Non-Archimedean Preferences Over Countable Lotteries

Jeffrey Sanford Russell

March 2020

Abstract

We prove a representation theorem for preference relations over countably infinite lotteries that satisfy a generalized form of the Independence axiom, without assuming Continuity. The representing space consists of lexicographically ordered transfinite sequences of bounded real numbers. This result is generalized to preference orders on abstract superconvex spaces.

1 Infinite Lotteries and Independence

Let a (discrete) **lottery** over a set of outcomes X be a function $\lambda : X \to [0, 1]$ such that $\sum_{x \in X} \lambda(x) = 1$. Let ΔX be the set of all such lotteries. Let a **preference order** be a total preorder $\lesssim$ on ΔX. (That is, $\lesssim$ is assumed to be transitive, reflexive, and complete in the sense that for each pair of lotteries, $\lambda \lesssim \mu$ or $\mu \lesssim \lambda$.)

A standard axiom of rational preference is the *Independence* axiom

(I) For any lotteries $\lambda, \lambda', \mu \in \Delta X$ and any probability $0 < p \leq 1$,

$$\lambda < \lambda' \qquad \text{iff} \qquad p \cdot \lambda + (1 - p) \cdot \mu < p \cdot \lambda' + (1 - p) \cdot \mu$$

Let an **indexed family of lotteries** be a function $\lambda : I \to X$, for any set I. For total preorders, Independence can be equivalently restated:[1]

---

[1] Hammond's (1998) label (D) for this principle stands for "dominance." But note that the principle is stronger than a natural analogous dominance principle for *outcomes*, since it applies generally to mixtures of *lotteries*. (See Buchak 2013, 162ff, for discussion of the normative import of this distinction.)

(D) For any finite set I of indices, any probability distribution $p \in \Delta I$, and any indexed families of lotteries $\lambda, \mu : I \to X$, if $\lambda_i \lesssim \mu_i$ for each $i \in I$, then

$$\sum_i p(i) \cdot \lambda_i \lesssim \sum_i p(i) \cdot \mu_i$$

If furthermore $\lambda_i < \mu_i$ for some i such that $p(i) > 0$, this preference is strict.

This naturally suggests a generalized principle, which simply drops the qualification "finite" (see Blackwell and Girshick 1954; Hammond 1998, sec. 8.4).[2]

(**D***) For any set I of indices, any probability distribution $p \in \Delta I$, and any families of lotteries $\lambda, \mu : I \to X$, if $\lambda_i \lesssim \mu_i$ for each $i \in I$, then

$$\sum_i p(i) \cdot \lambda_i \lesssim \sum_i p(i) \cdot \mu_i$$

If furthermore $\lambda_i < \mu_i$ for some $i \in I$ such that $p(i) > 0$, this preference is strict.

This very natural principle can be motivated by some of the same considerations as the original Independence principle and its close cousin, the Sure Thing Principle. For example, Savage (1954, sec. 2.7) writes:

> A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant to the attractiveness of the purchase. So, to clarify the matter for himself, he asks whether he would buy if he knew that the Republican candidate were going to win, and decides that he would do so. Similarly, he considers whether he would buy if he knew that the Democratic candidate were going to win, and again finds that he would do so. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event obtains … [E]xcept possibly for the assumption of simple ordering, I know of no other extralogical principle governing decisions that finds such ready acceptance.

The principle (D*) can be similarly motivated, by generalizing to questions with more than two possible answers. Let $E_1, E_2, \ldots$ be countably many mutually exclusive and jointly exhaustive events. If you prefer to buy if E_1 obtains, and you also

---

[2]Again, this label is Hammond's. Russell and Isaacs (forthcoming) call this principle "Countable Independence," to distinguish it more clearly from the weaker dominance principle concerning *outcomes*. The qualification "Countable" is included just because any discrete probability distribution $p \in \Delta I$ must have countable support.

prefer to buy if E_2 obtains, and so on, then this seems like a strong reason to buy, unconditionally. Suppose that you assign probability $\lambda(x)$ to the hypothesis that outcome x will result from buying (for each $x \in X$). Likewise, you assign conditional probability $\lambda_i(x)$ to the hypothesis that outcome x will result from buying *if* event E_i obtains. By the law of total probability, $\lambda = \sum_i p_i \cdot \lambda_i$, where p_i is the probability you assign to event E_i. So if your conditional and unconditional preferences between *acts* accord with your preferences between these associated *lotteries*, (D*) follows from the countable generalization of Savage's principle.[3]

When the set of outcomes is infinite, (D*) genuinely strengthens the finitary condition (D), and the equivalent Independence axiom (I). There are preference orderings on ΔX satisfying Independence which are representable by *unbounded* real utilities (see Fine 2008). But no such orderings satisfy (D*). This arises from well-known cases like the St. Petersburg paradox and the two-envelope paradox (Menger 1934; see also Arntzenius and McCarthy 1997, 44; for a striking related result see Seidenfeld, Schervish, and Kadane 2009).

One standard response to these puzzle cases is to impose the requirement that preferences can be represented by a bounded set of real numbers. But in fact, this is an overreaction. While (D*) is incompatible with preferences that are representable by *unbounded* real utilities, it is perfectly compatible with preferences that are not representable by *bounded* real utilities—because these preferences are not representable by real-valued utilities at all. A simple example is the lexicographic order on lotteries in $\Delta\{0, 1, 2\}$ with three possible outcomes. Preferences like these conflict with the standard *Archimedean Axiom* (or the *Continuity* axiom, which is equivalent in this context), which essentially rules out "infinite" or "infinitesimal" utilities. In the presence of the Archimedean Axiom, (D*) is indeed equivalent to representability by bounded real utilities (Blackwell and Girshick 1954; Hammond 1998, sec. 8.4 and 8.5).

