Expected Utility Theory under Non-Classical Uncertainty*

V. I. Danilov† and A. Lambert-Mogiliansky‡

December 27, 2008

Abstract

In this paper Savage's theory of decision-making under uncertainty is extended from a classical environment into a non-classical one. The Boolean lattice of events is replaced by an arbitrary ortho-complemented poset. We formulate the corresponding axioms and provide representation theorems for qualitative measures and expected utility. Then, we discuss the issue of beliefs updating and investigate a transition probability model. An application to a simple game context is proposed.

Keywords. Measurement, bet, non-classical probability, qualitative measure, transition probability, orthomodular poset.

1 Introduction

In this paper we propose an extension of the standard approach to decision-making under uncertainty in Savage's style from the classical model into the more general model of non-classical measurement theory. Formally, this means that we substitute the Boolean lattice with a more general orthoposet structure.

In order to provide a first line of motivation for our approach we turn back to Savage's theory in a very simplified version. In Savage [18], the issue is about the evaluation (or comparison) of "acts" with uncertain results. For simplicity we shall assume that the results can be evaluated (cardinally) in utils. But the results are uncertain, they depend on the unknown state of Nature.

*The financial support of the grant #NSh-6417.2006.6, School Support, is gratefully acknowledged.

†Central Economic Mathematical Institute, Russian Academy of Sciences, 47 Nakhimovski Prospect, 117418 Moscow, Russia. danilov@cemi.rssi.ru

‡PSE, Paris-Jourdan Sciences Economiques, Ecole Economique de Paris, alambert@pse.ens.fr

The classical approach formalizes this situation as follows. There exists a set S of states of nature, which may in principle occur. (For simplicity, we assume that the set S is finite.) An act is a function $f : S \to \mathbb{R}$. If the state $s \in S$ is realized, our agent receives a utility of $f(s)$ utils. But before hand it is not possible to say which state s is going to be realized and the agent has to choose among acts *before* he learns about the state s. This is the heart of the problem.

Among possible acts there are "constant" acts, i.e. acts with a result that is known before hand, independently of the state of nature. The constant act is described by a (real) number $c \in \mathbb{R}$. It is therefore natural to link an arbitrary act f with its "certainty equivalent" $CE(f) \in \mathbb{R}$ (such that our decision-maker is indifferent between the act f and the constant act which gives utility $CE(f)$). The first postulate of our simplified Savage model asserts the existence of the *certainty equivalent*:

- *S1.* There exists a certainty equivalent $CE : \mathbb{R}^S \to \mathbb{R}$ and for the constant act 1_S we have $CE(1_S) = 1$.

It is rather natural to require the monotonicity of the mapping CE:

- *S2.* If $f \leq g$ then $CE(f) \leq CE(g)$.

The main property we impose on CE is additivity[1]:

- *S3.* $CE(f + g) = CE(f) + CE(g)$ for any f and $g \in \mathbb{R}^S$.

Together with monotonicity the axiom $S3$ implies the linearity, that is $CE(\alpha f + \beta g) = \alpha CE(f) + \beta CE(g)$ for any $\alpha, \beta \in \mathbb{R}$. As a linear functional on the vector space $\mathbb{R}^X$, CE can be written in a form $CE(f) = \sum_{s \in S} \mu(s) f(s)$. By the axiom $S2$, $\mu \geq 0$; since $CE(1_S) = 1$ we have $\sum_s \mu(s) = 1$. Therefore $\mu(s)$ can be interpreted as the "probability"[2] for the realization of the state s. With such an interpretation, $CE(f)$ becomes the "expected" utility of the act f.

In the present paper we propose to substitute the Boolean lattice of events with a more general orthoposet. The move in that direction was initiated long ago, in fact with the creation of Quantum Mechanics (QM). The Hilbert space entered into the theory immediately, beginning with von Neumann [19] who proposes to use the lattice of projectors in the Hilbert space as the suitable model for QM instead of the classical

[1]The requirement in $S3$ looks very straightforward and ingenuous indeed. Savage himself and his followers preferred to appeal to the so-called "sure thing principle" and to derive additivity from other axioms. But the related considerations are not relevant to the point we make in this paper.

[2]Sometimes this probability is called subjective or personal, because it only expresses the likelihood that a specific decision-maker assigns to event s.

(Boolean) logic. In recent years, we have seen a number of works in social sciences and psychology that propose to use the quantum formalism to explain behavioral paradoxes and anomalies (see e.g., [4, 9, 10, 14, 15, 2]).

But most closely related to our paper are [6, 11]. In both papers the standard expected utility theory is transposed into a Hilbert space model. Lehrer and Shmaya write "We adopt a similar approach and apply it to the quantum framework. While classical probability is defined over subsets (events) of a state space, quantum probability is defined over subspaces of a Hilbert space. Gyntelberg and Hansen (2004) apply a general event-lattice theory (with axioms that resemble those of von Neumann and Morgenstern) to a similar framework." One could expect that Gyntelberg and Hansen truly would be working with general ortholattices. But no, they again work with a lattice of subspaces of a Hilbert space.

One of our objectives is to show that there is no need for a Hilbert space and that Savage's approach can just as well (and even easier) be developed within the frame of general orthoposets. This is not merely an attempt to generalize. Instead, our view is that the Hilbert space model is only one of the possible models for describing situations characterized by incompatible measurements. Moreover the Hilbert space model may not always be the most suitable model. According to [20] "Hilbert space, and Physics in general, are too tightly knit to serve as a criterion for less highly structured situation such as those occurring in the social sciences." In our paper [3] we describe the structure of such a general model.

For the sake of comparison with the Savage setup, we first develop the theory in a static context. That is we begin with an arbitrary orthoposet of events. We understand acts as "lotteries" whose results are governed by the outcomes of measurements and we use the term "bet" to refer to these lotteries. Our main theorem asserts that any reasonable preference on the set of bets can be represented as expected utility with respect to some belief, that is a probabilistic measure on the orthoposet of events.

Since non-classical phenomena are intimately linked to the impact of measurements on the state of the measured system, a static approach cannot be satisfactory. (If measurements do not change the state of the system, one can use Savage's classical paradigm.) A genuine theory of non-classical expected utility should apply to sequences of bets or measurements. In Sections 6 and 7 we discuss the issue of belief updating and propose a model which takes in account the impact of measurements on the state. In an illustration we show that the results in this paper are relevant to modelling interaction in simple games when a decision-maker faces a type indeterminate opponent, i.e. an agent whose type changes under the impact of decision-making.

In the Appendix we consider the conditions under which a qualitative measure can be represented by a quantitative measure.

2 The structure of events

A central notion in our decision theory under uncertainty is the notion of *measurement*. Intuitively, it is a (partial) resolution of uncertainty. One can perform a measurement on the "world" and obtain some information about "state of the world". More precisely, we learn the outcome of the measurement. Thus, every measurement M is characterized by a set $O(M)$ of possible outcomes.

A *bet* on the basis of a measurement M (or supported by a measurement M) is a mapping $f : O(M) \to \mathbb{R}$. If we perform the measurement M and an outcome $o \in O(M)$ is realized then the agent obtains $f(o)$ utils.

$\mathcal{M}$ denotes the set of admissible measurements. Correspondingly the set of bets is the (disjoint) union $\coprod_{M\in\mathcal{M}} \mathbb{R}^{O(M)}$. Our decision maker (DM) is assumed to have preferences defined on this set. In the classical setup all measurements can be collected into a single complete measurement. The set of its outcomes can be identified with the set S of the states of the world. We are interested in a more general situation, when there may be several essentially different measurements. To explain what we understand by 'essentially different measurements' we have to first introduce the notion of *derived* measurement.

Let M be a measurement with the set $O(M)$ of outcomes, and let $\varphi : O(M) \to O'$ be a mapping of sets. Then we can form a new (derived) measurement M' with the set of outcomes O'. It is constructed in the following way: we perform M and declare $\varphi(o)$ as the outcome of M' provided o is the outcome of M. Actually M' is merely a transformation of the results of M. In particular, we can perform a measurement and completely ignore its result. Denote this (ignored) derived measurement as TM.

We assume further that the set $\mathcal{M}$ is stable with respect to derived measurements, i.e., for any $M \in \mathcal{M}$, the measurements derived from M also belong to $\mathcal{M}$. Moreover we assume that $\mathcal{M}$ contains the trivial measurement T with a single outcome, i.e. no proper measurement is performed.

In general case there can be several (essentially different) incompatible measurements. Roughly speaking, measurements M and M' are incompatible if they cannot be performed simultaneously. More importantly, M and M' are incompatible if performing measurement M' can change the result of the measurement M compared with the outcome previously obtained when performing M. Therefore the issue of timing (i.e., of the sequence in which measurements are performed) becomes important. If all measurements are compatible, we can design a composite measurement and in the limit we get the finest "universal" measurement. But if some measurements are incompatible then such a universal measurement does not exist. Of course, measurements derived from of a common measurement are compatible.

