Continuous Linear Representation of Coherent Lower Previsions

SEBASTIAN MAASS
Universität Bremen, Germany

Abstract

This paper studies the possibility of representing lower previsions by continuous linear functionals. We prove the existence of a linear isomorphism between the linear space spanned by the coherent lower previsions and that of an appropriate space of continuous linear functionals. Moreover, we show that a lower prevision is coherent if and only if its transform is monotone. We also discuss the interpretation of these results and the new light they shed on the theory of imprecise probabilities.

Keywords

coherent lower previsions, Möbius transform, Choquet's theorem, Bishop-de Leeuw theorem, Dempster-Shafer-Shapley representation theorem

1 Introduction

The theory of imprecise probabilities especially that of coherent lower previsions has been designed to mathematically cope with subjective behavior in decision situations (cf. Walley [11]). It has evolved so extensively that coherent lower previsions have been repeatedly reinvented under different names like e.g. "coherent risk measures" (cf. Delbaen [4]) or "maxmin expected utility" (cf. Gilboa and Schmeidler [7]).

From an applicational and often also mathematical point of view nonlinear functionals like coherent lower previsions cannot as nice be handled as (monotone) continuous linear functionals. So, in this paper, we are interested in representing the former functionals by the latter. For nonadditive set functions such a representation is well-known as Dempster-Shafer-Shapley Representation Theorem in the discrete case or as Möbius transform in the general case (cf. Denneberg [5], Gilboa and Schmeidler [6] and Marinacci [10]).

The main steps of constructing such a transformed set function run as follows. First, given a totally monotone set function ν on an algebra $\mathcal{A}$, to every set $A \in \mathcal{A}$ is assigned a function $\tilde{A}$ on the extreme points of the convex set of normalized totally monotone set functions and defined by $\tilde{A}(\eta) := \eta(A)$. Since the extreme points are the filter games and therefore $\{0,1\}$-valued (cf. Choquet [3] p. 260 f.) all $\tilde{A}$ can be interpreted as characteristic functions of the sets $\{\eta \mid \eta(A) = 1\}$. Then, by different methods, it can be shown that there exists a bijective mapping from the set of totally monotone set functions to the set of (positive) measures on the σ-algebra generated by the $\tilde{A}$, $A \in \mathcal{A}$. Finally, this bijective mapping can be extended to the linear spaces each spanned by the respective class of set functions.

In this paper we will show that the main results of these theorems do not presuppose the functions being totally monotone set functions. Even a structured domain like an algebra is not necessary to obtain analogous results for coherent lower previsions. In our main theorem (Theorem 2) we provide a representation theorem for coherent lower previsions which contains results analogously to those sketched in the preceding paragraph for totally monotone set functions.

2 Preliminaries

Let Ω be a nonempty set, $B(2^\Omega)$ the linear space of bounded (w.r.t. the supremum norm) real-valued functions on Ω and $K \subset B(2^\Omega)$ be nonempty. To avoid laborious considerations of special cases, we will assume that there is at least one nonzero function in K. A *lower prevision* on K is a real-valued functional $\underline{P} : K \to \mathbb{R}$. A lower prevision $\underline{P}$ is called *coherent*, if $\underline{P}(f) \geq \sum_{i=1}^{n} \lambda_i \underline{P}(f_i) + \lambda_0$ whenever $f \geq \sum_{i=1}^{n} \lambda_i f_i + \lambda_0$ with $f, f_i \in K$, $\lambda_i > 0$, $\lambda_0 \in \mathbb{R}$, $n \in \mathbb{N}$. This definition is not the usual one (cf. Walley [11, Definition 2.5.1]) but it follows immediately from Proposition 3.1.2 (d) and Lemma 3.1.3 (b) in Walley's book and it will be of use to prove a functional being *not* coherent. Furthermore, this characterization of coherence can nicely be interpreted in the following way. As usual, Ω denotes a possibility space, K a set of *gambles*, i.e. positive or negative rewards depending on the uncertain state $\omega \in \Omega$. A lower prevision $\underline{P}$ of a gamble f is then the supremum buying price for f one is willing to pay. Since the system of buying prices have to fulfill some justified consistency properties, Walley introduced the notion of coherence which, using the characterization given above, means that, whenever a gamble f is dominating a portfolio of other gambles (possibly including a sure gain or loss λ_0) independently of the state ω, one should be willing to pay at least as much for f as one is willing to pay for the individual gambles included in the portfolio (not for the portfolio as whole - this would be considered as one gamble).