But while the combination of (D*) and the Archimedean Axiom is well-understood, it is also worthwhile to investigate the constraints imposed on preferences by (D*) in the general non-Archimedean case. This is for two reasons. First, while there are strong normative arguments for (D*), there are not such strong arguments in favor of the Archimedean Axiom; the latter may not be a genuine constraint on *rational* preference, but merely a technically convenient simplification (see for example Luce and Raiffa 1989, 27; Hájek and Nover 2008, 649ff). Second, this exploration helps us conceptually distinguish two very different kinds of "infinity problem" in decision theory. Some difficulties arise from *unbounded* utilities, like the St. Petersburg

---

[3]For a closely related "consequentialist" argument see Hammond (1998, sec. 8.4). For further discusssion see Russell and Isaacs (forthcoming, sec. 2).

puzzle and the two-envelope puzzle. Other difficulties arise from *infinite* utilities, like Pascal's wager (for discussion see Hájek 2003). If (D*) is a normative requirement on preferences, it turns out that this will spare rational agents the first kind of difficulty—but not the second.[4]

The principle (D*) imposes interesting structural constraints on preferences. In a sense we can make precise, this principle rules out *unbounded finite* utilities; but it is perfectly compatible with *infinite* utilities, and indeed, with infinitely many different "regimes" of infinite utilities. The structure of these "regimes" is highly constrained: in particular, (D*) rules out ascending chains of lotteries each of which is infinitely better than those before it.

In other words, these utility "regimes" or dimensions have the structure of *ordinals*. Classic representation theorems for non-Archimedean preferences (such as Fishburn 1971) represent utilities as lexicographically ordered *n*-tuples of real numbers $(u_1, u_2, \ldots, u_n)$, indexed by numbers up to some finite n.[5] But not all preferences satisfying (D*) have utility representations given by finite sequences: in general, some preferences require infinite sequences. Moreover, they can require quite *long* infinite sequences. Some infinite sequences are indexed by just the finite numbers $1, 2, 3, \ldots$. Other sequences have further dimensions beyond those, and may be indexed by *ordinal* numbers

$$1, 2, \ldots, \omega, \omega + 1, \omega + 2, \ldots, 2\omega, 2\omega + 1, \ldots, \omega^2, \omega^2 + 1, \ldots$$

and so on, as far as you like.[6] In principle, the length of the utility sequences we'll need is only limited by the cardinality of the set of outcomes X: if this is a large infinite set, then we may require very large infinite ordinal numbers. The key structural feature of the ordinals is that they are *well-ordered*. This tells us that for any two distinct sequences of real numbers indexed by ordinals, there is a *first* index at which they differ. Thus for any ordinal α, we can *lexicographically* order the sequences of real numbers indexed by the ordinals before α.

---

[4]Collapsing these issues might lead one to the natural but incorrect conclusion that the difficulties raised by the St. Petersburg game could be avoided by dropping the Archimedean Axiom, as suggested by Hájek and Nover (2008).

[5]Fishburn (1974) and Martínez-Legaz (1998) survey prior work on lexicographic utilities. Borie (2016) proves a utility representation theorem for non-Archimedean preferences without completeness. Fishburn's and Borie's theorems apply to *finite* mixture spaces, without considering constraints like (D*) on preference between infinite mixtures of lotteries.

[6]For an introduction to ordinals see for example Hrbacek and Jech (1999) or Potter (2004). Exactly how far the ordinals go is a deep question in the foundations of mathematics: it is equivalent to the question of what "large cardinals" there are, and many different answers to this question are consistent with the standard axioms of set theory, ZFC. But these axioms do tell us that for any set whatsoever (of any cardinality) there are enough ordinals to index that set.

Here we will give a utility representation theorem for preference orders that satisfy (D*), without assuming the Archidemean Axiom: preference orders that satisfy (D*) are precisely those that can be represented by *bounded lexicographic utilities*. This representation theorem extends the classic results of Hausner (1952) and Hausner and Wendel (1952) concerning non-Archimedean preference orders. Hausner and Wendel's results show that any preference order on lotteries that satisfies Independence is representable by a convex subset of a lexicographically ordered vector space. Our extension of their result shows that the infinitary generalization of Independence (D*) requires, first, that this lexicographically ordered vector space takes a special concrete form—namely, the Banach space of bounded sequences $\ell^\infty(\alpha)$ for some ordinal α—and second, that this convex subset is itself bounded (with respect to the ∞-norm). Restricting utilities to a bounded set of sequences blocks the construction of a St. Petersburg lottery. Nonetheless, this representation permits highly non-trivial "infinite utility" structure, since the space of utilites can have arbitrarily high dimension.

2 Representation Theorem

In this section we will take lotteries to be concrete probability distributions. In we will consider a more general approach where "lotteries" are elements of a space with an abstract countable mixture operation, called a *superconvex space*. As it turns out, this apparent additional generality does not actually add much: any ordering of such space is equivalent to an ordering on an ordinary lottery space.

We begin with some basic definitions.

Definition 1. (König 1986) Let V be a Banach space. A subset $C \subseteq V$ is called **σ-convex** iff C is closed under countable mixtures, in the following sense: for each $p \in \Delta\omega$ and sequence $x \in C^\omega$, the infinite sum

$$\sum_{i<\omega} p_i \cdot x_i$$

is well-defined and contained in C. We will also use the more compact notation $\Sigma(p, x)$ for this mixture.

If $C \subseteq V$ and $C' \subseteq V$ are both σ-convex sets, a function $f : C \to C'$ is **mixture-preserving** iff for each $p \in \Delta\omega$ and $x \in C^\omega$,

$$f\big(\Sigma(p, x)\big) = \sum_i p_i \cdot f(x_i)$$

The restriction of any continuous linear map to a σ-convex set is mixture-preserving. In particular, this holds for any *contraction* (a linear map $f : V \to V'$ such that $\|f(v)\|_{V'} \leq \|v\|_V$; equivalently, $f(BV) \subseteq BV'$).

