb{R}$ is called submodular if*

1. *For $X \leq 0$ we have that $\psi(X) \leq 0$.*
2. *If X and Y are bounded random variables then $\psi(X + Y) \leq \psi(X) + \psi(Y)$.*
3. *For $\lambda \geq 0$ and $X \in L^\infty$ we have $\psi(\lambda X) = \lambda\psi(X)$*

The submodular function is called translation invariant if moreover

4. *For $X \in L^\infty$ and $a \in \mathbb{R}$ we have that $\psi(X + a) = \psi(X) + a$.*

Definition 2.2'. *A mapping $\phi: L^\infty \to \mathbb{R}$ is called supermodular if*

1. *For $X \geq 0$ we have that $\phi(X) \geq 0$.*
2. *If X and Y are bounded random variables then $\phi(X + Y) \geq \phi(X) + \phi(Y)$.*
3. *For $\lambda \geq 0$ and $X \in L^\infty$ we have $\phi(\lambda X) = \lambda\phi(X)$*

The supermodular function is called translation invariant if moreover

4. *For $X \in L^\infty$ and $a \in \mathbb{R}$ we have that $\phi(X + a) = \phi(X) + a$.*

Remark. If ρ is a coherent risk measure and if we put $\psi(X) = \rho(-X)$ we get a translation invariant submodular functional. If we put $\phi(X) = -\rho(X)$, we obtain a supermodular functional. These notations and relations will be kept fixed throughout the paper.

Remark. Submodular functionals are well known and were studied by Choquet in connection with the theory of capacities , see [Ch]. They were used by many authors in different applications, see e.g. section 7 of this paper for a connection with game theory. We refer the reader to [Wa1] for the development and the application of the theory to imprecise probabilities and belief functions. These concepts are certainly not disjoint from risk management considerations. In [Wa2], P. Walley gives a

discussion of properties that may also be interesting for risk measures. In [Maa], Maaß gives an overview of existing theories.
The following properties of a translation invariant supermodular mappings ϕ, are immediate

(1) $\phi(0) = 0$ since by positive homogeneity: $\phi(0) = \phi(2 * 0) = 2\phi(0)$.
(2) If $X \leq 0$, then $\phi(X) \leq 0$. Indeed $0 = \phi(X + (-X)) \geq \phi(X) + \phi(-X)$ and if $X \leq 0$, this implies that $\phi(X) \leq -\phi(-X) \leq 0$.
(3) If $X \leq Y$ then $\phi(X) \leq \phi(Y)$. Indeed $\phi(Y) \geq \phi(X) + \phi(Y - X) \geq \phi(X)$.
(4) $\phi(a) = a$ for constants $a \in \mathbb{R}$.
(5) If $a \leq X \leq b$, then $a \leq \phi(X) \leq b$. Indeed $X - a \geq 0$ and $X - b \leq 0$.
(6) ϕ is a convex norm-continuous, even Lipschitz, function on L^∞. In other words $|\phi(X - Y)| \leq \|X - Y\|_\infty$.
(7) $\phi(X - \phi(X)) = 0$.

The following theorem is an immediate application of the bipolar theorem from functional analysis.

Theorem 2.3. *Suppose that $\rho\colon L^\infty(\mathbb{P}) \to \mathbb{R}$ is a coherent risk measure with associated sub(super)modular function ψ (ϕ). There is a convex $\sigma(\mathbf{ba}(\mathbb{P}), L^\infty(\mathbb{P}))$–closed set $\mathcal{P}_{\mathbf{ba}}$ of finitely additive probabilities, such that*

$$\psi(X) = \sup_{\mu \in \mathcal{P}_{\mathbf{ba}}} \mathbf{E}_\mu[X] \text{ and } \phi(X) = \inf_{\mu \in \mathcal{P}_{\mathbf{ba}}} \mathbf{E}_\mu[X].$$

Proof. Because $-\rho(X) = \phi(X) = -\psi(-X)$ for all $X \in L^\infty$, we only have to show one of the equalities. The set $C = \{X \mid \phi(X) \geq 0\}$ is clearly a convex and norm closed cone in the space $L^\infty(\mathbb{P})$. The polar set $C^\circ = \{\mu \mid \forall X \in C : \mathbf{E}_\mu[X] \geq 0\}$ is also a convex cone, closed for the weak* topology on $\mathbf{ba}(\mathbb{P})$. All elements in C° are positive since $L^\infty_+ \subset C$. This implies that for the set $\mathcal{P}_{\mathbf{ba}}$, defined as $\mathcal{P}_{\mathbf{ba}} = \{\mu \mid \mu \in C^\circ \text{ and } \mu(1) = 1\}$, we have that $C^\circ = \cup_{\lambda \geq 0} \lambda \mathcal{P}_{\mathbf{ba}}$. The duality theory, more precisely the bipolar theorem, then implies that $C = \{X \mid \forall \mu \in \mathcal{P}_{\mathbf{ba}} : \mathbf{E}_\mu[X] \geq 0\}$. This means that $\phi(X) \geq 0$ if and only if $E_\mu[X] \geq 0$ for all $\mu \in \mathcal{P}_{\mathbf{ba}}$. Since $\phi(X - \phi(X)) = 0$ we have that $X - \phi(X) \in C$ and hence for all μ in $\mathcal{P}_{\mathbf{ba}}$ we find that $\mathbf{E}_\mu[X - \phi(X)] \geq 0$. This can be reformulated as

$$\inf_{\mu \in \mathcal{P}_{\mathbf{ba}}} \mathbf{E}_\mu[X] \geq \phi(X).$$

Since for arbitrary $\varepsilon > 0$, we have that $\phi(X - \phi(X) - \varepsilon) < 0$, we get that $X - \phi(X) - \varepsilon \notin C$. Therefore there is a $\mu \in \mathcal{P}_{\mathbf{ba}}$ such that $\mathbf{E}_\mu[X - \phi(X) - \varepsilon] < 0$ which leads to the opposite inequality and hence to:

$$\inf_{\mu \in \mathcal{P}_{\mathbf{ba}}} \mathbf{E}_\mu[X] = \phi(X).$$

□

Remark on notation. From the proof of the previous theorem we see that there is a one–to–one correspondence between

(1) coherent risk measures ρ,
(2) the associated supermodular function $\phi(X) = -\rho(X)$,
(3) the associated submodular function $\psi(X) = \rho(-X)$,
(4) weak* closed convex sets of finitely additive probability measures $\mathcal{P}_{\mathbf{ba}} \subset \mathbf{ba}(\mathbb{P})$,
(5) $\|.\|_\infty$ closed convex cones $C \subset L^\infty$ such that $L^\infty_+ \subset C$.

The relation between C and ρ is given by

$$\rho(X) = \inf \{\alpha \mid X + \alpha \in C\}\,.$$

The set C is called the set of acceptable positions, see [ADEH2]. When we refer to any of these objects it will be according to these notations.

Remark on possible generalisations. In the paper by Jaschke and Küchler, [JaK] an abstract ordered vector space is used. Such developments have interpretations in mathematical finance and economics. In a private discussion with Kabanov it became clear that there is a way to handle transactions costs in the setting of risk measures. In order to do this, one should replace the space L^∞ of bounded real-valued random variables by the space of bounded random variables taking values in a finite dimensional space $\mathbb{R}^n$. By replacing Ω by $\{1, 2, \dots, n\} \times \Omega$, part of the present results can be translated immediately. The idea to represent transactions costs with a cone was developed by Kabanov, see [Ka].

Remark on the interpretation of the probability space. The set Ω and the σ-algebra $\mathcal{F}$ have an easy interpretation. The σ-algebra $\mathcal{F}$ for instance, describes all the events that become known at the end of an observation period. The interpretation of the probability $\mathbb{P}$ seems to be more difficult. The measure $\mathbb{P}$ describes with what probability events might occur. But in economics and finance such probabilities are subjective. Regulators of the finance industry might have a completely different view on probabilities than the financial institutions they control. Inside one institution there might be a different view between the different branches, trading tables, underwriting agents, etc.. An insurance company might have a different view than the reinsurance company and than their clients. But we may argue that the class of negligible sets and consequently the class of probability measures that are equivalent to $\mathbb{P}$ remains the same. This can be expressed by saying that only the knowledge of the events of probability zero is important. So we only need agreement on the "possibility" that events might occur, not on the actual value of the probability.
In view of this, there are two natural spaces of random variables on which we can define a risk measure. Only these two spaces remain the same when we change the underlying probability to an equivalent one. These two spaces are $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ and $L^0(\Omega, \mathcal{F}, \mathbb{P})$. The space L^0 cannot be given a norm and cannot be turned into a locally convex space. E.g. if the probability $\mathbb{P}$ is atomless, i.e. supports a random variable with a continuous cumulative distribution function, then there are no nontrivial (i.e. non identically zero) continuous linear forms on L^0, see [Ni]. The extension of coherent risk measures from L^∞ to L^0 is the subject of section 5.

3. The σ–additive Case

The previous section gave a characterisation of translation invariant submodular functionals (or equivalently coherent risk measures) in terms of finitely additive probabilities. The characterisation in terms of σ–additive probabilities requires additional hypotheses. E.g. if μ is a purely finitely additive measure, the expression $\phi(X) = \mathbf{E}_\mu[X]$ gives a translation invariant submodular functional. This functional, coming from a purely finitely additive measure cannot be described by a σ–additive probability measure. So we need extra conditions.

Definition 3.1. *The translation invariant supermodular mapping $\phi\colon L^\infty \to \mathbb{R}$ is said to satisfy the Fatou property if $\phi(X) \geq \limsup \phi(X_n)$, for any sequence, $(X_n)_{n\geq 1}$, of functions, uniformly bounded by 1 and converging to X in probability.*

Remark. Equivalently we could have said that the coherent risk measure ρ associated with the supermodular function ϕ satisfies the Fatou property if for the said sequences we have $\rho(X) \leq \liminf \rho(X_n)$.

Using similar ideas as in the proof of theorem 2.3 and using a characterisation of weak* closed convex sets in L^∞, we obtain:

Theorem 3.2. *For a translation invariant supermodular mapping ϕ, the following 4 properties are equivalent*

(1) *There is an $L^1(\mathbb{P})$-closed, convex set of probability measures $\mathcal{P}_\sigma$, all of them being absolutely continuous with respect to $\mathbb{P}$ and such that for $X \in L^\infty$:*

$$\phi(X) = \inf_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_\mathbb{Q}[X].$$

(2) *The convex cone $C = \{X \mid \phi(X) \geq 0\}$ is weak*, i.e. $\sigma(L^\infty(\mathbb{P}), L^1(\mathbb{P}))$ closed.*
(3) *ϕ satisfies the Fatou property.*
(4) *If X_n is a uniformly bounded sequence that decreases to X a.s., then $\phi(X_n)$ tends to $\phi(X)$.*

Proof. (2) $\Rightarrow$ (3) If C is weak* closed, then ϕ satisfies the Fatou property. Indeed if X_n tends to X in probability, $\|X_n\|_\infty \leq 1$ for all n and $\phi(X_n) \to a$, then $X_n - \phi(X_n) \in C$. Since C is weak* closed, the limit $X - a$ is in C as well. This shows that $\phi(X - a) \geq 0$ and this implies that $a \leq \phi(X)$. This proves the Fatou property.
(3) $\Rightarrow$ (2) If ϕ satisfies the Fatou property, then C is weak* closed. This is essentially the Krein-Šmulian theorem as used in a remark, due to Grothendieck, see [G], Supplementary Exercise 1, Chapter 5, part 4. Translated to our special case of C being a cone, this means that it is sufficient to check that $C \cap B_1$ is closed in probability (B_1 stands for the closed unit ball of L^∞). So let X_n be a sequence of elements in C, uniformly bounded by 1 and tending to X in probability. The Fatou property then shows that also $\phi(X) \geq 0$, i.e. $X \in C$.
(2) $\Rightarrow$ (1). This is not difficult and is done in exactly the same way as in theorem 2.3. But this time we define the polar C° of C in L^1 and we apply the bipolar theorem for the duality pair (L^1, L^∞). Because of C being weak* closed, this poses no problem and we define:

$$\begin{aligned} C^\circ &= \left\{ f \mid f \in L^1 \text{ and } \mathbf{E}_\mathbb{P}[fX] \geq 0 \text{ for all } X \in C \right\}, \\ \mathcal{P}_\sigma &= \{ f \mid d\mathbb{Q} = f\, d\mathbb{P} \text{ defines a probability and } f \in C^\circ \}. \end{aligned}$$

Of course we have

$$C^\circ = \cup_{\lambda\geq 0} \lambda \mathcal{P}_\sigma.$$

So we find a closed convex set of probability measures $\mathcal{P}_\sigma$ such that

$$\phi(X) = \inf \{ \mathbf{E}_\mathbb{Q}[X] \mid \mathbb{Q} \in \mathcal{P}_\sigma \}.$$

(1) $\Rightarrow$ (2), Fatou's lemma implies that every translation invariant submodular mapping, that is given by the inf over a set of probability measures, satisfies the Fatou property. Indeed for each $\mathbb{Q} \in \mathcal{P}_\sigma$ we get

$$\mathbf{E}_\mathbb{Q}[X] \geq \limsup \mathbf{E}_\mathbb{Q}[X_n] \geq \limsup \phi(X_n),$$

where X_n is a sequence, uniformly bounded by 1 and tending to X in probability. We omit the proof of the other implications. $\square$

Corollary 3.3. *With the notations and assumptions of theorem 2.3 and 3.2 we get that the set $\mathcal{P}_\sigma$ is $\sigma(\mathbf{ba}(\mathbb{P}), L^\infty(\mathbb{P}))$ dense in $\mathcal{P}_{\mathbf{ba}}$.*

Remark on notation. From the proof of the previous theorem we see that there is a one–to–one correspondence between

(1) coherent risk measures ρ having the Fatou property,
(2) closed convex sets of probability measures $\mathcal{P}_\sigma \subset L^1(\mathbb{P})$,
(3) weak* closed convex cones $C \subset L^\infty$ such that $L^\infty_+ \subset C$.