If K consists of characteristic functions then $\underline{P}$ can be interpreted as a set function and then is called a *coherent lower probability*. We have shown in [8] that the normalized exact games in cooperative game theory are the coherent lower probabilities. Simple examples of coherent lower probabilities are unanimity games, i.e. set functions u_A on an algebra $\mathcal{A}$ with $A \in \mathcal{A}$ and $u_A(B) = 1$ if $B \supset A$ and 0 else. If K is a linear space containing constant functions then $\underline{P}$ is a coherent lower prevision if and only if it is monotone, positively homogeneous, superadditive, normalized (i.e. $\underline{P}(1) = 1$) and constant additive (i.e. $\underline{P}(f+c) = \underline{P}(f)+\underline{P}(c)$). This characterization is almost equivalent to that of "coherent risk measures" (cf. Delbaen [4] and Maaß [9]). It is well-known (cf. Walley [11], Chapter 3) that every coherent lower prevision can be extended coherently to the linear space of all bounded real-valued functions. The minimum of all such extensions exists and is called the *natural extension*. Since coherence implies that every function $f \in K$ is mapped into the bounded interval $\inf f, \sup f$, $CLP(K)$ is contained in $BLP(K)$. Denote by $BLP(K)$ the linear space of all lower previsions on K which are bounded w.r.t. the operator norm $\|\cdot\|$, $\|\underline{P}\| := \sup_{f\in K, f\neq 0} \frac{|\underline{P}(f)|}{\|f\|_\infty}$, and by $CLP(K)$ the convex set of all coherent lower previsions on K.

The linear space $BLP(K)$ will additionally be considered as a topological space endowed with the topology $\mathcal{T}$ having as subbase the sets $B(\underline{P}, f, \varepsilon) := \{\underline{P}' \in BLP(K) \mid |\underline{P}'(f) - \underline{P}(f)| < \varepsilon\}$, with $\underline{P} \in BLP(K)$, $f \in K$, $\varepsilon > 0$. The definition of $\mathcal{T}$ is similar to that of the weak* topology and it is the smallest making all functions

$$\tilde{f} : BLP(K) \to \mathbb{R}, \quad \tilde{f}(\underline{P}) := \underline{P}(f)$$

continuous for all $f \in K$. The set of all such $\tilde{f}$ will be denoted by $\tilde{K}$, the linear space spanned by $\tilde{K}$ will be denoted by $\mathrm{span}(\tilde{K})$. The topology $\mathcal{T}$ is also known as the topology of pointwise convergence and, by definition of the product topology, $\mathcal{T}$ is identical with the relative topology of $BLP(K)$ as a subset of the product space $\Pi_{f\in K}\mathbb{R}_f$, $\mathbb{R}_f := \mathbb{R}$ for all $f \in K$.

We start with some topological results that will serve as technical basis for the following analysis.

Proposition 1 *Under the topology $\mathcal{T}$ the linear space $BLP(K)$ is a locally convex and Hausdorff topological linear space.*

Proof. We have to show that $\mathcal{T}$ possesses a base consisting of convex sets. Since convexity is preserved under forming intersections it suffices to show that the given subbase of $\mathcal{T}$ consists of convex sets. Therefore, suppose $\underline{P}_1, \underline{P}_2 \in B(\underline{P}, f, \varepsilon)$ with $\underline{P} \in BLP(K)$, $f \in K$ and $\varepsilon > 0$ and let $\lambda \in [0,1]$. Then

$$|\lambda\underline{P}_1(f) + (1-\lambda)\underline{P}_2(f) - \underline{P}(f)| \leq \lambda|\underline{P}_1(f) - \underline{P}(f)| + (1-\lambda)|\underline{P}_2(f) - \underline{P}(f)| < \varepsilon,$$

i.e. $B(\underline{P}, f, \varepsilon)$ is convex since $\underline{P}_1, \underline{P}_2$ and λ were chosen arbitrarily. Hence, all elements of the subbase are convex since $\underline{P}, f$ and ε were chosen arbitrarily. □

Proposition 2 *The unit ball in $(BLP(K), \|\cdot\|)$, $B := \{\underline{P} \in BLP(K) \mid \|\underline{P}\| \leq 1\}$, is $\mathcal{T}$-compact.*

Proof. Let $I := \Pi_{f\in K}[-1,1]$. By Tychonoff's Theorem, I is compact w.r.t. the product topology. Let $\tau : B \to I$ be the injective mapping $\tau(\underline{P}) := \Pi_{f\in K} \frac{\underline{P}(f)}{\|f\|_\infty}$. Since the sets $B(\underline{P}, f, \varepsilon) := \{\underline{P}' \in B \mid |\underline{P}'(f) - \underline{P}(f)| < \varepsilon\}$ with $\underline{P} \in B$, $f \in K$, $\varepsilon > 0$ form a subbase for the relative topology $\mathcal{T}_B$ of B generated by $\mathcal{T}$ and since $\{\Pi_{f\in K} U_f \mid U_f = \mathbb{R}\ \forall f \in K \setminus \{f'\}, U_{f'} =]x-\varepsilon, x+\varepsilon[, f' \in K, x \in \mathbb{R}, \varepsilon > 0[\}$ is a subbase of the product topology in $\mathbb{R}^K$, the images of the $\mathcal{T}_B$-subbase elements of $\mathcal{T}_B$ form a subbase of the relative product topology in $\tau(B)$. Thus τ is a homeomorphism between B endowed with the relative $\mathcal{T}$-topology, and $\tau(B)$ endowed with the relative product topology. Therefore, to prove that B is $\mathcal{T}$-compact, it suffices to show that B is $\mathcal{T}$-closed. This is easily done since for any $\underline{P} \in BLP(K)$ with $\|\underline{P}\| > 1$ there exist a $f \in K$ and a $\varepsilon > 0$ with $|\underline{P}(f)| > \|f\|_\infty + \varepsilon$ such that $B(\underline{P}, f, \varepsilon)$ is an open neighborhood of $\underline{P}$ disjoint from B, i.e. B is $\mathcal{T}$-closed. □