The outcomes of different (even incompatible) measurements can be related to one another, and the DM is assumed to know these relations.

Definition. We say that an outcome o (of a measurement M) *implies* outcome o' (of another measurement M') if performing M' immediately after (M, o) gives the outcome o' with certainty. We denote this relation as $o \Longrightarrow o'$.

We assume that the following two axioms are fulfilled:

M1. $o \Longrightarrow o$.

M2. If $o \Longrightarrow o'$ and $o' \Longrightarrow o''$ then $o \Longrightarrow o''$.

Due to axioms M1 and M2 the relation $\Longrightarrow$ on the set $O = \coprod_{M \in \mathcal{M}} O(M)$ of outcomes is a pre-order. In particular, the symmetric part $\approx$ of the relation is an equivalence relation. Let us denote $\mathcal{E} = O/\approx$ the corresponding factor-set; the elements of $\mathcal{E}$ will be called *events*. The relation $\Longrightarrow$ defines a (partial) order on the set $\mathcal{E}$, which will be denoted by $\leq$. Thus $\mathcal{E}$ is a poset, i.e. a partially ordered set. The event $O(T)$ (more precisely, its equivalent class) is the largest element of the poset; we denote it $\mathbf{1}$. We add formally a minimal element $\mathbf{0}$ representing the "impossible" event.

The set $\mathcal{E}$ possesses one more important structure, that of *ortho-complementation.*

For any outcome o there exists a dichotomous measurement with outcomes $(o, \bar{o})$, where $\bar{o}$ denotes the outcome opposite to o, i.e. "non-o". (Generally speaking, $\bar{o}$ depends on the choice of the measurement.) Let now o and o' be two outcomes (generally, of two different measurements M and M'). Say that o is *orthogonal* to o' (we write $o \perp o'$) if $o \Longrightarrow \bar{o}'$. Intuitively, this means that these outcomes exclude one another. This suggests that the orthogonality relation should be symmetric (apparently, it is irreflexive). We postulate it as the following axiom:

M3. $o \perp o'$ if and only if $o' \perp o$.

Remark. To clarify the meaning of this axiom, we reformulate it in two different ways. We say that an outcome o' *is possible* given an outcome o, and we write $o \rightarrow o'$, if o' is *not* orthogonal to o (or the relation $o \Longrightarrow \bar{o}'$ is false). One can easily check that axiom M3 is equivalent (provided M1 and M2) to the following axiom

M3′. If $o \rightarrow o'$ and $o' \Longrightarrow o''$ then $o \rightarrow o''$.

Another equivalent form of M3 is

M3″. If $o \rightarrow o'$ then $o' \rightarrow o$.

Of course, Axioms M3$'$ and M3$''$ also are disputable although they seem rather acceptable; in any case the interpretation is clear.

Axiom M3 can be rewritten as the statement: if $o \Longrightarrow o'$ then $\bar{o}' \Longrightarrow \bar{o}$. In particular, if o and o' are equivalent then $\bar{o}$ and $\bar{o}'$ are equivalent as well. This implies that the orthogonality relation $\perp$ can be carried over to the factor-set of events $\mathcal{E}$, where it is a symmetric and irreflexive relation. For an event e (represented by an outcome o) we denote $e^{\perp}$ the event represented by the outcome $\bar{o}$. The correctness of this definition follows from Axiom M3. Moreover, the event $e^{\perp}$ is the largest event orthogonal to e. Indeed, suppose that an outcome o' is orthogonal to an outcome o representing the event e. That is $o \Longrightarrow \bar{o}'$. Due to Axiom M3, we have $o' \Longrightarrow \bar{o}$ that is $e' \leq e^{\perp}$.

The following proposition gathers the properties of the operation $\perp$ on $\mathcal{E}$:

Proposition 1 *The operation $\perp$ is an anti-monotone involution of $\mathcal{E}$, and $e \vee e^{\perp} = 1$ for any event e.*

Proof. The fact that $\perp$ is an involution of (that is $(e^{\perp})^{\perp} = e$) is obvious by construction. Above we already noted that $e \leq e'$ implies $(e')^{\perp} \leq e^{\perp}$. Since there exists no outcome o' that implies both o and $\bar{o}$ we have $e \wedge e^{\perp} = \mathbf{0}$. Applying $\perp$ to this equality and using the anti-monotonicity of $\perp$, we get the equality $\mathbf{1} = \mathbf{0}^{\perp} = (e \wedge e^{\perp})^{\perp} = e^{\perp} \vee ((e^{\perp})^{\perp}) = e^{\perp} \vee e$. $\square$

Thus, the event set $\mathcal{E}$ (equipped with the order relation $\leq$ and the operation $\perp$) is an *orthoposet* (an ortho-complemented partially ordered set). It can be considered as the logic of the one-shot decision-making situation.

Let us return to the measurement M with outcomes $o \in O(M)$. (The same letters o will be used for the corresponding elements in the orthoposet of events $\mathcal{E}$.)

Lemma 1 *Let A be a subset of $O(M)$ with $\bar{A} = O(M) - A$, the complement of A in $O(M)$. Then the join $\bigvee_{o \in A} o$ exists and is equal to $(\bigvee_{o \in \bar{A}} o)^{\perp}$.*

Indeed, the join $\bigvee_{o \in A} o$ is the event "the outcome of measurement M belongs to A", whereas $\bigvee_{o \in \bar{A}} o$ is the opposite event. $\square$

This motivate the following notion.

Definition. An *Orthogonal Decomposition of the Unit* (briefly, an *ODU*) is a family $(a(i),\ i \in I)$ of events satisfying the following two conditions:

1) for any $J \subset I$ the join $\bigvee_{i \in J} a(i)$ exists; we denote it as $a(J)$.

2) $a(J)^{\perp} = a(I - J)$ for any $J \subset I$.

In particular, all events $a(i)$ are pairwise orthogonal, and $a(I) = \bigvee_{i \in I} a(i) = \mathbf{1}$. This justifies the term 'ODU'.

For instance, the single-element family $\mathbf{1}$ is an (trivial) ODU. For any $a \in \mathcal{E}$, the two-element family $(a, a^{\perp})$ is an ODU.

By Lemma 1 every measurement defines some ODU. For example, the trivial measurement T define the single-element ODU. Any two-element ODU $(a, a^{\perp})$ is realized by some (dichotomous) measurement. An ODU is called *admissible* if it is realized by some measurement $M \in \mathcal{M}$.

Thus, a decision situation defines an orthoposet of events $\mathcal{E}$ and a set $\mathcal{A}$ of admissible ODUs in the orthoposet. In the sequel, we shall assume that any ODU is admissible. The reader can easily extend the results to the general case.

We can now identify (to a certain extent) measurements with ODUs. One can (temporarily) forget about measurements and deal with arbitrary orthoposet $\mathcal{E}$. A *bet* is now a pair (α, f), where $\alpha = (a(i), i \in I(\alpha))$ is an ODU (the *basis* of the bet) and $f : I(\alpha) \to \mathbb{R}$ is a real-valued function (the *payoff*). The set of bets with fixed basis α is the vector space $F(\alpha) = \mathbb{R}^{I(\alpha)}$. The set F of all bets is the disjoint union of all $F(\alpha)$. Theorem 1 below asserts that any reasonable preference order on F is given by a probability measure on $\mathcal{E}$. Our next step will be to define a probability on an orthoposet.

3 Non-classical probability

The classical probability theory starts with a set S of elementary (mutually exclusive) events. Thereafter it moves over to general events. The next key concept is a "collection of mutually exclusive events". In the classical model this is simply a partition of the set X, that is a decomposition $S = A_1 \amalg \ldots \amalg A_n$. In our language events are represented by an event orthoposet $\mathcal{E}$. The collection of mutually exclusive events is replaced by the notion of an Orthogonal Decomposition of the Unit, ODU.

Let $\mathcal{E}$ be an orthoposet. To facilitate the presentation we impose the following finiteness requirement: any ODU is (essentially) finite. For example, this requirement is fulfilled if the orthoposet $\mathcal{E}$ is finite.

Definition. An *evaluation* on an orthoposet $\mathcal{E}$ is a mapping $\nu : \mathcal{E} \to \mathbb{R}_+$ such that $\nu(\mathbf{0}) = 0$ and $\nu(\mathbf{1}) = 1$. An evaluation ν is called

1) *monotone* if $\nu(a) \leq \nu(b)$ provided $a \leq b$;

2) *additive* if $\nu(a \vee b) = \nu(a) + \nu(b)$ for orthogonal events a and b. We write $a \oplus b$ instead of $a \vee b$ to emphasize that $a \perp b$;

3) *probabilistic* (or a probability) if $\sum_i \nu(a(i)) = 1$ for every ODU $(a(i),\ i \in I)$.