Example 1. For any nonempty set X, the set ΔX of discrete lotteries with outcomes in X is a σ-convex subset of the Banach space $\ell^1(X)$ (consisting of all functions $u : X \to \mathbb{R}$ for which the 1-*norm* $\|u\| = \sum_{x \in X} |u(x)|$ is finite).

Example 2. Let V be a Banach space. The **closed unit ball** in V

$$BV = \left\{ x \in V \mid \|x\| \leq 1 \right\}$$

is a σ-convex subset of V. In particular, let α be an ordinal, and consider the Banach space $\ell^\infty(\alpha)$ of length α bounded ordinal sequences of real numbers, with the ∞-norm $\|x\| = \sup_{i<\alpha} |x_i|$. The closed unit ball $B\ell^\infty(\alpha)$, which consists of all α-sequences of numbers in the interval $[-1, 1]$, is a σ-convex subset of $\ell^\infty(\alpha)$.

Definition 2. (a) A **preorder** on a set X is a reflexive and transitive relation $\lesssim$ on X. We say $x \sim y$ iff $x \lesssim y$ and $y \lesssim x$. We say $x < y$ iff $x \lesssim y$ but not $y \lesssim x$. A preorder is **total** iff for each $x, y \in X$, either $x \lesssim y$ or $y \lesssim x$. A **partial order** is a preorder which is also antisymmetric (if $x \sim y$ then $x = y$). A **total order** is a partial order which is also total.

(b) If $(X, \lesssim)$ and $(X', \lesssim')$ are preordered sets, a function $f : X \to X'$ is **monotone** iff for any $x \lesssim y$ in X, $f(x) \lesssim' f(y)$ in X'. It is **strictly monotone** iff the converse also holds, so for any $x, y \in X$, $x \lesssim y$ iff $f(x) \lesssim' f(y)$.

(c) A preorder on a vector space V is **linear** iff vector addition and multiplication by positive scalars are both monotone functions. That is, for any $x, y, z \in V$ and $a \in \mathbb{R}^+$, if $x \lesssim y$ then $x+z \lesssim y+z$ and $ax \lesssim ay$. A **preordered/partially ordered/totally ordered vector space** is a vector space equipped with a *linear* preorder/partial order/total order.

(d) If V and V' are preordered vector spaces, a function $f : V \to V'$ is **positive** iff for each $x \in V$ such that $x \gtrsim 0$, $f(x) \gtrsim 0$, and **strictly positive** iff additionally for each $x \in V$ such that $x > 0$, $f(x) > 0$.

It is straightforward to check that a linear map between preordered vector spaces is (strictly) positive iff it is (strictly) monotone.

Here is the main theorem we will prove:

Theorem 1. *For any set X, a total preorder on the set ΔX of discrete lotteries satisfies (D*) iff for some ordinal α there is a strictly monotone mixture-preserving function from ΔX to $B\ell^\infty(\alpha)$ with its lexicographic order.*

6

First, we will show that the order on lotteries can be represented in a totally ordered Banach space which a certain boundedness property. Second, we will show that we can define a strictly monotone contraction from any such ordered Banach space to $\ell^\infty(\alpha)$.

We begin by extending the order on lotteries to the surrounding Banach space $\ell^1(X)$.

Lemma 1. For any total preorder $\precsim$ on ΔX that satisfies the Independence axiom (I), there is a unique linear total preorder $\precsim^*$ on $\ell^1(X)$ such that

    (a) The embedding of $(\Delta X, \precsim)$ in $(\ell^1(X), \precsim^*)$ is strictly monotone; and
    (b) The function $h : \ell^1(X) \to \mathbb{R}$ such that $h(u) = \sum_{x \in X} u(x)$ is positive.

(In particular, for each lottery $\lambda \in \Delta X$, we have $h(\lambda) = 1$ and thus $\lambda >^* 0$.)

Proof. First, for any nonzero $u \in \ell^1(X)$ such that $h(u) = 0$, there exist $\lambda, \mu \in \Delta X$ such that $u = \|u\| \cdot (\lambda - \mu)$. For $x \in X$ let

$$u^+(x) = \begin{cases} u(x) & \text{if } u(x) > 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad u^-(x) = \begin{cases} -u(x) & \text{if } u(x) < 0 \\ 0 & \text{otherwise} \end{cases}$$

We have $0 = h(u) = \|u^+\| - \|u^-\|$ and $\|u\| = \|u^+\| + \|u^-\|$. It follows that $\lambda = 2u^+/\|u\|$ and $\mu = 2u^-/\|u\|$ are elements of ΔX such that $u = \|u\|(\lambda - \mu)$.

Next, we define the order on $\ell^1(X)$ lexicographically, first using the function h, and second using the order on lotteries. For $u \in \ell^1(X)$, let $u \succsim^* 0$ iff either of the following cases holds:

    (a) $h(u) > 0$, or
    (b) $h(u) = 0$ and $u = \|u\| \cdot (\lambda - \mu)$ for some $\lambda, \mu \in \Delta X$ such that $\lambda \succsim \mu$.

(In general, $v \succsim^* u$ iff $v - u \succsim^* 0$.) It is straightforward to check that (I) implies that this order is linear and satisfies the two conditions of the lemma.

For uniqueness, it suffices to check that, for any linear order $\precsim^*$ on $\ell^1(X)$, and any nonzero $u \in \ell^1(X)$, if $h(u) = 0$ and $u = \|u\|(\lambda - \mu)$, then $\lambda \succsim \mu$ iff $u \succsim^* 0$. Both directions follow from linearity.

$\square$

In order to get to a totally ordered Banach space, the natural thing to do next is quotient out the set I of vectors which are equivalent to zero. The difficulty is that in general the quotient space may not retain the structure of a Banach space: this holds only if the subspace I is *closed*. It does turn out that our axiom (D*) guarantees this, but this takes a little bit of showing.