We now give an answer to the question when a coherent risk measure can be given as the supremum of expected values, taken with respect to equivalent probability measures.

Definition 3.4. *The coherent risk measure ρ is called relevant if for each set $A \in \mathcal{F}$ with $\mathbb{P}[A] > 0$ we have that $\rho(-\mathbf{1}_A) > 0$. When using the associated submodular function, this means that $\psi(\mathbf{1}_A) > 0$ or when using the associated supermodular-function: $\phi(-\mathbf{1}_A) < 0$.*

It is easily seen, using the monotonicity of ρ, that the property of being relevant is equivalent to $\rho(X) > 0$ for each nonpositive $X \in L^\infty$ such that $\mathbb{P}[X < 0] > 0$. The economic interpretation of this property is straightforward.

Theorem 3.5. *For a coherent risk measure, ρ, that satisfies the Fatou property, the following are equivalent*

(1) *ρ is relevant.*
(2) *The set $\mathcal{P}^e_\sigma = \{\mathbb{Q} \in \mathcal{P}_\sigma \mid \mathbb{Q} \sim \mathbb{P}\}$ is non empty.*
(3) *The set $\mathcal{P}^e_\sigma = \{\mathbb{Q} \in \mathcal{P}_\sigma \mid \mathbb{Q} \sim \mathbb{P}\}$ is norm (i.e. L^1 norm) dense in $\mathcal{P}_\sigma$.*
(4) *There is a set $\mathcal{P}' \subset \mathcal{P}_\sigma$ of equivalent probability measures such that*

$$\psi(X) = \sup_{\mathbb{Q}\in\mathcal{P}'} E_\mathbb{Q}[X] \text{ and similarly } \phi(X) = \inf_{\mathbb{Q}\in\mathcal{P}'} E_\mathbb{Q}[X].$$

Proof. The other implications being trivial, we only show that (1) $\Rightarrow$ (2). This is a reformulation of the Halmos-Savage theorem [HS]. For convenience of the reader, let us briefly sketch the exhaustion argument. Since the set $\mathcal{P}_\sigma$ is norm closed and convex, the class of sets

$$\left\{\left\{\frac{d\mathbb{Q}}{d\mathbb{P}} > 0\right\} \mid \mathbb{Q} \in \mathcal{P}_\sigma\right\},$$

is stable for countable unions. It follows that up to $\mathbb{P}$-null sets, there is a maximal element. Because ρ is relevant, the only possible maximal element is Ω. From this it follows that there is $\mathbb{Q} \in \mathcal{P}_\sigma$ such that $\mathbb{Q} \sim \mathbb{P}$. $\square$

The following theorem characterises the coherent risk measures that satisfy a continuity property that is stronger than the Fatou property.

Theorem 3.6. *For a translation invariant supermodular mapping, ϕ, the following properties are equivalent*

1. *The set $\mathcal{P}_\sigma$ is weakly compact in L^1.*
2. *The sets $\mathcal{P}_{\mathbf{ba}}$ and $\mathcal{P}_\sigma$ coincide.*
3. *If $(X_n)_{n\geq 1}$ is a sequence in L^∞, uniformly bounded by 1 and tending to X in probability, then $\phi(X_n)$ tends to $\phi(X)$.*
4. *If $(A_n)_n$ is a increasing sequence whose union is Ω, then $\phi(\mathbf{1}_{A_n})$ tends to 1.*

Proof. Clearly (1) $\Leftrightarrow$ (2). (1) $\Rightarrow$ (3). If (1) holds, then the set $\mathcal{P}_\sigma$ is uniformly integrable (by the Dunford-Pettis theorem, see [DS] or [G]) and it follows that $\mathbf{E}_\mathbb{Q}[X_n]$ tends to $E_\mathbb{Q}[X]$ uniformly over the set $\mathcal{P}_\sigma$. This implies that $\phi(X_n)$ tends to $\phi(X)$. Clearly (3) $\Rightarrow$ (4). To prove (4) $\Rightarrow$ (1) observe that (4) implies that the set $\mathcal{P}_\sigma$ is uniformly integrable. Indeed, if B_n is a sequence of decreasing sets such that the intersection $\cap_n B_n$ is empty, then $\sup_{\mathbb{Q}\in\mathcal{P}_\sigma}\mathbb{Q}[B_n]\leq 1-\inf_{\mathbb{Q}\in\mathcal{P}_\sigma}\mathbb{Q}[B_n^c]$ and hence tends to 0. $\mathcal{P}_\sigma$ being convex and norm closed, this together with the (easy part of the) Dunford-Pettis theorem, implies that $\mathcal{P}_\sigma$ is weakly compact. $\square$

Example 3.7. This example shows that the property "$\phi(\mathbf{1}_{A_n})$ tends to zero for every decreasing sequence of sets with empty intersection", does not imply that ρ satisfies the 4 properties of the preceding theorem. It does not even imply that ρ, or ϕ, has the Fatou property. Take $(\Omega,\mathcal{F},\mathbb{P})$ big enough to support purely finitely additive probabilities, i.e. $L^\infty(\mathbb{P})$ is supposed to be infinite dimensional. Take $\mu\in\mathbf{ba}(\mathbb{P})$, purely finitely additive, and let $\mathcal{P}_{\mathbf{ba}}$ be the segment (the convex hull) joining the two points μ and $\mathbb{P}$. Because there is a σ–additive probability in $\mathcal{P}_{\mathbf{ba}}$, it is easily seen that $\rho(\mathbf{1}_{A_n})=-\phi(\mathbf{1}_{A_n})=-\inf_{\mathbb{Q}\in\mathcal{P}_{\mathbf{ba}}}(A_n)$ tends to zero for every decreasing sequence of sets with empty intersection. But clearly the coherent measure cannot satisfy the Fatou property since $\mathcal{P}_\sigma=\{\mathbb{P}\}$ is not dense in $\mathcal{P}_{\mathbf{ba}}$. 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 [YH]), there is a countable partition of Ω into sets $(B_n)_{n\geq 1}$ such that for each n, we have $\mu(B_n)=0$. Take now X an element in L^∞ such that $\mathbf{E}_\mathbb{P}[X]=0$ and such that $\mathbf{E}_\mu[X]=-1$. This implies that $\rho(X)=1$. For the sequence X_n we take $X_n=X\,\mathbf{1}_{\cup_{k\leq n}B_k}$. The properties of the sets B_n imply that $\mu=0$ on the union $\cup_{k\leq n}B_k$ and hence we have that $\rho(X_n)=\mathbf{E}_\mathbb{P}[X_n]$, which tends to 0 as n tends to ∞.

The next proposition characterises those coherent risk measures that tend to zero on decreasing sequences of sets.

Theorem 3.8. *For a coherent risk measure, ρ, the following are equivalent*

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

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 $\mu \in \mathcal{P}_{\mathbf{ba}}$ 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 3.9. *If K is a closed, weak* compact, convex set of finitely additive, nonnegative measures, such that $\delta = \inf\{\|\nu_p\| \mid \nu \in K\} > 0$, then there exists a decreasing (nonincreasing) 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

$$\inf\{\|\mu_p\| \mid \mu \in \mathcal{P}_{\mathbf{ba}}\} > 0.$$

We can therefore apply the lemma in order to get a contradiction to (1).
The proof that (2) and (3) are equivalent is left to the reader. $\square$

In the proof of the lemma, as well as in section 5, 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 Dellacherie–Meyer, [DM], page 404.

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 u is a bilinear function $u\colon E \times F \to \mathbb{R}$. For each $l \in L$ we suppose that the partial (linear) function $u(l, .)$ is continuous on F. Then we have that*

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

Proof of Lemma 3.9. Of course, we may suppose that for each $\mu \in K$ we have $\|\mu\| \leq 1$. If λ is purely finitely additive 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 $\mathbb{P}[A] \leq \varepsilon$ and such that $\mu(A) \geq \delta$. For each $\varepsilon > 0$ we now introduce the convex set, F_ε, of functions, $f \in L^\infty$ such that f is nonnegative, $f \leq 1$ and $\mathbf{E}_{\mathbb{P}}[f] \leq \varepsilon$. The preceding reasoning implies that

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

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

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

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

Example 3.10. In example 3.7, $\mathcal{P}_{\mathbf{ba}}$ contained a σ–additive probability measure. The present example is so that the properties of the preceding theorem 3.8 still hold, but there is no σ–additive probability measure in $\mathcal{P}_{\mathbf{ba}}$. In the language of the theorem 3.8 (2), 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 $\mathbb{F}$ denotes the convex weak*-closed set of all purely finitely additive probabilities μ. Such measures can be characterised as finitely additive probability measures such that $\mu(\{n\}) = 0$ for all $n \in \Omega$. This is also a quick way to see that that $\mathbb{F}$ is weak* closed. 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 $\mathcal{P}_{\mathbf{ba}}$ is the weak* closure of the set

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

The set is clearly convex and it defines a coherent risk measure, ρ. Since obviously $\sup\{\|\mu_a\| \mid \mu \in \mathcal{P}_{\mathbf{ba}}\} = 1$, the properties of theorem 3.8 hold. The difficulty consists in showing that there is no σ–additive measure in the set $\mathcal{P}_{\mathbf{ba}}$. Take an arbitrary element $\mu \in \mathcal{P}_{\mathbf{ba}}$. By the definition of the set $\mathcal{P}_{\mathbf{ba}}$ 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 $\nu^\alpha \in \mathbb{F}$, 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 $\sigma(l^1, c_0)$. 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 we have that 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 that, by taking subnets, 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 $\nu \in \mathbb{F}$.
(4) Of course $\sum_n \frac{|\lambda_n^\alpha - \kappa_n|}{(n+1)^2}$ tends to 0.

As a result we obtain that

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

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 $\mu(\Omega) = 1$.

4. Examples

The examples of this section will later be used in relation with VaR and in relation with convex games. The coherent measures all satisfy the Fatou property and hence are given by a set of probability measures. We do not describe the full set $\mathcal{P}_\sigma$, the sets we will use in the examples are not always convex. So in order to obtain the set $\mathcal{P}_\sigma$ we have to take the closed convex hull. We recall, see the remark after corollary 3.3, that there is a one–to–one correspondence between norm-closed convex sets of probability measures and coherent risk measures that satisfy the Fatou property.

Example 4.1. Here we take

$$\mathcal{P}_\sigma = \{\mathbb{Q} \mid \mathbb{Q} \ll \mathbb{P}\}.$$

The corresponding risk measure is easily seen to be $\rho(X) = \text{ess.}\sup(-X)$, i.e. the maximum loss. It is clear that using such a risk measure as capital requirement would stop all financial/insurance activities. The corresponding supermodular function is given by $\phi(X) = \text{ess.}\inf(X)$.