Proposition 3 *The set $CLP(K)$ is $\mathcal{T}$-compact in $BLP(K)$.*

Proof. Obviously, $CLP(K)$ is a subset of the $\mathcal{T}$-compact set B. So, it remains to prove that $CLP(K)$ is $\mathcal{T}$-closed. Suppose $\underline{P}$ is a noncoherent lower prevision. Then there exist $f, f_i \in K$, $\lambda_i > 0$, $\lambda_0 \in \mathbb{R}$, $i \in \{1, \ldots, n\}$ and $\varepsilon > 0$ with $f \geq \sum_{i=1}^n \lambda_i f_i + \lambda_0$ and $\underline{P}(f) + \varepsilon < \sum_{i=1}^n \lambda_i \underline{P}(f_i) + \lambda_0$. Setting $\varepsilon_i := \varepsilon/(2\sum_{k=1}^n \lambda_k)$, the set $B(\underline{P}, f, \frac{1}{2}\varepsilon) \cap \bigcap_{i=1}^n B(\underline{P}, f_i, \varepsilon_i)$ is an open neighborhood of $\underline{P}$ which is disjoint from $CLP(K)$. Hence, $CLP(K)$ is $\mathcal{T}$-compact. □

The main result of this paper will heavily base on the Bishop-de Leeuw Theorem (cf. Alfsen [1, Theorem I.4.14]) which, like Choquet's Theorem, belongs to a group of results generalizing the famous Krein-Milman Theorem. We recall that the Baire σ-algebra is the smallest σ-algebra for which all continuous real-valued functions are measurable, with, as usual, the Borel σ-algebra on the range space $\mathbb{R}$. Furthermore, denote by $\mathrm{ex}(X)$ the set of extreme points of X.

Theorem 1 (Bishop-de Leeuw) *Suppose E is a locally convex Hausdorff space over $\mathbb{R}$ and X a nonempty compact convex subset of E. Denote by $A(X)$ the linear space of continuous real-valued functions $a : X \to \mathbb{R}$ which are affine, i.e. $a(\lambda x + (1-\lambda)y) = \lambda a(x) + (1-\lambda)a(y)$ for $x, y \in X$, $0 \leq \lambda \leq 1$ and by $\mathcal{B}_0$ the Baire σ-algebra on X. Then for every $x \in X$ there exists a probability measure μ_x on the σ-algebra $\mathrm{ex}(X) \cap \mathcal{B}_0$, such that*

$$a(x) = \int a\, d\mu_x \quad \textit{for all } a \in A(X). \tag{1}$$

Generally, it is not possible to replace the Baire σ-algebra by the more usual Borel σ-algebra (cf. Alfsen [1, p. 39 f.]).

3 Main Results

In this section, we present the announced isomorphism between the linear space spanned by $CLP(K)$ and a linear space of continuous linear functionals and characterize the previsions in $CLP(K)$ by monotonicity of their transform. As a preparation, we start with a simple application of the Bishop-de Leeuw Theorem.

Lemma 1 *For every coherent lower prevision $\underline{P}$ on K there exists a probability measure $\mu_{\underline{P}}$ on the σ-algebra $\mathrm{ex}(CLP(K)) \cap \mathcal{B}_0$, such that*

$$\underline{P}(f) = \int \tilde{f}\, d\mu_{\underline{P}} \quad \text{for all } f \in K. \tag{2}$$

Proof. The assertion made in the lemma follows directly from Theorem 1 using Proposition 1 and 3 and from $\tilde{f} \in A(CLP(K))$ for all $f \in K$. □

We obviously have found that the continuous linear functional $\int \cdot\, d\mu_{\underline{P}}$ represents the coherent lower prevision $\underline{P}$ via the nonlinear application $f \mapsto \tilde{f}$. Unfortunately, the representing measure $\mu_{\underline{P}}$ needs not to be unique as the following example shows.

