Essays on Games and Decisions

Siddharth Chatterjee

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
Quantitative Economics

Indian Statistical Institute

July 2022

Advisor

Professor Arunava Sen

Economics and Planning Unit
Indian Statistical Institute, Delhi Center

To my teachers.

Abstract

In the first chapter, a solution concept for two–person zero–sum games is proposed with players' preferences only assumed to satisfy Independence. To each player, there is a set of *admissible* strategies assuring him minimum guarantees. Moreover, rationality requires players to reject non–admissible strategies from any further consideration. Additional knowledge assumptions allow iterated elimination of non–admissible strategies. This leads to a pair of strategy sets, one for each player, whose cross product are the *consideration equilibria.* Consideration equilibria always exist and include Nash equilibria if any. Further, consideration equilibria and Nash equilibria (or, minimax strategies) coincide if players' preferences additionally satisfy Continuity. Three examples are analysed for illustration.

The second chapter investigates the implications of additivity type axioms in economic theory. In several areas of microeconomic theory, axiomatic characterizations have been provided for the respective objects of study to possess lexicographic structures. We introduce the concept called *graded halfspace* which is an abstraction of "lexicographic structures". Then, we formulate and establish a geometric result called the *Decomposition Theorem.* This result characterizes graded halfspaces as the convex cones which are elements of some partition, of a given Euclidean space, consisting of a pair of mutually reflecting convex cones and a subspace. Thus, the Decomposition Theorem formalizes the following intuitive idea: an "object" defined over a convex "domain" is additive, if and only if, it has a lexicographic "structure". To illustrate this geometric approach, we present four applications ranging over decision theory, social choice, convex analysis and linear algebra.

In the third chapter, we consider *pre–norms* on the Euclidean space which are functions that satisfy the definition of a norm except that a vector and and its reflection through the origin may have different values. Then, we characterize those binary relations on the Euclidean space which admit a pre–norm as a (utility) representation. The notion of the *dual* of such a binary relation is introduced. For any such binary relation, its *second dual* — the dual of the dual — is identical to itself. Further, such a binary relation is *self dual* if and only if it is "spherical" — the Euclidean norm is a representation. The duality theory allows us to generalize the Hölder's inequality to arbitrary pre–norms. Binary relations which admit a norm as a representation are also characterized. We specialize our theory to characterize binary relations which admit a p–norm as a representation. Thus, the classical inequalities due to Minkowski and Hölder follow as corollaries of the general theory.

Acknowledgements

I begin with the professional acknowledgements. The chapters in this thesis have been shaped in large measure by the critical assessments and feedback of the following: Fuad Aleskerov, Bhavook Bhardwaj, Kalyan Chatterjee, In–Koo Cho, Pradeep Dubey, Bhaskar Dutta, Federico Echenique, Jeff Ely, Albin Erlanson, Sergiu Hart, Vijay Krishna, Rohit Kumar, Fabio Maccheroni, Debasis Mishra, Stephen Morris, Hervé Moulin, Vishal Kumar Rai, Marzena Rostek, Ariel Rubinstein, Arunava Sen, Shigehiro Serizawa and Jörgen Weibull.

In particular, Ariel Rubinstein has entertained my answers to his many incisive questions and my counter–questions, regarding the solution concept for two–player zero–sum games proposed in the first chapter, over a span of six months which has showed me how and why one must be constantly critical of one's own work. Moreover, Vijay Krishna and Ariel Rubinstein have each indicated how must I proceed to extend my analysis of the issues raised in the first two chapters. Furthermore, Federico Echenique has taken particular note of the duality theory of convex preferences in the third chapter.

Moreover, I have greatly benefitted from the many discussions with the participants of the following events: the 2020 and 2022 editions of the Winter School of the Econometric Society at the Delhi School of Economics, Asian Summer School in Economic Theory 2022 of the Econometric Society at the National University of Singapore, the Asian Meeting of the Econometric Society 2023 at the Indian Institute of Technology Bombay, Workshop in Microeconomic Theory 2021 by the Osaka University, and the Graduate Students Workshop 2021 by the Higher School of Economics at Moscow.

The reason which led me into the work done in these chapters is threefold. First, I have had the tremendous fortune of learning microeconomic theory from three experts — Eddie Dekel, Bhaskar Dutta and Arunava Sen. It is truly amazing—and it still baffles me—as to how different the approach can be on the same issue. Of course, this difference in approach has been a huge gain for me. Each of them showed us, what does it truly mean to *understand* something!

Secondly, Arunava Sen's infectious love for mathematical thinking in general, with its role in microeconomic theory in particular, shaped my thought over a span of a decade already. Thirdly, A. K. Shukla really started my engines on mathematics and physics when I was a high school student. Next, the more *personal* acknowledgements follow. The reader interested in the formal aspects of the thesis may skip the rest of this section without any loss of continuity.

I could not see it coming. It was Vijay Krishna who, over a single discussion on the first version of the second chapter, managed to force me to think hard, "What, if any, is the *new* contribution?". What is surprising is that he managed to do this while making me feel even more motivated than I earlier was. However, it is likely that he is not aware that he ends up having this effect!

At Indian Statistical Institute (henceforth, "ISI"), Debasis Mishra has ranks very high on a "multidimensional type space". He always addressed our questions on games, auctions and so on. His advice on courses and career plans have been precise. Moreover, he has always been there as a guardian whenever I fell ill or having been ε–close to some kind of trouble. No third party can ever tell that he does all of this. Does he do it consciously, or, that it comes naturally to him? I am not very sure at the moment!

Two of my professors, Debasis Kundu and Maneesh Thakur, pushed me to study probability and mathematics during my undergrad and masters even though I was not a math major. They made clear the claim by Professor K. R. Parthasarathy over lunch at the mess of ISI, "Mr. Chatterjee, stop asking the question: whether you *can* do it or not? The real question is: do you *want* to do it or not?"

Again at ISI, I found three friends in the last two years among the graduate students. Two of them managed to inspire me, just by demonstrating their grit in whatever they were doing, showing me the point of putting in the hours at the desk. The first has provided me with an endless supply of americano, cookies, pastries and so on. He gifted me a copy of *Game and Decisions* by LUCE & RAIFFA (1957) on my birthday while I was writing this thesis. Further, we are working on a paper, the basic result in which (the "Reduction Lemma") is rather very cute in our eyes. This person, who is also a natural stand up comedian, is Bhavook Bhardwaj!

The second person is Vilok Taori who has tolerated several annoying requests ranging from printouts to supplying me papers or books. Bhavook and Vilok taught me about one major subdiscipline of microeconomic theory each that is *not* connected to this thesis—revealed preference theory and matching! The third friend is Vishal Kumar Rai who works in macroeconomic theory. However, he tolerated me for some hundred hours as I talked about the material in this thesis which was still in a nascent stage. In addition, he has enlightened me about ideas in macroeconomics which otherwise I would have never hoped to understand. Finally, the masters student called Abhay Gupta surfaced as whip–smart collaborator. Surely, much more shall follow.

At home, my father, Protyush, showed that any adversity can be overcome with patience and grit as long as you are alive! He often asks, "So, what's the point of studying mathematics, probability or economics that you guys do?". I tell him that equilibria "explain" things that we see out there. My (elder) sister, Roopsha, always provided the shield against wordly concerns—financial or otherwise—without which pursuing my objectives would have been impossible.

When I was young, my father taught me counting and school geometry—parents are the first teachers. This brings me to my mother, Kakoli, who is a kid at heart yet a mentor but is also a friend and taught me the languages. She knew only Bengali but learnt Hindi, Sanskrit and English enough from my school teachers to teach us! She would be strict about breaking up any word in the right way to learn how to spell and pronounce it. She also taught Biology which was memory intensive. As a child, I doubted the value of the ability of memorise. Of course, I couldn't have been more wrong. Now, I spend hours talking to her over phone describing my ongoing work and she patiently listens to me while managing to show interest!

Now, I come to Arunava Sen! I joined ISI in 2011 to study probability theory. However, students would speak very highly of this professor who is an economist. I was naive, knowing nothing about economics, and believed that economics was something very vague. They challenged me to attend Arunava's lectures on "Social Choice Theory". I said to myself, "This is going to be sheer wrote memorization". Yet, I attended two lectures in which he taught us Arrow's Impossibility Theorem — he started off with the words, "We shall study binary relations that are complete and transitive!" and, I decided to attend his course on game theory. Half way into that course, I decided that I would pursue research in microeconomic theory.

While teaching the Expected Utility Theorem of von Neumann & Morgenstern (1944), he made an error while illustrating in the simplex the Independence axiom. That showed us, Independence alone implies existence of lexicographic expected utility representations — we were not aware of Hausner (1954). His "something" and "whatever" is always both entertaining and productive. However, three sentences stand out. First, "What ever you do, do it well!". Second, "Work hard!". Third, "Repetition is the key!". And, when I asked him whether he is satisfied with my work, he said, "It is very much to my taste . . . although I'm sure I didn't have much to do with it."!

New Delhi, 2023 Siddharth Chatterjee

Contents

Chapter 0

Introduction

This thesis is on some aspects of individual and collective decision making. Within the context of individual decision making, two particular themes receive focus. First, the Independence axiom of expected utility theory. Second, characterization of preferences—over the Euclidean space—which admit norms as utility representations. Within the context of collective decision making, one objective is to revisit the foundations of two–person zero–sum games, and the second is to explore the setting of Arrovian aggregation. Thus, the three chapters—in the order of their appearance—are entitled as follows:

1. Two–Person Zero–Sum Games without Expected Utility Preferences: A Proposal.
2. Additivity over Convex Domains is Equivalent to Lexicographic Structures.
3. Preferences with Norms as Representations.

At a methodological level, we investigate the implications of convexity and linearity for decision making problems. A brief overview of each of the three chapters follows.

Solutions concepts in game theory, such as Rationalizable Strategies and Nash Equilibrium, depend in part for their existence on the assumption that players' preferences satisfy Continuity. They also require some plausible behavioral assumption such as Independence. However, the axiom of Continuity is at best a technical condition.

We consider two–player games, where players' pure action sets are finite but they may play any mixed strategy. We assume that the preference $\succ_i$ of each player i, on the set of lotteries over all pure strategy tuples, satisfies Independence. In fact, we assume only a weaker version of the Independence axiom of von Neuman & Morgenstern (1944) which we propose in chapter 2. Then, we define a two–person game to be *zero–sum* if, one player's loss is another's gain:

$$p \succ_1 q \iff q \succ_2 p.$$

Next, we introduce the notion of "admissible set" of player i. A subset of strategies A_i for player i is said to satisfy property B if, for any strategy x_i of player i which is not in A_i, the following holds:

$$(x_1', x_2') \succ_i (x_1, x_2) \quad \text{for all } x_1' \in A_1 \text{ and all } x_2',$$

where x_j is any best response of player j to x_i. Thus, playing from A_i ensures some minimum guarantees to player i. Observe, this idea of minimum guarantees is embodied in the Minimax Strategies of von Neumann (1928). Note, the entire simplex of all mixed strategies of player i satisfies property B vacuously. We show that all subsets of the simplex which satisfy property B form a nest whose intersection is nonempty and also satisfies property B. In other words, there exists a unique smallest nonempty set of strategies which satisfies property B. We call it the *admissible set* of player i and denote it by A_i^1.

We place superscript of '1' to indicate that we shall now treat the admissible sets as if they are the simplices and obtain admissible subsets A_i^2 thereof. This is possible because admissible sets are shown to be convex and compact. Thus, to each player i there is a nested sequence $A_i^1 \supseteq A_i^2 \supseteq \ldots$ of compact convex sets obtained via the iterated eliminiation of non–admissible strategies. The rectangle of surviving strategy tuples $A_1^\infty \times A_2^\infty$ are the *consideration equilibria.* Such equilibria always exist and are interchangeable. Further, if Continuity holds additionally, then they are precisely the Minimax Strategies (or, Nash Equilibria) for which the Minimax Theorem holds.

Additivity type axioms are commonplace in economic theory. For instance, consider the axioms such as Independence in expected utility theory, Cardinal Measurability & Unit Comparability in the theory of interpersonal comparison of utilities in social choice and so on. These axioms are of normative or ethical appeal depending upon the context under consideration.

Often, in conjunction with some technical condition such as Continuity, additivity is shown to characterize some linear real–valued function. Some important examples are the Expected Utility Theorem of von Neumann & Morgenstern (1944), Generalized Utilitarianism of Harsanyi (1955) or d'Aspremont & Gevers (1977), and so on. Our objective is to drop the supporting technical conditions such as Continuity and to focus on the consequences of the additivity type axiom(s) alone. We find that additivity, when the domain is convex, is equivalent to a lexicographic structure.

As our first example, we revisit the classical result due to Hausner (1954) which says that preferences that satisfy Independence are characterized by the fact that they admit a lexicographic expected utility representation. We weaken the classical Independence axiom. To state our weakening, we first recall the original version. Suppose p, q and r are any three lotteries, and $\alpha \in (0, 1)$. Then,

$$p \succ q \iff \alpha \cdot p \oplus (1-\alpha) \cdot r \succ \alpha \cdot q \oplus (1-\alpha) \cdot r.$$

Then, our version of Independence can be stated as follows. Suppose p, q and r are any three lotteries. Then,

$$p \succ q \iff \big(\forall \alpha \in (0,1)\big)\big[\alpha \cdot p \oplus (1-\alpha) \cdot r \succ \alpha \cdot q \oplus (1-\alpha) \cdot r\big].$$

Observe, the " $\implies$ " part is the same. However, whereas the original version declares $p \succ q$ if $\alpha \cdot p \oplus (1-\alpha) \cdot r$ $\succ$–dominates $\alpha \cdot q \oplus (1-\alpha) \cdot r$ for even *one* $\alpha \in (0, 1)$, our version does *not.* The latter requires $\alpha \cdot p \oplus (1-\alpha) \cdot r$ to $\succ$–dominate $\alpha \cdot q \oplus (1-\alpha) \cdot r$ for *every* $\alpha \in (0, 1)$ in order to conclude that $p \succ q$. Thus, our axiom is logically weaker than the original Independence. However, we find that for binary relations that satisfy transitivity and completeness, our version is also necessary and sufficient for the existence of lexicographic expected utility representations. Thus, we achieve a logical strengthening of Hausner's theorem. Moreover, it is normatively more appealing.

After expected utility theory, we consider social choice theory. Here we obtain lexicographic extensions of Generalized Utilitarianism which were characterized by Harsanyi (1955). We achieve this under the key normative axiom which is Cardinal Measurability & Unit Comparability. Strengthening this axiom to Non–Comparability results in the following two characterizations. First, the additional assumption of Strong Pareto enforces serial dictatorships. Second, the milder additional assumption of Weak Pareto enforces weak dictators — Arrow's Impossibility Theorem. Of course, requiring Continuity and Weak Pareto additionally under the Unit Comparability assumption characterizes Generalized Utilitarianisms.

We next consider the problems of existence of linear representations for weak orders on convex subsets of the Euclidean space. Thus, we generalize Theorem 4.3.1 of Blackwell & Girshick (1954) to arbitrary convex subsets. Their axioms, namely Invariance and Continuity, achieve the characterization of linearly representable weak orders over arbitrary convex sets. However, Continuity and Invariance imply Convexity — every upper and lower contour set of the weak order is convex. Then, Invariance and Convexity characterize those weak orders which admit lexicographic extensions of linear representations.

Our last application is to obtain a simple proof of the characterization of finite dimensional ordered vector spaces over the reals due to Hausner & Wendel (1952). Before we close the overview of chapter 2, we must point out that our approach to the applications is via a common method. We introduce the notion of "graded halfspaces". Given any orthonormal collection of vectors, let the first "slice" be the open halfspace generated by the first given vector such that the origin is on the boundary of the halfspace. Then, the boundary is a subspace of dimension one less and contains the remaining given orthonormal vectors. Thus, we may recursively generate a list of slices with progressiveli collapsing dimensions. Then, the graded halfspace generated by the given orthonormal vectors is the union of these slices.

Graded halfspaces are an abstraction of lexicographic structures. For instance, the strict upper contour set of the standard lexicographic order on the Euclidean plane is a graded halfspace. We provide a geometric characterization of graded halfspaces which we call the Decomposition Theorem. It is the application of the this result that allows us to achieve the characterizations that we claimed in the above application domains. This result formalizes the qualitative claim: additivity over convex domains is equivalent to lexicographic structures.

AN OVERVIEW OF CHAPTER 3

In this chapter, we are concerned with the characterization of those weak orders on the Euclidean space which admit some norm as a utility representation. Of particualar importance are the Minkowski norms $\|\cdot\|_p$ which further contain as a special case the Euclidean norm $\|\cdot\|_2$. CHAMBERS & ECHENIQUE (2020) characterized preferences based on the Euclidean norm — "spherical preferences".

It is perhaps plausible to percieve their work as a response to the question that has manifested owing to decades of work in political economy and social choice in the context of spatial voting or voting over multiple issues. For instance, consider MCKELVEY & WENDELL (1976). In these applications, it has been assumed that individuals of the society have preferences which admit the Euclidean norm as a representation.

However, many authors such as WENDELL & THORSON (1974), BORDER & JORDAN (1983) and ZHOU (1991) have correctly argued that norms beyond the Euclidean are also equally important. Further, it has been established (see ENELOW ET AL. (1988) for instance) that empirical testing of the voting model equibrium analysis strongly depends on the correctness of specification of individuals' preferences. Thus, we find that obtaining decision theoretic foundation for arbitrary norms, and p–norms, is essential.

As a starting point, we generalize our question by introducing objects called "pre–norms". These are real–valued functions over the Euclidean space which satisfy all defining properties of norms except for the symmetry condition that a vector and its reflection through the origin must result in the same value. It is then immediate, if a weak order admits a pre–norm as a representation, it must satisfy Homotheticity, Convexity[1], increasing marginal returns (we call this, "Scale Monotonicty") and Continuity. Our first main result is that the converse is also true. The key is that these axioms imply, the weak lower contour sets are compact and contain the origin in their interior.

We obtain p–norms essentially by additionally requiring the axiom of Separability due to DEBREU (1959). We also develop a "duality theory" which is analogous to the relation of the Utility Maximixation Problem vs. Expenditure Minimization Problem in consumer choice. One of the key findings of "duality" is that a preference is dual to itself if and only if it is "spherical".

[1]All weak *lower* contour sets are convex

Blackwell, D. & M. A. Girshick (1954): *Theory of Games and Statistical Decisions* New York: John Wiley and Sons, Inc.

Border, K. C. & J. S. Jordan (1983): "Straightforward Elections, Unanimity and Phantom Voters", *The Review of Economic Studies*, vol. 50, no. 1, pp. 153–170.

Chambers, C. P. & F. Echenique (2020): "Spherical Preferences", *Journal of Economic Theory*, vol. 189, 105086.

d'Aspremont, C., & L. Gevers (1977): "Equity and te Informational Basis of Collective Choice", *Review of Economic Studies*, vol. 44, no. 2. pp. 199–209.

Debreu, G. (1959): "Topological Methods in Cardinal Utility", in K. J. Arrow, S. Karlin and P. Suppes (editors) *Mathematical Methods in the Social Sciences*, Stanford: Stanford University Press.

Enelow, J. M., N. R. Mendell & S. Ramesh (1988): "A Comparison of Two Distance Metrics through Regression", *The Journal of Politics*, vol. 50, no. 4, pp. 1057–1071.

Harsanyi, J. C. (1955): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility", *Journal of Political Economy*, vol. 63, no. 4, pp. 309–321.

Hausner, M. (1954): "Multidimensional Utilities", in R. M. Thrall, C. H. Coombs and R. L. Davis (editors) *Decision Processes*, New York: John Wiley & Sons Inc., 1954.

Hausner, M. & J. G. Wendel (1952): "Ordered Vector Spaces", *Proceedings of the American Mathematical Society*, vol. 3, no. 6, pp. 977–982.

McKelvey, R. D. & R. E. Wendell (1976): "Voting Equilibria in Multidimensional Choice Spaces", *Mathematics of Operations Research*, vol. 1, no. 2, pp. 144–158.

von Neumann, J. (1928): "Zur Theorie der Gesellschaftsspiele", *Math. Annalen*, vol. 100, pp. 295–320.

von Neumann, J. & O. Morgenstern (1944) *Theory of Games and Economic Behavior* New Jersey: Princeton University Press.

Wendell, R. E. & S. J. Thorson (1974): "Some Generalizations of Social Decisions under Majority Rule", *Econometrica*, vol. 42, no.

5, pp. 893–912.

Zhou, L. (1991): “Impossibility of Strategy–Proof Mechanisms in Economies with Pure Public Goods”, *The Review of Economic Studies*, vol. 58, no. 1, pp. 107–119.

Chapter 1

Two–Person Zero–Sum Games without Expected Utility Preferences: A Proposal

1. introduction

Two–person zero–sum games occupy a central position in game theory as they model situations of bilateral conflict. von Neumann (1928) published his Minimax Theorem which provides a basis for how players should play. This foundational result was established assuming existence of expected utility representations. However, expected utility representations exist if and only if preferences of the players satisfy the Independence axiom *and* the Archimedean property (or, Continuity) as was shown by von Neumann & Morgenstern (1944).

While Independence is normatively appealing in decision theory, Continuity is a technical condition needed for existence of numerical representations. Hausner (1954) showed that if Independence holds, then preferences admit lexicographic representations. Despite being non–Archimedean, lexicographic preferences are natural in modeling competing firms or bilateral trade—each party has multiple decision criteria and a priority over these.[2] For applications, see Chipman (1960, pp. 221), Fishburn (1970, pp. 110) and Thrall (1954). The Archimedean property is not applicable in such models.

[2]Concrete examples are presented in the subsection below which may be read at this stage.

Additionally, Thrall (1954) shows that the set of maximizers of such a preference over any convex and compact set is a convex and compact set. Using this, he argues, "This discussion illustrates the fact that non–Archimedean utilities are perfectly satisfactory for game theory". Several later writings, such as Ferguson (1958, pp. 20–21) and Luce & Raiffa (1957, pp. 27), indicate that this had become an accepted fact in game theory. For instance, Aumann (1964, pp. 453) writes, "It will still be possible to solve maximization problems and games under exactly the same conditions as before".

Unfortunately, Fishburn (1971) demonstrated that the Minimax Theorem does *not* hold for non–Archimedean preferences. Therefore, he concluded, "The impression remains that game theory without the Archimedean axiom is rather barren". Our contribution is to propose a solution concept, which we call the *consideration equilibrium*, for the class of all two–person zero–sum games. Its existence requires *only* the Independence axiom of the players' preferences. Further, consideration equilibria are precisely the Minimax strategies, which are also the Nash equilibria, when preferences additionally satisfy Continuity.

We briefly outline the solution concept. Let the two players be 1 and 2. Suppose, there is a set A_1 of mixed strategies of player 1 with the following property: if player 1 considers playing any x_1 not in A_1, then there is some play x_2 of his opponent such that playing any x_1' in A_1 instead of x_1, no matter what his opponent plays, is strictly preferred by player 1. Thus, strategies in A_1 assure some "minimum guarantee" for player 1 against any play of his opponent. The smallest such set of strategies, denoted A_1^*, shall be called *admissible*. It extends the notion of a minimum guarantee irrespective of the opponent's play which is the basis of the concept of *value* in the classical minimax theory due to von Neumann (1928).

Instead of the defining property of A_1 as above, we may consider the following property: if player 1 considers playing any x_1 not in A_1 and x_2 is player 2's best response against x_1, then player 1 strictly prefers that he plays any x_1' in A_1 where his opponent plays any best response. These two properties are equivalent. The set A_1^* of admissible strategies of player 1 is non–empty and unique. Likewise, there is a unique non–empty set of admissible strategies, say A_2^*, of player 2.

In order that $A_1^* \times A_2^*$ be a solution concept, the following property is desirable: for any (x_1, x_2) and (x_1', x_2') in $A_1^* \times A_2^*$, players are indifferent between (x_1, x_2) and (x_1', x_2'). Otherwise, what should players play from $A_1^* \times A_2^*$? Unfortunately, this property does *not* hold for $A_1^* \times A_2^*$. However, iterated elimination of non–admissible strategies ensures that this property holds. This elimination affords a justification along the lines of Bernheim (1984) and Pearce (1984).

We briefly outline the logic behind the elimination. First, A_1^* and A_2^* are convex and compact sets. Moreover, suppose that it is common knowledge between players 1 and 2 that each player i shall play from A_i^*. Then, it is *as if* the sets A_1^* and A_2^* are the simplices of all mixed strategies of players 1 and 2. That is, the "context" of consideration changes from *all* pairs of mixed strategies to those in (A_1^*, A_2^*). Thus, to justify the elimination it remains to argue: it is common knowledge between the players that each player i shall play from A_i^*.

Let player 2's conjecture about player 1's play be x_1. Suppose, x_1 is not in A_1^*. Thus, if player 1 knows that this is player 2's conjecture about player 1's play, then player 1 knows that player 2 will play some best response x_2. However, playing any x_1' in A_1^* is strictly preferred by player 1 when player 2 is to play x_2. This is known to player 2. Thus, if x_1 is not in A_1^*, then "player 1 shall play x_1" is *not* a plausible conjecture by player 2 about player 1's play. Hence, the elimination of non–admissible strategies is justified.

The *context* comprising $\Delta(S_1)$ and $\Delta(S_2)$—mixed strategy spaces of players 1 and 2—led to the admissible strategy sets A_1^* and A_2^*. Now, the context is the pair (A_1^*, A_2^*). Thus, there exist unique non–empty sets of admissible strategies A_1^{**} and A_2^{**}, for players 1 and 2, with respect to the context (A_1^*, A_2^*). Hence, a nest $\Delta(S_i) \supseteq A_i^* \supseteq A_i^{**} \supseteq \dots$ obtains for each player i. Denoting by A_i^∞ the intersection of A_i^*, A_i^{**}, ... etc., the set of *consideration equilibria* is $A_1^\infty \times A_2^\infty$. The solution concept thus embodies the following reasoning by the players.

> "Starting with all of our mixed strategies as the context, if you do not restrict your strategy considerations to your admissible set with respect to this context, *then* so will I thereby making you strictly worse than had you considered any strategy in your admissible set. Thus, we both must restrict our considerations to our admissible sets which, therefore, become the new context with respect to which we look for admissible sets thereof ... and so on. Hence, we must *not* consider strategy tuples which are not consideration equilibria. Further, each of us is indifferent between any two consideration equilibria. Hence, we may play any consideration equilibrium."

Our solution concept generalizes the theory of von Neumann in the following respects. First, consideration equilibria always exist and form a convex and compact set. Second, each player is indifferent between any two consideration equilibria. Third, if the game has a Nash equilibrium, it is also a consideration equilibrium. Fourth, if players' preferences admit expected utility representations, then consideration equilibria coincide with Nash equilibria.

The admissible sets and consideration equilibria are shown to arise as solutions to finite lists of linear programs. This is because a preference satisfies Independence if and only if it admits a lexicographic expected utility representation as shown, for instance, in Hausner (1954), Blume et al. (1989) and Chatterjee (2022).

The rest of the article is organised as follows. The framework is in section 2. Sections 3 and 4 present the concepts of admissible sets and consideration equilibria. Section 5 presents the applications. The comparative statics are presented in section 6. The procedure for the computation of admissible sets and consideration equilibria is described in section 7. Proofs omitted from the main text are supplied in the Appendix. We close this introduction by presenting some examples. However, their "solution" will be deferred until section 5. This is because the general framework and our solution concept shall have to be presented first as done in sections 2 to 4.

Some Examples

The objective of subsection is as follows. We substantiate the case, stated in the second paragraph of the overview above, that it is natural to write zero–sum games as models for situations of strategic interaction of agents whose preferences arise from a priority over multiple criteria. We do so by presenting three examples as follows. Note, in such games a single "numerical payoff" corresponding to an outcome is insufficient. However, lexicographic expected utilities are a natural choice to model such preferences of the players.

Example 1: Two firms 1 ("Player I") and 2 ("Player II") are about to engage in a competition (Figure 1). Firm 1 has two strategies which are "Execute a hostile price–cut" (T) or "Poach top talent of firm 2" (B). Also, firm 2 has two strategies which are "Counter firm 1's move to poach talent, if any" (L) or "Match firm 1's hostile price–cut, if any" (R). The firms may randomize over their respective pure strategies, or, they may even jointly randomize.

Each firm strictly prefers a higher market share than less. However, if two plays result in the same market share, then each firm is better off with a larger pool of top talent. Thus, each firm has a lexicographic preference. To describe such preferences over all joint randomizations, it is enough to specify "lexicographic payoffs" to each player for every possible play involving pure strategy tuples. Further, if the sum of firms' market shares and their total talent size can each be taken as a constant, then the game is zero–sum. Thus, it is enough to only specify to firm 1's lexicographic payoff for pure strategy tuples.

The ordered pair in each cell of Figure 1 represents firm 1's payoffs with the order reflecting the priority over the two criteria: (1) market share of the firm, and (2) if two plays lead to same market share of the firm, then size of the firm's top talent. Thus, the play (T, L) gives firm 1 an advantage in market share as T means "Execute hostile price–cut" but L means "Counter firm 1's move to poach talent, if any". That is, the first component of the ordered pair corresponding to the play (T, L) is 1. However, if firm 1 chooses B which means "Poach top talent of firm 2" or firm 2 chooses R which means "Match firm 1's hostile price–cut", then the first component of the corresponding ordered pair is 0 as firm 1 gains no advantage in market share.

			Player II	
			q	$1-q$
			L	R
Player I	p	T	$(1,0)$	$(0,0)$
	$1-p$	B	$(0,0)$	$(0,1)$

FIGURE 1: Two competing firms.

Moreover, the play (B, R) gives firm 1 an advantage in size of its top talent as B means "Poach top talent of firm 2" but R means "Match firm 1's hostile price–cut". Thus, the second component of the ordered pair corresponding to (B, R) is 1. However, firm 1 chooses T which means "Execute hostile price–cut" or firm 2 chooses L which means "Counter firm 1's move to poach talent, if any", then firm 1 has no advantage in its size of top talent. Thus, the second component of the corresponding ordered pairs are 0.

Observe, when firm 2 considers "Match firm 1's hostile price–put, if any" (the strategy R), it does not consider "Counter firm 1's move to poach talent, if any" (the strategy L). Further, if firm 1 considers deploying the strategy "Execute hostile price–cut", then it knows that firm 2 has the option to play "Match firm 1's hostile price–cut". Moreover, the strategies of the firms are such that whereas firm acts by *making a move*, firm 2 only acts by being *responsive*.