It is clear that 2) implies 3). In the classical (Boolean) case 3) implies 1) and 2), but that is not true in the general case (see Example 3 in Section 4 and Proposition 2 below).

For the sequel we need the following

Lemma 2 *Let ν be a probability, and $(a(i),\ i \in I)$ be an ODU. Then, for any $J \subset I$,*

$$\nu(\bigoplus_{i\in J} a(i)) = \sum_{i\in J} \nu(a(i)).$$

Proof. Since $(a(i), i \in I)$ is an ODU, we have the equality $\sum_{i\in I} \nu(a(i)) = 1$. On the other hand, the family $(a(J); a(i), i \in I - J)$ is an ODU as well. Therefore, we have the equality $\nu(a(J)) + \sum_{i\in I-J} \nu(a(i)) = 1$. Hence, $\nu(a(J)) = \sum_{i\in J} \nu(a(i))$. □

There exists an important case when everything simplifies and approaches the classical case. It is the case of orthomodular posets (OMP), that is orthoposets satisfying the property of *orthomodularity*: if $a \le b$ then $b = a \vee (b \wedge a^\perp)$. For example, every Boolean lattice is an OMP.

Proposition 2 *Let $\mathcal{E}$ be an OMP. Then any probabilistic evaluation on $\mathcal{E}$ is additive and monotone.*

Proof. Let ν be a probabilistic evaluation on $\mathcal{E}$. We first establish additivity. Suppose $a \perp b$ and pose $c = (a \oplus b)^\perp$. Since $(c, c^\perp)$ is an ODU, $\nu(c) + \nu(c^\perp) = 1$.

We assert that (a, b, c) is an ODU as well. To prove that we need to show that $a^\perp = b \oplus c$. Since a, b and c are pairwise orthogonal, $b \oplus c \le a^\perp$. By force of the property of orthomodularity we have that $a^\perp = (b \oplus c) \oplus (a^\perp \wedge (b \oplus c)^\perp)$. But $a^\perp \wedge (b \oplus c)^\perp = (a \vee b \vee c)^\perp = (a \oplus b)^\perp \wedge c^\perp = c \wedge c^\perp = \mathbf{0}$. Hence $a^\perp = b \oplus c$. Similarly $b^\perp = a \oplus c$. The equality $c^\perp = a \oplus b$ is satisfied by definition. Thus, the triplet (a, b, c) is an ODU.

Therefore we have the equality $\nu(a) + \nu(b) + \nu(c) = 1$. Hence $\nu(a \oplus b) = \nu(c^\perp) = 1 - \nu(c) = \nu(a) + \nu(b)$, which yields the additivity of ν.

Monotonicity follows trivially from the formula $b = a \oplus (b \wedge a^\perp)$, the additivity of ν and the inequality $\nu(b \wedge a^\perp) \ge 0$. □

Thus, in the orthomodular case a probability may also be defined as an additive evaluation. For reasons which will become clear in Section 5, we shall especially be interested in monotone probabilistic evaluation on E; $\Delta(E)$ denotes the set of such evaluations. If $\mathcal{E} = 2^X$ is a Boolean lattice, $\Delta(2^X)$ is the set of ordinary probabilities (or probability measures) on a (finite) set X.

4 Some examples

Here we consider a few examples illustrating the material of Section 2 and 3. In each example we present an event orthoposet $\mathcal{E}$ and describe the corresponding set $\Delta(\mathcal{E})$.

Example 1. Let S be a set. An event is an arbitrary subset of S, so that $\mathcal{E} = 2^S$. The order is defined by set-theoretical inclusion. For $A \subset X$ the opposite event $A^\perp = X - A$ is the set-theoretical complement. It is the classical situation.

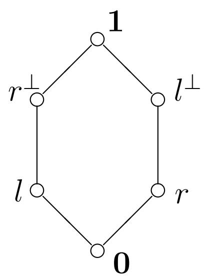
When S is a finite set, the probability set is the unit simplex with vertices in S.

Example 2. Let $\mathcal{H}$ be a (finite dimensional) Hilbert space (over the field of real or complex numbers). An event is a vector subspace in $\mathcal{H}$. $\perp$ is the orthogonal complementation with respect to the inner product. The orthoposet $Sub(\mathcal{H})$ of vector subspaces in $\mathcal{H}$ is an orthomodular ortholattice. This model is standard in Quantum Mechanics.

Due to Gleason's theorem, the probabilities on $Sub(\mathcal{H})$ are given by positive self-adjoint operators with the trace 1 ('density matrices'). If ρ is such an operator then the 'probability' of a subspace $L \subset \mathcal{H}$ is equal to the trace of the product ρP_L, where P_L is the operator of orthogonal projection on the subspace L.

We next consider several 'toy' examples.

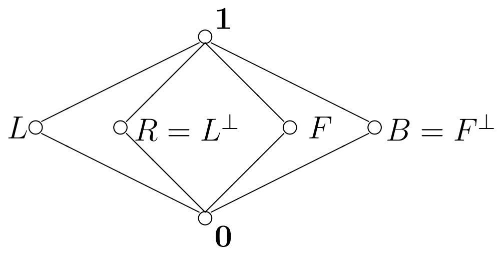
Example 3. The event orthoposet (in fact, an ortholattice) is depicted below



There are two ODUs: $(l, l^\perp)$ and $(r, r^\perp)$ (the pair (l, r) is not an ODU although r and l are orthogonal and $r \oplus l = \mathbf{1}$). The event l implies $r^\perp$, the event r implies $l^\perp$. However $r^\perp$ does not exclude $l^\perp$ (and vice versa). Therefore, given the event $l^\perp$, the measurement $(r, r^\perp)$ can yield both r and $r^\perp$ as the outcomes. The following sequence of outcomes r, $l^\perp$, $r^\perp$ is possible. Hence the measurements are incompatible.

In order to define a probability ν on the orthoposet is to give we only need two numbers $\nu(r)$ and $\nu(l)$, both between 0 and 1. Then $\nu(l^\perp) = 1 - \nu(l)$ and $\nu(r^\perp) = 1 - \nu(r)$. The probability is monotone if $\nu(l) + \nu(r) \leq 1$; the probability is additive if $\nu(l) + \nu(r) = 1$.

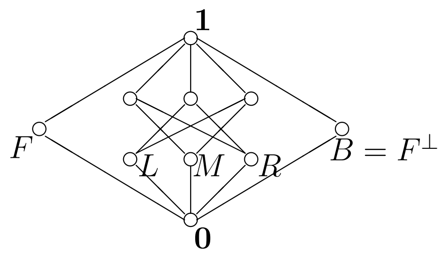
Example 4. Let us consider the orthomodular lattice (OML) depicted below



as the event poset. Here are two ODUs: (L, R) and (F, B), and two corresponding dichotomous (and incompatible) measurements.

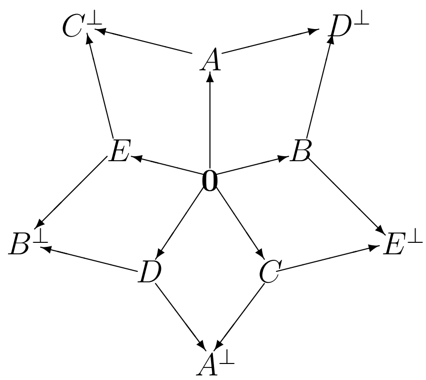
A probability here is given by two numbers $\nu(L)$ and $\nu(F)$, both between 0 and 1. Thus the set of probabilities is the square $[0, 1] \times [0, 1]$ (or $\Delta(\{L, R\}) \times \Delta(\{F, B\})$).

Example 5. Let us consider as the event orthoposet the following OML



There are two complete measurements with outcomes $\{F, B\}$ and $\{L, M, R\}$ (and the derived ones). The set of probabilities is the Cartesian product of the segment $\Delta(\{F, B\})$ and the simplex $\Delta(\{L, M, R\})$.

Example 6 *A variant of the Wright's pentagon* [20]. Consider the following orthoposet of events (we do not draw **1**)



We have five dichotomous ODUs: $(A, A^{\perp})$, $(B, B^{\perp})$, and so on. Each measurement is a question: A or not A? B or not B?, and so on.

Let us calculate monotone probabilities on this orthoposet. Such a probability ν is given by five numbers: $a = \nu(A)$, $b = \nu(B), ..., e = \nu(E)$. They must satisfy by 10 inequalities: $a \geq 0, ..., e \geq 0$, and $a + c \leq 1$, $b + d \leq 1$, ..., $e + b \leq 1$.