Lemma 2. For any preordered Banach space V, let

$$B^+V = \{x \in BV \mid x > 0\}$$

If $(\Delta X, \lesssim)$ satisfies (D*), then $B^+\ell^1(X)$ is σ-convex (with respect to the extended order $\lesssim^*$).

Proof. Let $p \in \Delta\omega$ and $x_1, x_2, \dots \in B^+\ell^1(X)$. Then $h(x_i) \geq 0$ for each i. If $h(x_i) > 0$ for any i, then

$$h(\Sigma(p, x)) = \sum_i p_i \cdot h(x_i) > 0$$

so in that case $\Sigma(p, x)$ is positive as well. Otherwise, $h(x_i) = 0$ for every $i \in \omega$. In that case, for each i there is some $\lambda_i > \mu_i$ in ΔX such that $x_i = \|x_i\|(\lambda_i - \mu_i)$. Since $\|x_i\| \leq 1$, the series $\sum_i p_i \|x_i\|$ converges to some value $a \in \mathbb{R}^+$. Then

$$\sum_i p_i \cdot x_i = a\Big(\sum_i \frac{p_i \|x_i\|}{a} \cdot \lambda_i - \sum_i \frac{p_i \|x_i\|}{a} \cdot \mu_i \Big)$$

By (D*) and the definition of $\lesssim^*$, this is positive.

$\square$

The following very simple fact really contains the core idea of the main theorem. The central "paradoxical" feature of the St. Petersburg lottery is that it is strictly preferred over every one of its outcomes. (D*) rules out this situation.

Lemma 3. Let V be a preordered Banach space such that B^+V is σ-convex. There is no $p \in \Delta\omega$ and $x_1, x_2, \dots \in B^+V$ such that

$$x_i < \Sigma(p, x) \quad \text{for each } i \in \omega$$

Proof. If there were such p and $x_1, x_2, \dots$, then the vectors $\frac{1}{2}(\Sigma(p, x) - x_i)$ would each be in B^+V. This would imply that B^+V also contains the mixture

$$\sum_i p_i \cdot \Big(\frac{1}{2}(\Sigma(p, x) - x_i) \Big) = \frac{1}{2}\Big(\Sigma(p, x) - \sum_i p_i \cdot x_i \Big) = 0$$

And this is false.

$\square$

This means that (D*) blocks the construction of St. Petersburg lotteries. There are two different versions of this. One kind of "improper mixture" arises from unbounded relatively finite utilities, and another arises from infinite ascending chains of relatively *infinite* utilities. Both kinds are captured by the following condition.

Lemma 4. Let V be a preordered Banach space such that B^+V is σ-convex. Then there is no **supergeometric** sequence $x_1, x_2, \dots \in B^+V$ such that $2x_i \lesssim x_{i+1}$ for each $i \in \omega$.

Proof. This is the St. Petersburg paradox. Suppose $x_1, x_2, \dots \in B^+V$ is a supergeometric sequence, and let

$$w = \sum_i 2^{-i} \cdot x_i$$

We will prove that this is an improper mixture. Consider the *truncated* sum

$$
\begin{aligned}
w_n &= \sum_{0 < i \leq 2^n + n} 2^{-i} x_i \\
&\gtrsim \sum_{n < i \leq 2^n + n} 2^{-i} x_i \quad = 2^{-n} \sum_{0 < j \leq 2^n} 2^{-j} x_{n+j} \\
&\gtrsim 2^{-n} \sum_{0 < j \leq 2^n} x_n \quad = x_n
\end{aligned}
$$

using linearity together with the facts that each x_i is positive and $x_n \lesssim 2^{-j} x_{n+j}$ for each j. Furthermore, w is the sum of w_n and a mixture of the remaining terms x_i, each of which are positive. Let $N = 2^n + n$. Then:

$$
\begin{aligned}
w &= w_n + \sum_{i > N} 2^{-i} \cdot x_i \\
&= w_n + 2^{-N} \sum_i 2^{-i} \cdot x_{N+i} \quad > w_n
\end{aligned}
$$

The last step uses our assumption that B^+V is σ-convex, which implies that the mixture of elements of B^+V is positive.

$\square$

Here is another way of putting this consequence.

Definition 3. A preordered vector space V is **nice** iff every norm-bounded subset of V is order-bounded.

Lemma 5. If V is a totally preordered Banach space and B^+V is σ-convex, then every subspace of V is nice.

Proof. Let U be a subspace of V. It suffices to show that B^+U has an order bound in U. If not, then for each $u \in U$ there is some $x \in B^+U$ such that $2u < x$. So in particular there is a supergeometric sequence $x_1, x_2, \ldots$ in B^+U (and thus in B^+V) such that $2x_i < x_{i+1}$ for each i.

$\square$

Now we will introduce a few useful concepts for working with non-Archimedean ordered vector spaces. (See Paulsen and Tomforde 2007; Emelyanov 2014.)

Definition 4. For a preordered vector space V, define

$$I(V) = \{x \in V \mid x \sim 0\}$$
$$N(V) = \{x \in V \mid \text{for some } v \in V, \text{for every } a \in \mathbb{R}^+, -av \lesssim x \lesssim av\}$$

Intuitively, $N(V)$ is the set of vectors which are infinitesimally small in comparision to some other vector. V is **Archimedean** iff $N(V) = I(V)$. V is **trivial** iff $I(V) = V$.

For any $v \in V$, define

$$F(v) = \{x \in V \mid \text{for some } a \in \mathbb{R}^+, -av \lesssim x \lesssim av \}$$

Intuitively, $F(v)$ is the set of vectors which are finite or infinitesimal in comparison to v. A vector $e \in V$ is an **order unit** iff $F(e) = V$.