Example 4.2. In this example we take for a given α, $0 < \alpha < 1$:

$$\mathcal{P} = \{\mathbb{P}[\,.\,|A] \mid \mathbb{P}[A] > \alpha\}.$$

It follows that

$$\rho(X) = \sup_{\mathbb{P}[A] > \alpha} \frac{1}{\mathbb{P}[A]} \int_A (-X)\, d\mathbb{P}.$$

This measure was denoted as WCM_α in [ADEH2]. Of course, if the space is atomless, then it doesn't matter if we use the condition $\mathbb{P}[A] \geq \alpha$ instead of the strict inequality $\mathbb{P}[A] > \alpha$. We remark that in this example the Radon-Nikodym derivatives $\frac{d\mathbb{Q}}{d\mathbb{P}}$ are bounded by $1/\alpha$ and hence the set $\mathcal{P}_\sigma$ is weakly compact in L^1. In the case where $\mathbb{P}$ is atomless, we get that the closed convex hull (taken in $L^1(\mathbb{P})$) of $\mathcal{P}$ is equal to

$$\mathcal{P}_\sigma = \left\{ f \mid 0 \leq f, \|f\|_\infty \leq \frac{1}{\alpha} \text{ and } \mathbf{E}_\mathbb{P}[f] = 1 \right\}.$$

The extreme points of this set are of the form $\frac{\mathbf{1}_A}{\mathbb{P}[A]}$ where $\mathbb{P}[A] = \alpha$, see [Lin].

Example 4.3. This example, as well as the next one, shows that although higher moments cannot be directly used as risk measures, there is some way to introduce their effect. For fixed $p > 1$ and $\beta > 1$, we consider the weakly compact convex set:

$$\mathcal{P}_\sigma = \left\{ \mathbb{Q} \mid \quad \left\| \frac{d\mathbb{Q}}{d\mathbb{P}} \right\|_p \leq \beta \right\}.$$

If $p = \infty$ and $\beta = 1/\alpha$, then we simply find back the preceding example. So we will suppose that $1 < p < \infty$. If we define $q = \frac{p}{p-1}$, the conjugate exponent, then we have the following result:

Theorem 4.4. *For nonnegative bounded functions X, we have that*

$$c\,\|X\|_q \leq \rho(-X) = \psi(X) \leq \beta\|X\|_q,$$

where $c = \min(1, \beta - 1)$.

Proof. The right hand side inequality is easy and follows directly from Hölder's inequality. Indeed for each $h \in \mathcal{P}_\sigma$ we have that

$$\mathbf{E}_{\mathbb{P}}[Xh] \leq \|h\|_p\,\|X\|_q \leq \beta\|X\|_q.$$

The left hand side inequality goes as follows. We may of course suppose that X is not identically zero. We then define $Y = \frac{X^{q-1}}{\|X\|_q^{q-1}}$. As well known and easily checked, we have that $\|Y\|_p = 1$. Also $\mathbf{E}_{\mathbb{P}}[XY] = \|X\|_q$. We now distinguish two cases:
Case 1: $(1 - \mathbf{E}_{\mathbb{P}}[Y]) \leq \beta - 1$. In this case we put $h = Y + 1 - \mathbf{E}_{\mathbb{P}}[Y]$. Clearly we have that $\mathbf{E}_{\mathbb{P}}[h] = 1$ and $\|h\|_p \leq \beta$. Of course we also have that $\mathbf{E}_{\mathbb{P}}[hX] \geq \mathbf{E}_{\mathbb{P}}[XY] = \|X\|_q$.
Case 2: $(1 - \mathbf{E}_{\mathbb{P}}[Y]) \geq \beta - 1$. (Of course this can only happen if $\beta \leq 2$). In this case we take

$$h = \alpha Y + 1 - \alpha\mathbf{E}_{\mathbb{P}}[Y] \text{ where } \alpha = \frac{\beta - 1}{1 - \mathbf{E}_{\mathbb{P}}[Y]}.$$

Clearly $\mathbf{E}_{\mathbb{P}}[h] = 1$ and $\|h\|_p \leq \alpha + 1 - (\beta - 1)\mathbf{E}_{\mathbb{P}}[Y]/(1 - \mathbf{E}_{\mathbb{P}}[Y]) \leq \beta$. But also $\mathbf{E}_{\mathbb{P}}[Xh] \geq \alpha\|X\|_q \geq (\beta - 1)\|X\|_q$, since $1 - \mathbf{E}_{\mathbb{P}}[Y] \leq 1$. □

Remark. It is easily seen that the constant c has to tend to 0 if β tends to 1.
If we take $p = 2$ we get a risk measure that is related to the $\|.\|_2$ norm of the variable. More precisely we find that, in the case $p = 2 = \beta$:

$$\|X\|_2 \leq \psi(X) = \rho(-X) \leq 2\|X\|_2$$

In insurance such a risk measure can therefore be used as a substitute for the standard deviation premium calculation principle. The use of coherent risk measures to calculate insurance premiums has also been addressed in the paper [ADEH2]. For more information on premium calculation principles, we refer to [Wan].

Remark. In section 5, we will give a way to construct analogous examples as the one in 4.3, but where the L^p space is replaced by an Orlicz space.

Example 4.5, Distorted measures. In this example we define directly the coherent risk measure. Section 7 will show that it is a coherent risk measure such that $\mathcal{P}_\sigma$ is weakly compact. We only define $\rho(-X) = \psi(X)$ for nonnegative variables X, the translation property is then used to calculate the value for arbitrary bounded random variables. The impatient reader can now check that the translation property is consistent with the following definition:

$$\psi(X) = \rho(-X) = \int_0^\infty (\mathbb{P}[X > \alpha])^\beta \, d\alpha.$$

The number β is fixed and is chosen to satisfy $0 < \beta < 1$. The exponent q is defined as $q = 1/\beta$. The reader can check that $\beta = 0$, gives us $\|X\|_\infty$, already discussed above. The value $\beta = 1$ just gives the expected value $\mathbf{E}_{\mathbb{P}}[X]$. As usual the exponent p is defined as $p = q/(q-1)$. The following theorem gives the relation between this risk measure and the finiteness of certain moments. We include a proof for the reader's convenience.

Theorem 4.6. *If X is a nonnegative and such that $\int_0^\infty (\mathbb{P}[X > \alpha])^\beta \, d\alpha < \infty$, then also $X \in L^q$. If for some $\varepsilon > 0$, $X \in L^{q+\varepsilon}$, then $\int_0^\infty (\mathbb{P}[X > \alpha])^\beta \, d\alpha < \infty$.*

Proof. If $k = \int_0^\infty \mathbb{P}[X > \alpha]^{1/q} d\alpha < \infty$, then for every $x \geq 0$ we have the inequality $x\mathbb{P}[X > x]^{1/q} \leq k$. This leads to $x^{q-1}\mathbb{P}[X > x]^{1/p} \leq k^{q-1}$. From this we deduce that

$$\begin{aligned}
\|X\|_q^q &= q \int_0^\infty x^{q-1}\mathbb{P}[X > x] \, dx \\
&= q \int_0^\infty x^{q-1}\mathbb{P}[X > x]^{1/p} \, \mathbb{P}[X > x]^{1/q} \, dx \\
&\leq qk^{q-1} \int_0^\infty \mathbb{P}[X > x]^{1/q} \, dx \\
&\leq qk^q.
\end{aligned}$$

This gives us $\|X\|_q \leq q^{1/q} k \leq e^{1/e} k \leq 1.5\, k$. The other implication goes as follows. If $X \in L^{q+\varepsilon}$ for $\varepsilon > 0$, then we have that $x^{q+\varepsilon}\mathbb{P}[X > x] \leq \|X\|_{q+\varepsilon}^{q+\varepsilon} = k$. This implies that

$$\begin{aligned}
\mathbb{P}[X > x] &\leq k x^{-(q+\varepsilon)} \text{ and hence} \\
\mathbb{P}[X > x]^{1/q} &\leq \|X\|_{q+\varepsilon}^{\left(1+\frac{\varepsilon}{q}\right)} x^{-\left(1+\frac{\varepsilon}{q}\right)} \text{ which gives} \\
\int_0^\infty \mathbb{P}[X > x]^{1/q} \, dx &\leq \|X\|_{q+\varepsilon}^{\left(1+\frac{\varepsilon}{q}\right)} \left(1 + \int_1^\infty x^{-\left(1+\frac{\varepsilon}{q}\right)} \, dx\right) \\
&\leq \|X\|_{q+\varepsilon}^{\left(1+\frac{\varepsilon}{q}\right)} \left(\frac{q+\varepsilon}{\varepsilon}\right).
\end{aligned}$$

As expected, the constant on the right hand side tends to ∞ if ε tends to 0. $\square$

We remark that in theorem 4.6, the converse statements do not hold. Indeed a nonnegative variable X such that for x big enough, $\mathbb{P}[X > x] = \frac{1}{(x \log x)^q}$ satisfies $X \in L^q$, but nevertheless $\int_0^\infty \mathbb{P}[X > x]^{1/q} \, dx = +\infty$. Also a nonnegative random variable such that, again for x big enough, $\mathbb{P}[X > x] = \frac{1}{x^q (\log x)^{2q}}$, satisfies $\int_0^\infty \mathbb{P}[X > x]^{1/q} \, dx < \infty$ but there is no $\varepsilon > 0$ such that $X \in L^{q+\varepsilon}$.

Remark. Distorted probability measures were introduced in actuarial sciences by Denneberg, see [De1].

Example 4.7. This example is almost the same as the previous one. Instead of taking a power function x^β, we can, as will be shown in section 7, take any increasing concave function $f \colon [0,1] \to [0,1]$, provided we assume that f is continuous, that $f(0) = 0$ and $f(1) = 1$. The risk measure is then defined, for X bounded nonnegative, as:

$$\psi(X) = \rho(-X) = \int_0^\infty f\left(\mathbb{P}[X > \alpha]\right) \, d\alpha.$$

The continuity of f (at 0) guarantees that the corresponding set $\mathcal{P}_\sigma$ is weakly compact. We will not make an analysis of this risk measure. Especially the behaviour of f at zero, relates this coherent risk measure to Orlicz spaces, in the same way as the x^β function was related to $L^{1/\beta}$ spaces.

Example 4.8. This example, not needed in the rest of the paper, shows that in order to represent coherent risk measures via expected values over a family of probabilities, some control measure is needed. 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. The class of Borel sets of first category, denoted by $\mathcal{N}$, forms a σ–ideal in $\mathcal{F}$. For a bounded function X defined on $[0,1]$ and Borel measurable, we define $\rho(X)$ as the "essential" supremum of $-X$. More precisely we define

$$\rho(X) = \min\{m \mid \{-X > m\} \text{ is of first category}\}.$$

Of course the associated submodular function is then defined as

$$\psi(X) = \min\{m \mid \{X > m\} \text{ is of first category}\}.$$

It is clear that $\rho(X)$ defines a coherent risk measure. It even satisfies the Fatou property in the sense that $\rho(X) \le \liminf \rho(X_n)$, where $(X_n)_{n\ge 1}$ is a uniformly bounded sequence of functions tending pointwise to X. If ρ were of the form

$$\rho(X) = \sup_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_{\mathbb{Q}}[-X],$$

where $\mathcal{P}_\sigma$ is a family of probability measures, then elements $\mathbb{Q}$ of the family $\mathcal{P}_\sigma$ should satisfy:

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

But if $\mathbb{Q}$ is a Borel measure that is zero on the compact sets of first category, then it is identically zero. However we can easily see that

$$\rho(X) = \sup_{\mu\in\mathcal{P}} \mathbf{E}_\mu[-X],$$

where $\mathcal{P}$ is a convex set of finitely additive probabilities on $\mathcal{F}$. The set $\mathcal{P}$ does not contain any σ–additive probability measure, although ρ satisfies some kind of Fatou property. Even worse, all elements in $\mathcal{P}$ are purely finitely additive.

5. Extension to the space of all measurable functions

In this section we study the problem of extending the domain of coherent risk measures to the space L^0 of all equivalence classes of measurable functions. We will focus on those risk measures that are given by a convex set of probability measures, absolutely continuous with respect to $\mathbb{P}$. We start with a negative result. The result is well known but for convenience of the reader, we give more or less full details.