Thus, we have the following questions. Is it the case that firm 1 ends up playing the strategy "Poach top talent of firm 2" (that is, B) and firm 2 ends up playing the strategy "Match firm 1's hostile price–cut" (that is, R)? In other words, is (B, R) an "equilibrium" of this game? If yes, is the "equilibrium" unique? ■

EXAMPLE 2: A financial institution ("Player I") and the rest of the financial market ("Player II") interact as follows. There is an asset A_1 about which the market is "Optimistic" (L) or "Pessimistic" (R), this market sentiment determines whether the value of A_1 will rise or fall. The institution guesses what the market feels about this asset. Also, there is another profitable asset A_2 which the financial institution either "acquires" or "does not acquire". The rest of the market has no control over the asset A_2's possession. Thus, the strategies of the financial institution are "Buy A_1 and buy A_2" (T) or "Short sell A_1" (B), where short selling is to bet against the asset A_2.

			Player II	
			q	$1-q$
			L	R
Player I	p	T	$(1,1)$	$(0,1)$
	$1-p$	B	$(0,0)$	$(1,0)$

FIGURE 2: Betting against the market.

If market participants are "Optimistic" then "Buy A_1 and buy A_2" pays off to the financial institution as A_1 is then valuable. However, if the other market participants are "Pessimistic", then "Buy A_1 and buy A_2" is worse for the financial institution as the asset A_1's valuation drops. Moreover, a limited quantity of the asset A_1 implies a loss to the other market participants if and only if it is a gain to the financial institution. Further, as regards asset A_2, the financial institution gains or not according as it plays "Buy A_1 and buy A_2" or "Short sell A_2", respectively. Again, the financial institution gains if and only if the rest of the market loses. Finally, both parties find profits or losses of trading in asset A_1 to be their top priority. The results of holdings of asset A_2 matter only when comparing two situations which lead to indifference as regards their profits from trading in asset A_1.

As in Example 1, each cell in Figure 2 represents the lexicographic payoffs to the financial institution for the corrseponding play of pure strategy tuples. Thus, if the financial institution plays "Buy A_1 and buy A_2" and the market plays "Optimistic", the payoffs to the financial institution are $(1,1)$ as A_1 becomes valuable, and A_2 is anyway valuable. Likewise, if the financial institution plays "Short sell A_1" and the market plays "Pessimistic", the payoffs to the financial institution are $(1,0)$ as A_1 loses value and A_2 is not acquired. For the remaining payoffs, note that the financial instution's guess is wrong.

Observe, the first components of the ordered pairs define a game of "matching pennies" which is known to have $(\frac{1}{2}T \oplus \frac{1}{2}B, \frac{1}{2}L \oplus \frac{1}{2}R)$ as the unique Nash equilibrium. This raises the following questions. Does the above game—as it is—have a Nash equilibrium? Is our solution concept able to predict some play in this game? If yes, then is indeed the prediction $(\frac{1}{2}T \oplus \frac{1}{2}B, \frac{1}{2}L \oplus \frac{1}{2}R)$? ■

EXAMPLE 3: The bilateral conflict between two nations 1 ("Player I") and 2 ("Player II") are defined by their strategies, and the resulting outcomes, as follows. There are three outcomes which, in the decreasing order of priority to nation 1, are the following:

1. "Have nuclear technologies".
2. "Surround 2 with allies".
3. "Achieve international collaborations if 2 does".

It is then plausible that nation 2's preferences are such that we have a zero–sum game. The strategy sets of nations 1 and 2 are $\{T, B\}$ and $\{L, M, R\}$, respectively. The description of each pure strategy is some combination of sentences from the following list:

$S_{\mathrm{I},1}$:= "Attempt to develop nuclear technologies".
$S_{\mathrm{I},2}$:= "Attempt to form allies that surround 2".
$S_{\mathrm{I},3}$:= "Do not make international collaborations".
$S_{\mathrm{I},4}$:= "Make international collaborations".
$S_{\mathrm{II},1}$:= "Trust that 1 will *not* develop nuclear technologies".
$S_{\mathrm{II},2}$:= "Enforce sanctions on 1".
$S_{\mathrm{II},3}$:= "Influence 1's potential allies that surround 2".
$S_{\mathrm{II},4}$:= "Do not make international collaborations".
$S_{\mathrm{II},5}$:= "Make international collaborations".

Then, the description of each pure strategy is as follows:

$$T := S_{\mathrm{I},1} \text{ and } S_{\mathrm{I},2} \text{ and } S_{\mathrm{I},3}.$$
$$B := S_{\mathrm{I},2} \text{ and } S_{\mathrm{I},4}.$$
$$L := S_{\mathrm{II},1} \text{ and } S_{\mathrm{II},3}.$$
$$M := S_{\mathrm{II},4} \text{ and (if } S_{\mathrm{I},1} \text{ then } [S_{\mathrm{II},2} \text{ and } S_{\mathrm{II},3}]).$$
$$R := S_{\mathrm{II},5} \text{ and (if } S_{\mathrm{I},1} \text{ then } S_{\mathrm{II},2}).$$

Thus, strategy B of 1 admits the interpretation, "Attempt to form allies that surround 2, and, do not make international collaborations".

Now, we come to the question of lexicographic payoffs under various plays by 1 and 2. For instance, consider the play (T, L). By the description of T and L, we have a conjuction of sentences $S_{\text{I},1}$ and $S_{\text{II},1}$ as part of the outcome. Then, the definition of $S_{\text{I},1}$ and $S_{\text{II},1}$ imply that 1 will face no hindrance in its attempt to develop nuclear technologies which is its top priority. As a result nation 1 gets a payoff of 1 as is reflected by the first component of the ordered triple in the cell in Figure 3 which corresponds to the pure strategy pair (T, L).

			Player II		
			q	r	$1-(q+r)$
			L	M	R
Player I	p	T	$(1,0,0)$	$(0,0,0)$	$(0,1,1)$
	$1-p$	B	$(0,0,0)$	$(0,1,1)$	$(0,1,0)$

FIGURE 3: Bilateral conflict.

Moreover, since there is conjunction of sentences $S_{\text{I},2}$ and $S_{\text{II},3}$ as well under the pair (T, L), it follows that though 1 attempts to form allies that surround 2, it fails because 2 influences those potential allies in this outcome. Since forming allies with those that surround 2 is 1's second priority, the second component of the ordered triple in the cell corresponding to (T, L) is 0. Further, the play (T, L) also involves the clause $S_{\text{I},3}$ which means that 1 does not form any international collaborations. As this is third in the priority of 1, we have 0 as the third component of the ordered triple in the cell corresponding to to the play (T, L). Having specified the lexicographic payoffs to 1 under the play (T, L), we observe that as the game is zero–sum the lexicographic payoffs to 2 thus stand specified in the obvious manner. Likewise, we obtain the remaining ordered triples in Figure 3.

Now, consider the play (B, R). Since $S_{\text{I},1}$ is not part of the definition of B, nation 1 will have no access to nuclear technologies. However, since B has $S_{\text{I},2}$ and R does not have $S_{\text{II},3}$, nation 2 will end up being surrounded by 1's allies. Moreover, since $S_{\text{I},4}$ is a part of B and $S_{\text{II},5}$ is a part of R, both the nations end up making international collaborations under (B, R). In particular, even though 1 is able to surround 2 with its allies, there is no further consequence of this to 2. Is it possible that (B, R) is an "equilibrium" of this game? If yes, then is it unique? Does this game admit some Nash equilibrium? ■

We now proceed to the general framework and the abstract theory.

2. FRAMEWORK

Let the set of players be $N \coloneqq \{1, 2\}$. Typically, we shall denote the two players by i and j. Each player i has a non–empty and finite set S_i of pure strategies. Let $\Delta(S_i)$ denote the set of all mixed strategies of player i each of which is a lottery over the set S_i. Also, $\Delta(S_1 \times S_2)$ is the set of all lotteries over $S_1 \times S_2$.

On several occassions, we shall talk of "randomly choosing one out of several lotteries". Given lotteries $p_1, \ldots, p_K \in \Delta(S_1 \times S_2)$ and a randomization device which randomly results in one out of K outcomes, where the kth outcome obtains with probability α_k, we have a compound lottery over $S_1 \times S_2$ by running the lottery p_k if the randomization device results in its kth outcome. This compound lottery shall be denoted by $\alpha_1 \cdot p_1 \oplus \alpha_2 \cdot p_2 \oplus \ldots \alpha_K \cdot p_K$ or $\oplus_{k=1}^{K} \alpha_k \cdot p_k$. We assume this compound lottery to be equivalent to the unique (simple) lottery which randomly selects a typical action pair $(s_1, s_2) \in S_1 \times S_2$ with probability $\sum_{k=1}^{K} \alpha_k p_k(s_1, s_2)$.

If $x_1 \in \Delta(S_1)$ and $x_2 \in \Delta(S_2)$ are mixed strategies of players 1 and 2, then we denote by (x_1, x_2) the lottery over $S_1 \times S_2$ which selects any (s_1, s_2) by *independently* selecting s_1 and s_2 according to x_1 and x_2, respectively. Therefore, the probability that the pair (s_1, s_2) obtains is $x_1(s_1)x_2(s_2)$. Now, if player 2's mixed strategy is x_2 but player 1 ramdomly selects his mixed strategy to be either x_1^* or x_1^{**}, with the probability of the former being α, then we essentially have the mixed strategy tuple $(\alpha \cdot x_1^* \oplus [1 - \alpha] \cdot x_1^{**}, x_2)$.

Each player i has a preference $\succsim_i$ which is a complete and transitive binary relation over $\Delta(S_1 \times S_2)$. Further, $\succsim_i$ satisfies our weakening (Theorem 2 of subsection 3.2 in Chatterjee [2022]) of Independence due to von Neumann & Morgenstern (1944).

Independence: *For any $p, q, r \in \Delta(S_1 \times S_2)$, $p \succ_i q$ if and only if*

$$\big(\forall \alpha \in (0,1)\big)\big[\; \alpha \cdot p \oplus [1-\alpha] \cdot r \succ_i \alpha \cdot q \oplus [1-\alpha] \cdot r \;\big].$$

A *two–person zero–sum game* is any tuple $\langle N, (S_i)_{i \in N}, (\succsim_i)_{i \in N} \rangle$ such that, for every $p, q \in \Delta(S_1 \times S_2)$:

$$p \succsim_1 q \iff q \succsim_2 p.$$

Thus, one player's loss is the other's gain. Further, if the preferences $\succsim_1$, $\succsim_2$ admit expected utility representations $u_1, u_2 : \Delta(S_1 \times S_2) \to \mathbb{R}$ respectively, then without loss of generality we have: $u_2 = -u_1$ if and only if $p \succsim_1 q \iff q \succsim_2 p$. Hence, we have a natural generalization of the classical definition of two–person zero–sum games.

3. ADMISSIBLE STRATEGIES

Admissible strategies shall be defined with respect to some context. A *context* is a pair $\langle C_1, C_2 \rangle$ where C_i is a non–empty and compact subset of $\Delta(S_i)$. Given a context $\langle C_1, C_2 \rangle$, for each player i, let $\mathscr{A}_i^G \langle C_1, C_2 \rangle$ be the class of non–empty closed $A_i \subseteq C_i$ with the following property.

PROPERTY G: *For every* $x_i \in C_i \setminus A_i$, *there exists* $x_j \in C_j$ *such that*

$$(x_i', x_j') \succ_i (x_i, x_j) \text{ for any } x_i' \in A_i \text{ and any } x_j' \in C_j.$$

Here, $\succ_i$ denotes "strict preference". Notice, C_i belongs to the class $\mathscr{A}_i^G \langle C_1, C_2 \rangle$ vacuously. For the interpretation of "$A_i \in \mathscr{A}_i^G \langle C_1, C_2 \rangle$", Figure 4 shows a context where $C_i = \Delta(S_i)$ for each player i.

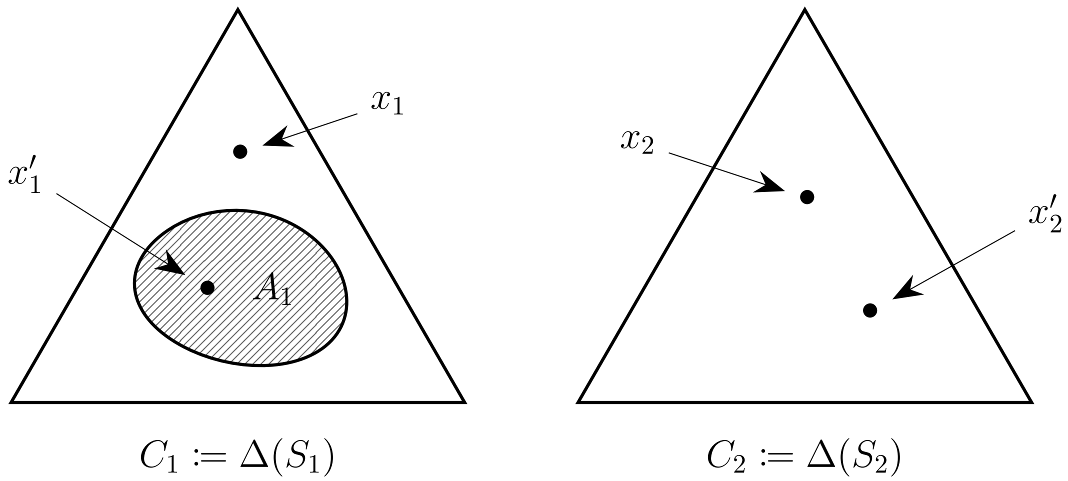


FIGURE 4: A set A_1 in $\mathscr{A}_1^G \langle \Delta(S_1), \Delta(S_2) \rangle$.

The mixed strategy x_1 of player 1 is not in $A_1 \subseteq \Delta(S_1)$. Also, x_1' is an arbitrary strategy in A_1. Thus, "$A_1 \in \mathscr{A}_1^G \langle \Delta(S_1), \Delta(S_2) \rangle$" holds, if and only if, there exists some mixed strategy x_2 of player 2 such that player 1 strictly prefers the play (x_1', x_2') to the play (x_1, x_2) for every possbile strategy x_2' of player 2. Thus, player 1 has some "minimum guarantees" if he considers playing strategies from A_1 irrespective of the strategy his opponent chooses to play.

This is on the lines of the minimax theory of VON NEUMANN (1928). In that theory, the maximin strategies assure players that they receive at least the value irrespective of their opponents play. However, this assured utility level is the highest that can be assured. To incorporate this additional feature, we consider the following.

DEFINITION 1: *A set* $A_i \subseteq C_i$ *is* admissible *with respect to the context* $\langle C_1, C_2 \rangle$ *if* $A_i \in \mathscr{A}_i^G \langle C_1, C_2 \rangle$, *and* $A_i \subseteq A_i'$ *for every* $A_i' \in \mathscr{A}_i^G \langle C_1, C_2 \rangle$.

That is, an admissible set in a context is one which is minimal in the sense of set–inclusion among all sets which satisfy property G in that context. For instance, if A_i' is a typical set that satisfies property G in the context $\langle C_1, C_2 \rangle$ and A_i is admissible, then the fact that $A_i \subseteq A_i'$ implies the following: for any $x_i \in A_i' \setminus A_i$, there exists $x_j \in C_j$ such that $(x_i', x_j') \succ_i (x_i, x_j)$ for all $x_i' \in A_i$ and all $x_j' \in C_j$. Thus, the "minimum guarantee" assured to player i by playing mixed strategies from his admissible set A_i is "as high as it can get".

To see the justification for player i to consider playing from his admissible set, we consider the following alternate perspective. For strategy $x_i \in C_i$ by player i, let $x_j \in C_j$ be a *best response in* C_j of player j if: $(x_i, x_j) \precsim_i (x_i, x_j')$ for every $x_j' \in C_j$. Given the context $\langle C_1, C_2 \rangle$, denote by $\mathscr{A}_i^D \langle C_1, C_2 \rangle$ the class of all non–empty compact sets $A_i \subseteq C_i$ with the following property.

PROPERTY B: *For every $x_i \in C_i \setminus A_i$ and any best response x_j in C_j, $(x_i', x_j') \succ_i (x_i, x_j)$ for any $x_i' \in A_i$ and any best response x_j' in C_j to x_i'.*

In contrast to property G which considered arbitrary beliefs by a player about his opponent, property D considers best responses in the context under consideration. The first result is as follows.

PROPOSITION 1: *For any context $\langle C_1, C_2 \rangle$, $\mathscr{A}_i^G \langle C_1, C_2 \rangle = \mathscr{A}_i^B \langle C_1, C_2 \rangle$.*

PROOF: Fix $A_i \in \mathscr{A}_i^G \langle C_1, C_2 \rangle$. Let $x_i \in C_i \setminus A_i$ and x_j be a best response in C_j of player j. For any arbitrary $x_i' \in A_i$, let x_j' be a best response in C_j of player j. By property G, there exists $x_j^* \in C_j$ such that $(x_i', x_j') \succ_i (x_i, x_j^*)$. Also, $(x_i, x_j) \succsim_j (x_i, x_j^*)$ as x_j is a best response in C_j. By defintion of two–person zero–sum game, $(x_i, x_j^*) \succsim_i (x_i, x_j)$. By transitivity of $\succsim_i$, $(x_i', x_j') \succ_i (x_i, x_j)$. Thus, $A_i \in \mathscr{A}_i^B \langle C_1, C_2 \rangle$.

Fix $A_i \in \mathscr{A}_i^B \langle C_1, C_2 \rangle$. Let $x_i \in C_i \setminus A_i$ and x_j be a best response in C_j. Fix an arbitrary $x_i' \in A_i$ and $x_j' \in C_j$. Thus, $(x_i', x_j^*) \succsim_j (x_i', x_j')$ where x_j^* is any best response in C_j to x_i'. Since the game is zero–sum, $(x_i', x_j') \succsim_i (x_i', x_j^*)$. By property B, $(x_i', x_j^*) \succ_i (x_i, x_j)$. Transitivity of $\succsim_i$ implies $(x_i', x_j') \succ_i (x_i, x_j)$. That is, $A_i \in \mathscr{A}_i^G \langle C_1, C_2 \rangle$. ■

Just as Definition 1 defines "admissibility" based on property G, it is possible to define an analogous notion based on property B. In light of the proposition above, both the notions must coincide. Henceforth, we shall write "$\mathscr{A}_i \langle C_1, C_2 \rangle$" for both "$\mathscr{A}_i^G \langle C_1, C_2 \rangle$" and "$\mathscr{A}_i^B \langle C_1, C_2 \rangle$". The following lemma asserts convexity of elements in $\mathscr{A}_i \langle C_1, C_2 \rangle$.

LEMMA 1: *Let $\langle C_1, C_2\rangle$ be a context. For any $i \in N$, if C_i is convex and $A_i \in \mathscr{A}_i\langle C_1, C_2\rangle$, then A_i is convex.*

PROOF: Assume $A_1 \in \mathscr{A}_1\langle C_1, C_2\rangle$ and suppose: A_1 is not convex. Thus, for some[3] $x_1^*, x_1^{**} \in A_1$ and $\alpha \in (0,1)$, $x_1^\alpha := \alpha \cdot x_1^* \oplus [1-\alpha] \cdot x_1^{**} \notin A_1$. Note, $x_1^\alpha \in C_1$ as C_1 is convex. Since $A_1 \in \mathscr{A}_1\langle C_1, C_2\rangle$, by property B of A_1 with respect to $\langle C_1, C_2\rangle$, there exists $x_2^* \in C_2$ such that:

$$(x_1, x_2) \succ_1 (x_1^\alpha, x_2^*) \text{ for all } (x_1, x_2) \in A_1 \times C_2.$$

In particular, $(x_1^*, x_2^*) \succ_1 (x_1^\alpha, x_2^*)$ and $(x_1^{**}, x_2^*) \succ_1 (x_1^\alpha, x_2^*)$ hold. By Independence, $(x_1^\alpha, x_2^*) \succ_1 (\alpha \cdot x_1^\alpha \oplus [1-\alpha] \cdot x_1^{**}, x_2^*)$ as $(x_1^*, x_2^*) \succ_1 (x_1^\alpha, x_2)$. Similarly, $(\alpha \cdot x_1^\alpha \oplus [1-\alpha] \cdot x_1^{**}, x_2^*) \succ_1 (x_1^\alpha, x_2^*)$ as $(x_1^{**}, x_2^*) \succ_1 (x_1^\alpha, x_2^*)$ by Independence. Then, the transitivity of $\succ_1$ implies $(x_1^\alpha, x_2^*) \succ_1 (x_1^\alpha, x_2^*)$. However, $\succ_1$ is asymmetric. Thus, we have a contradiction. Hence, our supposition must be wrong. Therefore, A_1 is convex. ■

For existence of admissible sets, consider the following result.

THEOREM 1: *Let $\langle C_1, C_2\rangle$ be any context. If $A_i, A_i' \in \mathscr{A}_i\langle C_1, C_2\rangle$, then $A_i \subseteq A_i'$ or $A_i' \subseteq A_i$. Further, admissible sets exist for each player which are unique, non–empty and compact. If C_1 and C_2 are convex, then so are the admissible sets.*

PROOF: Suppose, $A_i, A_i' \in \mathscr{A}_i\langle C_1, C_2\rangle$ are such that $A_i \setminus A_i' \neq \varnothing$ and $A_i' \setminus A_i \neq \varnothing$. Fix $x_i \in A_i \setminus A_i'$ and $x_i' \in A_i' \setminus A_i$. Since $x_i' \in C_i \setminus A_i$, $x_i \in A_i$ and $A_i \in \mathscr{A}_i\langle C_1, C_2\rangle$, there exists $x_j' \in C_j$ such that:

$$(x_i, x_j^*) \succ_i (x_i', x_j') \text{ for all } x_j^* \in C_j. \tag{1}$$

Moreover, since $x_i \in C_i \setminus A_i'$, $x_i' \in A_i'$ and $A_i' \in \mathscr{A}_i\langle C_1, C_2\rangle$, there exists $x_j \in C_j$ such that the following holds:

$$(x_i', x_j^*) \succ_i (x_i, x_j) \text{ for all } x_j^* \in C_j. \tag{2}$$

In particular, (1) implies $(x_i, x_j) \succ_i (x_i', x_j')$. Likewise, (2) implies $(x_i', x_j') \succ_i (x_i, x_j)$. Transitivity of $\succsim_i$ then implies $(x_i, x_j) \succ_i (x_i, x_j)$ which is a contradiction. Thus, we have established:

$$[A_i, A_i' \in \mathscr{A}_i\langle C_1, C_2\rangle] \implies [A_i \subseteq A_i' \text{ or } A' \subseteq A_i]. \tag{3}$$

[3]For $p, q \in \Delta(Z)$ and $\alpha \in (0,1)$, $\alpha \cdot p \oplus [1-\alpha] \cdot q \in \Delta(Z)$ is defined as the lottery over Z which selects with propbability $\alpha p(z) + [1-\alpha]q(z)$ any basic prize $z \in Z$.

Let[4] $A_i^* := \bigcap\{A_i : A_i \in \mathscr{A}_i\langle C_1, C_2\rangle\}$. Since each member of the class $\mathscr{A}_i\langle C_1, C_2\rangle$ is non–empty, (3) implies that the intersection of finitely many members of $\mathscr{A}_i\langle C_1, C_2\rangle$ is non–empty. Further, each element of $\mathscr{A}_i\langle C_1, C_2\rangle$ is compact. Thus, the set A_i^* is non–empty and compact. Clearly, $A_i^* \subseteq A_i$ for every $A_i \in \mathscr{A}_i\langle C_1, C_2\rangle$. Hence, to conclude that A_i^* is the unique admissible set for player i, with respect to the context $\langle C_1, C_2\rangle$, it is enough to argue that A_i^* satisfies property G.

For this, fix any $x_i \in C_i \setminus A_i^*$. Also, let $x_i' \in A_i^*$ and $x_j' \in C_j$ be arbitrary. Since $x_i \in C_i \setminus A_i^*$ and A_i^* is the intersection of all members of $\mathscr{A}_i\langle C_1, C_2\rangle$, there exists $A_i \in \mathscr{A}_i\langle C_1, C_2\rangle$ with $x_i \in C_i \setminus A_i$. Further, $x_i' \in A_i$ as $x_i' \in A_i^* \subseteq A_i$. By definition of $\mathscr{A}_i\langle C_1, C_2\rangle$, there exists $x_j \in C_j$ such that $(x_i', x_j') \succ_i (x_i, x_j)$. Thus, $A_i^* \in \mathscr{A}_i\langle C_1, C_2\rangle$. This proves: A_i^* is admissible with respect to the context $\langle C_1, C_2\rangle$.

It remains to argue: A_i^* is convex if C_1 and C_2 are convex. However, Lemma 1 shows that every element of $\mathscr{A}_1\langle C_1, C_2\rangle$ is convex. Further, A_1^* is the intersection of all elements of $\mathscr{A}_1\langle C_1, C_2\rangle$. Hence, A_1^* is convex. Symmetric arguments work for the admissible set of player 2. ■

That is, the collection $\mathscr{A}_i\langle C_1, C_2\rangle$ of sets satisfying property G (or B), with respect to the context $\langle C_1, C_2\rangle$, form a nest of compact sets whose intersection is the unique minimal element which also satisfies property G (or B). Define for each player i the set:

$$A_i^*\langle C_1, C_2\rangle := \bigcap \{S : S \in \mathscr{A}_i\langle C_1, C_2\rangle\}$$

Therefore, $A_i^*\langle C_1, C_2\rangle$ is *the* admissible set of player i. Thus, if players consider playing from $C_1 \times C_2$, then it makes sense for player i to restrict consideration to within the set $A_i^*\langle C_1, C_2\rangle$ as it is the minimal set satisfying property G (or B) with respect to $\langle C_1, C_2\rangle$.

4. CONSIDERATION EQUILIBRIA

The definition of the term "context" and theorem 1 imply that the pair of admissible sets with respect to a context form a context in its own right. However, if it makes sense for players to restrict consideration to admissible sets, then the pair of these admissible sets is *as if* the new context. Thus, players may further restrict their consideration to the resulting admissible sets with respect to this new context. That is, starting with the pair of simplices of all mixed strategies, players may consider eliminating non–admissible strategies iteratively.

[4]For any collection $\{A_\alpha : \alpha \in \mathscr{A}\}$ of sets, $\bigcap\{A_\alpha : \alpha \in \mathscr{A}\}$ is the intersection of its members.

Definition 2: *A sequence of contexts* $\{\langle C_{1,k}, C_{2,k}\rangle\}_{k\in\mathbb{N}}$ *is* tight *if,*

1. $C_{i,1} = \Delta(S_i)$ *for each* $i \in N$*, and*
2. *For every* $k \in \mathbb{N}$, $C_{i,k+1} = A_i^*\langle C_{1,k}, C_{2,k}\rangle$ *for each* $i \in N$.

If there exists a unique tight sequence of contexts $\{\langle C_{1,k}^*, C_{2,k}^*\rangle\}_{k\in\mathbb{N}}$*, then define the following pair of sets*:

$$A_i^\infty := \bigcap_{k=1}^{\infty} C_{i,k}^* \quad \text{for each } i \in N.$$

A consideration equilibrium *is any strategy tuple from* $A_1^\infty \times A_2^\infty$.

"Tightness" means iterated elimination of non–admissible strategies starting with the full simplices as the context. Thus, Figure 5 illustrates admissible sets, with respect to the present context, forming the next context as required by definition 2. Theorem 2 addresses questions of existence and structure of consideration equilibria.

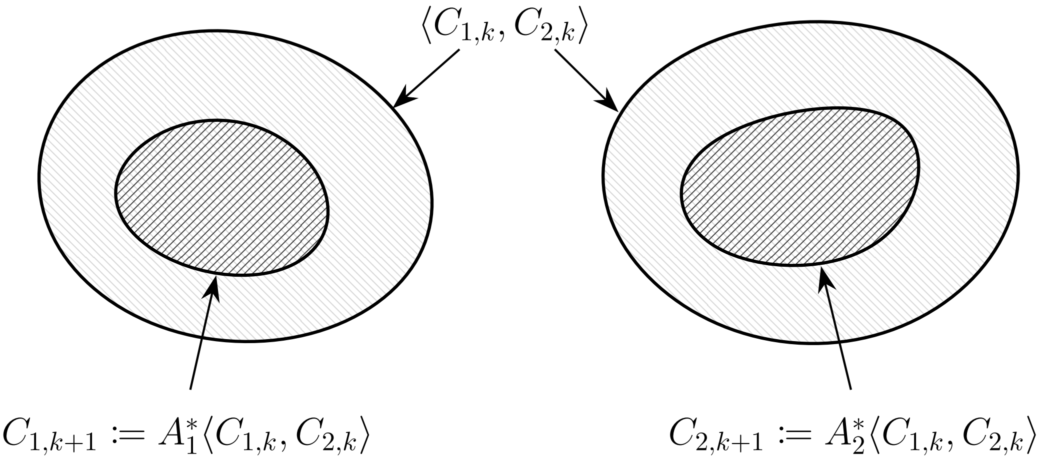


Figure 5: Next context as pair of admissible sets in present context.

Theorem 2: *There exists a unique tight sequence of contexts and the consideration equilibria form a unique, non–empty, compact and convex set. Further, if* $x_1^*, x_1^{**} \in A_1^\infty$ *and* $x_2^*, x_2^{**} \in A_2^\infty$*, then*:

$$(x_1^*, x_2^*) \sim_i (x_1^{**}, x_2^{**}) \text{ for each player } i.$$

Proof: A unique tight sequence of contexts exists by Theorem 1. Also, $C_{i,k+1}^*$ is a non–empty and compact subset of $C_{i,k}^*$ for each $k \in \mathbb{N}$. Thus, A_i^∞ is non–empty and compact. Additionally, convexity of $C_{i,k}^*$ for each $k \in \mathbb{N}$, and the definition of A_i^∞, implies the convexity of A_i^∞.

Thus, it remains to argue: if $x_1^*, x_1^{**} \in A_1^\infty$ and $x_2^*, x_2^{**} \in A_1^\infty$, then $(x_1^*, x_2^*) \sim_i (x_1^{**}, x_2^{**})$ for each player $i \in N$. This follows from Proposition 7 which is stated and proved in section 7. ■

Theorem 2 generalizes the result of von Neumann (1928) that all minimax strategies are interchangeable. In particular, the choice of one out of all consideration equilibria is not an issue.

Our justification for the solution concept is based on the role of property G or property B and minimality in the definition of admissible sets. Thus, given any context, players should restrict further attention to their admissible sets of strategies. The following result sharpens the basis for *not* considering non–admissible strategies. For this, it will be useful to keep figure 5 in perspective which illustrates a context and the corresponding admissible sets of the players.