The following facts are straightforward.

Lemma 6.   (a) For each $v \in V$, $I(V)$, $N(V)$, and $F(v)$ are subspaces of V.

    (b) V is partially ordered iff $I(V) = \{0\}$.

    (c) If e is an order unit for V and $e \in N(V)$, then V is trivial.

    (d) If V is a Banach space and e is an order bound for B^+V, then e is an order unit for V.

Lemma 7. If V is a totally preordered Banach space and B^+V is σ-convex, then $I(V)$ is closed.

Proof. Suppose $v \in V$ is not in $I(V)$: that is, $v \nsim 0$. We will show that v is not in the closure of $I(V)$: that is, there is some $\varepsilon > 0$ such that for each $u \in I(V)$, $\|v - u\| \geq \varepsilon$. Without loss of generality, assume $v \not\lesssim 0$. (Otherwise we can argue similarly regarding $-v$.) Consider the subspace $F(v)$. By Lemma 5, this subspace contains an order bound for $BF(v)$. This means there is some $a \in \mathbb{R}^+$ such that $x \lesssim av$ for each $x \in BF(v)$.

Now suppose $u \in I(V)$: that is, $u \sim 0$. Then $(v - u) \sim v$, which implies that $v - u \in F(v)$, and $(v - u)/\|v - u\|$ is in $BF(v)$. Thus:

$$\frac{v}{\|v - u\|} \sim \frac{v - u}{\|v - u\|} \lesssim av$$

Thus

$$\left(1 - a\|v - u\|\right)v \lesssim 0$$

By assumption $v \not\lesssim 0$, so we must have $\|v - u\| \geq 1/a$.

$\square$

Lemma 8. For any preordered vector space V, there is a strictly positive contraction from V onto some partially ordered vector space iff $I(V)$ is closed.

Proof. It is a standard fact that a subset of V is a closed subspace iff it is the kernel of a continuous linear surjection, and in particular a contraction $f : V \to V'$ (see for example Conway 1994, 70). We can define a partial order $\leq'$ on V' with respect to which f is strictly positive: let $0 \leq' f(x)$ iff $0 \lesssim' x$. The fact that $\ker f = I(V)$ implies that this is a well-defined partial order. (Compare Emelyanov 2014, sec. 1)

$\square$

Putting this much together completes the first part.

Proposition 1. If $\lesssim$ is a total preorder on ΔX that satisfies (D*), then there is a strictly monotone mixture-preserving function from ΔX to $B^+ V$ for some totally ordered Banach space V such that $B^+ V$ is σ-convex.

For the second part, we will prove a representation theorem for totally ordered Banach spaces of this type. The main idea of this part is that we can quotient out the infinitesimals from an ordered vector space. In the case of a non-trivial totally ordered *nice* vector space, the resulting quotient space is in fact isomorphic to $\mathbb{R}$. Thus if every subspace is nice, we can keep descending to smaller and smaller infinitesimal scales, representing each regime by real numbers, until we run out.

To prove this, we will appeal to the following fact, which is a variant of the Hahn-Banach Theorem for ordered vector spaces.

Lemma 9. (Paulsen and Tomforde 2007, Corollary 2.15) Let $(V, \leq)$ be a partially ordered vector space with order unit e. If E is a subspace of V containing e, then any positive linear functional $f : E \to \mathbb{R}$ may be extended to a positive linear functional $g : V \to \mathbb{R}$ whose restriction to E is f.

Corollary 1. If $(V, \leq)$ is a totally ordered nice vector space, then there is a positive contraction $f : V \to \mathbb{R}$ such that $\ker f = \mathcal{N}(V)$.

Proof. Let e be an upper bound for BV, and thus an order unit in V. Assume without loss of generality that $\|e\| \geq 1$. Let $E = \{ae \mid a \in \mathbb{R}\}$, and let $f(ae) = a$ for each element $ae \in E$. Then let g be a positive linear extension of f to V.

To show that g is a contraction, it suffices to show that for each $x \in BV$, $|g(x)| \leq 1$. This is clear, since in that case $-e \lesssim x \lesssim e$, g is positive, and $g(e) = 1$.

To show $\ker g = \mathcal{N}(V)$, first suppose $x \in \mathcal{N}(V)$. It follows that

$$-ae \lesssim x \lesssim ae \quad \text{for every } a \in \mathbb{R}^+$$

Thus since g is positive and extends f,

$$-a \leq g(x) \leq a \quad \text{for every } a \in \mathbb{R}^+$$

and thus $g(x) = 0$. Conversely, if $x \notin \mathcal{N}(V)$, then (since V is totally ordered) for some $a \in \mathbb{R}^+$, $x \gtrsim ae$ or $x \lesssim -ae$. So either $g(x) \geq a$ or $g(x) \leq -a$, and in either case $g(x) \neq 0$.

$\square$

Proposition 2. Let $(V, \leq)$ be a totally ordered Banach space such that B^+V is σ-convex. Then for some ordinal α, there is a contraction from V to $\ell^\infty(\alpha)$ which is strictly positive with respect to the lexicographic order on $\ell^\infty(\alpha)$.

Proof. The idea is to repeatedly apply Corollary 1, descending to smaller and smaller infinitesimal subspaces until we run out.

First we define a descending chain of subspaces of V by transfinite recursion, as follows.
$$V_0 = V$$
$$V_{i+1} = \mathcal{N}(V_i)$$
$$V_\lambda = \bigcap_{i < \lambda} V_i$$

In other words, V_i consists of the "level i infinitesimals (or smaller)." This sequence is strictly decreasing as long as V_i is non-trivial: each subspace V_i is nice, which implies

that it has an order unit, which implies that $\mathcal{N}(V_i) \subsetneq V_i$ unless V_i is trivial. So there must be some least ordinal α such that V_α is trivial. We'll consider the sequence of subspaces $(V_i)_{i<\alpha}$.