Theorem 5.1. *If the space $(\Omega,\mathcal{F},\mathbb{P})$ is atomless, then there is no real–valued coherent risk measure ρ on L^0. This means that there is no mapping*

$$\psi\colon L^0 \to \mathbb{R},$$

such that the following properties hold:

1. *If $X \ge 0$ then $\psi(X) \ge 0$.*
2. *Subadditivity: $\psi(X_1 + X_2) \le \psi(X_1) + \psi(X_2)$.*
3. *Positive homogeneity: for $\lambda \ge 0$ we have $\psi(\lambda X) = \lambda\psi(X)$.*
4. *For every constant function a we have that $\psi(a + X) = \psi(X) + a$.*

Proof. Suppose that such a mapping would exist, then it can be shown, exactly as in section 2, that $\psi(1) = 1$. Also we have that for $X \leq 0$, necessarily, $\psi(X) \leq 0$. By the Hahn–Banach theorem in its original form, see [Ban], there exists a linear mapping $f\colon L^0 \to \mathbb{R}$, such that $f(1) = 1$ and $f(X) \leq \psi(X)$ for all $X \in L^0$. If $X \geq 0$ then also $-f(X) = f(-X) \leq \psi(-X) \leq 0$. It follows that $f(X) \geq 0$ for all $X \geq 0$. So f is a linear mapping on L^0 that maps nonnegative functions into nonnegative reals. By a classic result of Namioka, see [Nam] (see below, somewhat hidden in the proof of Theorem 5.4, for an outline of a direct proof), it follows that f is continuous. But because the probability space is atomless, there are no non-trivial linear mappings defined on L^0, see [Ni]. This contradicts the property that $f(1) = 1$ and ends the proof. $\square$

Remark. Forrunning the definition of VaR, to be given in section 6, this theorem shows, indirectly, that VaR cannot be coherent. More precisely, since VaR satisfies properties 1,3 and 4 of the definition of coherent measure, VaR cannot be subadditive.

Remark. For completeness and to illustrate why a control in probability (or in L^0–topology) is too crude, we give a sketch of the proof that there are no nontrivial linear mappings defined on L^0, see [Ni] for the original paper. More precisely we show that if $C \subset L^0$ is absolutely convex, closed and has non–empty interior, then $C = L^0$. This implies the non–existence of non-trivial linear mappings. The proof can be given an interpretation in risk management. Indeed, the idea is to approximate arbitrary (bounded) random variables by convex combinations of random variables that are small in probability. We leave further interpretations to the reader. Because C has non-empty interior, it follows that 0 is in the interior of C. This implies the existence of $\varepsilon > 0$ such that for every set $A \in \mathcal{F}$, with $\mathbb{P}[A] \leq \varepsilon$, we have that $\alpha \mathbf{1}_A \in C$ and this for all scalars $\alpha \in \mathbb{R}$. Next, for $\eta > 0$ and $X \in L^\infty$ given, we take a partition of Ω into a finite number of sets $(A_i)_{i \leq N}$ such that

(1) each set A_i has measure $\mathbb{P}[A_i] \leq \varepsilon$.
(2) there is a linear combination Y of the functions $\mathbf{1}_{A_i}$ such that $\|X - Y\|_\infty \leq \eta$.

Since the convex combinations of functions of the form $\alpha_i \mathbf{1}_{A_i}$ exhaust all linear combinations of the functions $\mathbf{1}_{A_i}$, we find that $Y \in C$. Since C was closed we find that $X \in C$. Finally since L^∞ is dense in L^0, we find that $L^0 = C$.

The previous result seems to end the discussion on coherent measures to be defined on L^0. But economically it makes sense to enlarge the range of a coherent measure. The number $\rho(X)$ tells us the amount of capital to be added in order for X to become acceptable for the risk manager, the regulator etc. If X represents a very risky position, whatever that means, then maybe no matter what the capital added is, the position will remain unacceptable. Such a situation would then be described by the requirement that $\rho(X) = +\infty$. Since regulators and risk managers are conservative it is not abnormal to exclude the situation that $\rho(X) = -\infty$. Because this would mean that an arbitrary amount of capital could be withdrawn without endangering the company. So we enlarge the scope of coherent measures as follows

Definition 5.2. *A mapping $\rho\colon \to \mathbb{R} \cup \{+\infty\}$ is called a coherent measure defined on L^0 if*

(1) *If $X \geq 0$ then $\rho(X) \leq 0$.*
(2) *Subadditivity: $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.*

(3) *Positive homogeneity: for* $\lambda > 0$ *we have* $\rho(\lambda X) = \lambda\rho(X)$.
(4) *For every constant function* a *we have that* $\rho(a + X) = \rho(X) - a$.

The reader can check that the elementary properties stated in section 2 remain valid. Also it follows from item one in the definition that ρ cannot be identically $+\infty$. The subadditivity and the translation property have to be interpreted liberally: for each real number a we have that $a + (+\infty) = +\infty$.
We can try to construct coherent risk measures in the same way as we did in theorems 2.2 and 3.2. However this poses some problems. The first idea could be to define the risk measure of a random variable X as

$$\sup_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_{\mathbb{Q}}[-X].$$

This does not work since the random variable X need not be integrable for the measures $\mathbb{Q} \in \mathcal{P}_\sigma$. To remedee this we could try:

$$\sup\left\{\mathbf{E}_{\mathbb{Q}}[-X] \mid \mathbb{Q} \in \mathcal{P}_\sigma; X \in L^1(\mathbb{Q})\right\} \textit{ or}$$

$$\sup\left\{\mathbf{E}_{\mathbb{Q}}[-X] \mid \mathbb{Q} \in \mathcal{P}_\sigma; X^+ \in L^1(\mathbb{Q})\right\}.$$

Such definitions have the disadvantage that the set over which the *sup* is taken depends on the random variable X. This poses problems when we try to compare risk measures of different random variables. So we need another definition. The idea is to truncate the random variable X from above, say by $n \geq 0$. This means that first, we only take into account the possible future wealth up to a level n. We then calculate the risk measure, using the *sup* of all expected values and afterwards we let n tend to infinity. By doing so we follow a conservative viewpoint. High future values of wealth play a role, but their effect only enters through a limit procedure. Very negative future values of the firm may have the effect that we always find the value $+\infty$. This, of course, means that the risk taken by the firm is unacceptable. More precisely:

Definition 5.3. *For a given, closed convex set,* $\mathcal{P}_\sigma$, *of probability measures, all absolutely continuous with respect to* $\mathbb{P}$, *we define the associated support functional* $\rho_{\mathcal{P}_\sigma}$, *or if no confusion is possible,* ρ *as*

$$\rho(X) = \lim_{n\to+\infty} \sup_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_{\mathbb{Q}}\left[-(X \wedge n)\right].$$

For the associated submodular function, ψ and the associated supermodular function ϕ, we then get:

$$\psi(X) = \lim_{n\to-\infty} \sup_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_{\mathbb{Q}}\left[X \vee n\right] \text{ resp.}$$
$$\phi(X) = \lim_{n\to+\infty} \inf_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_{\mathbb{Q}}\left[X \wedge n\right].$$

Of course we need a condition to ensure that $\rho(X) > -\infty$, i.e. $\phi(X) < \infty$, for all $X \in L^0$. This is achieved in the following theorem

Theorem 5.4. *With the notation of definition 5.3 we have that the following properties are equivalent*

(1) *For each $X \in L^0$ we have that $\rho(X) > -\infty$*

(2) *For each $f \in L^0_+$ we have that*

$$\phi(f) = \lim_n \inf_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_\mathbb{Q}[f \wedge n] < +\infty$$

(3) *There is a $\gamma > 0$ such that for each A with $\mathbb{P}[A] \le \gamma$ we have*

$$\inf_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbb{Q}[A] = 0.$$

Under these conditions we have that ρ is a coherent risk measure defined on L^0.

Proof. We first prove that (1) $\Rightarrow$ (2). Let $f \in L^0_+$. Clearly

$$-\rho(f) = -\lim_n \sup_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_\mathbb{Q}\left[-(f \wedge n)\right] = \lim_n \inf_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_\mathbb{Q}\left[f \wedge n\right].$$

We next prove that (2) $\Rightarrow$ (3). If 3 would not hold then for each n we would be able to find A_n such that $\mathbb{P}[A_n] \le 2^{-n}$ and such that

$$\varepsilon_n = \inf_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbb{Q}[A_n] > 0.$$

Define now $f = \sum_n \frac{n\mathbf{1}_{A_n}}{\varepsilon_n}$. Because of the Borel–Cantelli lemma, f is well defined. Let us also take $m = n/\varepsilon_n$. Then of course we have that for each n and for the corresponding m:

$$\lim_{k\to\infty} \inf_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_\mathbb{Q}\left[f \wedge k\right] \ge \inf_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_\mathbb{Q}[f \wedge m] \ge \left(\frac{n}{\varepsilon_n}\right) \inf_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbb{Q}[A_n] \ge n\frac{\varepsilon_n}{\varepsilon_n} = n,$$

which contradicts 2.
Let us now show that (3) $\Rightarrow$ (1). For given X, let N be chosen so that $\mathbb{P}[X \ge N] \le \gamma$, where γ is given by (3). Since for each $n \ge N$ we have, by (3), that

$$\sup_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_\mathbb{Q}\left[-(X \wedge n)\right] = \sup_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_\mathbb{Q}\left[-(X \wedge N)\right] \ge -N,$$

we immediately get (1).
The last statement of the theorem is obvious. Positive homogeneity, subadditivity as well as the translation property are easily verified. □

Proposition 5.5. *The hypotheses of the previous theorem 5.4 are satisfied if for each nonnegative function $f \in L^0$, there is $\mathbb{Q} \in \mathcal{P}_\sigma$ such that $\mathbf{E}_\mathbb{Q}[f] < \infty$.*

The proof of the proposition is obvious. However we have more:

Theorem 5.6. *If $\mathcal{P}_\sigma$ is a norm closed, convex set of probability measures, all absolutely continuous with respect to $\mathbb{P}$, then the equivalent properties of theorem 5.4 are also equivalent with:*

(4) *For every $f \in L^0_+$ there is $\mathbb{Q} \in \mathcal{P}_\sigma$ such that $\mathbf{E}_\mathbb{Q}[f] < \infty$.*

(5) *There is a $\delta > 0$ such that for every set A with $\mathbb{P}[A] < \delta$, we can find an element $\mathbb{Q} \in \mathcal{P}_\sigma$ such that $\mathbb{Q}[A] = 0$.*

(6) *There is a $\delta > 0$, as well as a number K such that for every set A with $\mathbb{P}[A] < \delta$, we can find an element $\mathbb{Q} \in \mathcal{P}_\sigma$ such that $\mathbb{Q}[A] = 0$ and $\|\frac{d\mathbb{Q}}{d\mathbb{P}}\|_\infty \le K$.*

Remark. The proof of this theorem is by no means trivial, so let us first sketch where the difficulty is. We concentrate on (4). Suppose that for a given nonnegative function g, we have that

$$\lim_n \inf_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_{\mathbb{Q}}[g\wedge n] = \phi(g) < \infty.$$

This means that for every n we can find $f_n \in \mathcal{P}_\sigma$ such that

$$\mathbf{E}_{\mathbb{P}}\left[f_n\,(g\wedge n)\right] \le \phi(g) + 1.$$

The problem is that the sequence f_n does not necessarily have a weakly convergent subsequence. If however we could choose the sequence $(f_n)_n$ in such a way that it is uniformly integrable, then it is relatively weakly compact, a subsequence would be weakly convergent, say to an element $f \in \mathcal{P}_\sigma$ and a direct calculation then shows that also

$$\begin{aligned}\mathbf{E}_{\mathbb{P}}[fg] &= \lim_n \mathbf{E}_{\mathbb{P}}\left[f(g\wedge n)\right] = \lim_n \lim_k \mathbf{E}_{\mathbb{P}}\left[f_k(g\wedge n)\right]\\ &\le \limsup_k \mathbf{E}_{\mathbb{P}}\left[f_k(g\wedge k)\right] \le \phi(g) + 1 < +\infty.\end{aligned}$$

So in case $\mathcal{P}_\sigma$ is weakly compact, there is no problem. The general case however is much more complicated and requires a careful selection of the sequence f_n. The original proof consisted in constructing a sequence f_n in such a way that it became uniformly integrable. This was quite difficult and used special features from functional analysis. The present proof is much easier but in my viewpoint less transparent. It uses the Hahn–Banach theorem directly.