Example 1 *Let $\Omega = \{1,2,3\}$ and $\nu : 2^\Omega \to \mathbb{R}$ be the coherent lower probability defined by $\nu(A) := \frac{1}{2}$ iff $|A| = 2$ and $\nu(A) := 0$ iff $|A| < 2$. Then ν is an extreme point of the set of coherent lower probabilities on 2^Ω, $CLP(2^\Omega)$[1]. Suppose ν is a convex combination of two coherent lower probabilities ν_1 and ν_2 Obviously, $\nu_1(A) = \nu_2(A) = \nu(A)$ for all A with $\nu(A) \in \{0,1\}$, i.e. $|A| \neq 2$. Therefore, suppose $\nu_1(\{1,2\}) > \nu(\{1,2\}) = \frac{1}{2}$. By coherence of ν_1, $1_{\{1\}} \geq 1_{\{1,2\}} + 1_{\{1,3\}} - 1$ implies $\nu_1(\{1\}) \geq \nu_1(\{1,2\}) + \nu_1(\{1,3\}) - 1$ such that $\nu_1(\{1,3\}) < \frac{1}{2}$. Analogously, we conclude $\nu_1(\{2,3\}) < \frac{1}{2}$. The same argument applied to ν_2 implies that both ν_1 and ν_2 are at least for two of three sets A with $|A| = 2$ smaller than or equal to $\nu(A)$. Hence, $\nu_1 = \nu_2 = \nu$.*
Further on, it is easy to see that all unanimity games on 2^Ω are extreme points of $CLP(2^\Omega)$.
The coherent lower probability $\nu' : 2^\Omega \to \mathbb{R}$ defined by $\nu'(A) := \frac{1}{3}$ iff $|A| = 2$ and $\nu'(A) := 0$ iff $|A| < 2$ can be obtained by two different convex combinations of extreme points of $CLP(2^\Omega)$, $\nu' = \frac{1}{3}u_{\{1,2\}} + \frac{1}{3}u_{\{1,3\}} + \frac{1}{3}u_{\{2,3\}}$ and $\nu' = \frac{2}{3}\nu + \frac{1}{3}u_\Omega$. Since the coefficients of the extreme points used in the convex combinations are the masses of the transform $\mu_{\nu'}$ of ν', we obtain that uniqueness of the representing measure cannot be guaranteed.

To obtain uniqueness, we have to draw our attention to the integrals because for two representing measures $\mu_{\underline{P}}$ and $\mu'_{\underline{P}}$ of $\underline{P}$ we have, by Lemma 1,

$$\int \tilde{f}\, d\mu_{\underline{P}} = \int \tilde{f}\, d\mu'_{\underline{P}} \quad \text{for all } f \in K. \tag{3}$$

So, if we just restrict the continuous linear functional $\int \cdot\, d\mu_{\underline{P}}$ to the linear space $\mathrm{span}(\tilde{K})$ we get the desired uniqueness.

[1] It can be shown that ν is the only non-unanimity game in the set of extreme points of $CLP(2^\Omega)$.

The subsequent lemma (cf. Maaß [9, Proposition 6]) is mainly for technical use in the proof of the following theorem. As will be discussed after Lemma 3, it can be of practical use.

Lemma 2 *Let $\{\underline{P}_i\}_{i\in I}$ be a nonempty indexed set of coherent lower previsions on $K \subset B(2^\Omega)$ and the lower prevision $\underline{P}: B(2^I) \to \mathbb{R}$ be coherent. Then the functional*

$$K \to \mathbb{R}, \quad f \mapsto \underline{P}(i \mapsto \underline{P}_i(f)) \tag{4}$$

is a coherent lower prevision.

Proof. The functional defined in (4) is well defined since coherence of the $\underline{P}_i$ implies $-\infty < \inf f \leq \underline{P}_i(f) \leq \sup f < \infty$ such that the function $i \mapsto \underline{P}_i(f)$ is bounded for every $f \in K$. By considering the natural extensions $\underline{E}_i$ of $\underline{P}_i$, coherence is easily verified for the functional $B(2^\Omega) \to \mathbb{R}$, $f \mapsto \underline{P}(i \mapsto \underline{E}_i(f))$ by using the characterization of coherence on linear spaces, and therefore for its restriction to K as defined in (4). □

This rather abstract lemma can be used to prove results which were formulated as individual theorems in Walley's book (cf. Walley [11, 2.6.3 - 2.6.7] and Maaß [8, Corollary 4.2]). The following lemma generalizes one of these results, namely that convex combinations of coherent lower previsions are again coherent.

Lemma 3 *Let $X \subset CLP(K)$, $\mathcal{A}$ be a σ-algebra over X making all $\tilde{f}$ measurable and μ be a probability measure on $\mathcal{A}$. Then the lower prevision*

$$\underline{P}: K \to \mathbb{R}, \quad \underline{P}(f) := \int \tilde{f}\, d\mu \tag{5}$$

is coherent.