Theorem 3: *Let $\langle C_1, C_2 \rangle$ be any context. Assume i and j are the distinct players. Consider $x_i^* \in C_i$ and suppose the following hold*:

1. *j conjectures i will play x_i^*.*
2. *i knows j's conjecture.*
3. *j knows that i knows j's conjecture.*
4. *If $x_i^* \notin A_i^*\langle C_1, C_2 \rangle$, then*:
   (a) *j will play a best response in C_j to j's conjecture.*
   (b) *i knows* (a).
   (c) *i will play a best response in C_i to i's conjecture.*
   (d) *j knows* (b) and (c).

Then, j knows that i and j know $x_i^* \in A_i^*\langle C_1, C_2 \rangle$.

Proof: Let j conjecture that i plays x_i^*. Suppose that $x_i^* \notin A_i^*\langle C_1, C_2 \rangle$. By 4(a), j will play some best response C_j, say x_j, which i knows as 4(b) holds. However, $(x_i, x_j) \succ_i (x_i^*, x_j)$ for any $x_i \in A_i^*\langle C_1, C_2 \rangle$ by property B and Proposition 1. Hence, i shall play some $x_i^{**} \in A_i^*\langle C_1, C_2 \rangle$ as 4(c) holds. Since this is known to j as 4(d) holds, we have a contradiction to the fact that j's conjecture of i's play is x_i^*. Thus, $x_i^* \in A_i^*\langle C_1, C_2 \rangle$. By 1, j knows $x_i^* \in A_i^*\langle C_1, C_2 \rangle$. Also, i knows $x_i^* \in A_i^*\langle C_1, C_2 \rangle$ by 2. By 3, j knows that i knows $x_i^* \in A_i^*\langle C_1, C_2 \rangle$. Since j knows $x_i^* \in A^*\langle C_1, C_2 \rangle$, we have: j knows that j knows $x_i^* \in A_i^*\langle C_1, C_2 \rangle$. ■

A remark is in order. Suppose that x_i^* is j's conjecture about i's play. Further, let $\langle C_1, C_2 \rangle$ be any context such that x_i^* is in the corresponding admissible set $A_i^* \langle C_1, C_2 \rangle$. Then, assumption 4 of Theorem 3 holds vacuously. Thus, no logical inconsistency arises from the use of Theorem 3 to justify the proposed solution concept.

However, assumption 4 requires "maxmization" and its knowledge in only a *conditional* sense. The solution concept requires that each player takes the following stance about his strategic considerations with regard to his opponent's strategic considerations.

> "If you do not restrict your considerations given the context, *then* so will I. Then, if I think that your play will be outside of your admissible set, I too will play my best response to it within the context. Since I must then assume that your play is your best response in the context, I find that your play must be within your admissible set which is a contradiction."

Moreover, any such reasoning by a player is *irrelevant* if his conjecture lies in the admissible set, of his opponent, to begin with.

Thus, both players realize that their own conjecture about their opponent's play must be restricted to the admissible sets of their opponent. It is the implausibility of assuming maximization by the opponent *without* first restricting consideration to admimssible sets is what drives the iterated elimination in the solution concept. Pearce (1984) argues that Nash equilibrium is *not* the only sensible way for the players to behave based on rationality. Our point is that it makes sense for the players to only consider plausible conjectures while maximization. The following proposition asserts that if the game admits a Nash equilibrium, then it must be a consideration equilibrium.

Proposition 2: *Any Nash equilibrium is a consideration equilibrium.*

Proof: Let (x_1^*, x_2^*) be a Nash equilibrium. Clearly, for each $i \in N$, $x_i^* \in C_{i,1}^*$ as $C_{i,1}^* = \Delta(S_i)$. Suppose, there exists $k \in \mathbb{N}$ such that $(x_1^*, x_2^*) \in C_{1,k}^* \times C_{2,k}^*$ and $(x_1^*, x_2^*) \notin C_{1,k+1}^* \times C_{2,k+1}^*$. Assume, without loss of generality, $x_1^* \notin C_{1,k+1}^*$. Recall, $C_{i,k+1}^* = A^* \langle C_{1,k}^*, C_{2,k}^* \rangle$ for each $i \in N$. Thus, there exists $x_2 \in C_{2,k+1}^*$ such that: $(x_1, x_2^*) \succ_1 (x_1^*, x_2)$ for any $x_1 \in C_{1,k+1}^*$. Further, $(x_1^*, x_2^*) \succsim_2 (x_1^*, x_2)$ as (x_1^*, x_2^*) is a Nash equilibrium. Then, $(x_1^*, x_2) \succsim_1 (x_1^*, x_2^*)$ as the game is zero–sum. Thus, $(x_1, x_2^*) \succ_1 (x_1^*, x_2^*)$ for all $x_1 \in C_{1,k+1}^*$. However, this contradicts that fact that x_1^* is a best response in $\Delta(S_1)$ to x_2^*. Thus, $x_1^* \in C_{1,k}^*$ for every $k \in \mathbb{N}$. Hence, $x_1^* \in A_1^\infty$ by definition of A_1^∞. ■

The next proposition says that the proposed solution concept reduces exactly to the minimax strategies, which are also precisely the Nash equilibria, when preferences that satisfy the Independence axiom are additionally known to satisfy Continuity. Therefore, the concept of consideration equilibria indeed generalizes the classical theory.

PROPOSITION 3: *Suppose, players's preferences satisfy Independence and Continuity. Then, a strategy tuple is a Nash equilibrium*[5]*, if and only if, it is a consideration equilibrium.*

PROOF: Since $\succsim_i$ satisfies Independence and Continuity, the Theorem of VON NEUMANN & MORGENSTERN (1944) on existence of expected utility representations and the definition of two–person zero–sum game allow us to conclude: there exists $U_1, U_2 : \Delta(S_1 \times S_2) \to \mathbb{R}$ such that $U_2 = -U_1$, and U_i is an expected utility[6] that represents[7] $\succsim_i$ for each $i \in N$. Further, by the Minimax Theorem of VON NEUMANN (1928), there exists a *unique* value $v \in \mathbb{R}$ such that the sets:

$$M_1 := \{x_1^* \in \Delta(S_1) : U_1(x_1^*, x_2) \geq +v \text{ for all } x_2 \in \Delta(S_2)\}, \text{ and}$$
$$M_2 := \{x_2^* \in \Delta(S_2) : U_2(x_1, x_2^*) \geq -v \text{ for all } x_1 \in \Delta(S_1)\}$$

are the minimax strategies of the players 1 and 2, respectively. Note, the above description of M_1 and M_2 is equivalent to the more familiar one which is as follows: $(x_1^*, x_2^*) \in M_1 \times M_2$ if and only if,

$$(x_1^*, x_2) \succsim_1 (x_1^*, x_2^*) \succsim_1 (x_1, x_2^*) \text{ for all } (x_1, x_2) \in \Delta(S_1) \times \Delta(S_2)$$

Clearly, the set of all Nash equilibria is $M_1 \times M_2$. Without any loss of generality, we shall argue: $M_i = A_i^\infty$ for each $i \in N$.

Let $x_1 \in \Delta(S_1) \setminus M_1$. Thus, there exists $x_2 \in \Delta(S_2)$ such that $v > U_1(x_1, x_2)$. Fix any $x_1' \in M_1$ and $x_2' \in \Delta(S_2)$. Then, $U_1(x_1', x_2') \geq v$ holds. Since U_1 is a representation of $\succsim_1$, we have: $(x_1', x_2') \succ_1 (x_1, x_2)$. That is, M_1 satisfies property B with respect to $\langle \Delta(S_1), \Delta(S_2) \rangle$ as the context. Further, M_1 is convex. To see why, let $x_1^*, x_1^{**} \in M_1$. Fix an arbitrary $x_2 \in \Delta(S_2)$. Then, $U_1(x_1^*, x_2) \geq v$ and $U_1(x_1^{**}, x_2) \geq v$. If $\alpha \in (0,1)$, then[8] $U_1(\alpha \cdot x_1^* \oplus [1-\alpha] \cdot x_1^{**}, x_2) = \alpha U_1(x_1^*, x_2) + [1-\alpha] U_1(x_1^{**}, x_2)$ as U_1 is an expected utility. Hence, $U_1(\alpha \cdot x_1^* \oplus [1-\alpha] \cdot x_1^{**}, x_2) \geq v$ if $\alpha \in (0,1)$. Since $x_2 \in \Delta(S_2)$ is arbitrary, we have: $\alpha \cdot x_1^* \oplus [1-\alpha] \cdot x_1^{**} \in M_1$ for every $\alpha \in (0,1)$. That is, M_1 is convex.

[5]With Independence *and* Continuity, existence of a Nash equilibrium is guaranteed.

[6]The map $U : \Delta(Z) \to \mathbb{R}$ is an *expected utility* if: $U(p) = \sum_{z \in Z} p(z) U(z)$ for any $p \in \Delta(Z)$.

[7]The map $U : \Delta(Z) \to \mathbb{R}$ *represents* the preference $\succsim$ over $\Delta(Z)$ if: $p \succsim q \iff U(p) \geq U(q)$.

[8]For $p, q \in \Delta(Z)$ and $\alpha \in (0,1)$, $\alpha \cdot p \oplus [1-\alpha] \cdot q \in \Delta(Z)$ is defined as the lottery over Z which selects with propbability $\alpha p(z) + [1-\alpha] q(z)$ any basic prize $z \in Z$.

Further, we have: $M_1 \subseteq \Delta(S_1)$ is closed. To see this, fix an arbitrary $x_2 \in \Delta(S_2)$. Since the map U_1 is an expected utility and the set S_1 is finite, the map $x_1 \in \Delta(S_1) \mapsto U_1(x_1, x_2)$ is continuous. Thus, the set $M_1(x_2, v) := \{x_1 \in \Delta(S_1) : U_1(x_1, x_2) \geq +v\}$ is closed in $\Delta(S_1)$. Also, note that the following equality holds:

$$M_1 = \bigcap \big\{ M_1(x_2, v) : x_2 \in \Delta(S_2) \big\}.$$

Thus, $M_1 \subseteq \Delta(S_1)$ is closed. Since $\Delta(S_1)$ is compact, it follows that M_1 is compact. Thus, we have: $M_1 \in \mathscr{A}_1\langle \Delta(S_1), \Delta(S_2)\rangle$.

We now argue: $M_1 = A_1^*\langle \Delta(S_1), \Delta(S_2)\rangle$. Suppose, not! Theorem 1 implies that there exists non–empty, convex and compact $A_1 \subsetneq M_1$ which satisfies property B. Let $x_1 \in M_1 \setminus A_1$. Fix an arbitrary $x_2 \in \Delta(S_2)$. Thus, $U_1(x_1, x_2) \geq v$ by definition of M_1. Let $x_1' \in A_1$ and $x_2' \in M_2$. Then, $U_1(x_1', x_2') \geq v$ and $U_2(x_1', x_2') \geq -v$ by definition of M_1 and M_2, respectively. However, $U_2 = -U_1$ and $U_2(x_1', x_2') \geq -v$ implies $U_1(x_1', x_2') \leq v$. That is, $U_1(x_1', x_2') = v$. Thus, $U_1(x_1, x_2) \geq U_1(x_1', x_2')$. Since U_1 represents $\succsim_1$, we have: $(x_1, x_2) \succsim_1 (x_1', x_2')$. Since $x_2 \in \Delta(S_2)$ was arbitrary, we have a contradiction to property B of A_1. Thus, we have: $M_1 = A_1^*\langle \Delta(S_1), \Delta(S_2)\rangle$. Similarly, $M_2 = A_2^*\langle \Delta(S_1), \Delta(S_2)\rangle$.

By definition 2, $C_{i,1}^* = \Delta(S_i)$ and $C_{i,k+1}^* = A_i^*\langle C_{1,k}^*, C_{2,k}^*\rangle$ for all $k \in \mathbb{N}$. Thus, $M_i = C_{i,2}^*$. Since $U_2 = -U_1$ and U_i represents $\succsim_i$, if (x_1^*, x_2^*) and (x_1^{**}, x_2^{**}) are in $M_1 \times M_2$, then $(x_1^*, x_2^*) \sim_i (x_1^{**}, x_2^{**})$ for each $i \in N$. Thus, $C_{i,k}^* = C_{i,2}^*$ for all $k \geq 2$. To see why, assume $k \geq 2$ is such that $C_{i,k}^* = M_i$ for each $i \in N$. Suppose, $A_1 \subsetneq C_{1,k}^*$ is non–empty, compact, convex and satisfies property B with respect to the context $\langle C_{1,k}^*, C_{2,k}^*\rangle = \langle M_1, M_2\rangle$. Let $x_1 \in C_{1,k}^* \setminus A_1$ and $x_1' \in A_1$. Fix $x_2, x_2' \in M_2$ arbitrarily. Since $x_1 \in M_1$ and $x_2 \in M_2$, we have: $U_1(x_1, x_2) \geq +v$ and $U_2(x_1, x_2) \geq -v$. As $U_2 = -U_1$, $U(x_1, x_2) = v$. Similarly, $U_1(x_1', x_2') = v$. That is, $U_1(x_1, x_2) = U_1(x_1', x_2')$. Since U_1 represents $\succsim_1$, we have $(x_1, x_2) \sim_1 (x' _, x_2')$. This contradicts property B of A_1 with respect to the context $\langle M_1, M_2\rangle$. Thus, $C_{1,k}^* = A_1^*\langle C_{1,k}^*, C_{2,k}^*\rangle$. That is, $C_{1,k+1}^* = C_{1,k}^*$ which implies: $C_{1,k+1}^* = M_1$. By a similar argument, $C_{2,k+1}^* = M_2$ holds. As $C_{i,k}^* = M_i$ for all $k \geq 2$, by definition 2 we obtain: $M_i = A_i^\infty$ for any player $i \in N$. ■

We make one final remark. In the light of Proposition 3, Theorem 3 thus provides epistemic conditions for the classical solution concepts in the setting with continuous preferences. Further, it makes explicit the knowledge assumptions that are sufficient for consideration equilibria. This exercise is in the spirit of Aumann & Brandenburger (1995) and Polak (1999) for the Nash equilibrium. Ideas in the above proof are generalized in section 6 without assuming Continuity.

5. APPLICATIONS

We analyse the examples from subsection 1.1 to evaluate whether the predictions of play as the consideration equilibria are reasonable.

EXAMPLE 1: We revisit the example of two competing firms 1 and 2 which are labelled as “Player I” and “Player II”, respectively. We briefly recall the setup which is represented in Figure 1. Firm 1 has two strategies which are “Execute a hostile price–cut” (T) or “Poach top talent of firm 2” (B). Also, firm 2 has two strategies which are “Counter firm 1’s move to poach talent, if any” (L) or “Match firm 1’s hostile price–cut, if any” (R). Each firm strictly prefers a higher market share than less. However, if two plays result in the same market share, then each firm is beter off with a larger pool of top talent. We do not repeat the justification of the values in the ordered pairs.

			Player II	
			q	$1-q$
			L	R
Player I	p	T	$(1,0)$	$(0,0)$
	$1-p$	B	$(0,0)$	$(0,1)$

FIGURE 6: Two competing firms.

The first component of the ordered pair in each cell of Figure 6 specifies the value of the first Bernoullian u_I^1 of player I. Hence, for mixed strategies $p \cdot T \oplus (1-p) \cdot B$ and $q \cdot L \oplus (1-q) \cdot R$ (henceforth, simply referred as p and q) of players I and II, the resulting expected utility to player I is $u_{\text{I}}^1(p,q) = pq$. Similarly, the expected utility to player I according to the second Bernoullian is $u_{\text{I}}^2(p,q) = (1-p)(1-q)$. As the game is zero–sum, the corresponding expected utilities to player II are $u_{\text{II}}^1(p,q) = -pq$ and $u_{\text{II}}^2(p,q) = -(1-p)(1-q)$.

The initial context is $\langle C_{\text{I},1}^*, C_{\text{II},1}^* \rangle$ comprising of the full simplices $C_{\text{I},1}^* := \{p \in [0,1]\}$ and $C_{\text{II},1}^* := \{q \in [0,1]\}$. Based on u_{I}^1, the best response in $C_{\text{I},1}^*$ of player I to any $q > 0$ is $p = 1$ resulting in u_{II}^1 and u_{II}^2 expected utilities $-q$ and 0, respectively, to player II. Also, if $q = 0$ then any $p \in [0,1]$ results in u_{II}^1 and u_{II}^2 expected utilities 0 and $-(1-p)$, respectively, to player II. As $(u_{\text{II}}^1, u_{\text{II}}^2)$ is a lexicographic expected utility representation of player II’s preference, he must restrict all his further considerations to the singleton $\{q = 0\}$. Hence, player II’s admissible set given the present context is: $A_{\text{II}}^* \langle C_{\text{I},1}^*, C_{\text{II},1}^* \rangle = \{q = 0\}$.

To compute player I's admissible set $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$, with respect to the present context, we begin with the following observation. Suppose, some $0 < p_0 < 1$ does not belong to $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$. If $p > p_0$ is in $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$ then the best response to p, of player II in $C^*_{\mathrm{II},1}$, is $q = 0$ which results in u^1_{I} and u^2_{I} expected utilities of 0 and $1-p$ to player I. Likewise, the u^1_{I} and u^2_{I} expected utilities to player I are 0 and $1-p_0$ when player II plays his best response $q = 0$ in $C^*_{\mathrm{II},1}$ to p_0. As $p > p_0$, it follows that p cannot belong to $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$ because $(u^1_{\mathrm{I}}, u^2_{\mathrm{I}})$ is a lexicographic expected utility representation of player I's preference. That is, $p \in A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$ implies that $p < p_0$. Since $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$ is a non–empty convex and compact subset of $\{p \in [0,1]\}$, there exists $p_* < p_0$ such that $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle = [0, p_*]$. Now, consider the strategy $p = 0$ of player I. Clearly, any $q \in [0,1]$ results in u^1_{I} expected utility of 0 to player I. As $u^2_{\mathrm{II}}(p,q) = (1-p)(1-q)$ and $p = 0$, player I's u^2_{I} expected utility is 0 if player II plays $q = 1$. Since $p_0 < 1$, it follows that $p = 0$ belonging to $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle = [0, p_*]$ and $p_0 \notin A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle = [0, p_*]$ contradicts the fact that $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle = [0, p_*]$ satifies property B as required of the admissible set (this is Theorem 1). Thus, $0 < p < 1$ implies that $p \in A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle = [0, p_*]$. As the admissible set must be compact, we have: $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle = [0, p_*] = \{p \in [0,1]\}$.

According to definition 2, the pair of admissible sets with respect to the present context serve as the next context. Hence, we must now set $C^*_{\mathrm{I},2} \coloneqq \{p \in [0,1]\}$ and $C^*_{\mathrm{II},2} \coloneqq \{q = 0\}$. Since $C^*_{\mathrm{II},2}$ is a singleton, it follows that player II's admissible set with respect to the new context is $C^*_{\mathrm{II},2}$; that is, $A^*_{\mathrm{II}}\langle C^*_{\mathrm{I},2}, C^*_{\mathrm{II},2}\rangle = \{q = 0\}$. Since $C^*_{\mathrm{II},2} = \{q = 0\}$, the u^1_{I} expected utility of player I is 0 for any $p \in C^*_{\mathrm{I},2} = [0,1]$. Also, if $p = 0$ then player I's u^2_{I} expected utility is 1 because $q = 0$ is only strategy of player II in $C^*_{\mathrm{II},2}$. Further, if $p > 0$ then player I's u^2_{I} expected utility is $1-p$ which is strictly less than 1. Thus, the singleton $\{p = 0\}$ is the admissible set of player I with respect to the new context because $(u^1_{\mathrm{I}}, u^2_{\mathrm{I}})$ is a lexicographic expected utility representation of player I's preference; that is, $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},2}, C^*_{\mathrm{II},2}\rangle = \{p = 0\}$. Since both admissible sets are singletons, further iterations as required by definition 2 shall not lead to any elimination. Hence, the surviving sets are $A^\infty_{\mathrm{I}} = \{p = 0\}$ and $A^\infty_{\mathrm{II}} = \{q = 0\}$. That is, the strategy tuple (B, R) is the *unique* consideration equilibrium of the game.

Thus, we find that the unique consideration equilibrium involves firms 1 and 2 playing the strategies "Poach firm 2's top talent" and "Match firm 1's hostile price–cut, if any", respectively. Hence, firm 1 ends up poaching firm 2's top talent *but* does not execute hostile price–cuts ensuring that they equal market shares. Our observations, in the Introduction, are therefore confirmed. ■

Now, the above game admits a unique consideration equilibrium. Further, this equilibrium is arguably the *obvious* prediction one would make about reasonable play in such a situation. However, this is the first of a series of examples in FISHBURN (1971) to illustrate that Nash equilibrium may *not* exist in a game if the Archimedean property ceases to hold. Thus, Fishburn made the following remark.[9]

> "However, due to the lack of an equilibrium point, we can still find ourselves going in circles, as in pure strategy cycles of Archimedian zero–sum games with no pure–strategy equilibrium."

However, a consideration equilibrium exists and must therefore be free from the problem of "going in circles". This is because, in any further consideration, those strategies of the previous context which could have resulted in "going in circles" are eliminated because present consideration is limited only to admissible strategies. Recall that the admissible strategies of a player are those which serve him the best if his opponent were to play a best response (property B). The basis for restriction to only admissible strategies is mutual conditional threats of playing best responses in case the opponent does not restrict himself. We now proceed to analyse the second example.

EXAMPLE 2: Consider the game between the financial institution and the other participants of the financial market. The market is either "Optimistic" (L) or "Pessimistic" (R) about asset A_1 thereby determining its valuation as high or low. The financial institution guesses this by playing "Buy A_1 and buy A_2" (T) or "Short sell A_1" (B), where A_2 is a valuable asset which can be acquired or not only by the financial institution. Also, profits or losses from trades in A_1 are valued before that of A_2 by both the parties.

The mixed strategies for players I and II, as indicated in Figure 7, are $p \cdot T \oplus (1-p) \cdot B$ and $q \cdot L \oplus (1-q) \cdot R$. The first component of the ordered pair in each cell specifies the value of player I's first Bernoullian for the corresponding outcome. Likewise, for the second components. Thus, $u_{\mathrm{I}}^1(p,q) = pq + (1-p)(1-q)$ and $u_{\mathrm{I}}^2(p,q) = p$ are the first and second expected utilities of player I defining a lexicographic expected utility representation, $(u_{\mathrm{I}}^1, u_{\mathrm{I}}^2)$, of player I's preference. Because the game is zero–sum, $u_{\mathrm{II}}^1(p,q) = -[pq + (1-p)(1-q)]$ and $u_{\mathrm{II}}^1(p,q) = -p$ are the two expected utilities of player II.

[9]The first sentence of the last paragraph of section 3 of FISHBURN (1971).

			Player II	
			q	$1-q$
			L	R
Player I	p	T	$(1,1)$	$(0,1)$
	$1-p$	B	$(0,0)$	$(1,0)$

FIGURE 7: Betting against the market.

With $C^*_{\mathrm{I},1} := \{p \in [0,1]\}$ and $C^*_{\mathrm{II},1} := \{q \in [0,1]\}$. defining the initial context $\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$, we proceed to show that player I's admissible set with respect to this context is $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle = \{p = 1/2\}$. For any $p \in [0,1]$, let $Q^*(p)$ be the set of best responses in $C^*_{\mathrm{II},1}$ of player II. Thus, $Q^*(p) \subseteq Q^*_1(p) := \mathrm{argmin}_{q\in[0,1]} u^1_{\mathrm{I}}(p,q)$. Noting that $u^1_{\mathrm{I}}(p,q) = 2(p-1/2)q + (1-p)$, we obtain:

$$Q^*_1(p) = \begin{cases} 0 & \text{if } p > 1/2; \\ [0,1] & \text{if } p = 1/2; \\ 1 & \text{if } p < 1/2. \end{cases}$$

For every $p \in [0,1]$, evaluating $u^1_{\mathrm{I}}(p,q)$ for any $q \in Q^*_1(p)$, we have:

$$\min_{q\in[0,1]} u^1_{\mathrm{I}}(p,q) = \begin{cases} 1-p & \text{if } p > 1/2; \\ 1/2 & \text{if } p = 1/2; \\ p & \text{if } p < 1/2. \end{cases}$$

Since $1-p < 1/2$, we have $p_* := 1/2$ as the unique element in $\{p \in [0,1]\}$ such that, for every $p \neq p_*$, there exists $q \in Q^*(p) \subseteq C^*_{II,1}$ that satisfies:

$$u^1_{\mathrm{I}}(p_*, q') > u^1_{\mathrm{I}}(p,q) \text{ for all } q' \in C^*_{\mathrm{II},1}.$$

Since $(u^1_{\mathrm{I}}, u^2_{\mathrm{II}})$ is a lexicographic expected utility representation of player I's preference, the singleton $\{p = 1/2\}$ satisfies property B (or, G). By definition 1, it follows that $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle = \{p = 1/2\}$.

To compute player II's admissible set $A^*_{\mathrm{II}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$, we begin with the following observation. Define $p' := 1-p$ and $q'q$. Then, $u^1_{\mathrm{II}}(p,q) = p'q' + (1-p')(1-q') - 1$. That is, $u^1_{\mathrm{II}}(p,q) = u^1_{\mathrm{I}}(p',q') - 1$. Also, $u^1_{\mathrm{I}}(p',q') = u^1_{\mathrm{I}}(q',p')$. Thus, $u^1_{\mathrm{II}}(p,q) = u^1_{\mathrm{I}}(q',p') - 1$. Hence, by the previous argument, we have: $A^*_{\mathrm{II}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle = \{q = 1/2\}$. Note, in the analysis thus far, no appeal has been made to u^2_{I} or u^2_{II}.

Finally, as each of the two admissible sets is a singleton, further iterations as required by definition 2 lead to no updation of these sets. Thus, $A_{\mathrm{I}}^{\infty} = \{p = 1/2\}$ and $A_{\mathrm{II}}^{\infty} = \{q = 1/2\}$. Therefore, the pair $(\frac{1}{2}T \oplus \frac{1}{2}B, \frac{1}{2}L \oplus \frac{1}{2}R)$ is the *unique* consideration equilibrium. Observe, this analysis did *not* depend on the specification of u_{I}^2 or u_{II}^2. Note, our suspicions in the Introduction are indeed confirmed. ■

The reader may have noted that the game defined only by the first components of the ordered pairs of the cells in Figure 7 is the standard "chicken game". It is well–known that $(\frac{1}{2}T \oplus \frac{1}{2}B, (\frac{1}{2}L \oplus \frac{1}{2}R)$ is the unique Nash equilibrium (or, Minimax strategy tuple) of the chicken game. It seems *plausible* that if further levels in players' lexicographic expected utilities do not feature into the analysis of consideration equilibria of a game, then the consideration equilibria should coincide with Nash equilibria.[10] Such is indeed the case. Lastly, Example 2 does *not* admit any Nash equilibria as shown in FISHBURN (1971).

EXAMPLE 3: We now revisit the game describing the bilateral conflict of nations 1 and 2. Recall, 1 cares in decreasing order of priority about (1) having nuclear technologies, (2) surrounding 2 with its allies, and (3) making international collaborations. The situation is zero–sum and thus lexicographic payoffs of 2 stand specified the moment the same are specified for 1. Figure 8 is resulting matrix game indicating the payoff triples of 1 for every play of pure strategy tuples.

The interpretation of the matrix game, for the game illustrated in Figure 8, is the same as was in Examples 1 and 2 except for two differences. First, player II now has the pure strategy M available in addition to L and R. Second, each player has three expected utilities representing lexicographically his preference. Thus, the mixed strategies for players I and II, as illustrated in Figure 8, are $p \cdot T \oplus (1-p) \cdot B$ and $q \cdot L \oplus r \cdot M \oplus (1 - [q + r]) \cdot R$, respectively, where $p \in [0, 1]$ and the pair $(q, r) \in [0, 1]^2$ satisfies $q + r \leq 1$. Hence, the three expected utilities for player I, for this mixed strategy pair, are $u_{\mathrm{I}}^1(p\,; q, r) = pq$, $u_{\mathrm{I}}^2(p\,; q, r) = (1-p)r + [1-(q+r)]$ and $u_{\mathrm{I}}^3(p\,; q, r) = (1-p)r + p[1-(q+r)]$. As the game is zero–sum, the three expected utilities for player II are $u_{\mathrm{II}}^1(p\,; q, r) = -u_{\mathrm{I}}^1(p\,; q, r)$, $u_{\mathrm{II}}^2(p\,; q, r) = -u_{\mathrm{I}}^2(p\,; q, r)$ and $u_{\mathrm{II}}^3(p\,; q, r) = -u_{\mathrm{I}}^3(p\,; q, r)$. Hence, $(u_{\mathrm{I}}^1, u_{\mathrm{I}}^2, u_{\mathrm{I}}^3)$ and $(u_{\mathrm{II}}^1, u_{\mathrm{II}}^2, u_{\mathrm{II}}^3)$ are a lexicographic expected utility representations of the preferences of players I and II, respectively. Now, we proceed to analyse this game.

[10] This is not the same as Proposition 3. The point of that proposition is that consideration equilibria coincide with Nash equilibria if players' preference satisfy Independence *and Continuity*. Here, we are arguing that despite having discontinuous preferences, if players' higher expected utility levels do not feature in the analysis of the consideration equilibria, then the proposed solution concept should reduce to the classical solution concept

		Player II		
		q	r	$1-(q+r)$
		L	M	R
Player I	p T	$(1,0,0)$	$(0,0,0)$	$(0,1,1)$
	$1-p$ B	$(0,0,0)$	$(0,1,1)$	$(0,1,0)$

FIGURE 8: Bilateral conflict.

With $C^*_{\mathrm{I},1} := \{p \in [0,1]\}$ and $C^*_{\mathrm{II},1} := \{q \in [0,1]\}$, the initial context is $\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$. To compute players' admissible sets with respect to the this context, observe the following. If $p = 0$, then $u^1_{\mathrm{I}}(p\,;q,r) = 0$ for all $(q,r) \in [0,1]^2$ such that $q + r \leq 1$. Also, if $p > 0$, then minimization of the u^1_{I} expected utility implies $q = 0$ resulting in the u^1_{I} expected utility to be 0. Hence, for any $p > 0$, we have $u^2_{\mathrm{I}}(p\,;q,r) = 1 - pr$ and $u^3_{\mathrm{I}}(p\,;q,r) = r + p(1-2r)$ by the lexicographic process as q must be 0. Likewise, for $p = 0$, we have $u^2_{\mathrm{I}}(p\,;q,r) = 1 - q$ and $u^3_{\mathrm{I}}(p\,;q,r) = r$ for every $(q,r) \in [0,1]^2$ such that $q + r \leq 1$.