Proof. It is clear that (6) $\Rightarrow$ (5) $\Rightarrow$ (4), which in turn implies the properties (1),(2) and (3) of Theorem 5.4. So we only have to show that the properties (1),(2) and (3) imply property (6). Let $k > \frac{2}{\gamma}$ and let A with $\mathbb{P}[A] < \frac{\gamma}{2}$ be given. We will show that 3 implies 6. We suppose the contrary. So let us take $H_k = \{f \mid |f| \le k,\ f = 0 \text{ on } A\}$. If H_k and $\mathcal{P}_\sigma$ were disjoint we could, by the Hahn-Banach theorem, strictly separate the closed convex set $\mathcal{P}_\sigma$ and the weakly compact, convex set H_k. This means that there exists an element $X \in L^\infty$, $\|X\|_\infty \le 1$ so that

$$\sup\left\{\mathbf{E}[Xf] \mid f \in H_k\right\} < \inf\left\{\mathbf{E}_{\mathbb{Q}}[X] \mid \mathbb{Q} \in \mathcal{P}_\sigma\right\}.$$

We will show that this inequality implies that $\|X\mathbf{1}A^c\|_1 = 0$. Indeed if not, we would have $\mathbb{P}[\mathbf{1}_{A^c}|X| > \frac{2}{\gamma}\|X\mathbf{1}_{A^c}\|_1] \le \frac{\gamma}{2}$ and hence for each $\varepsilon > 0$ there is a $\mathbb{Q} \in \mathcal{P}_\sigma$ so that $\mathbb{Q}[A \cup \{|X| > \frac{2}{\gamma}\|X\mathbf{1}_{A^c}\|_1\}] \le \varepsilon$. This implies that the right side of the separation inequality is bounded by $\frac{2}{\gamma}\|X\mathbf{1}_{A^c}\|_1$. However, the left side is precisely $k\|X\mathbf{1}A^c\|_1$. This implies $k\|X\mathbf{1}_{A^c}\|_1 < \frac{2}{\gamma}\|X\mathbf{1}_{A^c}\|_1$, a contradiction to the choice of k. Therefore $X = 0$ 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. $\square$

Corollary 5.7. *Let the closed convex set of probability measures $\mathcal{P}_\sigma$ satisfy the (equivalent) conditions of theorem 5.6 (or 5.4). Then there is a constant K_0 such that all the sets, defined for $K \ge K_0$,*

$$\mathcal{P}_{\sigma,K} = \left\{f \mid f \in \mathcal{P}_\sigma \text{ and } \|f\|_\infty \le K\right\},$$

all satisfy the conditions of theorem 5.6.

Proof. Obvious

Corollary 5.8. *Let h be a strictly positive $\mathbb{P}$-integrable random variable. If $\mathcal{P}_\sigma$ satisfies the conditions of theorem 5.6, then there is a $\delta > 0$ as well as a constant K such that for each set A of measure $\mathbb{P}[A] < \delta$ we can find an element $f \in \mathcal{P}_\sigma$ such that $\|f/h\|_\infty \leq K$.*

Proof. It does not do any harm to normalise the function h. We therefore may suppose that the measure $d\mathbb{P}' = h\, d\mathbb{P}$ is a probability, equivalent to $\mathbb{P}$. When working with the measure $\mathbb{P}'$, the $\mathbb{P}$-densities $f \in \mathcal{P}_\sigma$ have to be replaced with the $\mathbb{P}'$-densities f/h. The corollary is now a rephrasing of the theorem. □

Example 5.9. The reader might ask whether the set

$$\mathcal{P}_{\sigma,\infty} = \{f \mid f \in \mathcal{P}_\sigma \text{ and } \|f\|_\infty < \infty\},$$

is dense in the set $\mathcal{P}_\sigma$. This example shows that, in general, it is not the case. The set $\mathcal{P}_\sigma$ is defined as follows. We start by taking an unbounded density, i.e. we start by taking a nonnegative random variable f, such that f is unbounded and $\mathbf{E}_\mathbb{P}[f] = 1$. Then we define:

$$\mathcal{P}_\sigma = \{\lambda h + (1-\lambda) f \mid 0 \leq h \leq 2, \mathbf{E}_\mathbb{P}[h] = 1, 0 \leq \lambda \leq 1\}.$$

Obviously, this set satisfies the properties of Theorem 5.6. It is also clear that the only bounded elements in $\mathcal{P}_\sigma$ are the elements h such that $\|h\|_\infty \leq 2$, i.e. where λ is taken to be 1. The unbounded function f will never be in the closure.

Remark 5.10. In view of the discussions in the beginning of this section and after having seen theorem 5.6, the reader might ask whether there was really a need to have this special limit procedure. More precisely, let $\mathcal{P}_\sigma$ be a closed convex set of probability measures, all absolutely continuous with respect to $\mathbb{P}$. Suppose that $\mathcal{P}_\sigma$ satisfies the conditions of Theorem 5.6, i.e. there is a $\delta > 0$ such that for each set A with $\mathbb{P}[A] \leq \delta$, we can find an element $\mathbb{Q} \in \mathcal{P}_\sigma$ with $\mathbb{Q}[A] = 0$. The risk measure ρ is now defined as

$$\rho(X) = \lim_n \sup_{\mathbb{Q} \in \mathcal{P}_\sigma} \mathbf{E}_\mathbb{Q}\left[-(X \wedge n)\right].$$

Is it true that also $\rho(X)$ is given by the expression:

$$\alpha(X) = \sup\left\{\mathbf{E}_\mathbb{Q}\left[-X\right] \mid \mathbb{Q} \in \mathcal{P}_\sigma \text{ and } X^+ \in L^1(\mathbb{Q})\right\}?$$

The answer is no, as the following counter–example shows. Fix an atomless probability space. Take a number $k > 1$ and let $\mathcal{K}$ to be the set of all probability densities g (with respect to $\mathbb{P}$) such that $\|g\|_\infty \leq k$. This set is a weakly compact set in $L^1(\mathbb{P})$. Fix now a probability density h that is unbounded. Let $\mathcal{P}_\sigma$ be the set of all convex combinations $\lambda h + (1-\lambda) g$ where $g \in \mathcal{K}$. This set is obviously closed and convex, it is even weakly compact (since $\mathcal{K}$ is). Since h is unbounded it is easy to find a $\mathbb{P}$-integrable random variable X, such that $\mathbf{E}_\mathbb{P}[X^+ h] = \mathbf{E}_\mathbb{P}[X^- h] = \infty$. It is easy to see that the only elements $\mathbb{Q} \in \mathcal{P}_\sigma$ such that $\mathbf{E}_\mathbb{Q}[X^+] < \infty$ (or $\mathbf{E}_\mathbb{Q}[X^-] < \infty$) are the ones in $\mathcal{K}$. Hence we get that $\alpha(X) \leq k\, \|X\|_{L^1(\mathbb{P})}$, whereas $\rho(X) = +\infty$.

Remark 5.11. The examples 4.1, 4.2 and 4.3 trivially satisfy the assumptions of theorem 5.6. The distorted measure of example 4.5 doesn't satisfy the assumptions of theorem 5.6. The more general case of example 4.7 satisfies the assumptions of theorem 5.6 if and only if $f(x) = 1$ for some $x < 1$. This will become clear after section 7.

The following examples show that Beppo Levi type theorems are false, even when the set $\mathcal{P}_\sigma$ is weakly compact.

Example 5.12. We use the same notation as in example 5.10. The probability space is supposed to be atomless. The set $\mathcal{P}_\sigma$ is the convex hull of the sets $\mathcal{K}$ of densities that are bounded by 2 and the unbounded density h. The random variable X is chosen so that $\mathbf{E}_\mathbb{P}[|X|] < \infty$ but $\mathbf{E}_\mathbb{P}[X^+ h] = \mathbf{E}_\mathbb{P}[X^- h] = +\infty$. The sequence X_k is defined as $X_k = \max(X, -k)$. Clearly X_k decreases to X. Now we have that for n big enough the quantity $E_\mathbb{P}[-(\max(X,-k) \wedge n)\, h] = \mathbf{E}_\mathbb{P}[h(X^- \wedge k)] - \mathbf{E}_\mathbb{P}[h(X^+ \wedge n)]$ is smaller then $-10\|X\|_{L^1(\mathbb{P})}$. We get that for n big enough the maximum is attained for an element of $\mathcal{K}$. It follows that for each k the quantity $\rho(X)$ is bounded by $2\|X\|_{L^1(\mathbb{P})}$. But as seen before $\rho(X) = +\infty$.

Example 5.13. This example is similar to the previous one but it does not have the drawback that $\rho(X) = +\infty$. We start with the same set $\mathcal{P}_\sigma$. This time we select X in such a way that $+\infty > \mathbf{E}_\mathbb{P}[X^-\, h] \geq 100\|X\|_{L^1(\mathbb{P})}$ and such that $\mathbf{E}_\mathbb{P}[X^+\, h] = +\infty$. The sequence X_k is defined as $X_k = -X^- + \mathbf{1}_{\{X>k\}}X^+$. Exactly as before, for each k the maximum expected value will be attained for an element in $\mathcal{K}$ and hence $\rho(X) \leq 2\|X\|_1$. But X_k decreases to $-X^-$ for which the risk measure gives $\rho(-X^-) \geq 100\|X\|_1$.

Example 5.14. This, as well as the following example deals with increasing sequences. Again we use the same set $\mathcal{P}_\sigma$. Let us take a nonnegative function f such that $\mathbf{E}_\mathbb{P}[fh] = +\infty$. Define $X = 0$ and $X_k = -f\mathbf{1}_{\{f>k\}}$. Clearly X_k increases to 0 but for each k we have $\rho(X_k) = +\infty$.

Example 5.15. In this example we fix the inconvenience that for each X_k the risk measure gives the value $+\infty$. This example is of a different nature and is related to example 4.3. But this time we use a non reflexive Orlicz space to define the set $\mathcal{P}_\sigma$. Let us start with some notation and some review of the definition of an Orlicz space (see [Nev] for more information on Orlicz spaces). All random variables will be defined on the space $[0,1]$ equipped with the usual Lebesgue measure. To construct an Orlicz space, we start with an increasing convex function $\Phi\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $\Phi(0) = 0$. Also we suppose that $\Phi(x)/x$ tends to $+\infty$ for $x \to +\infty$. The derivative $\varphi = \Phi'$ should also satisfy $\varphi(0) = 0$ and we suppose that φ is continuous and strictly increasing. (In a more general setup, not all these conditions are needed). The space L^Φ is now defined as the space of all (equivalence classes of) random variables X such that there exists $\lambda > 0$ with $\mathbf{E}_\mathbb{P}[\Phi(\lambda|X|)] < +\infty$. The norm is then defined as

$$\|X\|_\Phi = \inf\left\{\alpha > 0 \mid \mathbf{E}_\mathbb{P}\left[\Phi\left(\frac{|X|}{\alpha}\right)\right] \leq 1\right\}.$$

It can be shown that this is indeed a norm and that the space L^Φ is a Banach space. With each function Φ as above, we associate the dual function Ψ. This function has as its derivative the function which is the inverse function of φ. The spaces L^Φ and L^Ψ are in duality.
In our example we take $\Phi(x) = \frac{1}{2\log 2 - 1}\left((x+1)\log(x+1) - x\right)$. The expression $\frac{1}{2\log 2-1}$ is introduced in order to have $\Phi(1) = 1$ which together with Jensen's

inequality then implies that $\|X\|_1 \leq \|X\|_\Phi$ for all $X \in L^\Phi$. The associated function Ψ is then given by $\Psi(y) = \frac{1}{2\log 2-1}\left(e^{y(2\log 2-1)}-1\right)-y$. Young's inequality states that $\mathbf{E}_\mathbb{P}[XY] \leq 2\|X\|_\Phi\|Y\|_\Psi$. The dual of L^Φ is L^Ψ, but the dual norm is only equivalent (not equal) to $\|\cdot\|_\Psi$, in fact

$$\|\cdot\|_\Psi \leq \|\cdot\|_{(L^\Phi)^*} \leq 2\|\cdot\|_\Psi.$$

Before defining the set $\mathcal{P}_\sigma$, let us describe one of the pitfalls of Orlicz–space–theory. If we define the function $f(t) = \log(t^{-1})$ then the function f is in L^Ψ. But as can be verified by direct calculation, the norm $\|f - f\wedge k\|_\Psi$ remains bigger than $2\log 2 - 1 = \delta$. This means that the bounded functions are not dense in L^Ψ. This will in fact be the kernel of our example.
Let us now describe the set $\mathcal{P}_\sigma$. Exactly as in example 4.3, we take

$$\mathcal{P}_\sigma = \{h \mid h \geq 0, \mathbf{E}_\mathbb{P}[h] = 1 \text{ and } \|h\|_\Phi \leq 2\}.$$

From de la Vallée Poussin's theorem (on uniform integrability) it follows that $\mathcal{P}_\sigma$ is weakly compact. Let us now take the function $f = \log(t^{-1})$ as above and let $X = 0$ and $X_k = -f\mathbf{1}_{\{f>k\}}$. It follows from Young's inequality that $\rho(-f) < +\infty$. Clearly X_k increases to 0. Since $\|X_k\|_\Psi \geq \delta$, we get the existence of an element g_k such that $\|g_k\|_\Phi = 1$ and such that $\mathbf{E}_\mathbb{P}[g_k(-X_k)] \geq \delta/4$. Since $\mathbf{E}_\mathbb{P}[g_k] \leq \|g\|_{L^\Phi} \leq 1$ we get that the function $h_k = g_k + 1 - \mathbf{E}_\mathbb{P}[g_k]$ is in $\mathcal{P}_\sigma$. And of course we have that $\mathbf{E}_\mathbb{P}[h_k(-X_k)] \geq \delta/4$. All this shows that $\rho(X_k)$ does not decrease to zero.
The reader can check that the method used in the example above yields the following generalization of the result 4.4.