Proof. The integral $\int \cdot\, d\mu$ is of course coherent and applies to functions $X \to \mathbb{R}$, $\underline{P}' \mapsto \tilde{f}(\underline{P}') = \underline{P}'(f)$. Applying Lemma 2 yields the desired result. □

Before proceeding with the main issue of this paper, a possible application of Lemma 2 should be sketched. Suppose I is a nonempty set of persons assigning values in a coherent way to all gambles $f \in K$, i.e. $\{\underline{P}_i\}_{i\in I}$ is an indexed set of coherent lower previsions. Furthermore, suppose we also want to assign values coherently to all $f \in K$ just by incorporating the $\underline{P}_i$. Using the already cited well-known theorems (cf. Walley [11, 2.6.3 - 2.6.7]), we could take the lower envelope of all $\underline{P}_i$, $\inf_{i\in I} \underline{P}_i$, as our coherent lower prevision if we were very cautious. If we had certain opinions on the coherent lower previsions of all persons we also could assign weights λ_i to every $\underline{P}_i$ and take $\sum_{i\in I} \lambda_i \underline{P}_i$ as our coherent lower prevision (cf. Lemma 3). But using Lemma 2 we can go even further. We can assign weights $\mu(J)$ to "coalitions" $J \subset I$ in order to express that if certain persons agree on the evaluation of some gamble f this should count more than the evaluations of other persons. If this set function μ is supermodular then the Choquet integral $\int \cdot\, d\mu$ is coherent and, by Lemma 2, so is the lower prevision $f \mapsto \int (i \mapsto \underline{P}_i)\, d\mu$

By merely collecting the results from Lemma 1, the remarks following Example 1 (especially Equation (3)) and Lemma 3, we obtain the subsequent proposition which contains the essential mathematical part of the main theorem of this paper (Theorem 2).

Proposition 4 *The mapping*

$$\begin{aligned} CLP(K) &\to \Big\{ \big(\int \cdot\, d\mu \big)_{|\mathrm{span}(\tilde{K})} \mid \mu : \mathrm{ex}(CLP(K)) \cap \mathcal{B}_0 \to \mathbb{R} \textit{ probability measure} \Big\}, \\ \underline{P} &\mapsto \big(\int \cdot\, d\mu_{\underline{P}} \big)_{|\mathrm{span}(\tilde{K})} \end{aligned} \tag{6}$$

with $\underline{P}(f) = \int \tilde{f}\, d\mu_{\underline{P}}$ for all $f \in K$ is bijective.

We now expand this first result to the linear spaces spanned by the respective sets used in Proposition 4. Thus, denote by

$$V_1 := \Big\{ \lambda_1 \underline{P}_1 - \lambda_2 \underline{P}_2 \mid \lambda_1, \lambda_2 \geq 0, \underline{P}_1, \underline{P}_2 \in CLP(K) \Big\} \tag{7}$$

the linear space of functionals spanned by $CLP(K)$ and by

$$V_2 := \Big\{ \big(\int \cdot\, d\mu \big)_{|\mathrm{span}(\tilde{K})} \mid \mu : \mathrm{ex}(CLP(K)) \cap \mathcal{B}_0 \to \mathbb{R} \text{ of bounded variation} \Big\}. \tag{8}$$

the linear space of restricted integrals w.r.t. signed measures on $\mathrm{ex}(CLP(K)) \cap \mathcal{B}_0$ of bounded variation. Let V_1 be endowed with $\mathcal{T}_{V_1}$, the relative topology of V_1 generated by $\mathcal{T}$, i.e. the smallest topology making all $\tilde{f}$ restricted to V_1 continuous and let V_2 be endowed with $\mathcal{T}_{V_2}$, the weak* topology, i.e. the smallest topology making all natural embeddings $\tilde{\tilde{f}} : V_2 \to \mathbb{R}$, $\tilde{\tilde{f}}((\int \cdot\, d\mu)_{|\mathrm{span}(\tilde{K})}) := \int \tilde{f}\, d\mu$ continuous. Further on, let the norm $\| \cdot \|_{V_1}$ be defined by

$$\|\underline{P}\|_{V_1} := \inf\{\lambda_1 + \lambda_2 \mid \underline{P} = \lambda_1 \underline{P}_1 - \lambda_2 \underline{P}_2, \lambda_1, \lambda_2 \geq 0, \underline{P}_1, \underline{P}_2 \in CLP(K)\} \tag{9}$$

and the norm $\| \cdot \|_{V_2}$ be analogously to $\| \cdot \|_{V_1}$ defined by

$$\begin{aligned} &\| \big(\int \cdot\, d\mu \big)_{|\mathrm{span}(\tilde{K})} \|_{V_2} \\ &:= \inf \big\{ \lambda_1 + \lambda_2 \mid \big(\int \cdot\, d\mu \big)_{|\mathrm{span}(\tilde{K})} = \lambda_1 \big(\int \cdot\, d\mu_1 \big)_{|\mathrm{span}(\tilde{K})} - \lambda_2 \big(\int \cdot\, d\mu_2 \big)_{|\mathrm{span}(\tilde{K})}, \\ &\qquad \lambda_1, \lambda_2 \geq 0, \mu_1, \mu_2 \text{ probability measures} \big\}. \end{aligned} \tag{10}$$

We defer the easy but technical proof of $\| \cdot \|_{V_1}$ and $\| \cdot \|_{V_2}$ really being norms to the end of this section and just proceed with the main result.