Suppose, there exists $0 < p_0 < 1$ which does not belong to player I's admissible set $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$. Then, we argue: $p > p_0$ implies $p \notin A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$. Note, the U^2_{I} expected utility is minimized at $r = 1$ for both p and p_0 thereby resulting in expected utilities $1-p$ and $1-p_0$, respectively. As $p > p_0$ implies $1 - p < 1 - p_0$, if $p \in A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$, then we shall have a contradiction to the fact that admissible sets must satisfy property B. Hence, $p > p_0$ implies $p \notin A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$.

Since admissible sets must be non–empty, convex and compact by Theorem 1, it follows from the last conclusion: there exists a unique $0 \leq p_* < 1$ such that $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle = \{0 \leq p \leq p_*\}$. Now, the minimum u^2_{I} expected utility is 0 for the strategy $p = 0$ as is enforced by $q = 1$ which as argued in the previous paragraph can be considered for the case "$p = 0$" as per the lexicographic procedure. Thus, for any $p > p_*$, the minimum u^2_{I} expected utility, which is $1-p$, is strictly greater. This contradicts the fact that the admissible set must satisfy property B. Thus, our supposition that some $0 < p_0 < 1$ exists which does not belong to player I's admissible set must be wrong. Hence, the admissible set must include $\{0 < p < 1\}$. As admissible sets are compact, we have: $A^*_{\mathrm{I}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle = \{p \in [0,1]\}$.

We now compute player II's admissible set. Observe, for any $(q,r) \in A^*_{\mathrm{II}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$, it must be that $q = 0$ because the minimum u^1_{II} expected utility is 0 if $q = 0$, as enforced by any $p \in [0,1]$, in comparison to the minimum u^1_{II} expected utility of $-q$ if $q > 0$ enforced by $p = 1$.

Having concluded that $q = 0$, recall that the u^1_{II} expected utility is 0 for all $p \in [0,1]$ and $r \in [0,1]$. Also, $u^2_{\mathrm{II}}(p\,;q,r) = pr - 1$ and $u^3_{\mathrm{II}}(p\,;q,r) = 2p(r-1/2)-r$ for all $(p,r) \in [0,1]^2$ when $q = 0$. Then, the minimum u^2 expected utility is -1 for any $r \in [0,1]$ which is enforced by every $p \in [0,1]$ if $r = 0$ and by $p = 0$ if $r > 0$. Hence, $u^3_{\mathrm{II}}(p\,;q,r) = -p$ if $(q,r) = (0,0)$ and $u^3_{\mathrm{II}}(p\,;q,r) = -r$ if $(q,r) \in \{0\} \times (0,1]$.

Suppose, $0 < r_0 < 1$ is such that $(q,r) = (0,r_0) \notin A^*_{\mathrm{II}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$. Then, $r > r_0$ implies that the pair $(0,r)$ is not in $A^*_{\mathrm{II}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$. For otherwise, $u^3_{\mathrm{II}}(p\,;q,r) = -r < r_0 = u^3_{\mathrm{II}}(p\,;q,r_0)$ which would contradict the fact that admissible sets satisfy property B. Since an admissible set is also non–empty, compact and convex, it follows that $0 \le r_* < 1$ exists such that $A^*_{\mathrm{II}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle = \{(q,r) : q = 0\,; 0 \le r \le r_*\}$. However, the minimum u^3_{II} expected utility for $r = 0$ is -1 enforced by $p = 1$ and the minimum u^3_{II} expected utility for $r = 1$ is clearly -1. That is, the pair $(q = 0, r = 0)$ is in the admissible set and it is a strategy of player II which together with the strategy $p = 0$ of player I is an outcome which is indifferent, according to player II, to the outcome consituting the strategy $(q = 0, r = 1)$ by player II and the strategy $p = 1$ by I. This contradicts the fact that the admissible set satisfies property B. Hence, our supposition must be wrong. Thus, the set $\{(q,r) : q = 0\,; 0 < r < 1\} \subseteq A^*_{\mathrm{II}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle$. Since an admissible set is compact, we have: $A^*_{\mathrm{II}}\langle C^*_{\mathrm{I},1}, C^*_{\mathrm{II},1}\rangle = \{(q,r) : q = 0\,; 0 \le r \le 1\}$.

The new context is $\langle C^*_{\mathrm{I},2}, C^*_{\mathrm{II},2}\rangle$ where $C^*_{\mathrm{I},2} := \{p \in [0,1]\}$ and $C^*_{\mathrm{II},2} := \{(q,r) : q = 0\,; r \in [0,1]\}$. Thus, the game reduces to that in Figure 9.

			Player II	
			r	$1-r$
			M	R
Player I	p	T	$(0,0)$	$(1,1)$
	$1-p$	B	$(1,1)$	$(1,0)$

FIGURE 9: Bilateral conflict — the reduced game.

In this reduced game, player I's first and second expected utilities are $w^1_{\mathrm{I}}(p,r) = (1-p)r + (1-r) = 1 - pr$ and $w^2_{\mathrm{I}}(p,r) = (1-p)r + p(1-r)$, respectively. Since the game is zero–sum, player II's first and second expected utilities can be taken as $w^1_{\mathrm{II}}(p,r) = -w^1_{\mathrm{II}}(p,r)$ and $w^2_{\mathrm{II}}(p,r) = -w^2_{\mathrm{II}}(p,r)$, respectively. Thus, preferences of players I and II admit $(w^1_{\mathrm{I}}(p,r), w^2_{\mathrm{I}}(p,r))$ and $(w^1_{\mathrm{II}}(p,r), w^2_{\mathrm{II}}(p,r))$ as lexicographic expected utility representations, respectively.

If $p > 0$, the minimum w_{I}^1 expected utility of player I is $1 - p$ which is enforced by $r = 1$. However, the minimum w_{I}^1 expected utility of player I is 1 if $p = 0$ which is enforced by any $r \in [0, 1]$. Therefore, the singleton $\{p = 0\}$ satisfies property B with respect to the present context. Hence, $A_{\mathrm{I}}^* \langle C_{\mathrm{I},2}^*, C_{\mathrm{II},2}^* \rangle = \{p = 0\}$. Because this set is already a singleton, there shall be no updation in further iterations as demanded by definition 2. Hence, we conclude: $A_{\mathrm{I}}^\infty = \{p = 0\}$.

Next, $w_{\mathrm{II}}^1(p, r) = pr - 1$ and $w_{\mathrm{II}}^2(p, r) = 2p(r - 1/2) - r$ where $(p, r) \in [0, 1]^2$. Thus, by an argument identical to that in paragraphs 0 and 0, we have: $A_{\mathrm{II}}^* \langle C_{\mathrm{I},2}^*, C_{\mathrm{II},2}^* \rangle = \{r \in [0, 1]\}$. That is, there is no updation of player II's admissible set in this iteration. We proceed to the next iteration as follows.

Now, the context is $\langle C_{\mathrm{I},3}^*, C_{\mathrm{II},3}^* \rangle$ where $C_{\mathrm{I},3}^* \coloneqq \{p = 0\}$ and $C_{\mathrm{II},3}^* \coloneqq \{r \in [0, 1]\}$. It only remains to compute player II's admissible set with respect to this context. With $p = 0$, we have $w_{\mathrm{II}}^1(p, r) = -1$ and $w_{\mathrm{II}}^2(p, r) = -r$ for all $r \in [0, 1]$. Thus, player II's admissible set with respect to this context is $\{r = 0\}$ which is a singleton. Hence, we have: $A_{\mathrm{II}}^\infty = \{r = 0\}$. Since we had already concluded that $A_{\mathrm{I}}^\infty = \{p = 0\}$, it follows that the game has a *unique* consideration equilibrium which is the strategy tuple (B, R).

To interpret the final prediction, which is (B, R), we recall that the pure strategies B and R were defined as logical combinations of clauses $S_{\mathrm{I},1}$ to $S_{\mathrm{I},4}$ and $S_{\mathrm{II},1}$ to $S_{\mathrm{II},5}$, respectively. For convenience, we reproduce the descriptions of B and R as follows:

$$\begin{aligned} B &\coloneqq S_{\mathrm{I},2} \text{ and } S_{\mathrm{I},4}. \\ R &\coloneqq S_{\mathrm{II},5} \text{ and (if } S_{\mathrm{I},1} \text{ then } S_{\mathrm{II},2}). \end{aligned}$$

Further, the involved clauses are as follows:

$$\begin{aligned} S_{\mathrm{I},1} &\coloneqq \text{“Attempt to develop nuclear technologies”}. \\ S_{\mathrm{I},2} &\coloneqq \text{“Attempt to form allies that surround 2”}. \\ S_{\mathrm{I},4} &\coloneqq \text{“Make international collaborations”}. \\ S_{\mathrm{II},2} &\coloneqq \text{“Enforce sanctions on 1”}. \\ S_{\mathrm{II},5} &\coloneqq \text{“Make international collaborations”}. \end{aligned}$$

Thus, in the unique consideration equilibrium, nation 1 does not end up developing nuclear technologies but its allies surround nation 2. Moreover, both nations do form international collaborations. ■

As Fishburn (1971) shows, this game admits no Nash equilibrium. Note that these are *all* the examples in that article. Moreover, observe that lexicographic expected utilities *have* testable implications.

6. COMPARATIVE STATICS

In this section, we formalize and establish the following claim: consideration equilibrium makes "sharper predictions" than Nash equilibrium in the game obtained if agents' preferences are the "finest continuous coarsening" of their respective original preferences. In this statement, we have introduced the term "finest continuous coarsening" of a given preference which intuitively is *the* continuous preference which "best approximates" the given preference. Recall that S is the set $S_1 \times S_2$ of all pure strategy tuples in the two–person game. All preferences are defined over $\Delta(S)$. We begin with the following definition.

DEFINITION 3: *The preference* $\succsim^{**}$ refines *the preference* $\succsim^{*}$ *if,*

$$p \succ^{*} q \implies p \succ^{**} q.$$

For instance, consider $\succsim^{*}$ and $\succsim^{**}$ defined as follows. Fix $K \in \mathbb{N}$ and let $U_k : \Delta(S) \to \mathbb{R}$ be an expected utility for each $k \in \{1, \ldots, K\}$. Let $\succsim^{**}$ be the preference defined by:

$$p \succsim^{**} q \iff [U_1(p), \ldots, U_K(p)] \geq_L [U_1(q), \ldots, U_K(q)],$$

where $\geq_L$ is the lexicographic order over $\mathbb{R}^K$. Also, let $\succsim^{*}$ be defined as: $p \succsim^{*} q \iff U_1(p) \geq U_1(q)$. Then, the definition of $\geq_L$ implies $p \succ^{*} q \implies p \succ^{**} q$; that is, $\succsim^{**}$ refines $\succsim^{*}$. Observe, "refines" is a transitive binary relation over the class of all preferences on $\Delta(S)$.

DEFINITION 4: *Let* $\mathscr{P}$ *be any class of preferences over* $\Delta(S)$*. Then,* $\succsim^{**}$ *is the* finest *in* $\mathscr{P}$ *if* $\succsim^{**}$ *is in* $\mathscr{P}$ *and refines* $\succsim^{*}$ *for all* $\succsim^{*} \in \mathscr{P}$.

Having defined the term "refines", we say "$\succsim^{**}$ is *finer* than $\succsim^{*}$" or "$\succsim^{*}$ is *coarser* than $\succsim^{**}$" if $\succsim^{**}$ refines $\succsim^{*}$. Assume that Γ^{*} and Γ^{**} are the two–person zero–sum games $\langle N, (S_i)_{i \in N}, (\succsim_i^{*})_{i \in N} \rangle$ and $\Gamma^{**} := \langle N, (S_i)_{i \in N}, (\succsim_i^{**})_{i \in N} \rangle$, respectively. Observe the following.

PROPOSITION 4: $\succsim_1^{**}$ *refines* $\succsim_1^{*}$*, if and only if,* $\succsim_2^{**}$ *refines* $\succsim_2^{*}$.

PROOF: Assume that $\succsim_1^{**}$ refines $\succsim_1^{*}$. Let $p, q \in \Delta(S)$ be such that $p \succ_2^{*} q$. Since Γ^{*} is a zero–sum game, it follows that $q \succ_1^{*} p$. As $\succ_1^{**}$ refines $\succ_1^{*}$, we have $q \succ_1^{**} p$. Since Γ^{**} is a zero–sum game, it follows from $q \succ_1^{**} p$ that $p \succ_2^{**} q$. Thus, we have: $p \succ_2^{*} q \implies p \succ_2^{**} q$. That is, $\succsim_2^{**}$ refines $\succsim_2^{*}$. Therefore, we have shown: if $\succsim_1^{**}$ refines $\succsim_1^{*}$, then $\succsim_2^{**}$ refines $\succsim_2^{*}$. The converse follows by a symmetric argument. ■

The proposition justifies the use of the phrase "Γ^{**} refines Γ^{*}" to stand for the phrase "Γ^{*} and Γ^{**} are games where $\succsim_i^{**}$ refines $\succsim_i^{*}$ for some player i". Thus, we shall say "Γ^{**} is *finer* than Γ^{*}" or "Γ^{*} is *coarser* than Γ^{**}" if Γ^{**} refines Γ^{*}. Now, we are ready to state the basic comparative static result which is as follows.

THEOREM 4: *Let Γ^{*} and Γ^{**} be two–person zero–sum games and $\langle C_1, C_2 \rangle$ be any context. Suppose Γ^{**} refines Γ^{*}. Then, for each $i \in N$:*

$$A^{*}_{i,\Gamma^{**}}\langle C_1, C_2\rangle \subseteq A^{*}_{i,\Gamma^{*}}\langle C_1, C_2\rangle,$$

where $A^{*}_{i,\Gamma^{*}}\langle C_1, C_2\rangle$ and $A^{*}_{i,\Gamma^{**}}\langle C_1, C_2\rangle$ are admissible sets of player i in games Γ^{*} and Γ^{**}, respectively.

PROOF: Let $i \in N$ and consider an arbitrary $A_i \subseteq C_i$ such that A_i satisfies property G when the preferences of players 1 and 2 are $\succsim_1^{*}$ and $\succsim_2^{*}$, respectively. That is, fixing an arbitrary $x_i \in C_i$, there exists a $x_j \in C_j \setminus A_i$ such that: $(x_i', x_j') \succ_i^{*} (x_i, x_j)$ for all $(x_i', x_j') \in A_i \times C_j$. Since Γ^{**} refines Γ^{*}, it must be that $\succsim_i^{**}$ refines $\succsim_i^{*}$. Then, by definition 3, we have: $(x_i', x_j') \succ_i^{*} (x_i, x_j)$ implies $(x_i', x_j') \succ_i^{**} (x_i, x_j)$. Thus,

$$(x_i', x_j') \succ_i^{**} (x_i, x_j) \text{ for all } (x_i', x_j') \in A_i \times c_j.$$

That is, A_i satisfies property G with respect to the context $\langle C_1, C_2\rangle$ where the preferences of players 1 and 2 are $\succsim_1^{**}$ and $\succsim_2^{**}$, respectively. Hence, we have the following:

$$\mathscr{A}^{G}_{i,\Gamma^{*}}\langle C_1, C_2\rangle \subseteq \mathscr{A}^{G}_{i,\Gamma^{**}}\langle C_1, C_2\rangle,$$

where $\mathscr{A}^{G}_{i,\Gamma^{*}}\langle C_1, C_2\rangle$ and $\mathscr{A}^{G}_{i,\Gamma^{**}}\langle C_1, C_2\rangle$ are the classes of sets satisfying property G holds with respect to $\langle C_1, C_2\rangle$ corresponding to player i when his preferences are $\succsim_i^{*}$ and $\succsim_i^{**}$, respectively. Now, the definition of admissible sets implies the following:

$$A_{i,\Gamma^{*}} = \bigcap \left\{A_i \in \mathscr{A}^{G}_{i,\Gamma^{*}}\langle C_1, C_2\rangle\right\}, \text{ and}$$
$$A_{i,\Gamma^{**}} = \bigcap \left\{A_i \in \mathscr{A}^{G}_{i,\Gamma^{**}}\langle C_1, C_2\rangle\right\}.$$

Hence, the last set–inclusion implies: $A^{*}_{i,\Gamma^{**}}\langle C_1, C_2\rangle \subseteq A^{*}_{i,\Gamma^{*}}\langle C_1, C_2\rangle$. ■

Theorem 4 says the following: the finer are the players' preferences, the smaller[11] are their admissible sets with respect to any context.

[11] This is in terms of set–inclusion. That is, a set U is "smaller than" another set V iff $U \subseteq V$.

We return again to our discussion of an arbitrary preference $\succsim^{**}$ which refines another preference $\succsim^{*}$. Recall our terminology allows us to state this equivalently as $\succsim^{*}$ is coarser than $\succsim^{**}$. If in addition $\succsim^{*}$ is continuous[12], then $\succsim^{*}$ is a *continuous coarsening* of $\succsim^{**}$. For instance, consider the example following definition 3 where $\succsim^{**}$ is the preference which admits the lexicographic expected utility representation via the K–tuple of expected utilities $U_1, \ldots, U_K$, and $\succsim^{*}$ is the preference which admits U_1 as its expected utility representation. Since preferences that admit expected utility representations must be continuous by the the theorem due to von Neumann & Morgenstern (1944), it follows that $\succsim^{*}$ is a continuous coarsening of $\succsim^{**}$.

Now, for the arbitrary given preference $\succsim^{**}$, the preference which declares any two alternatives to be indifferent is trivially a continuous coarsening. Therefore, we want to formulate the notion of the "finest" among all continuous coarsenings of $\succsim^{**}$. Let $\mathscr{C}_{\succsim}$ be the class of all continuous coarsenings of any preference $\succsim$ which is non–empty as it contains the trivial preference. Also, recall the term "finest" from definition 4. We now claim that there exists a unique "finest continuous coarsening" of $\succsim$. Define $\succ^c$ and $\sim^c$ over $\Delta(S)$ as follows:

$$\succ^c := \bigcup \{ \succ^*: \succsim^* \in \mathscr{C}_{\succsim} \} \quad \text{and} \quad \sim^c := \bigcap \{ \sim^*: \succsim^* \in \mathscr{C}_{\succsim} \}.$$

Also, define $\succsim^c := \succ^c \bigcup \sim^c$. The key result is as follows.

Theorem 5: *$\succsim^c$ is the unique finest continuous coarsening of $\succsim$.*

The proof is supplied in subsection $A.1$ of the Appendix.[13] Consider the example of lexicographic expected utility preferences.

Corollary 1: *Suppose that $\succsim^{**}$ admits a lexicographic expected utility representation through $U_1, \ldots, U_K$ and assume U_1 is non–trivial. If $\succsim^{*}$ is the preference defined to have the expected utility U_1 as one of its representations, then $\succsim^{*}$ is the finest continuous coarsening of $\succsim^{**}$.*

The proof is in subsection $A.3$ of the Appendix but the intuition is as follows. The closure of any weak upper (lower) contour set of the preference $\succsim^{**}$ is a closed halfspace which is precisely the corresponding weak upper (lower) contour set of the preference $\succsim^{*}$.

[12] For any binary relation $\succsim$ over $\Delta(S)$, we shall follow the standard practice of denoting by $\succ$ and $\sim$ the strict and indifference components, respectively, of $\succsim$. Formally, their definitions are as follows: (1) $p \succ q$ iff ($p \succsim q$; not $q \succsim p$), and (2) $p \sim q$ iff ($p \succsim q$; $q \succsim p$). A preference $\succsim$ is *continuous* if, $p \succ q$ implies that there exists $\varepsilon > 0$ such that $p' \succ q'$ for every $p' \in B(p, \varepsilon)$ and $q \in B(q, \varepsilon)$. Here, $B(p, \varepsilon)$ is the open ball in $\Delta(S)$ of radius ε centered at p.

[13] However, we believe that Theorem 5 and its proof are of independent interest.

Notwithstanding the discussion of the above example, Theorem 5 is applicable for any general preference $\succsim$ to begin with. In particular, $\succsim$ may *not* satisfy Independence. The precise consequence of additionally assuming Independence of $\succsim$ is captured by the following.

PROPOSITION 5: *If the preference $\succsim$ satisfies Independence, then its finest continuous coarsening $\succsim^c$ also satisfies Independence.*

PROOF: Since $\succsim$ satisfies Independence,[14] for some $K \in \mathbb{N}$ there exists expected utilities $U_k : \Delta(S) \to \mathbb{R}$ for all $k \in \{1, \ldots, K\}$ such that:

$$p \succsim q \quad \text{iff} \quad [U_1(p), \ldots, U_K(p)] \geq_L [U_1(q), \ldots, U_K(q)].$$

Assume, without loss of generality, U_1 is not trivial. Thus, $\succsim^c$ admits U_1 as an expected utility representation. Therefore, $\succsim^c$ satisfies the Independence axiom. ■

Theorem 5 and Proposition 5 allow us to naturally talk of the "finest continuous coarsening" of any given two–person zero–sum game Γ in which player i's preference is $\succsim_i$ that satisfies Independence. Thus, the game Γ^c is the *finest continuous coarsening* of Γ if, each i's preference is $\succsim_i^c$ instead of $\succsim_i$. Then, we have the following.

PROPOSITION 6: *Suppose Γ is a two–person zero–sum game and let Γ^c be its finest continuous coarsening. Then, the set of consideration equilibria of Γ is a subset of the set of minimax strategies of Γ^c.*

PROOF: The set $A^{\infty}_{1,\Gamma} \times A^{\infty}_{2,\Gamma}$ of all consideration equilibria of Γ is a subset of $A^*_{1,\Gamma} \times A^*_{2,\Gamma}$ where $A^*_{i,\Gamma}$ is the admissible set of player i in the game Γ with respect to the context $\langle \Delta(S_1), \Delta(S_2) \rangle$. Further, the game Γ is finer than its finest continuous coarsening Γ^c. Let $M_{1,\Gamma^c} \times M_{2,\Gamma^c}$ be the set of all minimax strategy pairs of the game Γ^c. Further, let A^*_{i,Γ^c} be player i's admissible set in the game Γ^c with respect to the context $\langle \Delta(S_1), \Delta(S_2) \rangle$. Then, Proposition 3 implies that $A^*_{1,\Gamma^c} \times A^*_{2,\Gamma^c}$ since players' preferences in Γ^c satisfy Independence *and* Continuity. Moreover, Theorem 4 implies $A^*_{1,\Gamma} \times A^*_{2,\Gamma} \subseteq A^*_{1,\Gamma^c} \times A^*_{2,\Gamma^c}$. Thus, $A^*_{1,\Gamma} \times A^*_{2,\Gamma} \subseteq M_{1,\Gamma^c} \times M_{2,\Gamma^c}$ which completes the proof. ■

Thus, we have formalized the claim: consideration equilibria make finer predictions than Nash equilibria when players' preferences are the finest continuous coarsening of their respective original preferences.

[14]The statement that "a preference admits a lexicographic expected utility (LEU) representation, if and only if, it satisfies the Independence axiom" is provided in the next section.

7. COMPUTING EQUILIBRIA

The objective of this section is to characterize admissible sets with a view towards computation. We begin with a representation theorem for preferences which are assumed to satisfy only *our* Independence axiom. HAUSNER (1954) characterized the existence of lexicographic expected utility representations using the original Independence axiom; also see BLUME ET AL. (1989). Our axiom is weaker and the stronger characterization is in CHATTERJEE (2022).

To state this theorem, we introduce some concepts. Let Z be a finite non–empty set whose elements are the *basic prizes*. A *lottery* over Z is any map $p : Z \to [0,1]$ with $\sum_{z \in Z} p(z) = 1$. Let $\Delta(Z)$ be the set of all lotteries. Any map $U : \Delta(Z) \to \mathbb{R}$ is an *expected utility* (EU) if,[15] $U(p) = \sum_{z \in Z} p(z)U(z)$ for all $p \in \Delta(Z)$. If $\succsim$ is a preference over $\Delta(Z)$, then the list of some $K \in \mathbb{N}$ expected utilities $\langle U_k : k = 1, \ldots, K\rangle$ is a *lexicographic expected utility (LEU) representation* of $\succsim$ if:

$$p \succsim q \iff [U_1(p), \ldots, U_K(p)] \geq_L [U_1(q), \ldots, U_K(q)]$$

where $\geq_L$ is the lexicographic order over $\mathbb{R}^K$. Then, Hausner's theorem as adapted to this setting[16] can be stated as follows.

THEOREM (Existence of LEU Representations): *A preference satisfies Independence, if and only if, it admits an LEU representation.*

To apply the above theorem to our setting, we recall that basic prizes are all pure strategy tuples which constitute the set $S_1 \times S_2$. Since players' preferences $\succsim_1$ and $\succsim_2$ over $\Delta(S_1 \times S_2)$ are assumed to satisfy Independence, we obtain $K \in \mathbb{N}$ and an LEU representation $\langle U_{i,k} : k = 1, \ldots, K\rangle$ of $\succsim_i$ for each player $i \in N$ such that:

$$U_{2,k} = -U_{1,k} \quad \text{for every } k \in \{1, \ldots, K\}.$$

Note, the same K is used for each player as must be because the game is zero–sum. Further, the requirement that $U_{2,k} = -U_{1,k}$ for each $k \in \{1, \ldots, K\}$ is based on the fact that the game is zero–sum *and* because of the observation that any positive affine transformation of an expected utility is also an expected utility representing the same preference. The characterization of player i's admissible set $A_i^*\langle C_1, C_2\rangle$ with respect to any context $\langle C_1, C_2\rangle$ shall be casted in terms of the LEU representations of players' preferences.

[15]For the degenrate lottery $\delta_{z_*} \in \Delta(Z)$ with support $\{z_*\}$, we write $U(z_*)$ instead of $U(\delta_{z_*})$.

[16]HAUSNER (1954) considered abstract *mixture spaces*.

Before embarking on the characterization, we proceed to establish an "indifference property" of admissible sets. The insights are then generalized leading up to the desired characterization. Consider an arbitrary context $\langle C_1, C_2 \rangle$ such that C_1 and C_2 are convex. Also, let $k_* \in \{1, \ldots, K\}$ be the unique smallest element such that U_{i,k_*} is *not* a constant map over $C_1 \times C_2$ for some player i. With $A_i^*\langle C_1, C_2 \rangle$ as the admissible set of player i, consider the following definitions:

$$v^i := \max_{x_i \in C_i} \min_{x_j \in C_j} U_{i,k_*}(x_i, x_j), \tag{4}$$

$$v_*^i := \max_{x_i \in A_i^*\langle C_1, C_2\rangle} \min_{x_j \in C_j} U_{i,k_*}(x_i, x_j), \tag{5}$$

$$B^i := \big\{x_i \in C_i : U_{i,k_*}(x_i, x_j) \geq v^i \text{ for all } x_j \in C_j\big\}. \tag{6}$$

With the above definitions in place, the basic result is as follows.

PROPOSITION 7: *For each player i, $v^i = v_*^i$, $A_i^*\langle C_1, C_2\rangle \subseteq B^i$ and U_{i,k_*} is constant over $A_1^*\langle C_1, C_2\rangle \times A_2^*\langle C_1, C_2\rangle$.*

PROOF: Note that both v^i and v_*^i are well–defined real numbers. This rests on two observations. First, each of the two sets $A_i^*\langle C_1, C_2\rangle$ and C_i is non–empty and compact. Second, the map:

$$x_i \in \Delta(S_i) \mapsto \min_{x_j \in C_j} U_{i,k_*}(x_i, x_j)$$

is continuous. This follows from Berge's Theorem of Maximum.[17] To see why, note (a) the map $(x_i, x_j) \in \Delta(S_i) \times \Delta(S_j) \mapsto U_{i,k_*}(x_i, x_j)$ is continuous, and (b) the constant map $x_i \in \Delta(S_i) \mapsto C_j$ is a compact–valued and continuous correspondence.

Observe, $(x_i, x_j) \succ_i (x_i', x_j')$ if $U_{i,k_*}(x_i, x_j) > U_{i,k_*}(x_i', x_j')$ by the definition of k_*. Since $A_i^*\langle C_1, C_2\rangle \subseteq C_i$, it follows from (4) and (5) that $v^i \geq v_*^i$. We shall first argue: $v^i = v_*^i$. Suppose, $v^i > v_*^i$. By (4), there exists $x_i \in C_i \setminus A_i^*\langle C_1, C_2\rangle$ such that: $U_{i,k_*}(x_i, x_j) \geq v^i$ for every $x_j \in C_j$. Further, (5) implies that there exists $x_i' \in A_i^*\langle C_1, C_2\rangle$ and $x_j' \in C_j$ such that $U_{i,k_*}(x_i', x_j') = v_*^i$. Thus, $v^i > v_*^i$ implies:

$$(x_i, x_j) \succ_i (x_i', x_j') \text{ for every } x_j \in C_j.$$

Since $x_i \in C_i \setminus A_i^*\langle C_1, C_2\rangle$, the above conclusion is a contradiction to the fact that $A_i^*\langle C_1, C_2\rangle$ satisfies property G begin the admissible set with respect to the context $\langle C_1, C_2\rangle$. Thus, we have: $v^i = v_*^i$.

[17]See, for instance, Proposition $A4.7$ on page 476 of KREPS [2013].

Now, we shall show that $A_i^*\langle C_1, C_2\rangle \subseteq B^i$. Since $A^*\langle C_1, C_2\rangle$ is the smallest non–empty compact set that satisfies property G, it will be enough to argue: B^i is a non–empty and compact set that satisfies property G. From (6), observe that $B^i = \bigcap\{B^i(x_j) : x_j \in C_j\}$ where $B^i(x_j) := \{x_i \in C_i : U_{i,k_*}(x_i, x_j) \geq v^i\}$. By continuity of the map $x_i \in C_i \mapsto U_{i,k_*}(x_i, x_j)$, the set $B^i(x_j)$ is a closed subset of C_j. Thus, the compactness of C_j implies: B^i is compact. The non–emptiness of B^i follows from (4) and the following observation:

$$B^i = \{x_i \in C_i : \min_{x_j \in C_j} U_{i,k_*}(x_i, x_j) \geq v^i\}$$

We now argue: B^i satisfies property G. Fix an arbitrary $x_i \in C_i \backslash B^i$. From (6), it follows that there exists $x_j \in C_j$ such that $U_{i,k_*}(x_i, x_j) < v^i$. Further, consider an arbitrary $(x_i', x_j') \in B^i \times C_j$. Again, (6) implies that $U_{i,k_*}(x_i', x_j') \geq v^i$. That is, $U_{i,k_*}(x_i', x_j') > U_{i,k_*}(x_i, x_j)$. Hence, $(x_i', x_j') \succ_i (x_i, x_j)$ for all $(x_i', x_j') \in B^i \times C_j$. Thus, B^i satisfies property G. Therefore, $A_i^*\langle C_1, C_2\rangle \subseteq B^i$ holds.