Therefore $\Delta(\mathcal{E})$ is a convex polytope in a five-dimensional space. It has 12 vertices (extreme points). These are: 1) the point $0 = (0, ..., 0)$, 2) five points of the form $(1, 0, 0, 0, 0), (0, 1, 0, 0, 0), ..., (0, 0, 0, 0, 1)$ which can be identify with $A, B, ..., E$; 3) five points $(1, 1, 0, 0, 0), (0, 1, 1, 0, 0), ..., (1, 0, 0, 0, 1)$; and finally 4) the most interest point $\omega = (1/2, 1/2, 1/2, 1/2, 1/2)$. The last point ω is a pure state but it is not dispersion-free and therefore it has a non classical interpretation (for more details see [20]).

Example 7. Let us give one more striking example. To describe the event orthoposet we use the following Greechie diagram

THE DIAG GREECHIE cannot be read!

It consists of the set V of 15 vertices (depicted as black circles) and the collection L of 15 three-elements subsets in V called lines (depicted as smooth curves). We shall not explain how a general Greechie diagram (V, L) defines the corresponding ortholattice $\mathcal{E}$. Instead, we confine attention to this specific example. Let $\mathcal{E} = \{0, 1\} \coprod V \coprod V'$, where V' is a copy of V (for $x \in V$ its **"copy"** is denoted x'). We define an order on the set $\mathcal{E}$ as follows. **0** is the least and **1** is the greatest element; for $x, y \in V$, $x < y'$ if and only if $x \neq y$ and x and y are collinear (belong to the same line). An ortho-complementation is defined as follows: for $x \in V$ $x^\perp = x'$ and $x'^\perp = x$.[3]

The following property is important for us: if x, y, z is a line then the collection (x, y, z) is an ODU in the ortholattice E. Indeed, z is the only vertex which is collinear to both x and y. Therefore $x \vee y = z' = z^\perp$.

Let us now consider probabilities on $\mathcal{E}$. To give a probability one needs to give a function $\nu : V \to R_+$ such that $\nu(x) + \nu(y) + \nu(z) = 1$ for every collinear triple x, y, z. (The numbers $\nu(x')$ are defined from the relations $\nu(x) + \nu(x') = 1$.) We obtain 15 linear equations with 15 unknowns. Direct calculations show that there exists a unique solution to this system: $\nu(x) = 1/3$ for every $x \in V$. This is the main lesson from this example: the set $\Delta(\mathcal{E})$ consists of a unique (uniform) probability!

Note that Greechie [5] constructed a finite orthomodular lattice with no probability at all.

5 Non-classical expected utility theory

Let us return to the comparison of bets (given an event orthoposet $\mathcal{E}$). Recall that F denotes the set of all bets; $F = \coprod F(\alpha)$, where α runs over ODUs. We suppose that bets are compared using some certainty equivalent.

[3] One can also show that E is an orthomodular poset but we shall not use this property.

Definition. A *certainty equivalent* is a mapping $CE : F \to \mathbb{R}$ such that $CE(1) = 1$. Here 1 is any bet with payoffs equal to 1.

Let us start with an explicit formula for a possible CE. Assume that our DM has some belief about the state of the world, that is he has in his mind a monotone probabilistic evaluation μ on the event orthoposet $\mathcal{E}$. Then, for any bet f on the basis of ODU $\alpha = (a(i), i \in I(\alpha))$, he can compute the following number (the expected value of the bet f)

$$CE_\mu(f) = \sum_i \mu(a(i))f(i).$$

This is a certainty equivalent, and it possesses two nice properties.

First of all, it is a linear functional on every vector space $F(\alpha) = \mathbb{R}^{I(\alpha)}$. This is obvious from the formula above. The second property is monotonicity in some strong sense. To formulate this property, we need to introduce a dominance relation between bets. Intuitively a bet g dominates another bet f if g always yields higher payoff than f. But what does this precisely mean?

Let f be a bet on the basis of an ODU $\alpha = (a(i), i \in I)$, and g be a bet on the basis of another ODU $\beta = (b(j), j \in J)$. Remind that in Section 2 we have introduced the relation $\to$ of possibility. We shall say that an outcome $j \in J$ *is possible* provided an outcome $i \in I$ if $i \to j$ (or, what is the same, $a(i)$ is not orthogonal to $b(j)$); we denote this as $j \in P(i)$.

Definition. We say that g *dominates* f if $g(j) \geq f(i)$ for every $j \in P(i)$.

In other words, each time when the bet f gives the payoff $f(i)$, the bet g gives the no smaller payoff $g(j)$. In particular, suppose that f and g are two bets on the same basis α; then g dominates f if and only if $f \leq g$ as functions on $I(\alpha)$.

The second property of the certainty equivalent $CE = CE_\mu$ is the compatibility with the dominance relation.

Proposition 3 *If g dominates f then $CE(f) \leq CE(g)$.*

Proof. Let us order the elements of I according to decreasing values of f, so that $f(1) \geq f(2) \geq ... \geq f(n)$, $n = |I|$. Associate to f an auxiliary function $\hat{f} : [0,1] \to \mathbb{R}$ defined by the formula:

$$\hat{f}(t) = f(i), \text{ if } \mu(a(1)) + ... + \mu(a(i-1)) \leq t < \mu(a(1)) + ... + \mu(a(i)).$$

The function $\hat{f}$ is piece-wise constant, and $CE(f) = \int_0^1 \hat{f}(t)dt$.

For $i = 1, ..., n$, consider the subset $Q(i) = P(1) \cup ... \cup P(i)$ in J; this is the set of elements $j \in J$ which are possible at $1, ..., i$. Let us order the elements of J such that the first elements are in $Q(1)$, the next elements are in $Q(2)-Q(1)$, **and so on.** Again associate to g an auxiliary function $\hat{g} : [0,1] \to \mathbb{R}$ defined by the formula:

$$\hat{g}(t) = g(j), \text{ if } \mu(b(1)) + ... + \mu(b(j-1)) \leq t < \mu(b(1)) + ... + \mu(b(j)).$$

The function $\hat{g}$ is piece-wise constant, and $CE(g) = \int_0^1 \hat{g}(t)dt$. We prove Proposition 3 if we can show that $\hat{f} \leq \hat{g}$.

Claim 1 *$\mu(a(1)) + ... + \mu(a(i)) \leq \sum_{j \in Q(i)} \mu(b(j))$ for every $i = 1, ..., n$.*

We first complete the proof of Proposition 3 and then we prove the Claim. Suppose that t lies in the interval $\mu(a(1)) + ... + \mu(a(i-1)) \leq t < \mu(a(1)) + ... + \mu(a(i))$, where $\hat{f}(t)$ is equal to $f(i)$. By the Claim, we have $t < \sum_{j \in Q(i)} \mu(b(j))$. Therefore $\hat{g}(t)$ is greater or equal to the minimum of $g(j),\ j \in Q(i)$. Since, for any $j \in Q(i)$, we have $j \in P(i')$ for some $i' = 1, ..., i$, we conclude that $g(j) \geq \min(f(i'), i' = 1, ..., i) = f(i)$. Therefore $\hat{g}(t) \geq f(i) = \hat{f}(t)$. Thus we have proved that $\hat{g} \geq \hat{f}$ and consequently we have proved Proposition 3.

Proof of Claim 1. Accordingly Lemma 2, the left-hand-side of the inequality is equal to $\mu(a(1) \oplus ... \oplus a(i))$. Similarly, the right-hand-side is equal to $\mu(\bigoplus_{j \in Q(i)} b(j))$. Due to the monotonicity of μ, it is sufficient to prove the inequality

$$a(1) \oplus ... \oplus a(i) \leq \bigoplus_{j \in Q(i)} b(j) = b(Q(i)).$$

The set $Q(i)$ consists of all outcomes j which are possible at the outcomes $1, ..., i$ of the first measurement. Hence $b(Q(i)$ is the ortho-complement to the event $b(J - Q(i))$. But for any outcome j which is not possible at $1, ..., i$ we have $b(j) \perp a(1) \oplus ... \oplus a(i)$. Therefore $b(J - Q(i)) \perp a(1) \oplus ... \oplus a(i)$ and, consequently, $a(1) \oplus ... \oplus a(i) \leq b(J - Q(i))^{\perp} = b(Q(i))$. This proves the Claim and Proposition 3. □

Following Savage, we shall now go in the opposite direction. We shall show that if a certainty equivalent CE satisfies the two above properties then it is defined by a probabilistic evaluation on the orthoposet $\mathcal{E}$. More precisely, we consider two properties of CE. The first one refers to bets defined on a common basis.

(**Add**) CE is an additive function on every vector space $F(\alpha)$.