Theorem 2 *There is a linear isomorphism J^* between the linear spaces V_1 and V_2. The isomorphism is determined by the identity*

$$\underline{P}(f) = \int \tilde{f}\, d\mu \quad \text{for all } f \in K. \tag{11}$$

The isomorphism J^* is topological, i.e. a homeomorphism, between the topological spaces $(V_1, \mathcal{T}_{V_1})$ and $(V_2, \mathcal{T}_{V_2})$. The isomorphism J^* is isometric between the normed spaces $(V_1, \|\cdot\|_{V_1})$ and $(V_2, \|\cdot\|_{V_2})$. Moreover, $\underline{P}$ is coherent if and only if its transformed is monotone.

Proof. To prove that J^* is well defined, it suffices to show that for every $\underline{P} \in V_1$ there is a measure $\mu : \mathrm{ex}(CLP(K)) \cap \mathcal{B}_0 \to \mathbb{R}$ of bounded variation with $\underline{P}(f) = \int \tilde{f}\, d\mu$ for all $f \in K$ because uniqueness of the image is guaranteed by Equation (11). Suppose $\underline{P} = \lambda_1 \underline{P}_1 - \lambda_2 \underline{P}_2$ with $\lambda_1, \lambda_2 \geq 0$ and $\underline{P}_1, \underline{P}_2 \in CLP(K)$. Then, by Proposition 4, there exist probability measures μ_1, μ_2 on $\mathrm{ex}(CLP(K)) \cap \mathcal{B}_0$ satisfying $\underline{P}_1(f) = \int \tilde{f}\, d\mu_1$ and $\underline{P}_2(f) = \int \tilde{f}\, d\mu_2$ for all $f \in K$. Thus,

$$\underline{P}(f) = \lambda_1 \underline{P}_1(f) - \lambda_2 \underline{P}_2(f) = \lambda_1 \int \tilde{f}\, d\mu_1 - \lambda_2 \int \tilde{f}\, d\mu_2 = \int \tilde{f}\, d(\lambda_1 \mu_1 - \lambda_2 \mu_2) \tag{12}$$

for all $f \in K$, i.e. J^* is well defined. Injectivity of J^* directly follows from Equation (11) since $\underline{P}_1 \neq \underline{P}_2$, $\underline{P}_1, \underline{P}_2 \in V_1$, implies $\int \tilde{f}\, d\mu_1 \neq \int \tilde{f}\, d\mu_2$ for all $f \in K$ with $\underline{P}_1(f) \neq \underline{P}_2(f)$ and μ_1 resp. μ_2 satisfying Equation (11) for $\underline{P}_1$ resp. $\underline{P}_1$. Since, by the Hahn-Jordan Decomposition Theorem, every measure μ of bounded variation can be decomposed into a difference $\mu = \lambda_1 \mu_1 - \lambda_2 \mu_2$, $\lambda_i \geq 0$, μ_i probability measures, $i \in \{1, 2\}$, we obtain surjectivity of J^* simply by reading Equation (12) from right to left, again using Proposition 4. Linearity of J^* is rather obvious. So, we have shown that J^* is a linear isomorphism between the linear spaces V_1 and V_2.

By setting $X := K$ and $V := V_1$ in the subsequent Proposition 5, it follows immediately that J^* also is a homeomorphism between the topological spaces $(V_1, \mathcal{T}_{V_1})$ and $(V_2, \mathcal{T}_{V_2})$.

For proving isometry of J^*, we observe that any decomposition of $J^*(\underline{P})$, $J^*(\underline{P}) = \lambda_1 \big(\int \cdot\, d\mu_1\big)_{|\mathrm{span}(\tilde{K})} - \lambda_2 \big(\int \cdot\, d\mu_2\big)_{|\mathrm{span}(\tilde{K})}$, with $\lambda_1, \lambda_2 \geq 0$, μ_1, μ_2 probability measures, directly corresponds to a decomposition of $\underline{P}$ by Proposition 4, $\underline{P} = \lambda_1 \underline{P}_{\mu_1} - \lambda_2 \underline{P}_{\mu_2}$. Therefore, the infima in the respective definitions of $\|\cdot\|_{V_1}$ and $\|\cdot\|_{V_2}$ are taken over the same sets, i.e. $\|J^*(\underline{P})\|_{V_2} = \|\underline{P}\|_{V_1}$ for all $\underline{P} \in V_1$. $\square$

We now provide the deferred, fairly general proposition used in Theorem 2.[2]

[2] This proposition can also be used to prove that the isomorphism between the linear spaces respectively spanned by the totally monotone set functions and the signed bounded Borel measures (cf. Marinacci [10, Theorem 3]) is a homeomorphism. Marinacci proved homeomorphy only for the respective unit balls (w.r.t. the norm which is not compatible to the topology) instead of the whole spaces.