Proposition 5.16. *With the above notation for Young functions Φ with $\Phi(1) = 1$, and for*

$$\mathcal{P}_\sigma = \{h \mid h \text{ probability density on } \Omega \text{ with } \|h\|_\Phi \leq K\},$$

where $K \geq 1$, there is a constant $\delta = \min(K-1, 1) > 0$ such that for all nonpositive random variables X, we have

$$\delta\|X\|_\Psi \leq \rho(X) \leq 2K\|X\|_\Psi$$

Remark. It is of course a trivial consequence of the Beppo Levi theorem that the following holds:
If $(X_n)_{n\geq 1}$ is a sequence that is uniformly bounded from above and decreases to a function X, then $\rho(X_n)$ increases to $\rho(X)$.

Remark. By choosing the appropriate sets $\mathcal{P}_\sigma$, we can find measures that are related to entropy or to the Esscher premium calculation principles. Based on a suggestion by the author, premium calculation principles based on Orlicz norms were introduced in [HG]. However, the premium calculation principles studied there, do not satisfy the translation property.

6. The relation with VaR

The aim of this section is to give the relationship between coherent risk measures and the popular, although not coherent, measure VaR. We will restrict the analysis for risk measures that are defined on L^∞. The extension to the space L^0 is straightforward provided the properties of theorems 5.4 and 5.6 are satisfied. It is not easy to find a mathematically satisfactory definition of what is usually meant by VaR. Expressions such as "VaR summarizes the maximal expected loss within a given confidence interval" are hard to translate into a mathematical formula. The best we can do is to define VaR as a quantile of the distribution of the random variable. We start by defining what is meant by a quantile, see [EKM] for a discussion on how to estimate quantiles for extreme value distributions.

Definition 6.1. *If X is a real valued random variable, if $0 < \alpha < 1$, then we say that q is an α-quantile if $\mathbb{P}[X < q] \le \alpha \le \mathbb{P}[X \le q]$.*

It is easy to see that the set of quantiles forms a closed interval with endpoints q^-_α and q^+_α. These endpoints can be defined as

$$\begin{aligned} q^-_\alpha &= \inf\{q \mid \mathbb{P}[X \le q] \ge \alpha\} \\ q^+_\alpha &= \inf\{q \mid \mathbb{P}[X \le q] > \alpha\}. \end{aligned}$$

The fact that there are different values for a quantile will cause some troubles in the formulation of the theorems. Fortunately for a fixed variable X, the two quantiles coincide for all levels α, except on at most a countable set.
We also use the following

Definition 6.2. *The quantity $VaR_\alpha(X) = -q^+_\alpha(X)$ is called the value at risk at level α for the random variable X.*

We start with a characterisation of coherent risk measures that dominate VaR, this means that for every bounded random variable X, we have that $\rho(X) \ge VaR_\alpha(X)$.

Theorem 6.3. *A coherent risk measure ρ dominates VaR_α for bounded random variables, if and only if for each set B, $\mathbb{P}[B] > \alpha$ and each $\varepsilon > 0$, there is a measure $\mu \in \mathcal{P}_{\mathbf{ba}}$ such that $\mu(B) > 1 - \varepsilon$.*

Proof. We first prove necessity. Take $\varepsilon > 0$ and a set B such that $\mathbb{P}[B] > \alpha$. Since $VaR_\alpha(X) = 1$ for the random variable $X = -\mathbf{1}_B$, we conclude from the inequality $\rho \ge VaR_\alpha$, that there is a measure $\mu \in \mathcal{P}_{\mathbf{ba}}$ such that $\mu(B) \ge 1 - \varepsilon$.
For the sufficiency we take a random variable X as well as $\varepsilon > 0$ and we consider the set $B = \{X \le q^+_\alpha + \varepsilon\}$. By definition of VaR, we get that $\mathbb{P}[B] > \alpha$. There exists a measure $\mu \in \mathcal{P}_{\mathbf{ba}}$ with the property $\mu(B) \ge 1 - \varepsilon$. This gives the inequality

$$\rho(X) \ge \mathbf{E}_\mu[-X] \ge \mathbf{E}_\mu[-X\mathbf{1}_B] - \varepsilon\|X\|_\infty \ge (1-\varepsilon)(VaR_\alpha(X) - \varepsilon) - \varepsilon\|X\|_\infty.$$

Since the inequality holds for every $\varepsilon > 0$, we get the result $\rho \ge VaR_\alpha$. $\square$

Example 6.4. There are risk measures, satisfying the Fatou property, that dominate VaR_α but that do not dominate $-q^-_\alpha$. We will construct a coherent risk measure ρ and a set B of measure exactly equal to α such that $\sup_{\mu \in \mathcal{P}_\sigma} \mu(B) = 0$. This will do, since for $X = -\mathbf{1}_B$, we then have that $q^-_\alpha(X) = -1$. The construction goes as follows. As probability space we take the unit interval $[0, 1]$ with the Borel sigma algebra $\mathcal{F}$ and the Lebesgue measure, which we denote by $\mathbb{P}$. We will now describe a set of probability measures (or better a set of densities) $\mathcal{P}$. The set $\mathcal{P}_\sigma$ is then the closed convex hull of $\mathcal{P}$. For each set B of measure $\mathbb{P}[B] > \alpha$, we define the density

$$h_B = \frac{\mathbf{1}_{B\cap[0,1-\alpha]}}{\mathbb{P}\left[B\cap[0,1-\alpha]\right]}.$$

The set $\mathcal{P}$ is defined as the set

$$\mathcal{P} = \{h_B \mid \mathbb{P}[B] > \alpha\}.$$

It follows from the construction that the properties of theorem 6.3 are fulfilled and hence the corresponding risk measure $\rho(X) = \sup_{\mathbb{P}[B]>\alpha} \mathbf{E}_{\mathbb{P}}[(-X)h_B]$ dominates VaR_α. But $\rho(-\mathbf{1}_{[1-\alpha,1]}) = 0$ since all the densities are supported by $[0, 1-\alpha]$. As easily seen we also have that $q^-_\alpha\left(-\mathbf{1}_{[1-\alpha,1]}\right) = -1$ and hence ρ does not dominate the function $-q^-_\alpha$.
The example has a little disadvantage. The risk measure ρ is not relevant. This can be repaired by adding the measure $\mathbb{P}$ to the set $\mathcal{P}$. We get that ρ is relevant and we still have the inequality $\rho(-\mathbf{1}_{[1-\alpha,1]}) = \mathbb{P}\left[[1-\alpha, 1]\right] = \alpha < 1$.
We end the discussion of this example with the following result, proved in the same way as theorem 6.3.

Proposition 6.5. *For a coherent risk measure the following are equivalent*

1. *The risk measure is bigger than $-q^-_\alpha$.*
2. *For each set B, $\mathbb{P}[B] \geq \alpha$ and each $\varepsilon > 0$, there is a measure $\mu \in \mathcal{P}_{\mathbf{ba}}$ such that $\mu(B) > 1 - \varepsilon$.*

Corollary 6.6. *If $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless, if $\mathcal{P}_\sigma$ is weakly compact in L^1 and if ρ dominates VaR, then ρ also dominates $-q^-_\alpha$.*

Proof. If $\mathbb{P}[A] = \alpha$, then we simply take a decreasing sequence of sets A_n of measure $\mathbb{P}[A_n] > \alpha$, whose intersection is A. This can be done since the space is atomless. For each n we can find $h_n \in \mathcal{P}_\sigma$ such that $\mathbf{E}_{\mathbb{P}}[h_n \mathbf{1}_{A_n}] \geq 1 - 1/n$. Since the set $\mathcal{P}_\sigma$ is weakly compact, the sequence h_n has a weakly convergent subsequence, whose limit we denote by h. Clearly (by weak compactness, i.e. uniform integrability) we have that $\mathbf{E}_{\mathbb{P}}[h\mathbf{1}_A] = 1$. The corollary now follows. $\square$

Proposition 6.7. *If $\rho: L^\infty \to I\!R$ is a coherent risk measure that satisfies the Fatou property and is defined with the set $\mathcal{P}_\sigma$, if the extension $\rho: L^0 \to I\!R \cup \{+\infty\}$ defined as*

$$\rho(X) = \lim_n \sup_{\mathbb{Q}\in\mathcal{P}_\sigma} \mathbf{E}_{\mathbb{Q}}[-(X \wedge n)]$$

satisfies $\rho(X) > -\infty$ for all $X \in L^0$, if ρ dominates VaR for bounded random variables, then ρ dominates VaR for all random variables.

Proof. The proof of this proposition is obvious. $\square$

We now can give the first theorem on the relation between VaR and coherent risk measures.

Theorem 6.8. *For each bounded random variable X and each $\alpha, 0 < \alpha < 1$, we have that*

$$VaR_\alpha(X) = \min\{\rho(X) \mid \rho \geq VaR, \rho \textit{ coherent with the Fatou property}\}.$$

Since VaR is not coherent this shows that there is no smallest coherent risk measure that dominates VaR_α.

Proof. We only have to show that for X given, we can find a coherent risk measure that dominates VaR_α and with the property that $\rho(X) \leq VaR_\alpha(X)$. For each $\varepsilon > 0$, the set $C = \{X \leq q_\alpha^+ + \varepsilon\}$ has measure $\mathbb{P}[C] > \alpha$. But the definition of q_α^+ implies that $\mathbb{P}[X < q_\alpha^+] \leq \alpha$. It follows that the set $D = \{q_\alpha^+ \leq X \leq q_\alpha^+ + \varepsilon\}$ has strictly positive measure. Take now an arbitrary set B with measure $\mathbb{P}[B] > \alpha$. Either we have that $\mathbb{P}[B \cap C^c] \neq 0$, 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 $\mathbb{P}[X < q_\alpha^+] \leq \alpha$ 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 risk measure ρ is then defined as $\rho(Y) = \sup_{\mathbb{P}[B]>\alpha} \mathbf{E}_\mathbb{P}[(-Y)h_B]$. By theorem 6.3 we have that ρ dominates VaR but for the variable X we find that $\mathbf{E}_\mathbb{P}[(-X)h_B]$ is always bounded by $-q_\alpha^+$, i.e. $\rho(X) \leq VaR_\alpha(X)$. It follows that $\rho(X) = VaR_\alpha(X)$. $\square$

In order to prove the second theorem on the relation between VaR and coherent risk measures, we need a characterisation of atomless spaces. (Compare proposition 5.4 in [ADEH2], where another, but not unrelated, kind of homogeneity of the space Ω is used). The proof of the following proposition is left to the reader.