The second property compares bets with different bases and it is actually the statement in Proposition 3:

(**Dom**) If g dominates f then $CE(f) \leq CE(g)$.

Theorem 1 *Let a certainty equivalent CE possess the properties (**Add**) and (**Dom**). Then it has the form CE_μ for some (uniquely defined) monotone probability μ on the orthoposet $\mathcal{E}$.*

Proof. First of all we note that CE is monotone functional on every vector space $F(\alpha)$. Indeed, if $f \leq g$ then f is dominated by g and hence (due to (**Dom**)) $CE(f) \leq CE(g)$. Therefore CE is not only additive but a linear functional on $F(\alpha)$.

Now we construct explicitly the probability evaluation μ. To this aim we consider an event a and the bet 1_a on the basis of a dichotomous ODU $(a, a^\perp)$: 1_a is equal to 1 on a and 0 on $a^\perp$ ("unit bet on the event a"). Let $\mu(a) = CE(1_a)$. Obviously $\mu(a) \geq 0$.

Suppose now that $\alpha = (a(i), i \in I)$ is an arbitrary ODU. Let 1_i be the bet (on the basis of α) such that it is equal to 1 on i and equal to 0 on the other outcomes of the measurement. As it is easy to understand, this bet is equivalent (in the sense of the domination) to the bet $1_{a(i)}$. Therefore $CE(1_i) = \mu(a(i))$.

Consider now an arbitrary bet $f : I \to \mathbb{R}$ on the basis of α. Since $f = \sum_i f(i)1_i$ we conclude by the linearity that

$$CE(f) = \sum_i \mu(a(i))f(i) = CE_\mu(f).$$

In particular, if $f \equiv 1$ we obtain that $1 = CE(1) = \sum_i \mu(a(i))$. Therefore μ is a probabilistic valuation and $CE = CE_\mu$.

It remains to prove the monotonicity of μ.

Let $a \leq b$ be two events. Consider two bets: 1_a (on the basis of ODU $(a, a^\perp)$) and 1_b (on the basis of $(b, b^\perp)$). We note that 1_b dominates 1_a. Indeed, given the event a, only the event b is possible; the event $b^\perp$ is impossible because $a \perp b^\perp$. But $1_b(b) = 1 \geq 1 = 1_a(a)$. Thus 1_a is dominated by 1_b. Due to (**Dom**) $\mu(a) = CE(1_a) \leq CE(1_b) = \mu(b)$.

This completes the proof of Theorem 1. □

Thus, there is a natural bijection between the set of 'nice' certainty equivalents and the set $\Delta(\mathcal{E})$ of monotone probabilities. Any probability μ can be considered as a possible 'belief' of the DM, or as his 'prior'.

6 The updating problem

Up to now we have ignored the issue of the timing of consecutive measurements. Our decision situation was of a one-shot character; our DM only ranks bets, but no "real" resolution of uncertainty occurs. Here we would like to investigate a more dynamic situation, when the DM chooses a bet, the corresponding measurement is actually performed, the DM gets a payoff and (what can be more important) he obtains some new information about the realized (actualized) state of the world. The new information induces him to revise his beliefs.

In the classical situation we apply Bayes' rule for updating beliefs. In a non-classical setup with incompatible measurements, the performance of measurements also induces a revision of beliefs but the issue is more complicated as we next illustrate.

Let $\mathcal{E}$ be an event orthoposet, and let $\beta \in \Delta(\mathcal{E})$ be an initial belief of the DM. Suppose that a measurement is performed and that an event E occurs. Our DM should revise (update) the belief β, and replace it by some $\beta' = \beta(\cdot|E)$. According to which rule should β' be calculated? There is no general answer to this question. It depends on the DM's representation of the system.

To illustrate this point, consider the problem of updating in the simple situation of Example 4. There are two measurements LR and FB. Assume the initial belief assigns probabilities to the four events R, L, B, F as follows $\beta(R) = 1 - \beta(L)$ and $\beta(B) = 1 - \beta(F)$. Suppose we perform the measurement LR, and obtain outcome L. Of course, $\beta'(L) = 1$ and $\beta'(R) = 0$. But what about $\beta'(F)$ and $\beta'(B)$? The simplest is to assume that they do not change at all, and similarly when considering updating after a measurement FB. Such an assumption means that two operations measure independent (unrelated) properties of the system (or, possibly, properties of two independent systems). This is clearly a very specific and restrictive method of updating.

Already from this example we see a natural requirement on the revision of beliefs. Namely, $\beta(E|E) = 1$ and, as a consequence, $\beta(F|E) = 0$ if the event F is orthogonal to E. But what about $\beta(F|E)$ for general events F? We might try to use a variant of Bayes' rule posing, for an event F, $\beta(F|E) = \beta(F \wedge E)/\beta(E)$. This rule is sufficiently reasonable if $F \leq E$. Unfortunately, in the general case, it yields an evaluation which is *not* a probability.

In addition there can exist plain obstacles to updating. Recall that in Example 7 we have an orthoposet with a unique belief μ, and $\mu(x) \neq 1$ for every $x \in V$. This belief cannot be revised because there exist no (updated) beliefs compatible with any (complete) measurement!

This discussion shows that arbitrary orthoposets are too general to provide a framework for addressing the issue of updating when considering sequences of measurements. An appropriate model must contain not only an event orthoposet (and measurements as ODUs in the orthoposet) but also a rule for revising (updating) beliefs.

At this point it is convenient to bring into play the notion of a state of a measurable system. We wish to emphasize at once that it is a subtle and enigmatic notion. Is it an objective reality (the *ontic* point of view) or a state of our knowledge (the *epistemic* point of view)?[4] We think that both points of view have their merits. On the one hand, a state is a state of our knowledge, it is a belief. On the other hand, it reflects a real state of affairs, it is not only a pure product of our mind. In any case the notion of state is not a primitive one (in contrast, for example, with measurements or events); it is an element of a model which has to correctly describe a situation.

Below we present elements of a theory of belief updating based on the notion of state. This theory (originally proposed by Mielnik [12]) explicitly takes into account the impact of measurements on states.

7 Transition probability space

A *transition probability space* (TPS) is a pair (S, τ), where S is a set and τ is a mapping $S \times S \rightarrow [0, 1]$. The elements of S are interpreted as "pure" states of our system. The number $\tau(s, t)$ is interpreted as the probability for a transition from the state s to the state t under the impact of a suitable measurement. It is assumed that the function τ satisfy two axioms (later we shall add a third axiom):

T1) If $\tau(s, t) = 0$ then $\tau(t, s) = 0$;

T2) $\tau(s, t) = 1$ if and only if $s = t$.

When $\tau(s, t) = 0$ we say that s and t are *orthogonal* and write $(s \perp t)$. By force of axioms T1 and T2, the relation $\perp$ is symmetric and irreflexive. So that the pair $(S, \perp)$ is an orthospace (see [3]). For subsets A and A' of S we write $A \perp A'$ if $a \perp a'$ for every $a \in A$ and $a' \in A'$. $A^{\perp}$ denotes the set of states orthogonal to A.

A subset $A \subset S$ is called *orthogonal* if $a \perp b$ for different $a, b \in A$. A maximal (by inclusion) orthogonal subset is called an *orthobasis* of S (or simply an orthobasis). Any orthogonal set can be extended to an orthobasis.

We shall understand an orthobasis B as a complete (or finest) measurement. This measurement acts as follows. Let s be an initial state. Under impact of the measurement B, the state s transits (with the probability $\tau(s, b)$) into a new (or updated) state $b \in B$, and moreover we (or an observer) obtain the signal b. Thus, the set of outcomes of the measurement B is identified with B. To make it possible to treat the numbers $\tau(s, b)$ as a probabilities, we impose a third axiom:

T3) For any state $s \in S$ and any orthobasis B the sum $\sum_{b \in B} \tau(s, b)$ is equal to 1.

[4]Compare this with Penrose's dichotomy ([13], p. 309): "to believe in quantum mechanics or to take it seriously". In other words, is "the state vector a convenience, useful only for calculating probabilities for the results of 'measurements' performed on a system" or is it "an accurate mathematical description of a real quantum-level world"?

For simplicity, we shall assume that any orthobasis has finitely many elements. For example, this assumption is fulfilled if S is a finite set.

From the axiom T2 we see that if the system before the measurement B was in a state $b \in B$ then it remains in this state. That is our measurement satisfies the first-kindness property. Besides complete measurements we have to consider derived measurements. An outcome of such a measurement is an arbitrary orthogonal set. We shall see soon that events can be realized as subsets of S.