Proposition 5 *Let X be a nonempty set and V a linear space of real-valued functions on X. Define*

$$\begin{aligned}
\tilde{X} &:= \{\tilde{x} : V \to \mathbb{R} \mid \tilde{x}(v) := v(x), x \in X\}, && (13)\\
\tilde{V} &:= \{\tilde{v} : \tilde{X} \to \mathbb{R} \mid \tilde{v}(\tilde{x}) := \tilde{x}(v), v \in V\}, && (14)\\
\tilde{\tilde{X}} &:= \{\tilde{\tilde{x}} : \tilde{V} \to \mathbb{R} \mid \tilde{\tilde{x}}(\tilde{v}) := \tilde{v}(\tilde{x}), \tilde{x} \in \tilde{X}.\} && (15)
\end{aligned}$$

Endow $\mathcal{T}_V$ with the smallest topology on V making all $\tilde{x} \in \tilde{X}$ continuous and endow $\mathcal{T}_{\tilde{V}}$ with the smallest topology on $\tilde{V}$ making all $\tilde{\tilde{x}} \in \tilde{\tilde{X}}$ continuous.
Then $J : V \to \tilde{V}$, $v \mapsto \tilde{v}$ is a linear topological isomorphism.

Proof. Linearity and injectivity of J is easily verified by successively applying the definitions of $\tilde{V}$ and $\tilde{X}$. Additionally, by definition of $\tilde{V}$, J is surjective. For proving J being a homeomorphism it suffices to show that the elements of the respective subbase of $\mathcal{T}_V$, $\{\tilde{x}^{-1}(O) \mid O \subset \mathbb{R} \text{ open}\}$ and $\mathcal{T}_{\tilde{V}}$, $\{\tilde{\tilde{x}}^{-1}(O) \mid O \subset \mathbb{R} \text{ open}\}$, are mapped onto each other as preimages under J and J^{-1}. This follows almost directly from the above definitions since

$$\begin{aligned}
J^{-1}(\tilde{\tilde{x}}^{-1}(O)) &= J^{-1}(\{\tilde{v} \mid \tilde{\tilde{x}}(\tilde{v}) \in O\})\\
&= J^{-1}(\{\tilde{v} \mid \tilde{x}(v) \in O\})\\
&= J^{-1}(\{\tilde{v} \mid v \in \tilde{x}^{-1}(O)\})\\
&= \tilde{x}^{-1}(O)
\end{aligned}$$

and analogously $J(\tilde{x}^{-1}(O)) = \tilde{\tilde{x}}^{-1}(O)$. □

We end this section with a lemma proving the function $\|\cdot\|_{V_1}$ and $\|\cdot\|_{V_2}$ in fact being norms.

Lemma 4 *The functions $\|\cdot\|_{V_1}$ and $\|\cdot\|_{V_2}$ are norms on the respective spaces.*

Proof. Obviously, $\|0\|_{V_1} = 0$. Now suppose $\|\underline{P}\| = 0$ for a $\underline{P} \in V_1$. Since all $f \in K$ are bounded and since every coherent lower prevision maps f into the bounded interval $[\inf f, \sup f]$, we obtain $|\underline{P}(f)| < \varepsilon$ for every $\varepsilon > 0$, i.e. $\underline{P} = 0$. Further on, for all $\underline{P} \in V_1$ and $c \in \mathbb{R}$, $c \neq 0$,

$$\begin{aligned}
\|c\underline{P}\| &= \inf\left\{\lambda_1 + \lambda_2 \mid c\underline{P} = \lambda_1 \underline{P}_1 - \lambda_2 \underline{P}_2, \lambda_1, \lambda_2 \geq 0, \underline{P}_1, \underline{P}_2 \in CLP(K)\right\}\\
&= \inf\left\{|c|\tfrac{\lambda_1+\lambda_2}{|c|} \mid \underline{P} = \tfrac{\lambda_1}{c}\underline{P}_1 - \tfrac{\lambda_2}{c}\underline{P}_2, \lambda_1, \lambda_2 \geq 0, \underline{P}_1, \underline{P}_2 \in CLP(K)\right\}\\
&= |c| \inf\left\{\tfrac{\lambda_1}{|c|} + \tfrac{\lambda_2}{|c|} \mid \underline{P} = \tfrac{\lambda_1}{|c|}\underline{P}_1 - \tfrac{\lambda_2}{|c|}\underline{P}_2, \lambda_1, \lambda_2 \geq 0, \underline{P}_1, \underline{P}_2 \in CLP(K)\right\}\\
&= |c| \cdot \|\underline{P}\|.
\end{aligned}$$