Proposition 6.9. *For a probability space $(E, \mathcal{E}, \mathbb{Q})$ the following are equivalent*

1. *The space is atomless.*
2. *The space supports a random variable with a continuous distribution.*
3. *Every probability measure on $\mathbb{R}$ is the distribution of a random variable defined on E.*
4. *There is an i.i.d. sequence of random variables f_n such that $\mathbb{Q}[f_n = 1] = \mathbb{Q}[f_n = -1] = 1/2$.*
5. *Given a distribution ν on $\mathbb{R}$, there is an i.i.d. sequence f_n with distribution ν.*

Theorem 6.10. *Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless, suppose that ρ is a coherent risk measure that satisfies the Fatou property, dominates VaR_α and only depend on the distribution of the random variable, then $\rho \geq WCM_\alpha$ (see example 4.2 for a definition of WCM). In other words WCM_α is the smallest coherent risk measure, dominating VaR and being distribution invariant.*

Proof. We first show that

$$\rho(X) \geq \mathbf{E}_\mathbb{P}[-X \mid X \leq q_\alpha^+].$$

So let us take $\varepsilon > 0$. The set $A = \{X \leq q_\alpha^+ + \varepsilon\}$ has measure $\mathbb{P}[A] > \alpha$, by definition of q_α^+. The variable Y defined as

$$Y = X \text{ on } A^c \quad \text{and} \quad Y = \mathbf{E}_\mathbb{P}[X \mid A] \text{ on } A$$

satisfies $VaR_\alpha(Y) = \mathbf{E}_\mathbb{P}[-X \mid A]$ and hence $\rho(Y) \geq \mathbf{E}_\mathbb{P}[-X \mid A]$. We will now show that $\rho(Y) \leq \rho(X)$, which will then prove the statement $\rho(X) \geq \mathbf{E}_\mathbb{P}[-X \mid X \leq q_\alpha^+ + \varepsilon]$. To achieve this goal, we consider the atomless probability space $(A, \mathcal{F} \cap A, \mathbb{P}[\,.\mid A])$. On this space we consider an i.i.d. sequence of functions f_n each having the same distribution as X under the measure $\mathbb{P}[\,.\mid A]$. For each n we define the random variable

$$X_n = X \text{ on } A^c \quad \text{and} \quad X_n = f_n \text{ on } A.$$

The sequence X_n is a sequence of identically distributed random variables and the strong law of large numbers implies that $\frac{X_1+\dots X_n}{n}$ tends to Y almost surely. The convexity and the Fatou property then imply that

$$\rho(Y) \le \liminf \rho\left(\frac{X_1+\dots X_n}{n}\right) \le \liminf \frac{1}{n}\sum_1^n \rho(X_n) = \rho(X).$$

So we obtain that $\rho(X) \ge \mathbf{E}_\mathbb{P}[-X \mid X \le q_\alpha^+ + \varepsilon]$. If we let ε tend to zero, the right hand side converges to $\mathbf{E}_\mathbb{P}[-X \mid X \le q_\alpha^+]$ and we get the inequality

$$\rho(X) \ge \mathbf{E}_\mathbb{P}[-X \mid X \le q_\alpha^+].$$

Unfortunately the conditional expectation does not define a coherent risk measure. This is due to the fact that quantiles are not continuous with respect to any reasonable topology. So we still need an extra argument. Since the space is atomless we can, as can be shown easily, find a decreasing sequence of random variables $X + 1/n \ge Z_n \ge X$, tending to X in L^∞–norm and such that each Z_n has a continuous distribution. For each n we then have that

$$\rho(Z_n) \ge \mathbf{E}_\mathbb{P}[-Z_n \mid Z_n \le q_\alpha^+(Z_n)] = WCM_\alpha(Z_n).$$

If n tends to infinity the left hand side tends to $\rho(X)$. The right hand side tends to $WCM_\alpha(X)$. $\square$

Corollary 6.11. *Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless, suppose that ρ is a coherent risk measure that satisfies the Fatou property, dominates VaR_α and only depends on the distribution of the random variable, then ρ also dominates the function $-q_\alpha^-$. In particular WCM_α dominates $-q_\alpha^-$.*

Proof. We only have to show that WCM_α dominates $-q_\alpha^-$. This follows directly from corollary 6.6.

7. Convex Games and Comonotone Risk Measures

This section describes the relation between convex games and some coherent risk measures. The basic tool is Choquet integration theory, also called non linear integration. The use of non linear premium calculation principles was investigated by Denneberg [De1]. In [De2] the relation with Choquet integration theory is developed. In Delbaen, [D1], and Schmeidler, [Schm1], the reader can find the basics of convex game theory needed in this section.

Definition 7.1. *A 2-alternating, or strongly superadditive or supermodular, set function is defined as a function $w: \mathcal{F} \to \mathbb{R}_+$, that satisfies the property*

(1) $w(A \cap B) + w(A \cup B) \le w(A) + w(B)$.

If moreover

(2) $A = B$, $\mathbb{P}$ *a.s., implies* $w(A) = w(B)$,

we say that w is absolutely continuous with respect to $\mathbb{P}$. A game is defined as a function $v: \mathcal{F} \to \mathbb{R}_+$.
A convex game, defined on the space $(\Omega, \mathcal{F}, \mathbb{P})$ is defined as a function $v: \mathcal{F} \to \mathbb{R}_+$ such that

(1) $v(A \cap B) + v(A \cup B) \ge v(A) + v(B)$.

If moreover

(2) $A = B$, $\mathbb{P}$ *a.s., implies* $w(A) = w(B)$,

we say that v *is absolutely continuous with respect to* $\mathbb{P}$

Convex games v and 2-alternating functions w are related through the relations $v(\Omega) = w(\Omega)$ and $v(A) = v(\Omega) - w(A^c)$. The number $v(A)$ is interpreted as the minimum, the coalition A, has to get. If the relation between v and w is used to transform a convex game into a 2-alternating function, the quantity $w(A)$ is interpreted as the maximum the coalition A is allowed to get.

Remark. In this section, we will mostly require that the game or the 2-alternating set function is "absolutely continuous" with respect to $\mathbb{P}$. We need this property in order to use $\mathbb{P}$ as a control measure. See example 4.8 and example 7.9 below.

Definition 7.2. *The core* $\mathcal{C}(v)$, *of the game* v *is the set of all finitely additive nonnegative measures* $\mu \in \mathbf{ba}(\Omega, \mathcal{F})$, *such that* $\mu(\Omega) = v(\Omega)$ *and such that for all sets* $A \in \mathcal{F}$ *we have* $\mu(A) \geq v(A)$. *The* σ*-core,* $\mathcal{C}^\sigma(v)$ *is the set of all* σ*–additive measures in the core.*

Remark. In case the game v is absolutely continuous with respect to $\mathbb{P}$, it is easily seen that $\mathcal{C}(v)$ (resp. the σ–core $\mathcal{C}^\sigma(v)$) is actually a subset of $\mathbf{ba}(\Omega, \mathcal{F}, \mathbb{P})$ (resp. $L^1(\Omega, \mathcal{F}, \mathbb{P})$). Indeed if $\mathbb{P}[A] = 0$, then for $\mu \in \mathcal{C}(v)$ we necessarily have that $\mu(A^c) \geq v(A^c) = v(\Omega)$. This implies that $\mu(A) = 0$.
In [Schm1] and [D1] the reader can find the following properties of the core of a game v.

(1) The core of a convex game is non empty.
(2) The core is a weak* compact convex subset of $\mathbf{ba}$.
(3) If for each increasing sequence A_n, with union equal to Ω, the numbers $v(A_n)$ tend to 1, then the core is a weakly compact subset of L^1.
(4) For a bounded nonnegative function X the following equality holds (again we used the relation between v and the 2-alternating set function w):

$$\psi(X) = \sup_{\mu \in \mathcal{C}(v)} \mathbf{E}_\mu[X] = \int_0^\infty w(X > x)\, dx, \text{ and}$$

$$\phi(X) = \inf_{\mu \in \mathcal{C}(v)} \mathbf{E}_\mu[X] = \int_0^\infty v(X > x)\, dx.$$

(5) If $v(\Omega) = 1$, then $\rho(X) = \psi(-X) = \sup_{\mu \in \mathcal{C}(v)} \mathbf{E}_\mu[-X] = -\inf_{\mu \in \mathcal{C}(v)} \mathbf{E}_\mu[X]$, defines a coherent risk measure.

Definition 7.3. *Two functions* X *and* Y *defined on* Ω *are said to be comonotone if almost surely* $(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geq 0$ *for the product measure* $\mathbb{P} \otimes \mathbb{P}$ *on* $\Omega \times \Omega$.

Schmeidler, [Schm2] proved the following

Theorem 7.4. *The coherent risk measure* ρ *comes from a convex game* v, *i.e.*

$$\rho(X) = \psi(-X) = \sup_{\mu \in \mathcal{C}(v)} \mathbf{E}_\mu[-X],$$

if and only if ρ is comonotone, i.e. $\rho(X+Y) = \rho(X) + \rho(Y)$ for comonotone functions X and Y.

In [D1], section 12, an example of a convex game was given, using so called distorted measures. The example given there, corresponds to the function $f(x) = \frac{1-e^{-x}}{1-e^{-1}}$ of theorem 7.5 below. More precisely the following holds

Theorem 7.5. *If $f\colon [0,1] \to [0,1]$ is an increasing concave function such that $f(0) = 0$ and $f(1) = 1$, then the function $w(A) = f(\mathbb{P}[A])$ defines a 2-alternating set function. If f is continuous (at 0), then the core of the associated game v is a weakly compact set in L^1.*

Proof. We only have to show that the function w is 2-alternating, the rest follows from the general theorems stated above. As we will see below, the function f satisfies: for $0 \le x, y, z$ and $x + y + z \le 1$:

$$f(x+z) + f(y+z) \ge f(z) + f(x+y+z).$$

If we now take $x = \mathbb{P}[A \setminus B], y = \mathbb{P}[B \setminus A], z = \mathbb{P}[A \cap B]$, we get the desired result. The inequality for f can be proved as follows:

$$\begin{aligned} f(x+z) + f(y+z) &= f(x+z) + f(z) + \int_{]z,z+y]} f'(du) \\ &\ge f(x+z) + f(z) + \int_{]x+z,x+z+y]} f'(du) \text{ by concavity of } f \\ &\ge f(z) + f(x+y+z). \end{aligned}$$

□

Corollary 7.6. *For $0 \le \beta \le 1$, the function $f(x) = x^\beta$ defines a 2-alternating set function $w(A) = (\mathbb{P}[A])^\beta$. We also have that $\psi(X) = \int_0^\infty \mathbb{P}[X > x]^\beta \, dx$ for nonnegative functions X. The examples 4.2, 4.5 and 4.7 above therefore define a coherent measure. For $1 \ge \beta > 0$ we have that $\mathcal{P}_\sigma$ is weakly compact.*

Remark 7.7. We state without further proof that for $0 < \alpha < 1$ and for the function defined as $f(x) = x/\alpha$ if $x \le \alpha$ and $f(x) = 1$ for $x \ge \alpha$, the 2-alternating function $w(A) = f(\mathbb{P}[A])$ defines a convex game, that has as its core, the set $\mathcal{P}_\sigma$ of example 4.2. Therefore the risk measure WCM_α is comonotone. The set in example 4.3, cannot be obtained as the core of a convex game. We verify this using comonotonicity and under the extra assumption that the probability space is atomless. Let us, for $\beta > 1$, define

$$\mathcal{P}_\sigma = \left\{ f \mid 0 \le f, \mathbf{E}_\mathbb{P}[f] = 1, \mathbf{E}_\mathbb{P}[f^2] \le \beta^2 \right\}.$$

Take two sets $B \subset A$ such that $\mathbb{P}[B] < \mathbb{P}[A] = \frac{1}{\beta^2}$. If the risk measure were comonotone, then the same function $f \in \mathcal{P}_\sigma$ could be used to calculate $\rho(-\mathbf{1}_A)$ and $\rho(-\mathbf{1}_B)$. But as easily seen, the function $f = \frac{\mathbf{1}_A}{\mathbb{P}[A]}$ gives $\mathbf{E}_\mathbb{P}[f \mathbf{1}_A] = \rho(-\mathbf{1}_A)$, whereas for $-\mathbf{1}_B$ the optimal function is different from f. Indeed some functions of the form $f' = \lambda \mathbf{1}_B + (1-\lambda)$ with $0 < \lambda < 1$ yield a greater value. The reader can check, but this is not necessary to obtain the result, that the optimal function for $-\mathbf{1}_B$ is indeed of the form $\lambda \frac{\mathbf{1}_B}{\mathbb{P}[B]} + (1-\lambda)$, where λ is chosen so that $\|f'\|_2 = \beta$.

In [Pa], Parker investigates the existence of σ–additive elements in the core of a game. This problem is related to our section 3. Especially theorem 3.2 can be translated as follows

Theorem 7.8. *Let v be convex game as described in definition 7.1. Suppose that v is continuous from above, i.e. for each decreasing sequence of sets A_n with intersection A we have that $v(A) = \lim_n v(A_n)$. In that case the σ–additive elements of the core of v, the σ–core $\mathcal{C}^\sigma(v)$, is weak* dense, i.e. $\sigma(\mathbf{ba}, L^\infty)$, in the core $\mathcal{C}(v)$.*

Remark. We remark that the above theorem does not use regularity assumptions. In this sense the theorem generalises the results of [Pa].