Let A be an orthogonal set. We shall write $\tau(s, A) = \sum_{a \in A} \tau(s, a)$. Obviously, $\tau(s, A) \leq 1$ for any state s. We associate to the orthogonal set A its "envelope"

$$E(A) = \{s \in S, \tau(s, A) = 1\}.$$

The $E(A)$ consists of states s yielding an outcome from A when we perform any complete measurement B, $A \subset B$. Obviously, $A \subset E(A)$. This envelop allows to describe the "implication" relation $\Rightarrow$ between outcomes that we introduced in Section 2. Namely, $A \Rightarrow A'$ if $A \subset E(A')$.

Lemma 3 *Let A and A' be two disjoint orthogonal sets such that $A \cup A'$ is an orthobasis. For a state s the following assertions are equivalent:*
1) $s \in E(A)$;
2) $s \perp A'$;
3) $s \perp E(A')$.

Proof. $\tau(s, A) + \tau(s, A') = 1$. This gives the equivalence between 1) and 2). Let now t be an element of $E(A')$. By the just named equivalence, we have $t \perp A$, so that $A \cup t$ is an orthogonal set. Then $\tau(s, A \cup t) = \tau(s, A) + \tau(s, t) = 1 + \tau(s, A) \leq 1$, hence $\tau(s, t) = 0$ and $s \perp t$. Since t is an arbitrary element of $E(A')$, we obtain that $s \perp E(A')$. This proves the implication 1) $\Rightarrow$ 3). The implication 3) $\Rightarrow$ 2) is obvious. □

Corollary 1 $E(A) = E(A')^{\perp} = A^{\perp\perp}$ *for an orthogonal set A.*

Let now A and A' be two orthogonal sets in S. If $A \subset E(A')$ then $E(A) \subset E(A')$ by Corollary 1. This implies the transitivity of the implication relation $\Rightarrow$ so that axiom M2 is satisfied. Moreover, the two outcomes A and A' are equivalent if and only if $E(A) = E(A')$. Thus we have proved the following

Proposition 4 *The event orthoposet $\mathcal{E}$ is identified with the collection of subsets of the form $E(A)$, where A runs over orthogonal subsets in S.*

The event orthoposet $\mathcal{E}$ associated to a TPS (S, τ) has two nice property absent in the case of general orthoposets. First of all, if e and e' are orthogonal events then the join $e \vee e' = e \oplus e'$ exists. In other words, the orthogonal sums of events exist. The second property is the orthomodularity of the orthoposet; it was first proved by Pulmanova [16].

Up to now we have only weakly used transition probabilities and mainly appealed to the orthogonality relation $\perp$. We next turn to them more closely. Fix a state s and consider $\tau(s, \cdot)$ as a function on S. For brevity, we denote the function as $s : S \to \mathbb{R}_+$; $s(t) := \tau(s, t)$. Let $E = E(A)$ be an event associated to some orthogonal set A. Set $s(E) = \sum_{a \in A} s(a)$. We assert that this definition is correct and is depend only on the event E, not on A. Indeed, let $B \supset A$ be an orthobasis, and $A' = B - A$. Then $s(E) = 1 - \sum_{a' \in A'} s(a') = 1 - s(E(A')) = 1 - s(E^{\perp})$, and the right-hand-side depends on E. It is clear that the function s is an additive evaluation on the event orthoposet $\mathcal{E}$, that is a probability. Thus, every state s defines some probability (or belief) on $\mathcal{E}$ (denoted by the same letter s).

Theorem 2 *The mapping from S to $\Delta(\mathcal{E})$ is injective.* (This permits to consider S as a subset in $\Delta(S, \perp)$.) *Moreover, any state $s \in S$ is an extreme point of the convex hull $co(S)$ of the set S in $\Delta(\mathcal{E})$.*

Proof. Suppose that the states s and t realize one and the same probability, that is $s(\cdot) = t(\cdot)$ as functions on S. Let us choose s as the missing argument. We obtain the equality $1 = \tau(s, s) = \tau(t, s)$, which by axiom T2 implies that $s = t$. This proves the first statement.

Suppose now that a state $s \in S$ can be expressed as a convex combination $\sum \alpha_i s_i$ of states $s_i \in S$ with α_i strictly positive and summing to 1. Let us consider the value of this function at the element s. We have the equality

$$\tau(s, s) = \sum \alpha_i \tau(s_i, s).$$

The left hand side is equal to 1 by axiom T2. Therefore $\sum \alpha_i(1 - \tau(s_i, s) = \sum \alpha_i - \sum \alpha_i \tau(s_i, s) = 1 - 1 = 0$. Since $\tau(s_i, s) \leq 1$ and $\alpha_i > 0$, we conclude that all $\tau(s_i, s) = 1$. But this means that all $s_i = s$ which gives the second statement. $\square$

The first statement of Theorem 2 says that different states can be distinguished by some appropriate measurement. Moreover, it prompts to consider probabilities on $\mathcal{E}$ as a kind of 'generalized states'. The second statement justifies using the term 'pure state' for the elements of S. It is natural to consider convex combinations of pure states as "mixed" states. Theorem 1 tells that 'nice' certainty equivalents are given by probabilities. The results of this section suggest that a 'reasonable' belief is not an arbitrary probability, but should be a mixed state. Indeed, an initial prior could be an arbitrary probability, but after any (non-trivial) measurement the updated belief should be a mixed state.

Updating occurs in both the classical and non-classical case. The difference is that in the classical case only mixed state are revised and once you have reached a pure state (that is a state of complete information), there is no more updating. The distinguishing feature of the non-classical case is that even pure states change under the impact of a measurement and thus induce a revision of beliefs. When we perform a measurement on a pure state our information does not increase (it is already maximal), it simply changes.

We illustrate this point in a few examples below.

8 Examples revisited

Example 1 revisited. Let S be a set, and let $\tau(s,t) = 0$ for every distinct elements s, t of S. This define a TPS. The corresponding event orthoposet is the Boolean lattice 2^S, as in Example 1. This is the classical framework where measurements do not impact on pure states.

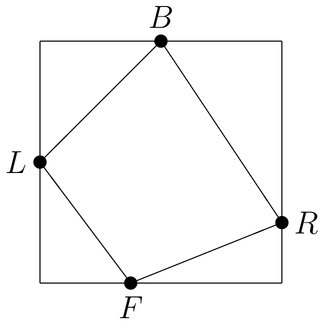
Example 2 revisited. Let $\mathcal{H}$ be a Hilbert space, and S be the set of one-dimensional vector subspaces in $\mathcal{H}$. For $s, t \in S$ let $\tau(s,t)$ be equal to $\cos^2(\varphi)$, where φ is the angle between the line s and t. The axioms T1 and N2 are obvious; T3 follows from Parseval equality. Thus we get a TPS. The corresponding orthoposet of event is the lattice of vector spaces in $\mathcal{H}$, as before. The updating formula is given by the well-known Luder-von Neumann's postulate (the projection principle).

Example 3 revisited. Example 3 cannot be represented by any TPS, because the orthoposet of events is not orthomodular.

Example 4 revisited. Example 4 can be realized as a TPS, moreover, in many different ways. Let $S = \{L, R, F, B\}$, and let τ be given by the following tableau (where $\tau(s,t)$ is placed in the intersection of s-row and t-column)

	L	R	F	B
L	1	0	y	$1-y$
R	0	1	y'	$1-y'$
F	x	$1-x$	1	0
B	x'	$1-x'$	0	1

These states are realized (in the square $\Delta(\mathcal{E})$) as it is shown in the figure below

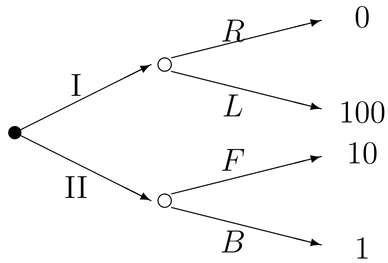


For example, the point R has the coordinates $(1, 1 - y')$ whereas the point F has the coordinate $(1 - x, 0)$. The set $\hat{S}$ of mixed states is the convex hull of the four points.

An illustration.

Let us illustrate with the help of Example 4 a difference between the classical and the non-classical models. To make the presentation a little more concrete, we formulate it in a game context. Our decision-maker faces uncertainty about the type (or preferences) of the agent whom he is interacting with. "The world" here is another decision-maker.[5] In that context the term type is equivalent to the term "state" when talking about arbitrary systems. A decision situation is measurement. The outcome of such a measurement, i.e., a choice, provides information about the type of the agent who made the decision.

Consider the following simple game. Player 1 (our DM) moves first, then player 2 moves and the game is over. The tree of the gave is depicted in the figure below



We only write the payoffs of player 1 and we consider this game from player 1's point of view. He has to choose between move I and II. Move I yields him a payoff of 0 or 100, and move II a payoff of 10 or 1 depending on the choice (type) of player 2 which is unknown to player 1. To make his decision our DM has to have a model of his opponent (or a model of the uncertainty).