Finally, we argue: U_{i,k_*} is constant over $A_1^*\langle C_1, C_2\rangle \times A_2^*\langle C_1, C_2\rangle$. We recall that $U_{2,k_*} = -U_{1,k_*}$. Since $C_1 \subseteq \mathbb{R}^{S_1}$ and $C_2 \subseteq \mathbb{R}^{S_2}$ are convex and compact, by the Minimax Theorem of von Neumann (1928):

$$\begin{aligned}\max_{x_2 \in C_2} \min_{x_1 \in C_1} U_{2,k_*}(x_1, x_2) &= \min_{x_1 \in C_1} \max_{x_2 \in C_2} U_{2,k_*}(x_1, x_2), \text{ and} \\ \min_{x_1 \in C_1} \max_{x_2 \in C_2} U_{2,k_*}(x_1, x_2) &= -\max_{x_1 \in C_1} \min_{x_2 \in C_2} U_{1,k_*}(x_1, x_2)\end{aligned}$$

where the latter follows trivially from $U_{2,k_*} = -U_{1,k_*}$. Combining the above with definitions of v^1 and v^2 as in (4), we obtain: $v^1 = -v^2$. Now, let $x_1 \in A_1^*\langle C_1, C_2\rangle$ and $x_2 \in A_2^*\langle C_1, C_2\rangle$ be arbitrary. Since $A_i^*\langle C_1, C_2\rangle \subseteq B^i$ for each i, it follows that $U_{1,k_*}(x_1, x_2) \geq v^1$ and $U_{2,k_*}(x_1, x_2) \geq v^2$. However, $U_{2,k_*} = -U_{1,k_*}$ implies $-v^2 \geq U_{1,k_*}(x_1, x_2)$. By $v^1 = -v^2$, $v^1 \geq U_{1,k_*}(x_1, x_2)$. Thus, $U_{1,k_*}(x_1, x_2) = v^1$. That is, U_{1,k_*} is constant over $A_1^*\langle C_1, C_2\rangle \times A_2^*\langle C_1, C_2\rangle$. A symmetric argument applies for U_{2,k_*}. This completes the proof. ■

To get some intuition, recall the definition of the admissible set. The key idea is to obtain the minimal set for a player such that if some strategy outside of that set is deployed then, for some play of the opponent, this player is strictly worse off than had he considered playing any strategy from within the set irrespective of what his opponent played. For an expected utility preference, this corresponds to von Neumann's value which is the best minimum guarantee to the player.

Proposition 7 asserts the contancy of the first non–trivial Bernoullian over the context in the resulting admissible set. However, to characterize the admissible set in question, it is necessary to "trim" further the resulting intermediate sets using the remaining Bernoullians in lexicographic expected utility representation. The description of these further "trimmings" is follows.

Without loss of generality, let $\langle C_1, C_2\rangle$ be any context that admits a unique smallest $k_* \in \{1, \ldots, K-1\}$ such that U_{i,k_*+1} is *not* a constant map over $C_1 \times C_2$ for some player i. We associate the list $\mathcal{M}\langle C_1, C_2\rangle$ consisting of pairs $\langle (B^j_k, v^j_k) : j \in N\rangle$ for each $0 \le k \le K - k_*$, where $B^j_k \subseteq C_j$ and $v^j_k \in \mathbb{R}$, which is iteratively defined as follows. Fix an arbitrary player i. Let $B^i_0 \coloneqq C_i$ and v^i_0 be constant value of the map U_{i,k_*} over $C_1 \times C_2$. Now, suppose that, for some $1 \le k \le K - k_*$, the pairs $\langle (B^j_l, v^j_l) : j \in N\rangle$ have already been defined for every $0 \le l < k$. Then, denote by $\mathbf{v}^i_k$ the list $\langle v^i_l : 0 \le l < k\rangle$ and define the set:

$$\Delta(x_i, \mathbf{v}^i_k) \coloneqq \big\{x_j \in C_j : U_{i,k_*+l}(x_i, x_j) = v^i_l \text{ for all } 0 \le l < k\big\} \tag{7}$$

for any $x_i \in B^i_{k-1}$. Then, define $v^i_k \in \mathbb{R}$ and $B^i_{k,\varepsilon} \subseteq B^i_{k-1}$ as follows:[18]

$$v^i_k \coloneqq \sup_{x_i \in B^i_{k-1}} \min_{x_j \in \Delta(x_i, \mathbf{v}^i_k)} U_{i,k_*+k}(x_i, x_j), \text{ and} \tag{8}$$

$$B^i_{k,\varepsilon} \coloneqq \big\{x_i \in B^i_{k-1} : \min_{x_j \in \Delta(x_i, \mathbf{v}^i_k)} U_{i,k_*+k}(x_i, x_j) \ge v^i_k - \varepsilon\big\}. \tag{9}$$

Also, define[19] $B^i_k \coloneqq \bigcap_{\varepsilon>0} \mathrm{cl}(B^i_{k,\varepsilon})$. Then, the following list:

$$\mathcal{M}\langle C_1, C_2\rangle = \langle (B^i_k, v^i_k) : 0 \le k \le K - k_* \,;\, i \in N\rangle$$

is unique, if it exists, with a nest $C_i = B^i_0 \supseteq B^i_1 \supseteq \ldots \supseteq B^i_{K-k_*}$ for each player i. We call $\mathcal{M}\langle C_1, C_2\rangle$ the *maxmin system* associated with the context $\langle C_1, C_2\rangle$. Admissible sets are characterized as follows.

Theorem 6: *Let $\langle C_1, C_2\rangle$ be a context with C_1 and C_2 convex. Then, the maxmin system $\mathcal{M}\langle C_1, C_2\rangle$ associated with $\langle C_1, C_2\rangle$ exists and is unique. Further, $B^i_{K-k_*} = A^*_i\langle C_1, C_2\rangle$ for each i.*

The proof of this result is technical and is, therefore, supplied in subsection $A.2$ of the Appendix. This concludes our presentation.

[18] "$U_{i,k}(x_i, x_j)$" stands for $U_{1,k}(x_1, x_2)$ or $U_{2,k}(x_1, x_2)$ according as (i, j) is $(1, 2)$ or $(2, 1)$.

[19] For any subset $A \subseteq \Delta(S_i)$, we shall indicate by $\mathrm{cl}(A)$ the closure of A relative to the topology on $\Delta(S_i)$ inherited from the standard topology of $\mathbb{R}^{S_i}$.

APPENDIX

A.1 *Proof of Theorem 5*

To make this subsection self–contained, we briefly recall the definitions and the claim. For any given preference $\succsim$ over $\Delta(S)$, let $\mathscr{C}_{\succsim}$ be the class of all continuous coarsenings of $\succsim$. Thus, a typical element of $\mathscr{C}_{\succsim}$ is any continuous preference $\succsim^*$ such that $\succsim$ *refines* $\succsim^*$; that is, $p \succ^* q \implies p \succ^* q$. Then, the binary relation $\succsim^c$ corresponding to $\succsim$ was defined as $\succsim^c := \succ^c \bigcup \sim^c$, where $\succ^c$ and $\sim^c$ were defined as:

$$\succ^c := \bigcup \{ \succ^* : \succsim^* \in \mathscr{C}_{\succsim} \} \quad \text{and} \quad \sim^c := \bigcap \{ \sim^* : \succsim^* \in \mathscr{C}_{\succsim} \}.$$

The result we prove here is Theorem 5 from section 6 restated as follows.

THEOREM 5: *$\succsim^c$ is the unique finest continuous coarsening of $\succsim$.*

We show that $\succsim^c$ is a preference which is continuous and is refined by $\succsim$. Further, we argue $\succsim^c$ refines every element in $\mathscr{C}_{\succsim}$. Finally, we prove that $\succsim^c$ is the unique such preference.

PROOF: The proof of Theorem 5 is organized via the following steps:

Step 1: We argue: the relations $\succ^c$ and $\sim^c$ are asymmetric and symmetric, respectively.[20] First, we show: $\succ^c$ is asymmetric. Assume $p \succ^c q$ holds. By definition of $\succ^c$, there exists $\succsim^* \in \mathscr{C}_{\succsim}$ such that $p \succ^* q$. By definition of $\mathscr{C}_{\succsim}$ and $\succsim^* \in \mathscr{C}_{\succsim}$, it follows that $\succsim$ refines $\succsim^*$. Thus, $p \succ^* q$ implies $p \succ q$. That is, $p \succ^c q$ implies $p \succ q$. Suppose $q \succ^c p$ holds. Then, we have $q \succ p$. However, this contradicts the asymmetry of $\succ$ because $\succsim$ is a preference. Hence, $p \succ^c q \implies$ not $q \succ^c p$. That is, $\succ^c$ is asymmetric. Second, we observe: $\sim^c$ is symmetric. This is because $\sim^c$ is the intersection of symmetric binary relations.

Step 2: We argue: $\succsim^c$ is complete. Suppose, $p, q \in \Delta(S)$ are such that neither $p \succsim^c q$ nor $q \succsim^c p$ hold. Thus, the definition of $\succsim^c$ implies none of $p \succ^c q$, $q \succ^c p$ or $p \sim^c q$ hold. Since $p \succ^c q$ does not hold, the definition of $\succ^c$ implies $p \succ^* q$ fails for all $\succsim^* \in \mathscr{C}_{\succsim}$. Also, since $p \sim^c q$ does not hold, the definition of $\sim^c$ implies $p \sim^* q$ for all $\succsim^* \in \mathscr{C}_{\succsim}$. Hence, $p \succsim^* q$ fails to hold for every $\succsim^* \in \mathscr{C}_{\succsim}$. However, each $\succsim^* \in \mathscr{C}_{\succsim}$ is a preference. Thus, failure of $p \succsim^* q$ implies $q \succ^* p$. Hence, the definition of $\succ^c$ requires $q \succ^c p$ which is a contradiction. Thus, $p \succsim^c q$ or $q \succsim^c p$ holds. That is, the relation $\succsim^c$ is complete.

[20]Formally, we wish to establish (1) ($p \succ^c q \implies$ not $q \succ^c p$), and (2) ($p \sim^c q \implies q \sim^c p$).

Step 3: We argue:[21] if P and I are respectively the asymmetric and symmetric components[22] of $\succsim^c$, then $P = \succ^c$ and $I = \sim^c$. We begin with some observations. Note, $P \bigcap I = \varnothing$ by definition of P and I. Also, $P \bigcup I = \succsim^c$ because $\succsim^c$ is complete as shown in step 2. Thus, $\{P, I\}$ partitions $\succsim^c$. Further, P and I are respectively asymmetric and symmetric. Next, $\succ^c$ and $\sim^c$ are disjoint. To see why, suppose $p \succ^c q$ and $p \sim^c q$ hold. By definition of $\succ^c$, there exists $\succsim^* \in \mathscr{C}_{\succsim}$ such that $p \succ^* q$. Also, by definition of $\sim^c$, $p \sim^c q$ holds. This contradicts the fact that $\succ^*$ and $\sim^*$ are disjoint as $\succsim^*$, being in $\mathscr{C}_{\succsim}$, is a preference. Also, $\succsim^c = \succ^c \bigcup \sim^c$ by definition of $\succsim^c$. Thus, $\{\succ^c, \sim^c\}$ partitions $\succsim^c$. Moreover, $\succ^c$ and $\sim^c$ are respectively asymmetric and symmetric from step 1. Hence, to complete the proof of the claim in this step it is enough to establish the following general result:

> Lemma: *Let X be a non–empty set. Suppose that $\mathfrak{A}_1, \mathfrak{A}_2$ are two asymmetric binary relations on X and $\mathfrak{S}_1, \mathfrak{S}_2$ are two symmetric binary relations on X such that $\mathfrak{A}_1 \bigcap \mathfrak{S}_1 = \varnothing = \mathfrak{A}_2 \bigcap \mathfrak{S}_2$ and $\mathfrak{A}_1 \bigcup \mathfrak{S}_1 = \mathfrak{A}_2 \bigcup \mathfrak{S}_2$. Then, $\mathfrak{A}_1 = \mathfrak{A}_2$ and $\mathfrak{S}_1 = \mathfrak{S}_2$.*

For proof, suppose $x_*, x^* \in X$ satisfy $(x_*, x^*) \in \mathfrak{A}_1$ and $(x_*, x^*) \notin \mathfrak{A}_2$. Then, $\mathfrak{A}_1 \bigcup \mathfrak{S}_1 = \mathfrak{A}_2 \bigcup \mathfrak{S}_2$ implies that $(x_*, x^*) \in \mathfrak{S}_2$. Because $\mathfrak{S}_2$ is symmetric, we have $(x^*, x_*) \in \mathfrak{S}_2$. Then, $\mathfrak{A}_1 \bigcup \mathfrak{S}_1 = \mathfrak{A}_2 \bigcup \mathfrak{S}_2$ implies $(x^*, x_*) \in \mathfrak{A}_1 \bigcup \mathfrak{S}_1$. However, $(x^*, x_*) \notin \mathfrak{A}_1$ because $(x_*, x^*) \in \mathfrak{A}_1$ and $\mathfrak{A}_1$ is asymmetric. Thus, we obtain $(x^*, x_*) \in \mathfrak{S}_1$. Then, the symmetry of $\mathfrak{S}_1$ implies that $(x_*, x^*) \in \mathfrak{S}_1$. Hence, we have $(x_*, x^*) \in \mathfrak{A}_1$ and $(x_*, x^*) \in \mathfrak{S}_1$. That is, $\mathfrak{A}_1 \bigcap \mathfrak{S}_1 \neq \varnothing$ which is a contradiction. Thus, our supposition is wrong. That is, $\mathfrak{A}_1 \subseteq \mathfrak{A}_2$. By a symmetric argument, we obtain $\mathfrak{A}_2 \subseteq \mathfrak{A}_1$. Hence, $\mathfrak{A}_1 = \mathfrak{A}_2$. Clearly, $\mathfrak{S}_1 = (\mathfrak{A}_1 \bigcup \mathfrak{S}_1) \setminus \mathfrak{A}_1$ as $\mathfrak{A}_1 \bigcap \mathfrak{S}_1 = \varnothing$. Similarly, $\mathfrak{S}_1 = (\mathfrak{A}_2 \bigcup \mathfrak{S}_2) \setminus \mathfrak{A}_2$. Thus, $\mathfrak{S}_1 = \mathfrak{S}_2$.

Step 4: We argue: $\succsim^c$ is transitive. Let $p, q, r \in \Delta(S)$ be such that $p \succsim^c q$ and $q \succsim^c r$. We are required to show that $p \succsim^c r$. From the definition of $\succsim^c$, it is enough to prove each of the following:

1. Cross–transitivity of $(\succ^c, \sim^c)$: $(\ p \succ^c q\ ;\ q \sim^c r\ ) \implies p \succ^c r$.
2. Cross–transitivity of $(\sim^c, \succ^c)$: $(\ p \sim^c q\ ;\ q \succ^c r\ ) \implies p \succ^c r$.
3. Transitivity of $\sim^c$: $(\ p \sim^c q\ ;\ q \sim^c r\ ) \implies p \sim^c r$.
4. Transitivity of $\succ^c$: $(\ p \succ^c q\ ;\ q \succ^c r\ ) \implies p \succ^c r$.

[21] This step justfies, $\succ^c$ and $\sim^c$ are indeed the asymmetric and symmetric components of $\succsim^c$.

[22] That is, (1) $pPq \iff (\ p \succsim^c q\ ;\ \text{not } q \succsim^c p\ )$, and (2) $pIq \iff (\ p \succsim^c q\ ;\ q \succsim^c p\ )$.

For 1., assume $p \succ^c q$ and $q \sim^c r$. The definition of $\succ^c$ and $p \succ^c q$ imply that there exists a preference $\succsim^* \in \mathscr{C}_{\succsim}$ such that $p \succ^* q$. Also, the definition of $\sim^c$ and $q \sim^c r$ imply $q \sim^* r$. Since $\succsim^*$ is a preference, cross–transitivity of $(\succ^*, \sim^*)$ holds. Thus, $p \succ^* q$ and $q \sim^* r$ imply $p \succ^* r$. Since $\succsim^* \in \mathscr{C}_{\succsim}$, the definition of $\succ^c$ and $p \succ^* r$ imply $p \succ^c r$. That is, $(\, p \succ^c q \; ; \; q \sim^c r \,) \implies p \succ^c r$ as required.

For 2., assume $p \sim^c q$ and $q \succ^c r$. The definition of $\succ^c$ and $q \succ^c r$ imply that there exists a preference $\succsim^* \in \mathscr{C}_{\succsim}$ such that $q \succ^* r$. Also, the definition of $\sim^c$ and $p \sim^c q$ imply $p \sim^* q$. Since $\succsim^*$ is a preference, cross–transitivity of $(\sim^*, \succ^*)$ holds. Thus, $p \sim^* q$ and $q \succ^* r$ imply $p \succ^* r$. Since $\succsim^* \in \mathscr{C}_{\succsim}$, the definition of $\succ^c$ and $p \succ^* r$ imply $p \succ^c r$. That is, $(\, p \sim^c q \; ; \; q \succ^c r \,) \implies p \succ^c r$ as required.

For 3., assume $p \sim^c q$ and $q \sim^c r$. Let $\succsim^* \in \mathscr{C}_{\succsim}$ be arbitrary. By definition of $\sim^c$ and $p \sim^c q$, $p \sim^* q$ holds. Similarly, we have $q \sim^* r$. As $\succsim^*$ is a preference, $\sim^*$ is transitive. Then, $p \sim^* q$ and $q \sim^* r$ imply $p \sim^* r$. As $\succsim^* \in \mathscr{C}_{\succsim}$ was arbitrary, we have $p \sim^c r$ by definition of $\sim^c$. That is, $(\, p \sim^c q \; ; \; q \sim^c r \,) \implies p \sim^c r$ as required.

For 4., assume $p \succ^c q$ and $q \succ^c r$. Suppose $p \succ^c r$ does not hold. Since $\succsim^c = \succ^c \bigcup \sim^c$ by definition and $\succsim^c$ is complete as shown in step 2, the supposition that $p \succ^c r$ does not hold implies that at least one $p \sim^c r$, $r \sim^c p$ or $r \succ^c p$ holds. Since $\sim^c$ is symmetric from step 1, we have: $p \sim^c r$ iff $r \sim^c p$. Moreover, if $r \sim^c p$ holds, then $q \succ^c r$ and cross–transitivity of $(\succ^c, \sim^c)$ imply $q \succ^c p$. However, this contradicts the asymmetry of $\succ^c$ as shown in step 1 because $p \succ^c q$ holds. Thus, neither $p \sim^c r$ nor $r \sim^c p$ holds. Hence, $r \succ^c p$ must hold.

Now, $p \succ^c q$ and the definition of $\succ^c$ imply that there exists $\succsim^* \in \mathscr{C}_{\succsim}$ such that $p \succ^* q$. Also, by definition of $\mathscr{C}_{\succsim}$, it must be that $\succsim$ refines $\succsim^*$. Thus, $p \succ^* q$ implies $p \succ q$. That is, $p \succ^c q$ implies $p \succ q$. Similarly, $q \succ^c r$ implies $q \succ r$. But $\succ$ is transitive as $\succsim$ is a preference. Hence, $p \succ^q$ and $q \succ r$ imply $p \succ r$. Now, recall we also have $r \succ^c p$ from the last paragraph. Also, $r \succ^c p$ implies $r \succ p$. Moreover, $\succ$ is asymmetric as it is the asymmetric component of the preference $\succsim$. But $p \succ r$ and $r \succ p$ constitute a contradiction to the asymmetry of $\succ$. Thus, our supposition that $p \succ^c r$ fails to hold must be wrong. Hence, we obtain $p \succ^c r$. That is, $(\, p \succ^c q \; ; \; q \succ^c r \,) \implies p \succ^c r$ as required. This completes the argument for transitivity of $\succsim^c$.

Step 5: We argue: $\succsim^c$ is continuous. Assume $p \succ^c q$ holds. Then, there exists $\succsim^* \in \mathscr{C}_{\succsim}$ such that $p \succ^* q$. By definition of $\mathscr{C}_{\succsim}$, $\succsim^*$ is a continuous preference. Then, $p \succ^* q$ implies that there exists $\varepsilon > 0$ such that: if $p' \in B(p, \varepsilon)$ and $q' \in B(q, \varepsilon)$, then $p' \succ^* q'$. Since $\succsim^* \in \mathscr{C}_{\succsim}$, the definition of $\succ^c$ implies: $p' \succ^c q'$ for every $p' \in B(p, \varepsilon)$ and $q' \in B(q, \varepsilon)$. Hence, $\succsim^c$ is a continuous preference.

Step 6: We argue: $\succsim$ refines $\succsim^c$. Assume $p \succ^c q$ holds. Then, by definition of $\succ^c$, there exists $\succsim^*$ such that $p \succ^* q$. Also, the definition of $\mathscr{C}_{\succsim}$ implies that $\succsim$ refines $\succsim^*$. Thus, $p \succ^* q$ implies $p \succ q$. Hence, $p \succ^c q \implies p \succ q$ holds. That is, $\succsim$ refines $\succsim^c$ as required.

Step 7: We argue: if $\succsim^* \in \mathscr{C}_{\succsim}$, then $\succsim^c$ refines $\succsim^*$. For this, assume $\succsim^* \in \mathscr{C}_{\succsim}$ and $p \succ^* q$. Then, the definition of $\succ^c$ implies $p \succ^q$. Thus, $p \succ^* q \implies p \succ^c q$ holds if $\succsim^* \in \mathscr{C}_{\succsim}$. Hence, from Definition 3, we obtain: if $\succsim^*$, then $\succsim^c$ refines $\succsim^*$.

Step 8: We argue: $\succsim^c$ is a finest continuous coarsening of $\succsim$. From steps 2 and 4, we have $\succsim^c$ is complete and transitive. That is, $\succsim^c$ is a preference. Also, from step 5, we have $\succsim^c$ is continuous. Moreover, step 6 shows that $\succsim$ refines the continuous preference $\succsim^c$. Thus, $\succsim^c \in \mathscr{C}_{\succsim}$ by the definition of the class $\mathscr{C}_{\succsim}$. That is, $\succsim^c$ is a continuous coarsening of $\succsim$. Finally, step 7 shows that $\succsim^c$ refines every continuous coarsening of $\succsim$. That is, $\succsim^c$ is finer than every continuous coarsening of $\succsim$. Hence, $\succsim^c$ is a finest continuous coarsening of $\succsim$.

Step 9: We argue: if $\succsim^1$ and $\succsim^2$ are finest continuous coarsenings of $\succsim$, then $\succsim^1$ and $\succsim^2$ coincide. For this, assume that each of $\succsim^1$ and $\succsim^2$ is a finest continuous coarsening of $\succsim$. Since $\succsim^1$ is a finest continuous coarsening of $\succsim$, it follows that $\succsim^1 \in \mathscr{C}_{\succsim}$. Moreover, $\succsim^2$ being a finest continuous coarsening of $\succsim$ must refine every element of $\mathscr{C}_{\succsim}$. Thus, $\succsim^2$ refines $\succsim^1$. That is, $p \succ^1 q \implies p \succ^2 q$ holds. Interchanging the positions of the superscripts "1" and "2", in this argument, leads to: $p \succ^2 q \implies p \succ^1 q$. Hence, we obtain: $p \succ^1 q \iff p \succ^2 q$.

Now, we argue: $p \succsim^1 q \iff p \succsim^2 q$. Suppose $p \succsim^1 q$ holds but $p \succsim^2 q$ does not hold. Since $\succsim^2$ is a preference, it follows that the failure of $p \succsim^2 q$ implies $q \succ^2 p$ holds. Further, $q \succ^2 p$ implies $q \succ^1 p$. However, by definition of $\succ^1$, $q \succ^1 p$ implies $p \succsim^1 q$ does not hold. Since we have a contradiction, our supposition must be wrong. Thus, $p \succsim^1 q \implies p \succsim^2 q$ holds. Interchanging the superscripts "1" and "2", in this argument, allows us to conclude: $p \succsim^2 q \implies p \succsim^1 q$. Hence, we obtain: $p \succsim^1 q \iff p \succsim^2 q$. That is, $\succsim^1$ and $\succsim^2$ coincide.

Observe that step 8 shows that $\succsim^c$ is *one* preference which is a finest continuous coarsening of the given preference $\succsim$. Moreover, step 9 shows that any two finest continuous coarsening of the given preference $\succsim$ must be identical. Thus, we have established: the preference $\succsim^c$ is the unique finest continuous coarsening of the preference $\succsim$. Thus, the proof of Theorem 5 is complete. ■

PROOF: Consider an arbitrary $1 \le k \le K - k_*$. By (8), each of the sets $B^i_{k,\varepsilon}$ is non–empty. Thus, $\{\mathrm{cl}(B^i_{k,\varepsilon}) : \varepsilon > 0)\}$ is a family of compact non–empty sets and satisfy the finite intersection property. Hence, B^i_k is a non–empty and compact subset of B^i_{k-1} which is clearly convex. Thus, to show that $A^*_i\langle C_1, C_2\rangle \subseteq B^i_k$, it is enough to show that B^i_k is admissible with respect to the context $\langle C_1, C_2\rangle$. This is because $A^*_i\langle C_1, C_2\rangle$ is the smallest admissible set in the context $\langle C_1, C_2\rangle$. For this, we must argue: B^i_k satisfies property B.

Assume, as part of the induction hypothesis, that B^i_{k-1} satisfies property[23] G. Now, fix $x_i \notin B^i_k$, $x'_i \in B^i_k$ and $x'_j \in C_j$ arbitrarily. If $x_i \notin B^i_{k-1}$, then from $B^i_k \subseteq B^i_{k-1}$ it follows that $(x'_i, x'_j) \succ_i (x_i, x_j)$ for some $x_j \in C_j$. Hence, assume $x_i \in B^i_{k-1}$. Because $x_i \in B^i_k \setminus B^i_{k-1}$, there exists $\varepsilon > 0$ such that $x_i \in B^i_{k-1} \setminus \mathrm{cl}(B^i_{k,\varepsilon})$. Since $B^i_{k,\varepsilon} \subseteq \mathrm{cl}(B^i_{k,\varepsilon})$, we have: $x_i \in B^i_{k-1} \setminus B^i_{k,\varepsilon}$. By (9), $\min_{x_j \in \Delta(x_i, \mathbf{v}^i_k)} U_{i,k_*+k}(x_i, x_j) < v^i_k - \varepsilon$. Then, there exists $x_j \in \Delta(x_i, \mathbf{v}^i_k)$ with $U_{i,k_*+k}(x_i, x_j) < v^i_k - \varepsilon$.

We may assume that $U_{i,k_*+l}(x'_i, x'_j) = v^i_l$ for each $0 \le l < k$; that is, $x'_j \in \Delta(x'_i, \mathbf{v}^i_k)$. For otherwise, there exists $0 \le l_* < k$ such that $U_{i,k_*+l}(x'_i, x'_j) = v^i_l$ for all $0 \le l < l_*$ and $U_{i,k_*+l_*}(x'_i, x'_j) > v^i_{l_*}$. This is so as B^i_l satisfies property B for all $0 \le l < k$ by the induction hypothesis and by definition (8). Then, $(x'_i, x'_j) \succ_i (x_i, x_j)$ anyway.

We argue: $U_{i,k_*+k}(x'_i, x'_j) > U_{i,k_*+k}(x_i, x_j)$. Choose $0 < \varepsilon' < \varepsilon$. Since $x'_i \in B^i_k$, it follows that $x'_i \in B^i_{k,\varepsilon'}$. Thus, (9) and $x'_j \in \Delta(x'_i, \mathbf{v}^i_k)$ imply $U_{i,k_*+k}(x'_i, x'_j) \ge v^i_k - \varepsilon'$. Since $\varepsilon' < \varepsilon$, $U_{i,k_*+k}(x'_i, x'_j) > U_{i,k_*+k}(x_i, x_j)$. Since $x_j \in \Delta(x_i, \mathbf{v}^i_k)$ and $x'_j \in \mathbf{v}^i_k$, we have: $(x'_i, x'_j) \succ_i (x_i, x_j)$. That is, B^i_k satisfies property G. Thus, $A^*_i\langle C_1, C_2\rangle \subseteq B^i_k$.

In particular, $A^*_i\langle C_1, C_2\rangle \subseteq B^i_{K-k_*}$ holds. However, we must also establish: $B^i_{K-k_*} = A^*_i\langle C_1, C_2\rangle$. Because the list $U_{i,k_*+1}, \ldots, U_{i,K}$ is a lexicographic expected utility representation of player i's preference over $C_1 \times C_2$, by arguments as above it is enough to establish the equality of the following two quantities:

$$v^i_* := \sup_{x_i \in A^*_i\langle C_1, C_2\rangle} \; \min_{x_j \in \Delta(x_i, \mathbf{v}^i_{K-k_*})} U_{i,K}(x_i, x_j), \text{ and} \tag{10}$$

$$v^i_{K-k_*} := \sup_{x_i \in B^i_{K-k_*-1}} \; \min_{x_j \in \Delta(x_i, \mathbf{v}^i_{K-k_*})} U_{i,K}(x_i, x_j), \tag{11}$$

where (11) is (9) of section (7) reproduced with $k := K - k_*$. However, since $A^*_i\langle C_1, C_2\rangle \subseteq B^i_{K-k_*}$, (10) and (11) imply: $v^i_{K-k_*} \ge v^i_*$.

[23]Recall, property G is equivalent to property B as shown in Proposition 1.