Finally, the triangle inequality holds because whenever $\underline{P} = \underline{P}_1 + \underline{P}_2$ with $\underline{P}, \underline{P}_1, \underline{P}_2 \in V_1$, $\underline{P}_1 = \lambda_{1,1}\underline{P}_{1,1} - \lambda_{1,2}\underline{P}_{1,2}$, $\underline{P}_2 = \lambda_{2,1}\underline{P}_{2,1} - \lambda_{2,2}\underline{P}_{2,2}$, $\underline{P}_{i,j} \in CLP(K)$

and $\lambda_{i,j} \geq 0$ with $i, j \in \{1,2\}$ then $\underline{P} = (\lambda_{1,1}\underline{P}_{1,1} + \lambda_{2,1}\underline{P}_{2,1}) - (\lambda_{1,2}\underline{P}_{1,2} + \lambda_{2,2}\underline{P}_{2,2})$ holds whereat $(\lambda_{1,1}\underline{P}_{1,1} + \lambda_{2,1}\underline{P}_{2,1}), (\lambda_{1,2}\underline{P}_{1,2} + \lambda_{2,2}\underline{P}_{2,2}) \in CLP(K)$. Therefore, $\|\underline{P}_1 + \underline{P}_2\| \leq \|\underline{P}'\| + \|\underline{P}''\|$. □

4 Summary, Outlook and Open Problems

In this paper we have presented a linear isomorphism between the linear space V_1, spanned by the coherent lower previsions on an arbitrary nonempty set K and an appropriate linear space V_2 of continuous linear functionals. Thereby, we have shown that the famous representation theorems for totally monotone set functions do not depend on this special class, not even on the structure of the domain.

For applications, we are heavily interested in transformations of coherent lower previsions that can practically be handled. It is well-known that the set of extreme points of the set of normalized totally monotone set functions on a finite algebra is finite and consists of all unanimity games which is a finite set. Therefore, every totally monotone set function on a finite domain can be represented as a convex combination of unanimity games. It remains as an open problem to determine the set of extreme points of $CLP(K)$ for a given K. Additionally, for possible application of Theorem 2, it remains as an open problem what condition K has to meet in order to make $\mathrm{ex}(CLP(K))$ finite.

Theorem 2 can be used to construct coherent lower previsions in the following way. After determining the extreme points of the convex set of coherent lower previsions any coherent lower prevision can be obtained by assigning weights to all extreme points. There is an analogous situation in Dempster-Shafer Theory where these weights are called "basic probability assignments". So, by working on the set of extreme points of $CLP(K)$ with linear functionals, things are getting easier and often more applicable.

Finally, we will outline why the transform given in Theorem 2 should not be called "Möbius transform" like in the case of totally monotone set functions. On an algebra $\mathcal{A}$ the zeta function can be expressed in terms of unanimity games, $\zeta(A,B) := u_A(B)$. In the case of considering totally monotone set functions on a finite algebra instead of coherent lower previsions (cf. Denneberg [5], Gilboa and Schmeidler [6] and Marinacci [10]), the integrand of Equation (11) is always a zeta function because the set of extreme points of the set of normalized totally monotone set functions consists of all unanimity games. This gives rise to call the two set functions appearing in the transformation equation the zeta transform resp., since the zeta function and the Möbius function are mutually inverse, the Möbius transform of the respective other set function. Since we have seen in Example 1 that the set of extreme points of $CLP(\mathcal{A})$ contains more than unanimity games the interpretation of using zeta functions can not be preserved such that the term "Möbius transform" can not be justified in our case.

Acknowledgements

The author is indebted to D. Denneberg for valuable discussions and to the anonymous referees for their helpful suggestions and comments.

References

[1] E. M. Alfsen. *Compact Convex Sets and Boundary Integrals*. Springer, Berlin, 1971.

[2] E. Bishop and K. de Leeuw. The Representation of Linear Functionals by Measures on Sets of Extreme Points. *Annales de l'Institut Fourier*, 9:305–331, 1959.

[3] G. Choquet. Theory of Capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953/54.

[4] F. Delbaen. Coherent Risk Measures on General Probability Spaces. 1–37 in K. Sandmann and P. J. Schönbucher. *Advances in Finance and Stochastics*, Springer, Berlin, 2002.

[5] D. Denneberg. Representation of the Choquet Integral with the σ-additive Möbius Transform. *Fuzzy Sets and Systems*, 92:139–156, 1997.

[6] I. Gilboa and D. Schmeidler. Canonical Representation of Set Functions. *Mathematics of Operations Research*, 20:197–212, 1995.

[7] I. Gilboa and D. Schmeidler. Maxmin Expected Utility with non-unique Prior. *Journal of Mathematical Economics*, 18:141–153, 1989.

[8] S. Maaß. Coherent Lower Previsions as Exact Functionals and their (Sigma-)Core. In *ISIPTA'01: Proceedings of the Second International Symposium on Imprecise Probabilitics and their Applications*, Shaker, Maastricht, 230–236, 2001.

[9] S. Maaß. Exact Functionals and their Core. *Statistical Papers*, 43:75–93, 2002.

[10] M. Marinacci. Decomposition and Representation of Coalitional Games. *Mathematics of Operations Research*, 21:1000–1015, 1996.

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

Sebastian Maaß is with the Department of Mathematics, Universität Bremen, Germany. E-mail: Sebastian.Maass@web.de