A classical model consists of four (pure) states (types) of the player 2: LF, LB,

[5]The idea that agent (represented by its preferences and beliefs) may be viewed as non-classical system was first proposed in [9]. The motivation for this approach is that a variety of empirical phenomena (see, for instance [2]), the so-called behavioral anomalies, can be explained when representing uncertainty about the type (preferences) of a decision-maker with a non-boolean ortholattice.

RF, and RB and a prior belief on this set. Suppose that player 1 knows nothing about the type of his opponent so he e.g., uses the uniform distribution (1/4,1/4,1/4,1/4) as his priors. This prompts our DM to choose the move I.

As an alternative we consider a non-classical TPS-model of the player 2. He has four (pure) states R, L, F, and B. Again he knows nothing about the state of the player 2 and his belief β is presented by the central point of the square, that is $\beta(L) = \beta(R) = \beta(F) = \beta(B) = 1/2$. Here too player 1 chooses the move I.

Thus, in the one-shot game we do not see any difference between the two models with respect to their recommendations to play. But imagine that the game is played three times in succession (with the same opponent). Suppose that the DM makes the move I, and his opponent makes the move R. In the classical model we conclude that the state of player 2 is the mixture $1/2(RF) + 1/2(RB)$. So at the second stage of the game the DM chooses the move II. In the non-classical model we conclude that the state of the player 2 is the pure state R. So that in the second stage the DM should also choose the move II. Again we see no difference between the two models.

Imagine, however, that player 2's move at the second stage is B. The classical model yields that the actual state of the player 2 is RB, unfortunately. At the third stage he should choose the move II and expect 1 util. In contrast, the non-classical model tells us that the current state of the player 2 is B. And if player 1 chooses the move I, his expected payoff (or its CE) is 50 utils which is more than the expected payoff of the move II. So we see that here the recommendations of the two models differ and so do the expected payoffs.

Which one of the two models is the correct one (if any of them) is ultimately an empirical question.

9 Conclusion

In this paper we show that Savage's theory of decision-making under uncertainty can be formulated in terms of a very general algebraic structure called an orthoposets instead of the more restrictive Boolean algebra. Our results shed new light on the generality of the Savage's approach. They also extends it so as to allow considering decision situations where the payoff relevant uncertainty pertains to non-classical objects. In this section we want to discuss some limitations of our approach.

Savage arguments are formulated in a static context. In a static context a classical state space can represent uncertainty even when measurements are incompatible provided they have disjoint outcome sets (as in Example 4). It is therefore legitimate to question the practical value of the proposed non-classical generalization. Our response is that the non-classical representation of uncertainty becomes truly valuable when we consider a dynamic situations, i.e. a situation when a series of decisions under uncertainty is to be made.

In a classical world, properties pre-exist the measurements, they are only revealed by measurements. As the decision-maker proceeds in the series of decisions, properties of the world (type characteristics of the agent) become known to him. The decision-maker, with a classical representation in mind, makes his next decisions on the basis of updated beliefs according to the Bayes' rule. But if the system is non-classical performing measurement alters state of the system. Bayes' rule which presupposes that pure states remain unchanged is not longer appropriate. In section 8 we demonstrated that already in a simple case the classical and the non-classical representation of uncertainty yield distinct recommendations for decision-making.

Appendix. Qualitative Measures

In this paper we model uncertainty by an orthoposet of events. Therefore we may talk of smaller or larger probability for the realization of events. In this appendix, we focus on the qualitative relation corresponding to the "more (or less) likely than" relation between events.

Definition. A *qualitative measure* on an orthoposet $\mathcal{E}$ is a binary relation (of "likelihood") $\preceq$ on $\mathcal{E}$ satisfying the following two axioms:

QM1. $\preceq$ is complete and transitive.

QM2. Let $a \preceq b$ and $a' \preceq b'$. Then $a \oplus a' \preceq b \oplus b'$ (recall that it means that $a \perp a'$ and $b \perp b'$). The last inequality is strict if at least one of the first inequalities is strict.[6]

We shall call a (quantitative) *measure* on orthoposet $\mathcal{E}$ an arbitrary ortho-additive mapping $\mu : \mathcal{E} \to \mathbb{R}$; the positivity does not assumed. Each measure μ defines (generates) the qualitative measure $\preceq = \preceq_\mu$: $a \preceq b$ if and only if $\mu(a) \leq \mu(b)$. In this appendix we are interested by the question as to when a qualitative measure can be generated by a quantitative measure (or when there exists a *probabilistic sophistication*). For simplicity we shall assume that the orthoposet $\mathcal{E}$ is finite. Even in the classical context the answer is generally negative ([8]), however. Therefore, in order to obtain a positive answer, we have to impose some additional conditions strengthening QM2. Here we consider a condition generalizing the classical "cancellation condition" (see [11]). We prefer to call it "hyper-acyclicity".

[6] The special case of QM2 when $a' = b'$ is referred to in [11] as de Finetti axiom.

Definition. A binary relation $\preceq$ on $\mathcal{E}$ is said to be *hyper-acyclic* if the following condition holds:

Assume that we have a finite collection of pairs (a_i, b_i) such that $a_i \preceq b_i$ for all i. If $\sum \mu(a_i) = \sum \mu(b_i)$ for every measure μ on $\mathcal{E}$ then $b_i \preceq a_i$ for all i .

Clearly, if a qualitative relation $\preceq$ is generated by a measure μ then it is hyper-acyclic. The main result of this section (and the analog of Theorem 1 in [11]) asserts that for finite ortholattice the reverse is true.

Theorem 3 *Let $\preceq$ be a hyper-acyclic weak order on a finite orthoposet $\mathcal{E}$. Then $\preceq$ is generated by some measure on $\mathcal{E}$.*

Clearly, if the relation $\preceq$ is monotone (that is $a \preceq b$ for $a \leq b$), then any measure μ generating $\preceq$ is also monotone. If, in addition, $\mathbf{0} \prec \mathbf{1}$ then $\mu(\mathbf{1}) > 0$; dividing the measure μ by $\mu(\mathbf{1})$ we can assume that μ is a normed measure. Thus, the measure μ is a monotone probability.

the outtine of the proof is as follows. We embed the orthoposet $\mathcal{E}$ into some vector space V and identify linear functionals on V with measures on $\mathcal{E}$. With the qualitative measure $\preceq$ we construct a subset $P \subset V$ and show that 0 does not belong to the convex hull of P. The separability theorem then guarantees the existence of a linear functional on V (hence a measure on $\mathcal{E}$) which is strictly positive on P. It is easy to show that this measure generates the relation $\preceq$.

The proof proceeds in several steps**.**

1. *Construction of the vector space V*. Denote $\mathbb{R} \otimes \mathcal{E}$ the vector space generated by the set $\mathcal{E}$. It consists of (finite) formal expressions of the form $\sum_i r_i a_i$, where $r_i \in \mathbb{R}$ and $a_i \in \mathcal{E}$. Let K be the vector subspace in $\mathbb{R} \otimes \mathcal{E}$ generated by expressions $a \oplus b - a - b$ (recall that $a \oplus b$ means that $a \oplus b = a \vee b$ and $a \perp b$.) Finally, $V = V(\mathcal{E})$ is the quotient space $\mathbb{R} \otimes \mathcal{E}$ by the subspace K, $V = (\mathbb{R} \otimes \mathcal{E})/K$.

The orthoposet $\mathcal{E}$ naturally maps into V; the image $1 \cdot a$ of an event $a \in \mathcal{E}$ is denoted a as well. Any linear functional l on V restricted to $\mathcal{E}$ gives a valuation on $\mathcal{E}$. Since $l(a \oplus b - a - b) = l(a \oplus b) - l(a) - l(b) = 0$, the valuation l is additive, that is a measure on the orthoposet $\mathcal{E}$. Conversely, let l be a measure on $\mathcal{E}$. We extend it by linearity to $\mathbb{R} \otimes \mathcal{E}$ assuming $l(\sum r_i a_i) = \sum r_i l(a_i)$. By force of additivity, l yields 0 for elements of the form $a \oplus b - a - b$, that is l vanishes on the subspace K. Therefore l factors through V and is obtained from a linear functional defined on V. We just proved

Proposition 5 *The vector space of measures on $\mathcal{E}$ is identified with the vector space V^* of linear functional on V.*

Remark. The canonical mapping $\mathcal{E} \to V(\mathcal{E})$ can be considered as the universal measure on the orthoposet $\mathcal{L}$. It is injective if and only if the orthoposet $\mathcal{E}$ is an OMP.