To show that $v^i_{K-k_*} = v^i_*$ holds, suppose: $v^i_{K-k_*} > v^i_*$. Then, from definitions (10) and (11), there exists $x^*_i \in B^i_{K-k_*-1} \setminus A^*_i\langle C_1, C_2\rangle$ such that, for every $x_i \in A^*_i\langle C_1, C_2\rangle$, the following holds:

$$\min_{x_j \in \Delta(x^*_i, \mathbf{v}^i_{K-k_*})} U_{i,K}(x^*_i, x_j) \quad > \min_{x_j \in \Delta(x_i, \mathbf{v}^i_{K-k_*})} U_{i,K}(x_i, x_j). \tag{12}$$

Let $x^*_j \in \Delta(x^*_i, \mathbf{v}^i_{K-k_*})$, $x_i \in A^*_i\langle C_1, C_2\rangle$ and $x_j \in \Delta(x_i, \mathbf{v}^i_{K-k_*})$ satisfy:

$$\begin{aligned} U_{i,K}(x^*_i, x^*_j) &= \min_{x_j \in \Delta(x^*_i, \mathbf{v}^i_{K-k_*})} U_{i,K}(x^*_i, x_j), \text{ and} \\ U_{i,K}(x_i, x_j) &= \min_{x_j \in \Delta(x_i, \mathbf{v}^i_{K-k_*})} U_{i,K}(x_i, x_j). \end{aligned}$$

Therefore, inequality (12) implies $U_{i,K}(x^*_i, x^*_j) > U_{i,K}(x_i, x_j)$. Moreover, $U_{i,k_*+l}(x^*_i, x^*_j) = v^i_l = U_{i,k_*+l}(x_i, x_j)$ for every $0 \leq l < K - k_*$ because $x^*_j \in \Delta(x^*_i, \mathbf{v}^i_{K-k_*})$ and $x_j \in \Delta(x_i, \mathbf{v}^i_{K-k_*})$ (see (7) in section 7). Also, $U_{i,l}(x^*_i, x^*_j) = U_{i,l}(x_i, x_j)$ for all $1 \leq l \leq k_*$ because $U_{i,l}$ is constant over $C_1 \times C_2$ for all $l \leq k_*$ by the definition of k_*. Then, since the list $U_{i,1}, \ldots, U_{i,K}$ is a lexicographic expected utility representation of player i's preference $\succsim_i$, we obtain: $(x^*_i, x^*_j) \succ_i (x_i, x_j)$. However, note that x^*_j is a best response of player j in C_j to x^*_i. Thus, we have a contradiction to the fact that $A^*_i\langle C_1, C_2\rangle$ satisfies property B. ■

A.3 *Proof of Corollary 1*

PROOF: Let $\succsim^{**}$ be a preference over $\Delta(S)$ that admits the following lexicographic expected utility representation via the expected utility functions $U_1, \ldots, U_K$ with U_1 is non–constant. Also, let $\succsim^*$ be the finest continuous coarsening of $\succsim^{**}$ which exists and is unique by Theorem 5. Fix an arbitrary p in the (relative) interior of $\Delta(S)$. Then, the closure of weak upper contour set $U_{\succsim^{**}}(p)$ of p according to $\succsim^{**}$ is the intersection of $\Delta(S)$ and the closed halfspace $H(p, U_1)$ passing through p and orthogonal to U_1. Since $\succsim^*$ is the finest continuous coarsening of $\succsim^{**}$, it follows that $\Delta(S) \bigcap H(p, U_1)$ is subset of the weak upper contour set $U_{\succsim^*}(p)$ of p according to $\succsim^*$. A similar set–containment must hold with respect to the weak lower contour sets of p. Thus, it is a necessary condition on $\succsim^*$ that it admits U_1 as one of its expected utility representations. For sufficiency, observe that $\succsim^*$ if defined such that U_1 is one of its expected utility representations, then $\succsim^*$ must be continuous and it must be refined by $\succsim^{**}$. ■

REFERENCES

AUMANN, R. J. (1964): "Utility Theory without the Completeness Axiom", *Econometrica*, vol. 30, pp. 445–462.

AUMANN, R. J. & A. BRANDENBURGER (1995): "Epistemic Conditions for Nash Equilibrium", *Econometrica*, vol. 63, pp. 1161–1180.

BERNHEIM, B. D. (1984): "Rationalizable Strategic Behavior", *Econometrica*, vol. 52, pp. 1007–1028.

BLUME, L., A. BRANDENBURGER & E. DEKEL (1989): "An Overview of Lexicographic Choice under Uncertainty", *Annals of Operations Research*, vol. 19, pp. 231–246.

CHATTERJEE, S. (2022): "Additivity over Convex Domains is equivalent to Lexicographic Structures", Chapter 2 of this PhD Thesis.

CHIPMAN, J. S. (1960): "The Foundations of Utility", *Econometrica*, vol. 28, pp. 193–224.

FERGUSON, C. E. (1958): "An Essay on Cardinal Utility", *Southern Economic Journal*, vol. 25, pp. 11–23.

FISHBURN, P. C. (1970) *Utility Theory for Decision Making* New York: John Wiley & Sons, Inc.

FISHBURN, P. C. (1971): "On the Foundations of Game Theory: The case for non–Archimedean Utilities", *International Journal of Game Theory*, vol. 1, pp. 65–71.

HAUSNER, M. (1954): "Multidimensional Utilities", in R. M. Thrall, C. H. Coombs and R. L. Davis (editors) *Decision Processes* New York: John Wiley & Sons Inc.

LUCE, R. D. & H. RAIFFA (1957) *Games and Decisions* New York: John Wiley & Sons, Inc.

PEARCE, D. G. (1984): "Rationalizable Strategic Behavior and the Problem of Perfection", *Econometrica*, vol. 52, pp. 1029–1050.

POLAK, B. (1999): "Epistemic Conditions for Nash Equilibrium, and Common Knowledge of Rationality", *Econometrica*, vol. 67, pp. 673–76.

THRALL, R. M. (1954): "Applications of Multidimensional Utility Theory", in R. M. Thrall, C. H. Coombs and R. L. Davis (editors) *Decision Processes* New York: John Wiley & Sons Inc.

VON NEUMANN, J. (1928): "Zur Theorie der Gesellschaftsspiele", *Math. Annalen*, vol. 100, pp. 295–320.

VON NEUMANN, J. & O. MORGENSTERN (1944) *Theory of Games and Economic Behavior* New Jersey: Princeton University Press.

Chapter 2

Additivity over Convex Domains is Equivalent to Lexicographic Structures

1. introduction

In a number of models covering disparate areas such as decision theory, social choice theory and linear algebra, axioms variously labelled additivity, independence and invariance are used. They are typically deployed in conjunction with a continuity axiom in order to establish fundamental results such as the Expected Utility Theorem and Utilitarianism. While the additivity or independence or invariance axioms are restrictions on behavior and aggregation, continuity is a technical assumption often with no independent justification.

Our goal is to investigate the consequences of dropping the continuity axiom entirely and to focus exclusively on the additivity type axioms. We show that in convex domains, additivity is equivalent to "lexicographic structures" — loosely speaking the application of the lexicographic criterion. The key to our approach is a geometric result which we call the Decomposition Theorem for Graded Halfspaces. By applying this result to the aforementioned areas, we are able to refine and extend existing results.

We briefly describe our findings for the application domains that we consider. Our first application domain is expected utility theory. The classical result due to von Neumann & Morgenstern (1944) is the Expected Utility Theorem. They introduced the Independence axiom which requires of any preference on lotteries, over a finite set of basic prizes, the following: for any three lotteries p, q, r and any $\alpha \in (0,1)$, $p \succ q$ holds *if*, and only if, the α–randomization of p and r is strictly preferred according to $\succ$ over the α–randomization of q and r. Then, Independence and Continuity was shown to characterize preferences which admit an expected utility representation.

However, Continuity is a technical assumption whereas Independence is a plausible assumption on decision making behavior. Hausner (1954) showed that Independence alone characterizes preferences that admit a lexicographic expected utility representation. Moreover, the lexicographic criterion is natural as a model of decision makers in many contexts. For instances, applications in portfolio theory are discussed extensively in Fishburn (1969, 1974). Thus, Hausner's result is both sharp and useful for economic modelling.

However, notice the "if" implication of the Independence axiom. It requires that the preference relation declares $p \succ q$ even if *one* α exists such that the α–randomization over p and r dominates via $\succ$ the α–randomization over q and r. We weaken the Independence axiom, stated above, as follows. The ranking between p and r will be concluded to be $p \succ r$ if *every* α–randomization over p and r dominates the corresponding α–randomization over q and r. The "only if" part of the original Independence is retained as such.

We believe our axiom to be normatively more appealing. Subsection 3.2 provides a full dicsussion. Further, our version of Independence is logically weakly weaker. Moreover, we introduce *affine local orders* in subsection 3.2 which are binary relations on the simplex. They satisfy our Independence axiom but may not be complete. However, by additionally requiring completeness, our axiom implies the existence of lexicographic expected utilities (Theorem 2 of subsection 3.1) thereby strengthening Hausner's result. Blume et al. (1991*a*) use Hausner's theorem to extend Anscombe & Aumann (1963) to lexicographic probabilities used in Blume et al. (1991*b*) for a theory of equilibrium selection in games via "higher order theories".

Despite its normative appeal, the classical Independence axiom has received criticisms in the decision theory literature due to the failure of the Expected Utility Hypothesis to accommodate Allais type paradoxes. However, Nielsen & Rehbeck (2022) experimentally find that people learn to follow Independence. In section 3.3, we sharpen the analysis of Segal (2023) in this direction.

Our second application domain is social choice theory. In the welfarist approach, to arrive at a social ranking over alternatives from individual preferences an aggregator considers only the vector of utilities associated with alternatives. One prominent class of such rules is Generalized Utilitarianism. Any rule in this class is defined by a system of weights—one for each individual—such that an alternative a socially dominates another alternative b, if and only if, the weighted sum of individual utilities from a is at least as high as the weighted sum of individual utilities from b. For a social welfare functional to satisfy Welfarism, the axioms of Binary Independence of Irrelevant Alternatives and Pareto Indifference must hold. Moreover, ethical assumptions such as Weak Pareto or Strong Pareto are also considered for the classification of various aggregators.

In addition to these Welfarism assumptions, the characterization of various aggregators are in part based on assumptions about how the aggregators process information inherent in the profile of individual utility functions. In particular, the questions of interest are (1) "whether the rule processes only the ordinal component or the cardinal component of individual preferences?", and (2) "to what degree does the rule assume individuals' utilities to be comparable?".

One such assumption is Cardinal Measurability & Unit Comparablity (CMUC). The classical result of HARSANYI (1955) is that any rule which satisfies the Welfarism axioms, Weak Pareto and CMUC, in conjunction with Continuity, must be a Generalized Utilitarianism. Of course, the converse also holds. Note that the assumptions in Harsanyi's characterization—except for Continuity—are principles of an ethical and normative nature. We find that CMUC in conjunction with the Welfarism axioms characterizes lexicographic extensions of Generalized Utilitarianisms — there exists a *list* of weight systems which is used according to the lexicographic criterion.

Lexicographic extensions fail to satisfy Continuity, if and only if, the list of weight systems has more than one element. Thus, the additional assumption of Continuity simply collapses the lexicographic extension to *almost* a Generalized Utilitarianism. We say "almost" because according to our definition of lexicographic extensions, weights may be negative. Thus, Harsanyi's result follows as a corollary when one just additionally assumes Weak Pareto.

We next strengthen the measurability–comparability requirement to Cardinal Measurability & Non–Comparability (CMNC). Our result is that the Welfarism axioms, Strong Pareto and CMNC characterize Serial Dictatorships. Moreover, weakening Strong Pareto to Weak Pareto characterizes the weak dictatorships. Thus, we are able to recover the Impossibility Theorem due to ARROW (1963).

Our third application domain is linear representations. The problem considered by BLACKWELL & GIRSHICK (1954) is as follows: what subclass of complete and transitive binary relations on $\mathbb{R}^n$ admit some "linear representation"? A linear representation is a mapping of all vectors x in $\mathbb{R}^n$ to corresponding numbers $\lambda \cdot x$, where λ is some fixed vector and "$u \cdot v$" is the standard inner product, such that:

$$x \succsim y \iff \lambda \cdot x \geq \lambda \cdot y.$$

The fundamental result is Theorem 4.3.1 (in their book) which is known as the Blackwell–Girshick Theorem. They introduce an axiom called Invariance which says, $x \succsim y$ iff $x + z \succsim y + z$. Further, they consider the axiom of Monotonicity which requires: $x >> y$ implies $x \succ y$. Then, their result characterizes complete and transitive binary relations (orders) on $\mathbb{R}^n$ which admit linear representations with positive λ as those which satisfy—in conjunction with Continuity—the axioms of Monotonicity and Invariance.

We briefly indicate the role of this result in applications. The result was developed in BLACKWELL & GIRSHICK (1954) to study two–person zero–sum games with obvious focus on the Minimax Theorem. Moreover, this result was used in statistical decision theory to study the class of minimax estimators from the point of view of a game between a statistician and nature. However, since the publication of this result, it has become a prominent tool in microeconomic theory. For instance, D'ASPREMONT & GEVERS (1977, 2002) and ROBERTS (1980*a*–*c*) contain several characterization theorems in social choice theory based on the Blackwell–Girshick Theorem.

Howover, the Blackwell–Girshick Theorem was originally developed for orderings over the *full* Euclidean space and requires Monotonicity in its proof in an essential way. One class of problems in the theory of mechanism design that has called for generalizations of this theorem to restricted domains is the characterization of dominant strategy incentive compatible mechanisms which are positive affine maximizers. The fundamental result of ROBERTS (1979) has been improved upon in MISHRA & SEN (2012) for which the latter authors extend the classical result to any open convex subset of $\mathbb{R}^n$.

Our contribution in the context of Blackwell–Girshick Theorem is twofold. First, we provide a generalization of the theorem to *arbitrary* convex subsets of $\mathbb{R}^n$ using only Invariance and Continuity. The only price paid is that λ may be negative — additionally assuming Monotonicity recovers non–negativity. Note, convex subsets may ail to be open or closed, and their closure may have an empty interior. Also, there are convex sets which are not Lebesgue measurable.

Our second contribution is a generalization of the classical result when Continuity is dropped. We consider Convexity as an assumption on the ordering. Convexity of the ordering requires every weak upper and lower contour set to be a convex subset of the ambient convex space. Continuity and Invariance imply Covexity. However, the converse does not hold. In fact, our characterization result shows that an ordering satisfies Invariance and Convexity, if and only if, the ordering admits a representation which is the lexicographic extension of linear representations. Again, we develop this result in the setting where the ambient space is an arbitrary convex subset of some $\mathbb{R}^n$.

Our fourth, and last, application is to linear algebra. A finite dimensional ordered vector space V is a vector space which is isomorphic to some $\mathbb{R}^n$ and has an order $\succ$ defined over it such that $\succ$ is "compatible" with vector space operations. For instance, if $x, y \in V$ are such that $x \succ y$ and the scalar $\alpha > 0$ then $\alpha x \succ \alpha y$. Further, if $x \succ y$ then $x + z \succ y + z$. A lexicographic function space is the space of all real–valued functions on $[n] := \{1, \ldots, n\}$ endowed with the linear order $\succ_n$ which makes it an ordered vector space such that only those functions on $[n]$ dominate the constant function which is zero on $[n]$ whose first non–zero value is positive.

Hausner & Wendel (1952) showed that every n–dimensional ordered vector space V admits a an ordered basis which makes V linearly isomorphic to $\mathscr{L}_n$ by preserving the order structure. This characterization of ordered vector spaces is often known as the Hausner–Wendel Theorem, and it plays a fundamental role in mathematics. Moreover, this result is the basis of the characterization of lexicographic expected utilities in Hausner (1954). We are able to provide a short proof of the Hausner–Wendel Theorem.

In each of the above applications, the "object" of study is defined over a convex "domain" and it satisfies some "additivity" property. While the "object" in all of these applications have been shown to possess a "*lexicographic* structure", observe, the limited role of Continuity type axioms when assumed additionally. Therefore, a natural conjecture in qualitative terms is as follows:

> Is "additivity" of an "object" over a convex "domain" *equivalent* to the "object" possessing a "lexicographic structure"?

The answer is in the affirmative! Formally, we shall introduce the concept of "graded halfspace" which is an abstract representation of any "lexicographic structure". Then, we state and prove what we call the "Decomposition Theorem" which characterizes graded halfspaces. This shall be a formal expression of the above statement.

We briefly explain the concept of a “graded halfspace”. Consider a finite dimensional vector space over the reals. Any open halfspace whose boundary contains the origin shall be called a *slice* of this vector space. Pick any slice of the given vector space. Then, the boundary of this slice is a subspace with dimension one less. Pick any a slice of this subspace. Thus, we have a halfspace, of the boundary of the previous slice, which is open relative to the topology inherited by the boundary of the first slice from the ambient vector space. The union of the resulting subsets, with a prespecified number of iterations of this procedure, is a *graded halfspace*. For instance, in the lexicographic order on the two–dimensional Euclidean plane, the strict upper contour set of the origin is a graded halfspace having two slices.

We now outline the statement of the Decomposition Theorem. For this, we begin by observing that a graded halfpace is a convex cone (not containing the origin). For any given subset of the ambient vector space, let its *reflection* be the subset obtained by reflecting through the origin every point of the set. Observe that the reflection of a graded halfspace is also a graded space. For instance, in the example with the lexicographic order over the two–dimensional Euclidean plane, the strict lower contour set of the origin is also a graded halfspace and it is the reflection of the strict upper contour set.

Mutually reflecting graded halfspaces must be disjoint. Moreover, the deletion of these graded halfspaces, from the ambient space, leaves a subspace. In the present example, the indifference set of the origin remains after deletion of the strict upper and lower contour sets from the two–dimensional Euclidean plane. Clearly, the indifference set of the origin, according to the lexicographic order, is a subspace of the ambient two–dimensional Euclidean plane.

That is, a graded halfspace and its reflection are a pair of mutually reflecting convex cones which together with a subspace form a partition of the ambient vector space. Our Decomposition theorem asserts that the *converse* also holds. The statement is as follows:

> Decomposition Theorem — *The cones in any partition of an Euclidean space, consisting of a pair of mutually reflecting convex cones and a subspace, is a graded halfspace.*

This statement formalizes our qualitative conjecture as follows. First, graded halfspaces are an abstraction of “lexicographic structures” of “objects”. Second, “additivity” yields the convex cones and subspaces on which the Decomposition Theorem applies. Additional qualifications, which are “mutually reflecting” and “partition”, make the connection tight enough as required by graded halfspaces.

The rest of the article is organized as follows. Section 2 presents the Decomposition Theorem and its proof sketch. Each of sections 3 to 6 consider one application domain. Proofs omitted from the main text are presented in the Appendix. We close this section with some coments about the background.

The Background

To characterize the "lexicographic structure", the usual method is to inductively invoke the Separating Hyperplane Theorem(s) because of the two ingredients—the convex "domain" and the "additive" object to be characterized. However, no *precise* connection beyond this is shared by these characterizations. For instance, compare KRANTZ ET AL. (1971), BLUME ET AL. (1989) and YOUNG (1975). Notwithstanding this, some *vague* connection has been suggested. Consider the following words[24] from FISHBURN (1970).

> "The purpose of this section is to note an *affinity* between additive utilities and lexicographic utilities."

They exemplify the above intuition. In fact, in FISHBURN (1969), the expected utility theories of VON NEUMANN & MORGENSTERN (1944) and ANSCOMBE & AUMANN (1963) are extended to the multivariate setup *but* not to a theory of lexicographic expected utilities. However, the Decomposition Theorem makes the connection precise.

2. DECOMPOSITION THEOREM

2.1 *Framework and Main Result*

To state the Decomposition Theorem, we must first define the concept of "graded halfspaces". We begin with some preliminaries. Any $C \subseteq \mathbb{R}^m$ is a[25] *cone* if, $\alpha\mathbf{x} + \beta\mathbf{y} \in C$ for any $\mathbf{x}, \mathbf{y} \in C$ and $\alpha, \beta > 0$. For any subspace $W_* \subseteq \mathbb{R}^m$, let $\mathbf{U}_* \coloneqq \langle \mathbf{u}_*^k \in W_* : k = 1, \ldots, K \rangle$ be a list of orthonormal vectors. For each $k \in \{1, \ldots, K_*\}$, let $\mathbf{U}_*^k$ be the set of $\mathbf{w} \in W_*$ such that $\langle \mathbf{u}_*^l, \mathbf{w} \rangle = 0$ for all $l < k$ and $\langle \mathbf{u}_*^k, \mathbf{w} \rangle > 0$. We call $\mathbf{U}_*^k$ the kth *slice* generated by the vectors in $\mathbf{U}_*$.

[24] See the opening line of section 4.3 on page 48.

[25] Notice, what we simply call a "cone" is often called a "*convex* cone".

DEFINITION 1: *The* graded halfspace *induced by* $\mathbf{U}_*$*, denoted by* $H_{\mathbf{U}_*}$*, is the union of slices generated by* $\mathbf{U}_*$.

That is, $H_{\mathbf{U}_*} = \bigcup_{k=1}^{K} \mathbf{U}_*^k$. For illustration, consider Figure 1 in which we take $\mathbb{R}^2$ as the subspace W_* (of, say $\mathbb{R}^3$). There is a list of two orthonormal vectors $\mathbf{U}_* = (\mathbf{u}_*^1, \mathbf{u}_*^2)$. The shaded region $\mathbf{U}_*^1$ is the open halfspace, in W_*, of all vectors which make an acute angle with respect to $\mathbf{u}_*^1$. Thus, $\mathbf{U}_*^1$ is the first slice generated by $\mathbf{U}_*$.

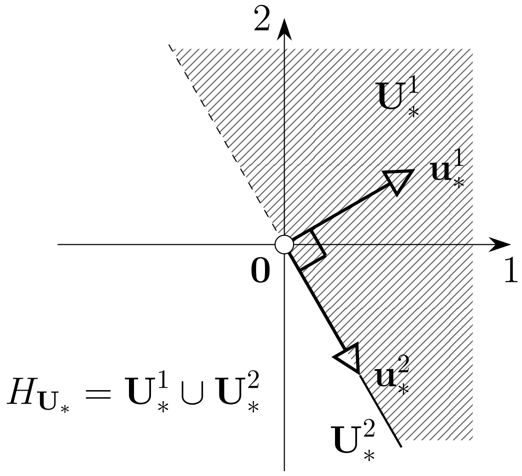


FIGURE 1: A Graded Halfspace.

The second slice $\mathbf{U}_*^2$ is the set of all vectors which are orthogonal to $\mathbf{u}_*^1$ and make an acute angle with respect to $\mathbf{u}_*^2$. That is, $\mathbf{U}_*^2$ is the ray without the origin along the direction of $\mathbf{u}_*^2$. Observe, the second slice is an open halfspace of the boundary of the first slice which in turn is an open halfspace of the ambient subspace W_*. The number of vectors in the list $\mathbf{U}_*$ can be anything up to the dimension of W_*. Note, the strict upper (or, lower) contour sets of the origin $\mathbf{0}$, with respect to any lexicographic preference over $\mathbb{R}^2$, must be a graded halfspace.

Notice that any graded halfspace is (convex) cone. For any $A \subseteq W_*$, let $-A := \{\mathbf{x} \in W_* : -\mathbf{x} \in A\}$. That is, $-A$ is the "reflection through the origin" (henceforth, "reflection") of the set A. Observe that the reflection of the graded halfpsace $H_{\mathbf{U}_*}$, induced by the vectors in $\mathbf{U}_*$, is the graded halfspace $H_{-\mathbf{U}_*}$ induced by the list of reflected vectors $-\mathbf{U}_* := \langle \mathbf{u}_*^k : k = 1, \ldots, K \rangle$. That is, $H_{-\mathbf{U}_*} = -H_{\mathbf{U}_*}$. Thus, $H_{\mathbf{U}_*}$ and $H_{-\mathbf{U}_*}$ are a pair of mutually reflecting cones and are *disjoint.*

As can be seen from Figure 1, since $\mathbf{0} \notin H_{\mathbf{U}_*}$ and $H_{-\mathbf{U}_*} = -H_{\mathbf{U}_*}$, we have $\mathbf{0} \notin H_{-\mathbf{U}_*}$. Moreover, $\mathbf{0}$ is the *only* point in W_* which is not in at least one of $H_{\mathbf{U}_*}$ or $H_{-\mathbf{U}_*}$. Note, $\{\mathbf{0}\}$ is a subspace of W_*. This situation is perfectly general: $W_* \setminus (H_{\mathbf{U}_*} \cup H_{-\mathbf{U}_*}) = O_{\mathbf{U}_*}$ is the subspace orthogonal to the given list of vectors $\mathbf{U}_*$.

In fact, in the above example, if $\mathbf{U}_* = (\mathbf{u}_*^1)$ instead, then the graded halfspaces $H_{\mathbf{U}_*}$ and $H_{-\mathbf{U}_*}$ are the open halfspaces, in $W_* = \mathbb{R}^2$, that consist of all vectors which make an acute angle with the vectors $\mathbf{u}_*^1$ and $\mathbf{u}_*^2$, respectively. Then, $W_* \setminus (H_{\mathbf{U}_*} \cup H_{-\mathbf{U}_*})$ is the subspace $O_{\mathbf{U}_*}$ of vectors perpendicular to $\mathbf{u}_*^1$.

Thus, given any list of orthonormal vectors $\mathbf{U}_*$ in the subspace $W_* \subseteq \mathbb{R}^m$, the graded halfspaces $H_{\mathbf{U}_*}$ and $H_{-\mathbf{U}_*}$ are a pair of mutually reflecting cones such that the triple $(H_{\mathbf{U}_*}, H_{-\mathbf{U}_*}, O_{\mathbf{U}_*})$, where $O_{\mathbf{U}_*}$ is the subspace of W_* orthogonal to $\mathbf{U}_*$, is a *partition* of the ambient space W_*. Our *Decomposition Theorem* asserts the *converse.*

THEOREM 1: *Let W_* be a subspace of $\mathbb{R}^m$. Let U_*, V_* be nonempty cones in W_* and S_* be a subspace of W_* such that (U_*, V_*, S_*) form a partition of W_* and $V_* = -U_*$. Then, with $K := \dim(W_*) - \dim(S_*)$, there exists a unique list $\mathbf{U}_* \equiv \langle \mathbf{u}_*^k : k = 1, 2, \ldots, K\rangle$ of orthonormal vectors in W_* such that $U_* = H_{\mathbf{U}_*}$, $V_* = -H_{\mathbf{U}_*}$ and $S_* = O_{\mathbf{U}_*}$.*

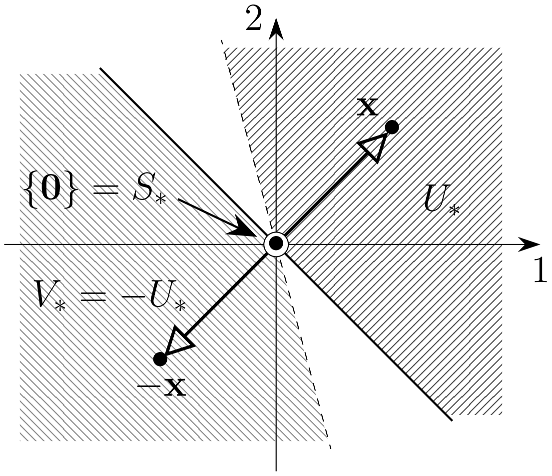


FIGURE 2: A triple (U_*, V_*, S_*).

That is, each cone in the partition, via a subspace and a pair of mutually reflecting convex cones, of a vector space must be a graded halfspace. For intuition, consider Figure 2 which shows two cones, U_* and V_*, not containing the origin. The cones U_* and V_* are reflections of each other: $\mathbf{x} \in U_*$ iff $-\mathbf{x} \in V_*$. Notice, each of the cones has an open ray as part of it but another closed ray which is not part of it. Further, $S_* = \{\mathbf{0}\}$ is a subspace such that (U_*, V_*, S_*) is a triple of pairwise disjoint non–empty subsets of $W_* := \mathbb{R}^2$. However, (U_*, V_*, S_*) fails to be a *partition* of W_*. But one way to turn this triple into a partition is to "expand" the cones U_* and V_* while maintaining the property $V_* = -U_*$ such that the "white spaces" in Figure 2 are "filled out". Then, U_* becomes the graded halfspace shown in Figure 1!

2.2 *Sketch of the Proof*

This subsection gives a technical overview of Theorem 1. The reader interested in applications may skip it without any loss of continuity. Theorem 1 rests on two geometric lemmas which are presented below. However, some elementary mathematical preliminaries are needed for their statement. We begin by stating these preliminaries.

Let $\mathscr{T}_{\mathbb{R}^m}$ be the standard topology on $\mathbb{R}^m$. For any $W_* \subseteq \mathbb{R}^m$, let $\mathscr{T}_{W_*} := \{W_* \cap A : A \in \mathscr{T}_{\mathbb{R}^m}\}$ be the subspace topology on W_*. The set $B_{\|\cdot\|}^{W_*}(\mathbf{w}, \varepsilon) := \{\mathbf{w}' \in W_* : \|\mathbf{w}' - \mathbf{w}\| < \varepsilon\}$, where $\mathbf{w} \in W_*$ and $\varepsilon > 0$, is the open ball relative to W_* centered on $\mathbf{w}$ with radius ε. If the "ambient space" $(W_*, \mathscr{T}_{W_*})$ is clear from the context, the qualifiers "relative to W_*" or "relative to the subspace topology of W_*" shall be often dropped. We shall also abuse some notation as specified next. Let $A \subseteq W_*$. $A^c := W_* \setminus A$ is the complement of A relative to W_*. Further, A°, $\bar{A}$, A' and ∂A are the interior, closure, limit points and boundary of A, respectively, relative to $\mathscr{T}_{W_*}$.

Lemma 1: *Let W_* be a subspace of $\mathbb{R}^m$ and T_* be a proper subspace of W_*. Then $W_* \setminus T_*$ is path–connected, if and only if, the codimension of T_* in W_* is higher than* 1.

The intuition behind the above result is as follows. W_* is isomorphic to a Euclidean space of dimension at most n as it is a linear subspace of $\mathbb{R}^m$. Now, if a hyperplane is deleted from an Euclidean space, then clearly the resulting set is not path connected as it is the union of two disjoint open halfspaces. However, if the deleted proper linear subspace is not a hyperplane then, for any two points in the resulting set, there is a path joining them that "goes around" the deleted subspace. The key result used to prove Theorem 1 is as follows.

Lemma 2: *Let W_* be a linear subspace of $\mathbb{R}^m$. If U_*, V_* are nonempty cones in W_* and S_* a linear subspace of W_*, with (U_*, V_*, S_*) forming a partition of W_* and $V_* = -U_*$, then there exists a unique $\mathbf{u} \in W_*$ such that $\|\mathbf{u}\| = 1$ and the following hold:*

1. $\bar{U}_* \cap \bar{V}_* = \{\mathbf{w} \in W_* : \langle \mathbf{u}, \mathbf{w} \rangle = 0\}$.
2. $U_*^\circ = \{\mathbf{w} \in W_* : \langle \mathbf{u}, \mathbf{w} \rangle > 0\} = -V_*^\circ$.
3. $\partial U_* = \bar{U}_* \cap \bar{V}_* = \partial V_*$.
4. $S_* \subseteq \bar{U}_* \cap \bar{V}_*$.
5. ∂U_* *is a subspace of W_* with codimension* 1.