For each non-zero vector $x \in V$, there is some $i < \alpha$ such that $x \in V_i \smallsetminus V_{i+1}$. Intuitively, this is the "scale" of x: the largest utility regime to which x makes a non-infinitesimal contribution. (We know $x \notin V_\alpha$, since $x > 0$ and V_α is trivial. So there is some $i < \alpha$ such that $x \notin V_{i+1}$. So there is some *least* ordinal i such that $x \notin V_{i+1}$. If i is a successor ordinal, this immediately implies $x \in V_i$. If i is a limit ordinal, then V_i is the intersection of all previous successor stages, each of which contains x, and so again $x \in V_i$.)

Next, we can define each component of the function $f : V \to \ell^\infty(\alpha)$. Let $i < \alpha$. By Corollary 1, there is a positive contraction $g : V_i \to \mathbb{R}$ such that $\ker g = \mathcal{N}(V_i) = V_{i+1}$. By the Hahn-Banach Extension Theorem (see for example Conway 1994, Theorem 6.2, p. 78), g can be extended to a contraction $f_i : V \to \mathbb{R}$ defined on the whole space. Note that this extension f_i is not generally *positive*, but we do have

$$f_i(x) > 0 \qquad \text{for each } x \in V_i \smallsetminus V_{i+1} \text{ such that } x > 0$$

(In this case we know $f_i(x) = g(x) \geq 0$ because g is positive, and $f_i(x) \neq 0$ because $x \notin \ker g$.)

The resulting linear map $f : V \to \ell^\infty(\alpha)$ with these components is a contraction: by definition, $\|f(x)\|_\infty = \sup_i |f_i(x)|$, and by construction $|f_i(x)| \leq \|x\|$ for each i.

To show that f is strictly positive, consider any $x \in V$ such that $x > 0$. Let i be the "scale" of x such that $x \in V_i \setminus V_{i+1}$. Since $x \in V_i$ and the sequence is nested, for all $j < i$ we have $x \in V_{j+1} = \ker f_j$, which means $f_j(x) = 0$. Since $x \in V_i \smallsetminus V_{i+1}$ and $x > 0$, we also have $f_i(x) > 0$. Thus $f(x) > 0$ according to the lexicographic ordering on $\ell^\infty(\alpha)$.

$\square$

This essentially completes the proof of one direction Theorem 1. For the other direction we just need a couple additional facts, which are straightforward to check.

Lemma 10. Let C and C' be σ-convex sets. If $(C', \lesssim')$ satisfies (D*) and there is a strictly monotone mixture-preserving function $f : C \to C'$, then $(C, \lesssim)$ satisfies (D*) as well.

Lemma 11. The lexicographic order on $B\ell^\infty(\alpha)$ satisfies (D*).

13

Proof of Theorem 1. One direction follows from Lemma 1, Proposition 1, and Proposition 2. The converse follows from Lemma 10 and Lemma 11.

□

3 Superconvex Spaces

So far we have considered the concrete case of preference orders over *lottery* spaces: the space of all discrete probability distributions over a given set of outcomes. But we can alternatively consider a more abstract notion of a discrete mixture space in which the "lotteries" are not given as probability distributions, but rather are just some objects equipped with an abstract "discrete mixture" operation. This kind of structure is called a **superconvex space** (or *superconvex module*) (König 1986; Pümplun 1991; Nörtemann 2002). In general, there are superconvex spaces which are not isomorphic to σ-convex subsets of Banach spaces. Even so, it turns out that any *ordering* of a general superconvex space is equivalent, in a natural sense, to some ordering of a lottery space (Lemma 12). So we can still apply the same representation theorem even in this more general setting.

For any set X, let $\delta : X \to \Delta X$ be the function that takes each $x \in X$ to the lottery that assigns probability one to x.

Definition 5. A **superconvex space** is a set S together with an operation $\Sigma : \Delta\omega \times S^\omega \to S$ with the following properties:

(a) For any sequence of elements $x \in S^\omega$ and any index $i \in \omega$,

$$\Sigma(\delta(i), x) = x_i$$

(b) For any probability distribution $p \in \Delta\omega$, any sequence of probability distributions $q \in (\Delta\omega)^\omega$, and any sequence of elements $x \in S^\omega$,

$$\sum_i p(i) \cdot \Big(\sum_j q_i(j) \cdot x_j \Big) = \sum_j \Big(\sum_i p(i) \cdot q_i(j) \Big) \cdot x_j$$

Principle (a) says that putting all of the weight on the ith element just gives you that same element x_i. Principle (b) corresponds to simplification of compound lotteries: a mixture of mixtures of elements of S can be "flattened" to an equivalent single-level mixture.

14

Here is an alternative way of describing superconvex structure, which is a bit more elegant, since it eliminates the arbitrary indices. Instead of the operation Σ which takes an indexed sequence of elements and a distribution over the set of indices, we can define an operation $\sigma : \Delta S \to S$ which takes a distribution directly over the elements of the superconvex space, and mixes those elements together with the appropriate weights. Intuitively, an element of ΔS plays the role of a (discrete) random variable taking values in S, and σ plays the role of an expectation operator.

For any function $f : X \to Y$, let $\Delta f : \Delta X \to \Delta Y$ be the obvious "lifted" function that preserves mixtures. (That is, for any lottery $\lambda \in \Delta X$, the lottery $\Delta f(\lambda) \in \Delta Y$ assigns to $y \in Y$ the total probability that λ assigns to $f^{-1}(y)$.)