2. *Construction of the set P.* Let $\preceq$ be a binary relation on $\mathcal{E}$; as usual, $\prec$ denote the strict part of $\preceq$. By the definition, $P = P(\preceq)$ consists of (finite) expressions of the form $\sum_i (a_i - b_i)$, where $b_i \preceq a_i$ for all i and $b_i \prec a_i$ for some i. (P is empty if the relation $\prec$ is empty, that is if all elements in $\mathcal{E}$ are equivalent relatively to $\preceq$.) We note also that P is stable with respect to the addition.

3. Suppose now that a relation $\preceq$ is hyper-acyclic. Note that the hyper-acyclicity of $\preceq$ means precisely that 0 does not belongs to P.

Proposition 6 *If the relation $\preceq$ is hyper-acyclic then* 0 *does not belong to the convex hull $co(P)$ of P in V.*

Proof. Assume that 0 is a convex combination of elements of P, $0 = \sum_i r_i p_i$, where $p_i \in P$, $r_i \geq 0$, and $\sum_i r_i = 1$. By Caratheodory's theorem we can assume that the p_i are affinely independent (and therefore the coefficients r_i are uniquely defined). We assert that in this case the coefficients are *rational* numbers.

It would be simplest to say that the set P is defined over the field of rational numbers. But it is not so easy to provide a precise meaning to it. For that purpose we choose and fix some subset $L \subset \mathcal{E}$, such that its image in V is a basis of this vector space. We also choose a subset M of expressions of the form $a \oplus b - a - b$, which constitute a basis of the subspace K. The union of L and M is a basis of the vector space $\mathbb{R} \otimes \mathcal{E}$. On the other hand side, $\mathcal{E}$ is a basis of $\mathbb{R} \otimes \mathcal{E}$ as well. Since elements of $L \cup M$ are rational combinations of elements of the $\mathcal{E}$, elements of $\mathcal{E}$, in turn, can be rationally expressed in terms of $L \cup M$. In particular, the images of elements of $\mathcal{E}$ in V are rational combinations of elements from L. All the more, the elements $p_i \in P$ can be rationally expressed in terms of L. Therefore (see, for example, Proposition 6 in [1], Chap. 2, § 6) that 0 can be expressed rationally through p_i. Since the coefficients r_i are defined uniquely, they are rational numbers.

Now the proof can be easily completed. We have an equality $0 = \sum_i r_i p_i$, where $p_i \in P$ and r_i are rational numbers (not all equal to zero). Multiplying with a suitable integer we may consider r_i themselves as integers. Since P is stable with respect to addition, we obtain that $0 \in P$, in the contradiction with hyper-acyclicity of the relation $\preceq$.

4. Together with Separation theorem of convex sets (see [17]) the results above imply the existence of a (non-trivial) linear functional μ on V, which is non-negative

on P. But we need strict positivity on P. To obtain the strict positivity we show that (in the case of a finite orthoposet $\mathcal{E}$) the convex hull of P is a polyhedron.

Let us introduce some notations. A denotes the set of expression $a-b$, where $a \succ b$. B denotes the set of rays of the form $\mathbb{R}_+(a-b)$, where $a \succeq b$. Finally, Q is the convex hull of $A \cup B$ in V. By definition, Q consists of elements of the form

$$q = \alpha_1(a_1 - b_1) + ... + \alpha_n(a_n - b_n) + \beta_1(c_1 - d_1) + ... + \beta_m(c_m - d_m), \qquad (*)$$

where $a_i, b_i, c_j, d_j \in \mathcal{E}$ (more precisely, belong to their image in V), $a_i \succ b_i$ for any i, $c_j \succeq d_j$ for any j, α_i, β_i are nonnegative, and $\sum_i \alpha_i = 1$.

Proposition 7 *The convex hull of P coincides with Q.*

Proof. It is clear from the definitions that any element of P belongs to Q. By the convexity of Q, the convex hull of P is also contained in Q.

It remains to show the converse, that any element q of Q belongs to the convex hull of P. For that (appealing to the convexity of $co(P)$) we can assume that q has the form in $(*)$ with n and m equal to 1, that is

$$q = (a - b) + \beta(c - d),$$

where $a \succ b$, $c \succeq d$ and $\beta \geq 0$. If β is an integer, it is clear that $q \in P$. In general case β is a convex combination of two nonnegative integers β_1 and β_2; then q is the corresponding convex combination of two points $(a-b)+\beta_1(c-d)$ and $(a-b)+\beta_2(c-d)$ both belonging to P.

Corollary 2 *Assume that an orthoposet $\mathcal{E}$ is finite. Then $co(P)$ is a polyhedron.*

In fact, in this case the sets A and B are finite. Therefore (see [17], theorem 19.1) Q is a polyhedra.

Thus, if 0 does not belong to the convex hull of P (see Proposition 2) then there exists a linear functional μ on V which is strictly positive on P. As we shall see, this immediately provides us with a proof of Theorem 1.

5. *End of the proof.* The assertion in the theorem is trivially true if all elements of $\mathcal{E}$ are equivalent to each other. Therefore we can assume that there exists at least one pair (a, b) such that $a \succ b$. Let μ be a linear functional on V (we may consider μ as a measure on the ortholattice $\mathcal{E}$) strictly positive on P. We assert that this measure generates the relation $\preceq$.

Let us suppose $c \succeq d$. Since for any integer positive number n the element $(a - b) + n(c - d)$ belongs to P, we have $\mu(a) - \mu(b) > n\ (\mu(d) - \mu(c))$ for any n. This implies $\mu(d) \leq \mu(c)$. Conversely, let us suppose $\mu(c) \geq \mu(d)$ for some $c, d \in \mathcal{E}$. We have to show that $c \succeq d$. If this is not the case then, by completeness of the relation $\succeq$, we have $d \succ c$. But then $d - c$ belongs to P and $\mu(d - c) = \mu(d) - \mu(c) > 0$, which contradicts to our first assumption. This completes the proof of Theorem 3.

References

[1] Bourbaki N. (1962) *Algebra*, Hermann, Paris.

[2] Busemeyer J.R., Wang Z. and Townsend J.T. "Quantum Dynamics of Human Decision-Making" *Journal of Mathematical Psychology* 50, 220-241.

[3] Danilov V.I. and A. Lambert-Mogiliansky (2006) Measurable Systems and Behavioral Sciences, forthcoming in *Mathematical Social Sciences*. See also xxx.lanl.gov/physics/0604051.

[4] Deutcsh D. (1999) Quantum theory of probability and decisions, *Proceedings of the Royal Society of London*, A455, 3129

[5] Greechie R.J. (1971) Orthomadular lattices admitting no states. *Journal of Combinatorial Theory* 10, 119-132.

[6] Gyntelberg J. and F. Hansen (2004) Expected utility theory with "small worlds". FRU Working Papers 2004/04, Univ. Copenhagen, Dep. of Economics.

[7] Khrennikov A. Yu (2007) "A Model of Quantum-like decision-making with application to psychology and Cognitive Sciences" http://arhiv.org/abs/0711.1366

[8] Kraft C.U., Pratt J.W., and Seidenberg A. (1959) Intuitive probability on finite sets, *Ann. Math. Stat.* 30, 408-419.

[9] Lambert-Mogiliansky A., S. Zamir, and H. Zwirn (2008) "Type-indeterminacy - A Model for the KT-(Kahneman and Tversky)-man", *forthcoming in the Journal of Mathematical Psychology.*

[10] La Mura P. (2005) Decision Theory in the Presence of Risk and Uncertainty. *mimeo* Leipzig Graduate School of Business.

[11] Lehrer E. and Shmaya E. (2005) A Subjective Approach to Quantum probability. *Proceedings of the Royal Society A* 462.

[12] Mielnik B. (1968) Geometry of Quantum States. Comm. *Math. Phys.* 9, 55-80.

[13] Penrose R. (1994) *Shadows of the Mind*, Oxford Univ. Press, New York.

[14] Pitowsky I. (2003) Betting on the outcomes of measurements. *Studies in History and Philosophy of Modern Physics* 34, 395-414. See also xxx.lanl.gov/quant-ph/0208121

[15] Pitowsky I. (2005) Quantum mechanics as a theory of probability, arXiv:quant-ph/0510095.

[16] Pulmanova S. (1986) Transition probability spaces, *Journal of Mathematical Physics* 27, 1791-1795.

[17] Rockafeller R.T. (1970) *Convex Analysis*, Princeton University Press, Princeton.

[18] Savage L. (1954) *The Foundations of Statistics.* John Wiley, New York.

[19] von Neumann J. (1932) *Mathematische Grunlagen der Quantummechanik.* Springer-Verlag, Berlin.

[20] Wright R. (1978) The state of the Pentagon, in *Mathematical Foundation of Quantum Theory* (ed. A. R. Marlow), Academic Press New York San Fransisco London.