The key insights underlying Lemma 2 are as follows. As U_* and V_* are (convex) cones, so must be $U_*^\circ, \overline{U}_*$, V_*° and $\overline{V}_*$. Moreover, $V_* = -U_*$ implies $V_*^\circ = -U_*^\circ$ and $\overline{V}_* = -\overline{U}_*$. Thus, $T_* := \overline{U}_* \cap \overline{V}_*$ is a cone with $T_* = -T_*$. Hence, T_* is a subspace of W_*. Since (U_*, V_*, S_*) partitions W_* and S_* is a subspace of W_*, S_* is a *proper* subspace of W_* implying that the cones U_*° and V_*° are non–empty. Further, $T_* = \partial U_* = \partial V_*$ and S_* is a subspace of T_*. Then, $(U_*^\circ, V_*^\circ, T_*)$ partitions W_*. Since U_*° and V_*° are cones, they are path–connected.

However, $U_*^\circ \cup V_*^\circ = W_* \setminus T_*$ is *not* connected as $T_* = \partial U_* = \partial V_*$. Then, lemma 2 implies that the codimension of the subspace T_* in W_* is 1. Thus, orthogonal projection of any vector from U_*° onto T_* when normlized to unit length, say $\mathbf{u}$, satisfies:

$$T_* = I_* := \{\mathbf{w} \in W_* : \langle \mathbf{u}, \mathbf{w} \rangle = 0\}.$$

Also, $P_* := \{\mathbf{w} \in W_* : \langle \mathbf{u}, \mathbf{w} \rangle > 0\}$ and $N_* := \{\mathbf{w} \in W_* : \langle \mathbf{u}, \mathbf{w} \rangle < 0\}$ are cones such that (P_*, N_*, I_*) partitions W_*. Since $T_* = I_*$, it follows that $U_*^\circ \cup V_*^\circ = P_* \cup N_*$. As each of U_*°, V_*°, P_* and N_* is a cone with $U_*^\circ \cap V_*^\circ = \varnothing = P_* \cap N_*$, $P_* \cap U_*^\circ \neq \varnothing$ implies $U_*^\circ = P_*$ and $V_*^\circ = N_*$. Thus, $U_*^\circ = \{\mathbf{w} \in W_* : \langle \mathbf{u}, \mathbf{w} \rangle > 0\} = -V_*^\circ$ which is point 2 claimed by the lemma. Also, $\overline{U}_* \cap \overline{V}_* = T_* = \{\mathbf{w} \in W_* : \langle \mathbf{u}, \mathbf{w} \rangle = 0\}$ as claimed in 1 of the lemma. We have already seen that $T_* = \partial U_* = \partial V_*$ and $S_* \subseteq T_*$ which are points 3 and 4, respectively. Since T_* has codimension 1 in W_* and $T_* = \partial U_*$, point 5 is established. That is, $\mathbf{u}$ is the vector which must exist as claimed by Lemma 2.

Theorem 1 is proved via induction on the dimension of subspace W_*. For this, we begin with a linear subspace $W_* \subseteq \mathbb{R}^m$ and a partition of it (U_*, V_*, S_*) as in the hypothesis of Theorem 1. Then, Lemma 2 gives a vector $\mathbf{u} \in W_*$ with $\|\mathbf{u}\| = 1$ such that $\partial U_* = \partial V_*$ is the subspace of W_* which is perpendicular to $\mathbf{u}$. This follows from parts 1 and 3 of Lemma 2. By parts 3 and 4, it follows that the subspace ∂U_* contains the subspace S_*. Then, the construction of the list $\mathbf{U}_*$, and the graded halfspace $H_{\mathbf{U}_*}$, proceeds as follows.

We set $\mathbf{u}_*^1 := \mathbf{u}$ and take U_*° as the first open halfspace of the graded halfspace $H_{\mathbf{U}_*}$ as in the conclusion of Theorem 1. Notice, U_*° is indeed an open halfspace of W_* is guaranteed by part 1 of Lemma 2. Now take ∂U_* as the next linear subspace whose dimension is exactly one less than that of W_* by part 5 of Lemma 2. With $S_* \subseteq \partial U_* = \partial V_*$ and that $\mathbf{u}_*^1$ is perpendicular to ∂U_*, the induction proceeds until the remaining set is the original linear subspace S_* itself. The formal proofs of the two lemmas and Theorem 1 are provided in subsections A.I.1–3 of the Appendix. We must point out that these arguments do *not* appeal to the "Separating Hyperplane Theorem(s)".

3. EXPECTED UTILITY THEORY

3.1 *Lexicographic Expected Utilities*

We introduce a normatively appealing weakening of the Independence axiom of VON NEUMANN & MORGENSTERN (1944). In conjuction with transitivity and completeness, it characterizes lexicographic expected utilities (Theorem 2) strengthening the result of HAUSNER (1954). In subsection 3.3, implications to decision theory are discussed.

Let Z be a finite and non–empty set whose elements are the *basic prizes*. A *lottery* is any map $p : Z \to \mathbb{R}_+$ such that $\sum_{z \in Z} p(z) = 1$. Let $\mathscr{L}(Z)$ be the set of all lotteries. For $p, q \in \mathscr{L}(Z)$ and $0 \leq \alpha \leq 1$, $\alpha \cdot p \oplus (1 - \alpha) \cdot q$ is the *compound lottery* that randomly results in either p or q with probabilities α or $1 - \alpha$, respectively. This compound lottery shall be identified with the lottery over basic prizes that selects any $z \in Z$ randomly with probability $\alpha p(z) + (1 - \alpha) q(z)$.

A *preference* is any complete and transitive binary relation $\succsim$ over $\mathscr{L}(Z)$. An *expected utility* (EU) is any map $u : \mathscr{L}(Z) \to \mathbb{R}$ that satisfies the following: if $p, q \in \mathscr{L}(Z)$ and $\alpha \in [0, 1]$ then,

$$u(\alpha \cdot p \oplus [1 - \alpha] \cdot q) = \alpha u(p) + (1 - \alpha) u(q).$$

Let $\mathscr{E}(Z)$ be the set of all EUs over $\mathscr{L}(Z)$. If $\succsim$ is a binary relation, its *expected utility* (EU) *representation* is any $u \in \mathscr{E}(Z)$ such that:

$$p \succsim q \quad \text{iff} \quad u(p) \geq u(q).$$

Let the asymmetric and symmetric components of $\succsim$ be denoted by $\succ$ and $\sim$, respectively. The Independence axiom is as follows.

INDEPENDENCE–0: *Let* $p, q, r \in \mathscr{L}(Z)$ *and* $\alpha \in (0, 1)$. *Then*:

$$p \succ q \quad \textit{iff} \quad \alpha \cdot p \oplus (1 - \alpha) \cdot r \succ \alpha \cdot q \oplus (1 - \alpha) \cdot r.$$

This axiom, and the following *Archimedean* property, hold for any binary relation $\succsim$ which admits an EU representation.

ARCHIMEDEAN: *If* $p, q, r \in \mathscr{L}(Z)$ *satisfy* $p \succ q \succ r$ *then, there exists* $\alpha, \beta \in (0, 1)$ *such that* $\alpha \cdot p \oplus (1 - \alpha) \cdot r \succ q \succ \alpha \cdot p \oplus (1 - \alpha) \cdot r$.

The first milestone of expected utility theory is,

VON NEUMANN–MORGENSTERN THEOREM: *A binary relation* $\succsim$ *is a preference that satisfies Independence–0 and the Archimedean property, if and only if, it admits an expected utility representation.*

In Herstein & Milnor (1953), the above theorem is generalized in two ways. First, they introduced the abstract notion of a "mixture set" over which the binary relation $\succsim$ is defined—the set of lotteries is but one example. We shall restrict attention only to the set of lotteries. Second, they relaxed Independence to the following.

Independence–1: *Let* $p, q, r \in \mathscr{L}(Z)$ *and* $\alpha \in (0,1)$. *Then*:

$$p \succ q \quad \textit{implies} \quad \alpha \cdot p \oplus (1-\alpha) \cdot r \succ \alpha \cdot q \oplus (1-\alpha) \cdot r.$$

Thus, the second milestone of expected utility is,

Herstein–Milnor Theorem: *A binary relation* $\succsim$ *is a preference that satisfies Independence–1 and the Archimedean property, if and only if, it admits an expected utility representation.*

Observe, this result is an improvement over the first because the reverse implication required by Independence–0 has been dropped in Independence–1. It is possible to see how this result improves upon the first in another manner. For this, consider the following.

Independence–2: *Let* $p, q, r \in \mathscr{L}(Z)$ *and* $\alpha \in (0,1)$. *Then*:

1. *If* $p \succ q$ *then* $\alpha \cdot p \oplus (1-\alpha) \cdot r \succ \alpha \cdot q \oplus (1-\alpha) \cdot r$, *and*
2. *If* $p \sim q$ *then* $\alpha \cdot p \oplus (1-\alpha) \cdot r \sim \alpha \cdot q \oplus (1-\alpha) \cdot r$.

When the binary relation $\succsim$ is complete, Independence–2 turns out to be equivalent to Independence–0. Also, Independence–1 is simply obtained by dropping the second of the two implications assumed in the statement of Independence–2.

The forms of "Independence" are regarded normatively appealing from a decision theoretic point of view. However, the Archimedean property (equivalently, Continuity) is harder to justify beyond serving as a techinical condition. Of course, this axiom is accepted widely as a technical condition without which there is no hope for *any* well–behaved numerical representation. Notwithstanding the widespread use of the Archimedean property, there is a class of preferences which does not satisfy the Archimedean property and yet is perfectly natural as a model of the decision maker. These preferences admit "*lexicographic* expected utility" representations. It is natural when the decision maker can be envisioned as one who decides using multiple criteria with a priority over these criteria. We now define the concept of lexcigraphic expected utility representations of preferences.

A *lexicographic expected utility* (LEU) *representation* of the binary relation $\succsim$ is any K–tuple of EUs $\langle u_k \in \mathscr{E}(Z) : k = 1, \ldots, K \rangle$ satisfying:

$$p \succsim q \quad \text{iff} \quad [u_1(p), \ldots, u_K(p)] \geq_L [u_1(q), \ldots, u_K(q)],$$

where $\geq_L$ is lexicographic order over $\mathbb{R}^K$. A binary relation that admits an LEU representation must be a preference that satisfies each of the above versions of Independence. However, they may fail to satisfy the Archimedean property. Thus, the third milestone in expected utility theory is the following result from HAUSNER (1954).

HAUSNER'S THEOREM: *A preference $\succsim$ satisfies Independence–2, if and only if, it admits a lexicographic expected utility representation.*

Hausner proved this theorem in the setting of mixture spaces based on the characterization by HAUSNER & WENDEL (1952) of ordered real linear spaces. Observe that just Independence–1, in addition to the Archimedean property though, is sufficient for the existence of EU representations according to the Herstein–Milnor Theorem. Thus, the insight of Hausner that the strengthening as Independence–2 alone is sufficient for existence of LEU representations is remarkable. As has been observed[26] by Peter C. Fisburn:

> "In the major development in lexicographic expected utility, Hausner [63] assumes . . . the following hold(s):
>
> $$A2'.\ x \sim y \implies \lambda x + (1-\lambda) z \sim \lambda y + (1-\lambda) z."$$

With this background in place, we introduce the following axiom.

INDEPENDENCE–3: *Let $p, q, r \in \mathscr{L}(Z)$. Then, $p \succ q$ if and only if:*

$$\big(\, \forall \alpha \in (0,1) \,\big) \big[\, \alpha \cdot p \oplus (1-\alpha) \cdot r \succ \alpha \cdot q \oplus (1-\alpha) \cdot r \,\big].$$

Then, our main result in this section is as follows.

THEOREM 2: *A preference $\succsim$ satisfies Independence–3, if and only if, it admits a lexicographic expected utility representation.*

In the next subsection, we argue that Independence–3 is normatively appealing and it is logically strictly weaker than either Independence–0 or Independence–2. In subsections 3.4 and 3.5, we obtain Theorem 2 from the Decomposition Theorem (that is, Theorem 1).

[26] See paragraph 5 in page 1464 of FISHBURN (1974).

We elaborate on the strength and normative appeal of Independence–3 as an axiom. For this, we begin with Independence–0. Consider the lotteries p, q and r. Also, fix any α in $(0, 1)$. Note, the interpretation $\mathfrak{I}_0$ of $\alpha \cdot p \oplus (1 - \alpha) \cdot r$ and $\alpha \cdot q \oplus (1 - \alpha) \cdot r$ is as follows:

> "The lottery $\alpha \cdot p \oplus (1 - \alpha) \cdot r$ results by tossing a coin, whose probability of showing "heads" is α, to choose one of p or r according as it shows "heads" or "tails". Likewise, the lottery $\alpha \cdot q \oplus (1 - \alpha) \cdot r$ is implemented using the *same* coin."

Then, the forward implication required by Independence–0 affords an interpretation $\mathfrak{I}_1$ which is as follows:

> "If the coin toss leads to "heads", comparing $\alpha \cdot p \oplus (1 - \alpha) \cdot r$ with $\alpha \cdot q \oplus (1 - \alpha) \cdot r$ tantamounts to comparing p with q. If the toss leads to "tails", comparing $\alpha \cdot p \oplus (1 - \alpha) \cdot r$ with $\alpha \cdot q \oplus (1 - \alpha) \cdot r$ tantamounts to comparing r with itself. However, the probability α of "heads" is *strictly* positive! Thus, if p is strictly preferred to q then $\alpha \cdot p \oplus (1 - \alpha) \cdot r$ *must be* strictly preferred to $\alpha \cdot q \oplus (1 - \alpha) \cdot r$."

As $\mathfrak{I}_1$ is plausible, the conditional with $p \succ q$ as hypothesis in each version of Independence has normative appeal. Thus, the normative defense of Independence–1, in particular, is accomplished. However, the comparison of the remaining implications remains. The interpretation $\mathfrak{I}_2$ of the reverse implication of Independence–0 is as follows:

> "Pick an arbitrary coin with a *given* probability α of showing up "head" in a toss. If the toss leads to "heads", comparing $\alpha \cdot p \oplus (1 - \alpha) \cdot r$ with $\alpha \cdot q \oplus (1 - \alpha) \cdot r$ tantamounts to comparing p with q. If the toss leads to "tails", comparing $\alpha \cdot p \oplus (1 - \alpha) \cdot r$ with $\alpha \cdot q \oplus (1 - \alpha) \cdot r$ tantamounts to comparing r with itself. However, α is *strictly* positive! Thus, if $\alpha \cdot p \oplus (1 - \alpha) \cdot r$ is strictly preferred to $\alpha \cdot q \oplus (1 - \alpha) \cdot r$ then p *must be* strictly preferred to q."

Observe, $\mathfrak{I}_2$ requires that p be strictly preferred to q even if *one* coin, with a given probability α of "heads", results in $\alpha \cdot p \oplus (1 - \alpha) \cdot r$ being strictly preferred to $\alpha \cdot q \oplus (1 - \alpha) \cdot r$. Now, consider the reverse implication of Independence–3 whose interpretation $\mathfrak{I}_3$ follows.

"Suppose the lottery $\alpha \cdot p \oplus (1-\alpha) \cdot r$ is strictly preferred to $\alpha \cdot q \oplus (1-\alpha) \cdot r$ for *every* coin whose probability α of showing up "heads" in a toss is strictly positive. Then, this strict preference must be attributed to a strict preference for p over q."

A comparison of $\mathfrak{I}_2$ and $\mathfrak{I}_3$ points out the following. First, $\mathfrak{I}_3$ holds whenever $\mathfrak{I}_0$ holds. It follows that logically Independence–0 is at least as strong as Independence–3. Second, Independence–3 is arguably more appealing than Independence–0 to a decision maker from a normative point of view. For comparing Independence–3 with Independence–0, one approach involves the following observation.

PROPOSITION 1: *Assume that $\succsim$ is a complete binary relation. Then, Independence–0 and Independence–2 are equivalent.*

PROOF: Let $\succsim$ be complete. With $p, q, r \in \mathscr{L}(Z)$ and $\alpha \in (0,1)$, let $s := \alpha \cdot p \oplus (1-\alpha) \cdot r$ and $t := \alpha \cdot q \oplus (1-\alpha) \cdot r$. Assume $p \sim q$. By Independence–0, $s \succ t$ implies $p \succ q$. As $\succ$ and $\sim$ are disjoint, $s \succ t$ is false. Similarly, $t \succ s$ does not hold. Since $\succsim$ is complete, $s \sim t$ holds. Thus, Independence–0 implies Independence–2.

Assume $s \succ t$. By Independence–2, $p \sim q$ implies $s \sim t$. As $\succ$ and $\sim$ are disjoint, $s \succ t$ implies $p \sim q$ does not hold. As $\succsim$ is complete, either $p \succ q$ or $q \succ p$ holds. By Independence–2, if $q \succ p$ then $t \succ s$. Then, $s \succ t$ contradicts the asymmetry of $\succ$. Hence, $p \succ q$ holds. Thus, Independence–2 implies Independence–0. ■

Thus, if the binary relation $\succsim$ is complete, logically Independence–2 is at least as strong as Independence–3. Moreover, it is arguable, for some decision makers, that the second implication in the statement of Independence–2 is a strong assumption.

To see this, we may change the point of view by requiring that the decision maker is modelled by an asymmetric binary relation $\succ$ over $\mathscr{L}(Z)$ as the *primitive*. Further, $\sim$ shall mean absence of $\succ$. Formally, we now *define* $\sim$ over $\mathscr{L}(Z)$ as follows:

$$p \sim q \quad \text{iff} \quad \big(\text{ not } p \succ q \text{ ; not } q \succ p \text{ }\big).$$

Notice, $\sim$ is symmetric. Then, $\succsim$ defined as $\succ \cup \sim$ is complete. This establishes the formal equivalence between the two approaches where one has $\succsim$ as the primitive and the other has $\succ$ as the primitive. Also observe, $\succsim$ is transitive iff $\succ$ is negatively–transitive.[27]

[27] Let R be a binary relation over X. Then, R is *transitive* if, $(xRy \text{ ; } yRz) \implies xRz$. Also, R is *negatively–transitive* if, $(\text{not } xRy \text{ ; not } yRz) \implies \text{not } xRz$.

From this point of view, consider a decision maker who is able to rank lotteries p and q according to $\succ$ if they are "close enough" but *not* if they are "far part". Then, the following may hold:

$$p \succ q \implies \big[\, \alpha \cdot p \oplus (1-\alpha)\cdot r \succ \alpha \cdot q \oplus (1-\alpha)\cdot r \,\big]$$

for *all* $\alpha \in (0,1)$ but the implication:

$$p \sim q \implies \big[\, \alpha \cdot p \oplus (1-\alpha)\cdot r \sim \alpha \cdot q \oplus (1-\alpha)\cdot r \,\big]$$

will *fail* to hold if $\alpha \in (0,1)$ is "small enough". Thus, Independence–2 ceases to hold. However, note that the following implication may still continue to hold for such a decision maker:

$$\big(\, \forall \alpha \in (0,1) \,\big)\big[\, \alpha \cdot p \oplus (1-\alpha)\cdot r \succ \alpha \cdot q \oplus (1-\alpha)\cdot r \,\big] \implies p \succ q.$$

This is because if the hypothesis in the above conditional holds, the lotteries p and q must be "close enough". That is, Independence–3 has appeal for such decision makers. Thus, a rigorous formulation of such binary relations will imply that Independence–3 is (1) logically *strictly* weaker and (2) more normatively appealing than Independence–0 and Independence–2 under the assumption of completeness. To fix ideas, we begin by presenting a simple example.

Example 1: Let $Z = \{z_1, z_2\}$ have two basic prizes. Fix $\theta \in (0, 1/\sqrt{2})$. A lottery $p \in \mathscr{L}(Z)$ is any map $p : Z \to \mathbb{R}_+$ such that $p(z_1)+p(z_2) = 1$. Define[28] the binary relations $\succ_\theta$ and $\sim_\theta$ over $\mathscr{L}(Z)$ as follows:

$$\begin{aligned} p \succ_\theta q \quad &\text{iff} \quad \big(\, \|p-q\|_2 \le \theta \;;\; p(z_1) > q(z_1) \,\big), \text{ and} \\ p \sim_\theta q \quad &\text{iff} \quad \big(\, \text{not } p \succ_\theta q \;;\; \text{not } q \succ_\theta p \,\big). \end{aligned}$$

Let $\succsim_\theta$ as $\succ_\theta \cup \sim_\theta$. Observe, $\succ_\theta$ is asymmetric and $\sim_\theta$ is symmetric. Note, $\succsim_\theta$ is complete. Also, notice the following:

$$p \sim_\theta q \quad \text{iff} \quad \big(\, \|p-q\|_2 > \theta \;\text{ or }\; p = q \,\big).$$

Let $p_*, q_*, r_* \in \mathscr{L}(Z)$ satisfy $p_*(z_1) = 1$, $q_*(z_1) = 0$ and $r_*(z_1) = 1/2$. Note, $\|p_* - q_*\|_2 = \sqrt{2} > \theta$. Thus, $p_* \sim_\theta q_*$. Also, let $\alpha_* := \theta/\sqrt{2}$ and pick any $\alpha \in (0, \alpha_*]$. Let $s_* := \alpha \cdot p_* \oplus (1-\alpha)\cdot r_*$ and $t_* := \alpha \cdot q_* \oplus (1-\alpha)\cdot r_*$. Thus, $s_*(z_1) = (1+\alpha)/2 = t_*(z_2)$ and $s_*(z_2) = (1-\alpha)/2 = t_*(z_1)$. Note, $s_*(z_1) > t_*(z_1)$ as $\alpha > 0$. Also, $\|s_* - t_*\|_2 = \alpha\sqrt{2} \le \theta$ as $\alpha \le \alpha_*$. Hence, $s_* \succ_\theta t_*$. That is, $\alpha \cdot p_* \oplus (1-\alpha)\cdot r_* \succ_\theta \alpha \cdot q_* \oplus (1-\alpha)\cdot r_*$. Therefore, $p_* \sim_\theta q_*$ implies: $\succsim_\theta$ does *not* satisfy Independence–2.

[28] We denote by $\|\cdot\|_2$ the *Euclidean norm* over $\mathbb{R}^Z$. Thus, $\|p-q\|_2 := \big(\sum_{z \in Z} |p(z) - q(z)|^2\big)^{1/2}$.

However, $\succsim_\theta$ satisfies Independence–3. For this, consider arbitrary lotteries p, q and r in $\mathscr{L}(Z)$ that satisfy the following:

$$\alpha \cdot p \oplus (1-\alpha) \cdot r \succ_\theta \alpha \cdot q \oplus (1-\alpha) \cdot r \quad \text{for all } \alpha \in (0,1).$$

Let $s_\alpha := \alpha \cdot p \oplus (1-\alpha) \cdot r$ and $t_\alpha := \alpha \cdot q \oplus (1-\alpha) \cdot r$ for any $\alpha \in [0,1]$. Note, $s_\alpha \succ_\theta t_\alpha$ implies $\|s_\alpha - t_\alpha\|_2 \leq \theta$. Observe, $\|s_\alpha - t_\alpha\|_2 = \alpha \|p-q\|_2$. As $s_\alpha \succ_\theta t_\alpha$ for all $\alpha \in (0,1)$, we have $\|p-q\|_2 \leq \theta$. Further, $s_\alpha \succ_\theta t_\alpha$ implies $s_\alpha(z_1) > t_\alpha(z_1)$. Note, $s_\alpha(z_1) = \alpha p(z_1) + (1-\alpha) r(z_1)$ and $t_\alpha(z_1) = \alpha q(z_1) + (1-\alpha) r(z_1)$. Thus, if $\alpha \in (0,1)$ then: $s_\alpha(z_1) > t_\alpha(z_1)$ iff $p(z_1) > q(z_1)$. Since $s_\alpha \succ_\theta t_\alpha$ for all $\alpha \in (0,1)$, we have $p(z_1) > q(z_1)$. Then, $\|p-q\|_2 \leq \theta$ and $p(z_1) > q(z_1)$ imply $p \succ_\theta q$. That is,

$$\big(\, \forall \alpha \in (0,1) \,\big)\big[\, \alpha \cdot p \oplus (1-\alpha) \cdot r \succ_\theta \alpha \cdot q \oplus (1-\alpha) \cdot r \,\big] \implies p \succ_\theta q.$$

For the converse, let p, q and r be lotteries with $p \succ_\theta q$. Pick an arbitrary $\alpha \in (0,1)$. Let $s_\alpha := \alpha \cdot p \oplus (1-\alpha) \cdot r$ and $t_\alpha := \alpha \cdot q \oplus (1-\alpha) \cdot r$. Note, $p \succ_\theta q$ implies $\|p-q\|_2 \leq \theta$. Since $\|s_\alpha - t_\alpha\|_2 = \alpha \|p-q\|_2$, we have: $\|s_\alpha - t_\alpha\|_2 \leq \theta$ for all $\alpha \in (0,1)$. Further, $p \succ_\theta q$ implies $p(z_1) > q(z_1)$. Since $s_\alpha(z_1) = \alpha p(z_1) + (1-\alpha) r(z_1)$ and $t_\alpha(z_1) = \alpha q(z_1) + (1-\alpha) r(z_1)$, we obtain: $s_\alpha(z_1) > t_\alpha(z_1)$ for all $\alpha \in (0,1)$. This proves the converse. That is, $\succsim_\theta$ satisfies Independence–3. ■

The binary relation $\succsim_\theta$ constructed in the above example satisfies Independence–3 but *not* Independence–2. Further, note that $\succsim_\theta$ is complete. Thus, by Proposition 1, $\succsim_\theta$ does *not* satisfy Independence–0. Moreover, observe that the following holds.

PROPOSITION 2: *Assume that $\succsim$ is a binary relation over $\mathscr{L}(Z)$. Then, Independence–0 implies Independence–3.*

PROOF: Let $p, q, r \in \mathscr{L}(Z)$ satisfy: $\alpha \cdot p \oplus (1-\alpha) \cdot r \succ \alpha \cdot q \oplus (1-\alpha) \cdot r$ for all $\alpha \in (0,1)$. Fix an arbitrary $\alpha_* \in (0,1)$. Then, $\alpha_* \cdot p \oplus (1-\alpha_*) \cdot r \succ \alpha_* \cdot q \oplus (1-\alpha_*) \cdot r$ holds. By Independence–0, $p \succ q$ follows. That is, the reverse implication of Independence–3 holds.

To establish the forward implication of Independence–3, let $p \succ q$ and $\alpha \in (0,1)$ be arbitrary. Then, $\alpha \cdot p \oplus (1-\alpha) \cdot r \succ \alpha \cdot q \oplus (1-\alpha) \cdot r$ by Independence–0. Since $\alpha \in (0,1)$ is *arbitrary*,

$$\big(\, \forall \alpha \in (0,1) \,\big)\big[\, \alpha \cdot p \oplus (1-\alpha) \cdot r \succ \alpha \cdot q \oplus (1-\alpha) \cdot r \,\big]$$

holds. That is, the forward implication of Independence–3 holds. Hence, Independence–0 implies Independence–3. ■

Propositions 1 and 2 together establish that, under the assumption of completeness, Independence–0 and Independence–2 are equivalent to each other but are logically at least as strong as Independence–3. However, Example 1 shows that Independence–0 and Independence–2 are in fact *strictly* stronger than Independence–3. While Example 1 has served its formal purpose, it is desirable to have a more general class of such binary relations which are in addition plausible models of decision makers. With this aim, we proceed as follows.

Definition 2: *An* affine screening criterion *is any non–constant map* $f : \mathscr{L}(Z) \to \mathbb{R}$ *such that*:

$$f(\alpha \cdot p \oplus [1-\alpha] \cdot q) = \alpha f(p) + [1-\alpha] f(q)$$

for any $p, q \in \mathscr{L}(Z)$ *and* $\alpha \in [0,1]$.

The numerical value $f(p)$ for the lottery p, by the screening function f, is *as if* a psychological "cost" incurred by the decision maker due to the contemplation necessary for comparing an arbitrary lottery to a reference lottery. The additional requirement of an"affine structure" on f captures "expected values" for random choice between lotteries. Denote by $\mathscr{F}$ the set of all affine screening criteria.

Definition 3: *A* filter *is any map* $\vartheta : \mathscr{F} \to \mathbb{R}_{++}$ *such that*:

$$f' = \alpha f + \beta \quad \textit{implies} \quad \vartheta(f') = \alpha \vartheta(f).$$

for any $f, f' \in \mathscr{F}$ *and* $(\alpha, \beta) \in \mathbb{R}_{++} \times \mathbb{R}$.

The answer to "Is $f(q) \leq f(p) + \vartheta(f)$?" dictates the feasibility of contemplation about q given the reference p. Suppose $f' = \alpha f + \beta$. Note, "$f'(q) \leq f'(p) + \vartheta(f')$" is equivalent to "$f(q) \leq f(p) + \vartheta(f)$", if and only if, $\vartheta(f') = \alpha \vartheta(f)$. Define R_ϑ over $\mathscr{L}(Z)$ as:

$$pR_\vartheta q \quad \text{iff} \quad \big(\, \forall f \in \mathscr{F} \,\big)\big[\, f(p) \leq f(q) + \vartheta(f) \,\big].$$

Also, let S_ϑ be the relation on $\mathscr{L}(Z)$ defined as:

$$pS_\vartheta q \quad \text{iff} \quad \big(\, \exists r \in \mathscr{L}(Z) \,\big)\big[\, pR_\vartheta r \;;\; qR_\vartheta r \,\big].$$

Note, S_ϑ is symmetric. An *affine order* is a total order $\succ_0$ on $\mathscr{L}(Z)$ such that $\succsim_0$ satisfies[29] Independence–3.

[29]Note, $p \sim_0 q$ iff $\big($ not $p \succ_0 q$; not $q \succ_0 p \big)$. Moreover, $\succsim_0$ is defined as $\succ_0 \cup \sim_0$.