Proof. For simplicity and without loss of generality we may suppose that $v(\Omega) = 1$. If X is a nonnegative bounded random variable then we can write

$$\phi(X) = \inf_{\mu \in \mathcal{C}(v)} \mathbf{E}_\mu[X] = \int_0^\infty v(X > x)\, dx.$$

If X_n is a sequence of random variables, $0 \le X_n \le 1$, and tending a.s. to a random variable X, then for each $\varepsilon > 0$ we can write

$$\begin{aligned}
\phi(X) = \int_0^1 v(\{X > x\})\, dx &\ge \int_0^1 \lim_n v\left(\cup_{m \ge n}\{X_m > x + \varepsilon\}\right)\, dx \\
&\ge \int_0^1 \limsup_m v(\{X_m > x+\varepsilon\})\, dx \ge \limsup_m \int_0^1 v(\{X_m > x+\varepsilon\})\, dx \\
&\ge \limsup_m \int_0^1 v(\{X_m > x\})\, dx - \varepsilon \ge \limsup_m \phi(X_m) - \varepsilon.
\end{aligned}$$

The first inequality is an immediate consequence from the inclusion $\{X > x\} \supset \cap_n \cup_{m \ge n} \{X_m > x + \varepsilon\}$ and the continuity of v from above. The second inequality follows from monotonicity of v, the third inequality is Fatou's lemma and the fourth inequality follows via a change of variables $x + \varepsilon$ into x. Since ε was arbitrary we get that ρ satisfies the Fatou property. Theorem 3.2 can now be applied. $\square$

Example 7.9 or example 3.8 again. Let us take $\Omega = [0, 1]$ and $\mathcal{F}$ the Borel sigma algebra on $[0, 1]$. The game v is defined as follows $v(A) = 1$ if A is a set of second category, i.e. contains a dense G_δ set. Otherwise we put $v(A) = 0$. The reader can check that v is a convex game and that v is continuous from above. However if μ is an element of the core, we must have that $\mu(A) = 1$ if A is a set of second category. From this it follows that $\mu(A) = 0$ if the set A is of first category. Exactly as in example 3.8, this shows that μ cannot be σ–additive. It follows that $\mathcal{C}^\sigma(v) = \emptyset$. This shows that some kind of control measure is needed.

In [Pa], Parker refers to [Bi] for an example of a convex game, continuous from above, and such that the σ–core is empty. The example in [Bi] is based on Ulam's work, [Ul]. This work is based on the continuum hypothesis, more precisely it is based on the hypothesis that the continuum is not a weakly measurable cardinal. See also [Ha], remark at the end of exercise 3 of chapter 3, section 16, for a related problem. The present example 7.9, is not based on such assumptions of set theory. The discussions in [Pa] and [Bi] show that statements such as in theorem 7.8, quickly turn into problems of set theory.

The following theorem generalises theorem 7 in [Pa].

Theorem 7.10. *Let v be a convex game defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose that v is continuous from above and absolutely continuous with respect to $\mathbb{P}$. Then for every* *finite chain of elements of $\mathcal{F}$, say $A_1 \supset A_2 \cdots \supset A_n$, there is an element $f \in \mathcal{C}^\sigma(v)$ such that for all $k \le n$ we have $\int_{A_k} f\, d\mathbb{P} = v(A_k)$*

Proof. Without loss of generality we may suppose that for $k \le n-1$ we have that $\mathbb{P}[A_k] > \mathbb{P}[A_{k+1}] > 0$. Also we suppose that $v(\Omega) = 1$. Let g be defined as $g = \sum_{1\le k\le n} \mathbf{1}_{A_k}$. The σ–core $\mathcal{C}^\sigma(v)$ is a closed convex bounded set of $L^1(\mathbb{P})$. Take now $1/4 > \varepsilon > 0$, but otherwise arbitrary. By the Bishop–Phelps theorem, [Di] and [BP], there is a function $h \in L^\infty$ such that $\|g-h\|_\infty < \varepsilon$ and such that $\int hf\, d\mathbb{P}$ attains its minimum on the σ–core. Let f be such that the minimum is attained. Since the minimum is attained and since the σ–core is weak* dense in the core (by theorem 7.8), we have that

$$\int hf\, d\mathbb{P} = \int_0^\infty v(h > x)\, dx.$$

Also since f is in the σ–core and hence in the core, we therefore have that for almost every x, $\int_{\{h>x\}} f\, d\mathbb{P} = v(h > x)$. By the choice of ε, we have that for $k - 3/8 \le x \le k - 1/4$ the sets $\{h > x\}$ and A_k coincide and this implies that for every k, $\int_{A_k} f\, d\mathbb{P} = v(A_k)$. $\square$

Corollary. *If g is $\mathcal{F}$ measurable and if g takes only a finite number of values, then there is $f \in \mathcal{C}^\sigma(v)$ such that*

$$\int gf\, d\mathbb{P} = \min_{\mu\in\mathcal{C}(v)} \mu[g].$$

Remark. The above theorem generalises theorem 7 in [Pa] since it does not use topological regularity of the game v. However there is no hope to generalise the result to infinite chains, compare theorem 5 in [Pa] and lemma 2, corollary 3 of section 2 in [D1]. Indeed such a generalisation would then imply, by the James' characterisation of weakly compact sets, see [Di], that the σ–core is weakly compact. This is true if and only if v is also continuous from below at Ω (and therefore also at every set in $\mathcal{F}$).

8. Some explicit examples

We conclude this paper with some explicit examples that show how different the risk measures can be. Since only the order of convergence is important we do not bother about the fact that distorted probability measures do not satisfy the relations of theorem 5.4. The reader can easily adapt the analysis.

Example 8.1 The one sided normal distribution:. In this case we suppose that $X = -f$, where $\mathbb{P}[f \ge x] = \sqrt{\frac{2}{\pi}} \int_x^{+\infty} e^{-u^2/2}\, du$ for $x \ge 0$. For $x_0 > 0$ given and big enough, we have that $\mathbf{E}[f \mid f \ge x_0]$ is approximately equal to x_0. It follows that if $x_0 = VaR_\alpha$, the quantities Value at Risk and the WCM_α almost coincide. See [EKM] for the details of the calculation. The risk measures of the form $\rho_\beta(X) = \int_0^\infty \mathbb{P}[f \ge x]^\beta\, dx$ behave differently. A quick calculation shows that $\rho_\beta(X)$ is of the order $\frac{1}{\sqrt{\beta}}$. Of course, since normal variables are unbounded, we must have that $\lim_{\beta\to 0} \rho_\beta(X) \to \|X\|_\infty = \infty$. But the order of convergence is only a square root.

Example 8.2 The case of the exponential distribution. Here we assume that $\mathbb{P}[f \geq x] = e^{-x}$ for $x \geq 0$. Since in this case we have that $\mathbf{E}[f \mid f \geq x_0] = x_0 + 1$, it turns out that the Value at Risk and the WCM measures differ by just one unit. The distorted measures $\rho_\beta(X) = 1/\beta$ tend (of course) to ∞ but this time faster than in the case of a normal distribution.

Example 8.3 The case of Pareto like distributions. Here we assume that for $x \geq 0$ we have $\mathbb{P}[f \geq x] = \frac{1}{(x+1)^\alpha}$ for some fixed $\alpha > 0$. Calculus shows us that $\mathbf{E}[f \mid f \geq x_0] = \frac{\alpha+1}{\alpha-1} x_0$ if $\alpha > 1$ and equals ∞ in case $\alpha \leq 1$. This means that WCM is $\frac{\alpha+1}{\alpha-1}$ as big as VaR. For the distorted measures of example 4.5 we find $\rho_\beta(X) = \frac{1}{\alpha\beta-1}$ if $\beta > 1/\alpha$ and ∞ if not. This means that for β small, the risk measure ρ_β gives us ∞ as a result. This coincides with the intuition that for small α the risk involved in Pareto distributions is too high to find an insurance. What small means is subject to personal risk aversion.

References

[ADEH1] Artzner, Ph., F. Delbaen, J.-M. Eber, and D. Heath (1997), *Thinking Coherently*, RISK **10**, November, 68–71.

[ADEH2] Artzner, Ph., F. Delbaen, J.-M. Eber, and D. Heath (1998), *Coherent Measures of Risk*, submitted.

[Ban] Banach, S. (1932), *Théorie des opérations linéaires*, Monografje Matematyczne **1**, Warsawa.

[Bi] Bierlein (1962), *Über die Fortsetzung von Wahrscheinlichkeiten*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **1**, 28–46.

[BP] Bishop, E and Phelps, R.R. (1961), *A proof that every Banach space is subreflexive.*, Bull. Amer. Math. Soc. **67**, 97–98.

[Ch] Choquet, G (1953), *Theory of capacities*, Ann. Inst. Fourier **5**, 131–295.

[D1] Delbaen, F. (1974), *Convex Games and Extreme Points*, Journ. Math. Anal. Appli. **45**, 210–233.

[D2] Delbaen, F. (2000), *Coherent Risk Measures*, Lecture Notes of the Cattedra Galileiana at the Scuola Normale di Pisa, March 2000.

[De1] Denneberg, D. (1989), *Verzerrte Wahrscheinlichkeiten in der Versicherungsmathematik, quantilabhängige Prämienprinzipen.*, Mathematik-Arbeitspapiere, 34, Universität Bremen.

[De2] Denneberg, D. (1997), *Non-Additive Measure and Integral*, Kluwer, Dordrecht.

[Di] Diestel, J. (1975), *Geometry of Banach Spaces-Selected Topics.*, Lecture Notes in Mathematics **485,**, Springer Verlag, Berlin Heidelberg New York.

[DS] Dunford, N. and Schwartz, J. (1958), *Linear Operators, Vol I*, Interscience.

[EKM] Embrechts, P., C. Klüppelberg, and T. Mikosch, (1997), *Modelling Extremal Events*, Springer, New York.

[G] Grothendieck, A. (1973), *Topological Vector Spaces*, Gordon and Breach, New York.

[Ha] Halmos, P (1950), *Measure Theory*, Van Nostrand, Princeton, N.J..

[HG] Haezendonck, J. and M. Goovaerts (1982), *A new premium calculation principle based on Orlicz norms*, Insurance Math. Econom. **1**, 41–53.

[HS] Halmos, P. and Savage, L.J. (1949), *Application of the Radon-Nikodym theorem to the theory of sufficient statistics*, Ann. Math. Stat. **20**, 225–241.

[JaK] Jaschke, S and Küchler, U. (19919), *Coherent Risk Measures, Valuation Bounds, and (μ, σ) Portfolio Optimisation*, http://wws.mathematik.hu-berlin.de/ jaschke/papers/ (preliminary version).

[Ka] Kabanov, Y. (1999), *Hedging and liquidation under transaction costs in currency markets*, Finance and Stochastics **3**, 237–248.

[Lin] Lindenstrauss, J. (1966), *A short proof of Lyapounoff's theorem*, J. of Math. Mech. **4**, 971–972.

[Ma] Maaß, S. (1999) Diplomarbeit Universität Bremen.

[Nam] Namioka, I. (1957), *Partially Ordered Linear Topological Spaces* **24**, Mem. Amer. Math. Soc..

[Nev] Neveu, J. (1975), *Discrete Parameter Martingales*, North Holland, Amsterdam.

[Ni] Nikodym, O. (1931), *Contribution à la théorie des fonctionnelles linéaires en connection avec la théorie de la mesure des ensembles abstraits*, Mathematica **5**, 130–141.

[Pa] Parker, J. (1991), *The Sigma-Core of a Cooperative Game*, Manus. Math. **70**, 247–253.

[Ru] Rudin, W. (1973), *Functional Analysis*, Mc Graw Hill, New York.

[Schm1] Schmeidler, D. (1972), *Cores of Convex Games*, J. Math. Anal. Appli. **40**, 214–225.

[Schm2] Schmeidler, D. (1986), *Integral Representation without Additivity*, Proc. Amer. Math. Soc. **97**, 255–261.

[Ul] Ulam, S.M. (1930), *Zur Maßtheorie in der allgemeinen Mengenlehre*, Fund. Math. **16**, 141–150.

[Wa1] Walley, P. (1991), *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall, London.

[Wa2] Walley, P. (1999), *Towards a Unified Theory of Imprecise Probability, paper presented at the 1st International Symposium on Imprecise Probabilities and Their Applications, Ghent, Belgium June 29–July 2*, This paper is available via ftp://ensmain.rug.ac.be/pub/isipta99/055.pdf.

[Wan] Wang, Shaun (1998), *Comonotonicity, correlation order and premium principles*, Insurance Math. Econom. **22**, 235–242.

[YH] Yosida, K and Hewitt, E. (1952), *Finitely additive measures*, Trans. Amer. Math. Soc. **72**, 46–66.

Departement für Mathematik, Eidgenössische Technische Hochschule Zürich, CH 8092 Zürich, Schweiz
E-mail address: `delbaen@math.ethz.ch`