Proposition 3. Let S be a superconvex space. For any distribution $\lambda \in \Delta S$, there is a countable sequence $x \in S^\omega$ such that $\lambda = \sum_i \lambda(x_i) \cdot \delta(x_i)$. Let $\sigma : \Delta S \to S$ be

$$\sigma(\lambda) = \sum_i \lambda(x_i) \cdot x_i$$

This is well-defined (it does not depend on the choice of sequence x), and we have the two identities

$$\sigma(\delta(x)) = x \qquad \text{for each } x \in S$$
$$\sigma(\sigma(\kappa)) = \sigma(\Delta\sigma(\kappa)) \quad \text{for each } \kappa \in \Delta\Delta S$$

More succinctly:

$$\sigma\,\delta = 1_S$$
$$\sigma\,\sigma_{\Delta S} = \sigma\,\Delta\sigma$$

(Here juxtaposition means function composition, 1_S is the identity function on S, and $\sigma_{\Delta S} : \Delta\Delta S \to \Delta S$ is the mixture operation on ΔS.)

We can sum this up abstractly using resources from category theory (see Pümplun 1991): superconvex spaces are the Eilenberg-Moore algebras of the "lottery monad" Δ.[7]

Any σ-convex subset of a Banach space is a superconvex space. In general, there are other kinds of superconvex spaces as well. For example, any countably complete semi-lattice is a superconvex space, with the operation

$$\Sigma(p, x) = \bigvee_{i : p(i) > 0} x_i$$

---

[7] This "lottery monad" consists of the functor that takes each set X to the set of distributions ΔX and each function $f : X \to Y$ to the mixture-preserving function Δf, together with the two natural transformations $\delta : 1 \to \Delta$ and $\sigma_\Delta : \Delta^2 \to \Delta$. (Compare the "Giry monad" of Giry 1982.) An Eilenberg-Moore algebra of this monad is precisely a pair of a set S and an operation $\sigma : \Delta S \to S$ satsifying the two conditions in Proposition 3.

(Consider, for example, a bounded set of integers with the maximum operation. This is clearly not a σ-convex subset of any Banach space.) There are also hybrid spaces that combine "lattice-like" and "vector-like" behavior (for interesting examples see Fritz 2015). Even so, we will show that any preorder on one of these general spaces can be represented perfectly well by some preorder on a lottery space.

We can straightforwardly extend the idea of a mixture-preserving function to abstract superconvex spaces. (Pümplun (1991) calls such functions *superaffine.*) Here we require a slightly more general notion, which preserves mixtures modulo indifference.

Definition 6. If S and S' are preordered superconvex spaces, a function $f : S \to S'$ is **essentially mixture-preserving** iff for each $p \in \Delta\omega$ and $x \in S^\omega$,

$$f\big(\Sigma(p, x)\big) \sim \sum_i p_i \cdot f(x_i)$$

Equivalently,

$$f(\sigma(\lambda)) \sim \sigma(\Delta f(\lambda)) \qquad \text{for each } \lambda \in \Delta S$$

Or more succinctly,

$$f\sigma \sim \sigma\,\Delta f$$

(where for $g, h : X \to S$ we write $g \sim h$ iff $gx \sim hx$ for each $x \in X$).

Lemma 12. For any superconvex space S and any preorder $\lesssim$ on S, there is a unique preorder $\lesssim^*$ on ΔS with respect to which the function $\delta : S \to \Delta S$ is strictly monotone and essentially mixture-preserving.

Proof. For any $\lambda, \mu \in \Delta S$, define

$$\lambda \lesssim^* \mu \qquad \text{iff} \qquad \sigma(\lambda) \lesssim \sigma(\mu)$$

For $x, y \in S$,

$$\delta(x) \lesssim^* \delta(y) \qquad \text{iff} \qquad x = \sigma(\delta(x)) \lesssim \sigma(\delta(y)) = y$$

So δ is strictly monotone. For δ to be essentially mixture-preserving requires $\delta\sigma \sim^* \sigma_{\Delta S}\,\Delta\delta$. So it suffices to show:

$$\sigma\delta\sigma = \sigma\,\sigma_{\Delta S}\,\Delta\delta$$

Using the identities $\sigma\delta = 1_S$ and $\sigma\,\sigma_{\Delta S} = \sigma\,\Delta\sigma$, and the functor identities $\Delta(fg) = (\Delta f)(\Delta g)$ and $\Delta 1_S = 1_{\Delta S}$, we have

$$\sigma\delta\sigma = \sigma = \sigma\,\Delta(\sigma\delta) = \sigma\,\Delta\sigma\,\Delta\delta = \sigma\,\sigma_{\Delta S}\,\Delta\delta$$

For uniqueness, note first that in general $\sigma_{\Delta X} \Delta\delta = 1_{\Delta X}$, which can be straightforwardly checked.[8] Therefore, since δ is essentially mixture-preserving,

$$1_{\Delta S} = \sigma_{\Delta S} \Delta\delta \sim^* \delta\,\sigma$$

Thus, if δ is also strictly monotone, for any $\lambda, \mu \in \Delta S$ we have

$$\lambda \lesssim^* \mu \qquad \text{iff} \qquad \delta\sigma(\lambda) \lesssim^* \delta\sigma(\mu) \qquad \text{iff} \qquad \sigma(\lambda) \lesssim \sigma(\mu)$$

$\square$

Lemma 13. If $(S, \lesssim)$ satisfies (D*), then so does $(\Delta S, \lesssim^*)$.

Proof. Suppose $\lambda_i \lesssim^* \mu_i$ for each $i \in \omega$; so $\sigma\lambda_i \lesssim \sigma\mu_i$. The operation $\sigma : \Delta S \to S$ is itself mixture-preserving (this is what the identity $\sigma\sigma_{\Delta S} = \sigma\,\Delta\sigma$ tells us). So for any $p \in \Delta\omega$ we have

$$\sigma\Sigma(p, \lambda) = \sum_i p_i \cdot \sigma\lambda_i \lesssim \sum_i p_i \cdot \sigma\mu_i = \sigma\Sigma(p, \mu)$$

And thus $\Sigma(p, \lambda) \lesssim^* \Sigma(p, \mu)$. The case where we also have strictly $\lambda_i < \mu_i$ for some i such that $p_i > 0$ goes the same way.

$\square$

We can now put this fact together with the representation theorem for lottery spaces, to apply that theorem to arbitrary superconvex spaces.

Corollary 2. A totally preordered superconvex space $(S, \lesssim)$ satisfies (D*) iff there is a strictly monotone mixture-preserving function from S to $B\ell^\infty(\sigma)$, for some ordinal α.

References

Arntzenius, Frank, and David McCarthy. 1997. "The Two Envelope Paradox and Infinite Expectations." *Analysis* 57 (1). Oxford University Press: 42–50.

Blackwell, David, and M. A. Girshick. 1954. *Theory of Games and Statistical Decisions*. New York: John Wiley & Sons, Inc. https://onlinelibrary.wiley.com/doi/abs/10.1002/nav.3800010313.

---

[8]This is one of the monad identity laws, since $\sigma_\Delta : \Delta^2 \to \Delta$ is the monadic multiplication operation.

Borie, D. 2016. "Lexicographic Expected Utility Without Completeness." *Theory and Decision* 81 (2): 167–76. https://doi.org/10.1007/s11238-015-9523-y.

Buchak, Lara. 2013. *Risk and Rationality*. Oxford University Press.

Conway, John B. 1994. *A Course in Functional Analysis*. 2nd edition. New York: Springer.

Emelyanov, Eduard Yu. 2014. "Archimedeanization of Ordered Vector Spaces." *arXiv:1406.3657 [Math]*, June. http://arxiv.org/abs/1406.3657.

Fine, Terrence L. 2008. "Evaluating the Pasadena, Altadena, and St Petersburg Gambles." *Mind* 117 (467). Oxford University Press: 613–32.

Fishburn, Peter C. 1971. "A Study of Lexicographic Expected Utility." *Management Science* 17 (11): 672–78. https://doi.org/https:// doi.org/10.1287/mnsc.17.11.672.

Fishburn, Peter C. 1974. "Lexicographic Orders, Utilities and Decision Rules: A Survey." *Management Science* 20 (11): 1442–71. https://www.jstor.org/stable/2629975.

Fritz, Tobias. 2015. "Convex Spaces I: Definition and Examples." *arXiv:0903.5522 [Math]*, October. http://arxiv.org/abs/0903.5522.

Giry, Michèle. 1982. "A Categorical Approach to Probability Theory." In *Categorical Aspects of Topology and Analysis*, edited by B. Banaschewski, 68–85. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer. https://doi.org/10.1007/BFb0092872.

Hammond, Peter J. 1998. "Objective Expected Utility: A Consequentialist Perspective." In *Handbook of Utility Theory*, edited by Peter J. Hammond, Salvador Barberà, and Christian Seidl. Vol. I. Kluwer Academic.

Hausner, M. 1952. "Multidimensional Utilities." *Rand Corp*, August.

Hausner, M., and J. G. Wendel. 1952. "Ordered Vector Spaces." *Proceedings of the American Mathematical Society* 3 (6): 977–82. https://doi.org/10.2307/2031746.

Hájek, Alan. 2003. "Waging War on Pascal's Wager." *Philosophical Review* 112 (1). Duke University Press: 27–56.

Hájek, Alan, and Harris Nover. 2008. "Complex Expectations." *Mind* 117 (467). Oxford University Press: 643–64.

Hrbacek, Karel, and Thomas Jech. 1999. *Introduction to Set Theory, Revised and Expanded*. 3 edition. New York: Marcel Dekker.

König, Heinz. 1986. "Theory and Applications of Superconvex Spaces." In *North-Holland Mathematics Studies*, 122:79–118. Elsevier. https://doi.org/10.1016/S0304-0208(08)71950-9.

Luce, R. Duncan, and Howard Raiffa. 1989. *Games and Decisions: Introduction and Critical Survey*. Revised ed. edition. New York: Dover Publications.

Martínez-Legaz, Juan E. 1998. "Lexicographic Utility and Orderings." In *Handbook of Utility Theory*, edited by Salvador Barberà, Peter J. Hammond, and Christian Seidl, 1: Principles:345–69. Dordrecht: Kluwer Academic Publishers.

Menger, Karl. 1934. "Das Unsicherheitsmoment in der Wertlehre." *Zeitschrift für Nationalökonomie* 5 (4): 459–85. https://doi.org/10.1007/BF01311578.

Nörtemann, Stefan. 2002. "The Hahn–Banach Theorem for Partially Ordered Totally Convex, Positively Convex and Superconvex Modules." *Applied Categorical Structures* 10 (4): 417–29. https://doi.org/10.1023/A:1016390813177.

Paulsen, Vern, and Mark Tomforde. 2007. "Vector Spaces with an Order Unit." *arXiv:0712.2613 [Math]*, December. http://arxiv.org/abs/0712.2613.

Potter, Michael. 2004. *Set Theory and Its Philosophy: A Critical Introduction*. Oxford University Press.

Pümplun, Dieter. 1991. "Banach Spaces and Superconvex Modules." In *Mathematics and Theoretical Physics*, edited by Minaketan Behara, Rudolf Fritsch, and Rubens G. Lintz, 323–38. Walter de Gruyter.

Russell, Jeffrey Sanford, and Yoaav Isaacs. Forthcoming. "Infinite Prospects." *Philosophy and Phenomenological Research*.

Savage, Leonard J. 1954. *The Foundations of Statistics*. Wiley Publications in Statistics.

Seidenfeld, Teddy, Mark J. Schervish, and Joseph B. Kadane. 2009. "Preference for Equivalent Random Variables: A Price for Unbounded Utilities." *Journal of Mathematical Economics* 45. Research Showcase @ Cmu: 329–40.