Definition 4: *The* affine local preorder *induced by the filter* ϑ *and the affine order* $\succ_0$ *is the binary relation* $\succ_\vartheta$ *over* $\mathscr{L}(Z)$ *such that*:

$$p \succ_\vartheta q \quad \textit{iff} \quad \big(p \neq q \; ; \; pS_\vartheta q \; ; \; p \succ_0 q \big).$$

The affine local order *induced by the affine preorder* $\succ_\vartheta$ *is the binary relation* $\succsim_\vartheta$ *which is* $\succ_\vartheta \cup \sim_\vartheta$ *where* $\sim_\vartheta$ *is as follows*:

$$p \sim_\vartheta q \quad \textit{iff} \quad \big(\textit{not } p \succ_\vartheta q \; ; \; \textit{not } q \succ_\vartheta p \big).$$

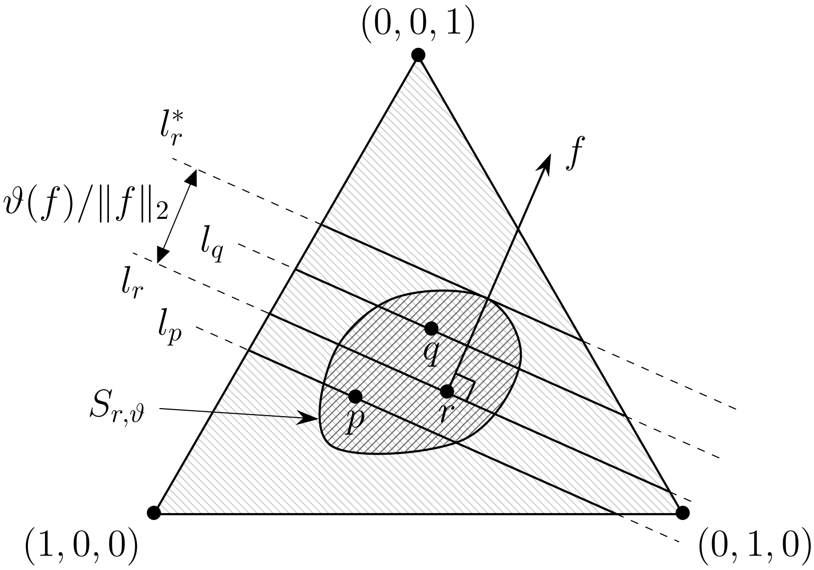


Figure 3: An affine *local* order.

Some remarks are in order. First, Theorem 2 implies that $\succ_0$ is an affine order, if and only if, there exists EU maps $u_1 \dots, u_{|Z|}$ such that (1) $u_1(z) = 1$ for all $z \in Z$, (2) $(u_1, \dots, u_{|Z|})$ are linearly independent as vectors in $\mathbb{R}^Z$, and (3) the following holds:

$$p \succ_0 q \quad \text{iff} \quad [u_1(p), \dots, u_{|Z|}(p)] >_L [u_1(q), \dots, u_{|Z|}(q)],$$

where $>_L$ is the strict component of the lexicographic order $\geq_L$ on $\mathbb{R}^{|Z|}$. Observe, (2) is critical for $\succ_0$ to be a total order.

Second, to see what definition 4 entails, consider Figure 3. Each affine screening criterion f defines a family of parallel straight lines, with f perpendicular to them, with f a constant on each. For instance, l_r and l_r^* restricted to $\mathscr{L}(Z)$ are the sets $\{q' \in \mathscr{L}(Z) : f(q') = f(r)\}$ and $\{q' \in \mathscr{L}(Z) : f(q') = f(r) + \vartheta(f)\}$, respectively. Let $S_{r,\vartheta}$ be the subset of lotteries p which satisfy $f(p) \leq f(r) + \vartheta(f)$ for *every* f. Thus, $pS_\vartheta q$ because p and q are in $S_{r,\vartheta}$. Then, $p \succ_\vartheta q$ iff $p \succ_0 q$. Observe, $\succ_0$ is a *total* order over $\mathscr{L}(Z)$ but $\succ_\vartheta$ is *local* in nature.

Notice, the set $S_{r,\vartheta}$ is shown to be compact in Figure 3. This need not be so for an arbitrary filter ϑ. However, a "continuity" requirement on ϑ is sufficient to ensure the compactness of the resulting set $S_{r,\vartheta}$ for any lottery r. To formulate this notion of "continuity", we begin by specifying a natural notion of convergence for seuences of affine screening criteria. For this, consider any $\mathscr{F}$–valued sequence (f_n) and any f_* in $\mathscr{F}$. Then, we say that (f_n) *converges* to f_* if:

$$\lim_{n\to\infty} f_n(p) = f_*(p) \quad \text{for every } p \in \mathscr{L}(Z).$$

We shall write "$f_n \to f_*$" for the phrase "f_n converges to f_*". Then, a filter ϑ is *continuous* if, $\lim_{n\to\infty} \vartheta(f_n) = \vartheta(f_*)$ for every $\mathscr{F}$–valued sequence (f_n) and f_* in $\mathscr{F}$ satisfying $f_n \to f_*$. The set $S_{r,\vartheta}$ is compact, for any lottery r, if ϑ is continuous. For any $\kappa > 0$ and any filter ϑ, let the map $\kappa \cdot \vartheta$ from $\mathscr{L}(Z)$ to $\mathbb{R}_{++}$ be defined as follows:

$$[\kappa \cdot \vartheta](f) \coloneqq \kappa\vartheta(f) \quad \text{for all } f \in \mathscr{F}.$$

PROPOSITION 3: *Let ϑ be a filter and $\succ_0$ be an affine order on $\mathscr{L}(Z)$. If $\succsim_\vartheta$ is the affine local order induced by ϑ and $\succ_0$ then*:

1. *$\succ_\vartheta$ is acyclic.*
2. *$\succsim_\vartheta$ satisfies Independence–3.*
3. *If ϑ is continuous then there exists $\kappa_\vartheta > 0$ such that $\succsim_{\kappa\cdot\vartheta}$ violates Independence–0 and Independence–2 for all $\kappa \in (0, \kappa_\vartheta)$.*

Propositions 1, 2 and 3 show that Independence–3 is indeed strictly weaker, under completeness, than Independence–0 or Independence–2. Thus, our characterization (that is, Theorem 2) of preferences which admit lexicographic expected utility representations is stronger than Hausner's theorem. Moreover, affine local orders are *not* covered by the class of binary relations which admit "coalitional expected multi–utility representations". The latter is the most general class of binary relations satisfying Independence–2 as was shown in HARA ET AL. (2019). The proof of Proposition 3 is in section A.II.1 of the Appendix.

We close this subsection with one remark. While our reason for introducing the class of affine local orders has been to show that our version of Independence is strictly weaker than the classical version, we believe that they are natural as models of decision makers. The recent work in choice via screening sets or attention filters, as considered in MANZINI & MARIOTTI (2007, 2014) and MASATLIOGLU ET AL. (2012) for instance, motivates this point of view.

3.3 *Decision Theoretic Implications*

The Independence axiom of von Neumann & Morgenstern (1944), further investigated in Marschak (1950), Samuelson (1952) and Herstein & Milnor (1953), is at the foundation of expected utility theory. However, the Expected Utility Hypothesis has been criticised due to the preference reversals as in Allais (1953). Therefore, some authors have weakened Independence; see Chew (1953), Chew et al. (1987), Chew et al. (1991), Dekel (1986) and Quiggin (1982) for instance. Some authors such as Machina (1982) have abandoned it altogether notwithstanding its normative appeal.

However, few authors have considered retaining Independence but relaxing other axioms such as completeness, transitivity or Continuity. For instance, Aumann (1962) relaxed completeness to characterize "one–way" representations and Hausner (1954) relaxed Continuity. More recently, completeness was relaxed by Dubra et al. (2004) to obtain a sharper "two–way" characterization. Further, Hara et al. (2019) considered adding these axioms progressively but their analysis largely retains the completeness axiom.

Allais type paradoxes have highlighted the "preference reversal phenomenon"; see Grether & Plott (1979), Holt (1986), Karni & Safra (1987), Pommerehne et al. (1982), Slovic & Lichtenstein (1983) and Tversky et al. (1990) for instance. Further, violations of transitivity have been investigated by Tversky (1969), Loomes et al. (1991) and Regenwetter et al. (2011) for instance. The original findings of these two strands in the literature have been questioned and re–examined later.

In Azrieli et al. (2018), a theoretical analysis has been provided of how experiments must be conducted for testing the validity or violations of axioms such that incentive and other problems are properly taken into account. Based on this analysis, Nielsen & Rehbeck (2022) found in their experiments that violations of assumptions such as transitivity or Independence are "mistakes" by individuals which they correct once explained. This suggests that perhaps preference reversals and Allais type paradoxes should be re–examined by conducting experiments designed along the above lines.

Even if Allais type paradoxes persist, then it is the Expected Utility Hypothesis but not just the Independence axiom which comes under question. This point has been emphasized by Uzi Segal for instance. The Reduction Axiom was relaxed in Segal (1988, 1990). Moreover, a weaker and non–testable version of Independence has been proposed in Segal (2023) which together with Continuity (and Monotonicity) is equivalent to the Expected Utility Hypothesis. In what follows, we sharpen the analysis in Segal (2023) via our Theorem 2.

Let us briefly recall Segal's analysis. He considers a complete and transitive binary relation $\succsim$, with $\succ$ and $\sim$ as its asymmetric and symmetric components, respectively. Further, he introduces the following weakening of the classical Independence axiom.

Weak Independence–0: *For every $p, q, r \in \mathscr{L}(Z)$, if $p \sim q$ then*:

$$\big(\exists\, \alpha \in (0,1)\,\big)\big[\,\alpha \cdot p \oplus (1-\alpha)\cdot r \sim \alpha \cdot q \oplus (1-\alpha)\cdot r\,\big].$$

Observe, this axiom is non–testable. Further, Continuity is non–testable but Monotonicity is testable. Segal proves the following.

Theorem (Segal, 2023): *Let $\succsim$ satisfy completeness and transitivity. Then, $\succsim$ satisfies Monotonicity, Continuity and Weak Independence–0, if and only if, $\succsim$ admits an EU representation.*

With this result in place, he argues that Allais type paradoxes imply a violation of the Expected Utility Hypothesis. However, this does not violate Weak Independence–0 but does falsify the combination of all the assumptions in the above theorem. In particular, the non–testability of Weak Independence–0 anyway makes it irrefutable. Furthermore, this axiom retains the normative appeal of classical Independence.

However, Continuity is another non–testable axiom in the above theorem and note that the conjunction of more than one non–testable axioms can result in testable implications. Furthermore, Continuity is an axiom which is not in the spirit of Independence — the latter being a "cancellation" property whereas the former is a "regularity" condition with technical motivations. Our objective will be to sharpen Segal's conclusion but based only on "cancellation" type axioms. To this end, we begin by introducing the following axiom.

Weak Independence–1: *For every $p, q, r \in \mathscr{L}(Z)$, if $p \succ q$ then*:

$$\big(\exists\, \alpha \in (0,1)\,\big)\big[\,\alpha \cdot p \oplus (1-\alpha)\cdot r \succ \alpha \cdot q \oplus (1-\alpha)\cdot r\,\big].$$

Three observations follow. First, this axiom is a Segal type version of Independence. Second, it is non–testable. Third, the *conjunction* of Weak Independence–0 and Weak Independence–1 (henceforth, "Weak Independence") is also non–testable. To see why, observe that the asymmetry of $\succ$ and the symmetry of $\sim$ implies that at most one of $p \succ q$ or $p \sim q$ holds for any $p, q \in \mathscr{L}(Z)$. Thus, there is no instance where the antecedents in the implications of Weak Independence–0 and Weak Independence–1 hold simultaneously. In other words, when one axiom binds, the other does does not.

Dropping Continuity necessitates some other axiom that retains its flavor just enough so that lexicographic expected utility (LEU) representations exist which also suffer from Allais type preference reversals. Further, we constrain such axioms to be "cancellation" type statements. One such axiom is as follows.

Global Monotonicity: *For every $p, q, r \in \mathscr{L}(Z)$, if $p \succ q$ then*:

$$\big(\exists\, \alpha \in (0,1)\big)\big[\,\alpha \cdot p \oplus (1-\alpha)\cdot r \succ \alpha \cdot q \oplus (1-\alpha)\cdot r\,\big]$$
$$\Downarrow$$
$$\big(\forall\, \alpha \in (0,1)\big)\big[\,\alpha \cdot p \oplus (1-\alpha)\cdot r \succ \alpha \cdot q \oplus (1-\alpha)\cdot r\,\big].$$

Observe, this axiom is testable. The result is as follows.

Theorem 3: *Assume $\succsim$ satisfies completeness and transitivity. Then, $\succsim$ satisfies Weak Independence and Global Monotonicity, if and only if, $\succsim$ admits an LEU representation.*

Proof: Necessity of the axioms is obvious. For sufficiency, assume $\succsim$ satisfies Weak Independence and Global Monotonicity in addition to completeness and transitivity. We argue: $\succsim$ satisifies Independence–3. Then, Theorem 2 (subsection 3.1) completes the proof.

First, fix any $p, q, r \in \mathscr{L}(Z)$ such that $p \succ q$. Since $p \succ q$, Weak Independence–1 implies $\alpha \cdot p \oplus (1-\alpha)\cdot r \succ \alpha \cdot q \oplus (1-\alpha)\cdot r$ for some $\alpha \in (0,1)$. Then, $p \succ q$ and Global Monotonicity implies:

$$\big(\forall\, \alpha \in (0,1)\big)\big[\,\alpha \cdot p \oplus (1-\alpha)\cdot r \succ \alpha \cdot q \oplus (1-\alpha)\cdot r\,\big].$$

To complete the proof, we now fix any $p, q, r \in \mathscr{L}(Z)$ such that the above statement holds. We must argue: $p \succ q$. Suppose, not! Thus, either $q \succ p$ or $p \sim q$ holds by completeness. If $q \succ p$ holds then Weak Independence–1 implies, there exists $\alpha \in (0,1)$ such that $\alpha \cdot q \oplus (1-\alpha)\cdot r \succ \alpha \cdot p \oplus (1-\alpha)\cdot r$. This contradicts the asymmetry of $\succ$. Hence, $p \sim q$ must hold. Then, Weak Independence–0 implies, there exists $\alpha \in (0,1)$ such that $\alpha \cdot p \oplus (1-\alpha)\cdot r \sim \alpha \cdot q \oplus (1-\alpha)\cdot r$. However, this is also a contradiction because $\succ$ is asymmetric and $\sim$ is symmetric. Thus, our supposition must be wrong. ■

Observe, Global Monotonicity is testable and recall Weak Independence is not. Further, both are "cancellation" properties inherited from classical Independence. Then, Allais type paradoxes may refute Global Monotonicity but not Weak Independence. Thus, Theorem 3 dissects Independence into irrefutable and refutable components.

However, this raises the following question: what aspect of Continuity, together with Weak Independence and Monotonicity as in Segal (2023), condenses Global Monotonicity? For an answer, two further "cancellation" type axioms are introduced as follows.

Inward Monotonicity: *For every* $p, q, r \in \mathscr{L}(Z)$ *and for every* $\alpha_* \in (0,1)$, *if* $p \succ q$ *then*:

$$\big[\,\alpha_* \cdot p \oplus (1-\alpha_*) \cdot r \succ \alpha_* \cdot q \oplus (1-\alpha_*) \cdot r\,\big]$$
$$\Downarrow$$
$$\big(\forall\, \alpha \in (\alpha_*, 1)\,\big)\big[\,\alpha \cdot p \oplus (1-\alpha) \cdot r \succ \alpha \cdot q \oplus (1-\alpha) \cdot r\,\big].$$

Consider the implication in the above axiom. The universal quantifier is in its consequent (as opposed to its antecedent). Hence, this axiom is testable. The second axiom is as follows.

Outward Monotonicity: *For every* $p, q, r \in \mathscr{L}(Z)$ *and for every* $\alpha_* \in (0,1)$, *if* $p \succ q$ *then*:

$$\big(\forall\, \alpha \in (\alpha_*, 1)\,\big)\big[\,\alpha \cdot p \oplus (1-\alpha) \cdot r \succ \alpha \cdot q \oplus (1-\alpha) \cdot r\,\big]$$
$$\Downarrow$$
$$\big[\,\alpha_* \cdot p \oplus (1-\alpha_*) \cdot r \succ \alpha_* \cdot q \oplus (1-\alpha_*) \cdot r\,\big].$$

In contrast to the Inward Monotonicity, this axiom is non–testable as the universal quantifier now is in the antecedent (as opposed to the consequent) of the implication. The result is as follows.

Theorem 4: *Assume* $\succsim$ *is complete and transitive. Then,* $\succsim$ *satisfies Weak Independence, Inward Monotonicity and Outward Monotonicity, if and only if,* $\succsim$ *admits an LEU representation.*

Proof: Necessity of the axioms is obvious. For sufficiency, let $\succsim$ be complete and transitive, and satisfies Weak Independence, Inward Monotonicity and Ouward Monotonicity. We argue: $\succsim$ satisfies Global Monotonicity. Then, Theorem 3 completes the proof.

Fix any $p, q, r \in \mathscr{L}(Z)$ such that $p \succ q$, and assume there exists $\alpha_1 \in (0,1)$ such that $\alpha_1 \cdot p \oplus (1-\alpha_1) \cdot r \succ \alpha_1 \cdot q \oplus (1-\alpha_1) \cdot r$. That is, $A := \{\alpha \in (0,1) : \alpha \cdot p \oplus (1-\alpha) \cdot r \succ \alpha \cdot q \oplus (1-\alpha) \cdot r\}$ is nonempty. Let $\alpha_* := \inf A$. Suppose $\alpha_* > 0$. By Inward Monotonicity, $(\alpha_*, 1) \subseteq A$. Outward Monotonicity implies $[\alpha_*, 1) = A$. Then, Weak Independence–1 implies, $\alpha < \alpha_*$ for some $\alpha \in A$ which contradicts $\alpha_* = \inf A$. Thus, $\alpha_* = 0$ proving Global Monotonicity. ■

All the axiom systems we have considered thus far are only as strong as the conjunction of classical Independence with completeness and transitivity. Further, every axiom which has been introduced is of the "cancellation" type which ensures that they retain the normative appeal of Independence. However, Global Monotonicity is too strong to be compatible with the Allais paradox. In particular, those nonlinear expected utility preferences characterized in DEKEL (1986) which are consistent with the Allais paradox violate this axiom. Further, while Inward Monotonicity is weaker than Global Monotonicty, it is also subject to the same criticism. Notice, no such axiom appears in Segal's theorem. Therefore, we introduce the following axiom.

AFFINE CONTINUITY: *For any $p, q, r \in \mathscr{L}(Z)$ and any $\alpha_* \in (0,1)$, if $p \succ q$ and $\alpha_* \cdot p \oplus (1-\alpha_*) \cdot r \succ \alpha_* \cdot q \oplus (1-\alpha_*) \cdot r$ then there exists $\varepsilon > 0$ such that the following holds*:

$$\alpha > 1 - \varepsilon \quad \Longrightarrow \quad \alpha \cdot p \oplus (1-\alpha) \cdot r \succ \alpha \cdot q \oplus (1-\alpha) \cdot r.$$

Observe, this is a non–testable axiom and is weaker than standard Continuity. Observe, it is satisfied by all preferences characterized in DEKEL (1986). In particular, it is not refuted by the Allais paradox. Further, this axiom is compatible with LEU preferences but Continuity is not. Then, we obtain the following result.

THEOREM 5: *Assume $\succsim$ is complete and transitive. Then, $\succsim$ satisfies Weak Independence, Affine Continuity and Outward Monotonicity, if and only if, $\succsim$ admits an LEU representation.*

PROOF: Necessity of the axioms is obvious. For sufficiency, it is enough to show that Inward Monotonicity holds for then Theorem 4 implies the claim. So, fix any $p, q, r \in \mathscr{L}(Z)$ and $\alpha_* \in (0,1)$ such that $p \succ q$ and $\alpha_* \cdot p \oplus (1-\alpha_*) \cdot r \succ \alpha_* \oplus (1-\alpha_*) \cdot r$. Let $A \subseteq [\alpha_*, 1)$ be the set of those α such that the following holds:

$$\alpha' > \alpha \quad \Longrightarrow \quad \alpha' \cdot p \oplus (1-\alpha') \cdot r \succ \alpha' \cdot q \oplus (1-\alpha') \cdot r.$$

Affine Continuity implies A is nonempty. Let $\alpha_{**} := \inf A$. Clearly, $\alpha_{**} \geq \alpha_*$. The proof is complete if $\alpha_{**} = \alpha_*$. Suppose, $\alpha_{**} > \alpha_*$. Then, Outward Monotonicity implies $\alpha_{**} \cdot p \oplus (1-\alpha_{**}) \cdot r \succ \alpha_{**} \cdot q \oplus (1-\alpha_{**}) \cdot r$. By Weak Independence, $\alpha^\dagger \cdot p \oplus (1-\alpha^\dagger) \cdot r \succ \alpha^\dagger \cdot q \oplus (1-\alpha^\dagger) \cdot r$ for some $\alpha^\dagger \in (0, \alpha_{**})$. Thus, Affine Continuity implies the existence of $\varepsilon > 0$ such that $\alpha_{**} - \varepsilon \in A$. This contradicts the fact that $\alpha_{**} = \inf A$. Hence, $\alpha_{**} = \alpha_*$ proving Inward Monotonicity. ■

Since LEU preferences are equivalent to classical Independence as shown by HAUSNER (1954), it follows that the above system of axioms is not stronger that classical Independence. Note, each axiom is non–testable and is a "cancellation" type statement. Further, all axioms except Weak Independence are satisfied by every preference characterised in DEKEL (1986). However, preferences that admit LEU representations are refuted by the Allais paradox.

Theorem 5 allows us to characterize preferences that admit expected utility representations. In particular, we replace "monotonicty" as required by SEGAL (2023) with the following weaker axiom.

EXCLUSIVITY: *For any $p, q \in \mathscr{L}(Z)$ and any $\alpha \in (0,1)$,*

$$\neg\big(\, p \sim q \,\big) \quad \implies \quad \neg\big(\, \alpha \cdot p \oplus (1-\alpha) \cdot q \sim q \,\big).$$

Notice that Exclusivity is testable. The result is as follows.

THEOREM 6: *A binary relation satisfies completeness, transitivity, Weak Independence–0, Exclusivity and Continuity, if and only if, it admits an expected utility representation.*

PROOF: Necessity of the axioms is obvious. Our strategy for sufficiency will be as follows. We show that Weak Independence–1 is implied by Exclusivity, Continuity and Weak Independence–0. Observe, Affine Continuity follows from Continuity. If Outward Monotonicity is shown to hold, then Theorem 5 implies that the preference admits an LEU representation. Note, the only LEU preferences satisfying Continuity are those which admit expected utility representations. Hence, it is enough to argue that Outward Monotonicity follows from the axioms. This shall be done through the following steps.

Step 1 — We shall show that Weak Independence–0 and Continuity imply the following: for any $p, q \in \mathscr{L}(Z)$,

$$p \sim q \implies \big(\forall\, \alpha \in (0,1)\big)\big[\, p \succ \alpha \cdot p \oplus (1-\alpha) \cdot q \sim q \,\big].$$

Suppose $p \sim q$ and $\alpha^{\dagger} \in (0,1)$ satisfy $\alpha^{\dagger} \cdot p \oplus (1-\alpha^{\dagger}) \cdot q \succ p$. Let $\mathscr{I}$ be the class of all intervals $I \subseteq [0,1]$ which contain $\alpha^{\dagger}$ and satisfy: $\alpha \cdot p \oplus (1-\alpha) \cdot q \succ p$ for all $\alpha \in I$. Let I_* be the union of the intervals in $\mathscr{I}$. Thus, I_* is the maximal element in $\mathscr{I}$ according to set–inclusion. Continuity implies I_* has a nonempty interior. Let $\alpha_* := \inf I_*$ and $\alpha^* := \sup I_*$. Thus, $\alpha_*, \alpha^* \in [0,1]$ satisfy $\alpha_* < \alpha^*$ and $I_* = [\alpha_*, \alpha^*]$. Define $p_* := \alpha_* p \oplus (1-\alpha_*) \cdot q$ and $q_* := \alpha^* p \oplus (1-\alpha^*) \cdot q$. Note, $\alpha \cdot p_* \oplus (1-\alpha) \cdot q_* \succ p$ for all $\alpha \in (0,1)$.

Continuity then implies $p_* \succsim p$ and $q_* \succsim p$. If at least one of $p_* \succ p$ or $q_* \succ p$ holds, then Continuity would imply a contradiction to the maximality of I_* in $\mathscr{I}$. Hence, $p_* \sim p$ and $q_* \sim p$. Transitivity implies (1) $p_* \sim q_*$, and (2) $\alpha \cdot p_* \oplus (1-\alpha) \cdot q_* \succ p_*$ for all $\alpha \in (0,1)$. Since $\succ$ is asymmetric but $\sim$ is symmetric, $p_*, q_* \in \mathscr{L}(Z)$ violate:

$$p_* \sim q_* \implies \big(\exists\, \alpha \in (0,1)\big)\big[\, p_* \sim \alpha \cdot p_* \oplus (1-\alpha) \cdot q_* \sim q_* \,\big].$$

However, this contradicts Weak Independence–0. showing that our supposition is impossible. Similarly, there does not exist $p, q \in \mathscr{L}(Z)$ and $\alpha^\dagger \in (0,1)$ such that $p \sim q$ and $p \succ \alpha^\dagger \cdot p \oplus (1-\alpha^\dagger) \cdot q$. Hence, we have established the claim made in this step.

Step 2 — We shall show that Weak Independence–0, Exclusivity and Continuity imply: for any $p, q \in \mathscr{L}(Z)$,

$$p \succ q \implies \big(\forall\, \alpha \in (0,1)\big)\big[\, p \succ \alpha \cdot p \oplus (1-\alpha) \cdot q \succ q \,\big].$$

Notice, the claim is that Weak Independence–1 holds. Fix any $p, q \in \mathscr{L}(Z)$. Suppose $\alpha_* \cdot p \oplus (1-\alpha_*) \cdot q \succ p$ for some $\alpha_* \in (0,1)$. Then, by Continuity, $p \succ q$ implies the existence of $\alpha' \in (0, \alpha_*)$ such that $\alpha' \cdot p \oplus (1-\alpha') \cdot q \sim p$. By step 1, $\alpha \cdot p \oplus (1-\alpha) \cdot q \sim p$ for all $\alpha \in [\alpha', 1]$. Note, $\alpha_* \in [\alpha', 1]$. Thus, $\alpha_* \cdot p \oplus (1-\alpha_*) \cdot q \sim p$. But $\succ$ is asymmetric and $\sim$ is symmetric. Thus, we have a contradiction. Hence, $p \succsim \alpha \cdot p \oplus (1-\alpha) \cdot q$ for all $\alpha \in (0,1)$. Similarly, $\alpha \cdot p \oplus (1-\alpha) \cdot q \succsim q$ for all $\alpha \in (0,1)$. That is, the following holds:

$$p \succsim \alpha \cdot p \oplus (1-\alpha) \cdot q \succsim q \quad \text{for every } \alpha \in (0,1).$$

Fix an arbitrary $\alpha \in (0,1)$. Let $r \coloneqq \alpha \cdot p \oplus (1-\alpha) \cdot q$. Note, $p \succ q$ implies $p \sim q$ fails. Then, Exclusivity implies $r \sim p$ fails. Since $p \succsim r$, we obtain: $p \succ r$. Further, $p \succ q$ implies $q \sim p$ fails. Then, Exclusivity implies $r \sim q$ fails. Since $r \succsim q$, we obtain: $r \succ q$. Thus, we have: $p \succ r \succ q$. Finally, note that $\alpha \in (0,1)$ is arbitrary.

Step 3 — We establish Outward Monotonicity. Fix $p, q, r \in \mathscr{L}(Z)$ and $\alpha_* \in (0,1)$. Assume $\alpha \cdot p \oplus (1-\alpha) \cdot r \succ \alpha \cdot q \oplus (1-\alpha) \cdot r$ for all $\alpha \in (\alpha_*, 1]$. Let $p_* \coloneqq \alpha_* \cdot p \oplus (1-\alpha_*) \cdot r$ and $q_* \coloneqq \alpha_* \cdot q \oplus (1-\alpha_*) \cdot r$. Then, Continuity implies $p_* \succsim q_*$. We must argue: $p_* \succ q_*$. Suppose not! Thus, $p_* \sim q_*$ holds. First, we rule out some cases.

If $p \sim r$ and $r \succ q$, then steps 1 and 2 imply $p_* \sim p$ and $q_* \sim q$. Thus, $p \succ q$ implies $p_* \succ q_*$. Hence, the case $p \sim r \succ q$ is ruled out. Similarly, the case $p \succ r \sim q$ is ruled out. Also, the case $p \succ r \succ q$ is ruled out by step 2 and transitivity.

Thus, the cases that remain are (1) $r \succ p \succ q$ and (2) $p \succ q \succ r$. Henceforth, assume $r \succ p \succ q$ (the other case is symmetric).

Suppose $p = \alpha^\dagger \cdot r \oplus (1-\alpha^\dagger) \cdot q$ for some $\alpha^\dagger \in (0,1)$. Clearly, such an $\alpha^\dagger$ is unique. Define $\alpha_1 := \alpha_* \alpha^\dagger + (1-\alpha_*)$. Notice, $\alpha_1 \in (0,1)$ as $\alpha^\dagger, \alpha_* \in (0,1)$. Observe that $p_* = \alpha_1 \cdot r \oplus (1-\alpha_1) \cdot q$. Then, $r \succ q$ and step 2 imply $p_* \succ q$. Define $\alpha_2 := (1-\alpha_*)/[\alpha_* \alpha^\dagger + (1-\alpha_*)]$. Note, $\alpha^\dagger, \alpha_* \in (0,1)$ implies $\alpha_2 \in (0,1)$. Further, observe that $q_* = \alpha_2 \cdot p_* \oplus (1 - \alpha_2 \cdot q$. Since $p_* \succ q$, step 2 implies $p_* \succ q_*$ contradicting $p_* \sim q_*$. Thus, $p \neq \alpha \cdot r \oplus (1-\alpha) \cdot q$ for all $\alpha \in (0,1)$.

Suppose $q = \alpha^\dagger \cdot p \oplus (1-\alpha^\dagger) \cdot r$ for some $\alpha^\dagger \in (0,1)$. Clearly, $\alpha^\dagger$ is unique. Recall, $p_* = \alpha_* \cdot p \oplus (1-\alpha_*) \cdot r$ where $\alpha_* \in (0,1)$. By $r \succ p$ and step 2, we have $r \succ p_*$. Observe that $q_* = \alpha^\dagger \cdot p_* \oplus (1-\alpha^\dagger) \cdot r$. By $r \succ p_*$ and step 2, we obtain $q_* \succ p_*$ which contradicts $p_* \sim q_*$. Thus, we have: $q = \alpha \cdot p \oplus (1-\alpha) \cdot r$ for every $\alpha \in (0,1)$.

Now, suppose $r = \alpha^\dagger \cdot p \oplus (1-\alpha^\dagger) \cdot q$ for some $\alpha^\dagger \in (0,1)$. By $p \succ q$ and step 2, we have $p \succ r$ which contradicts $r \succ p$. Thus, we have: $r \neq \alpha \cdot p \oplus (1-\alpha) \cdot q$ for all $\alpha \in (0,1)$. Denote by Δ_0 the simplex of all lotteries $\alpha_1 \cdot p \oplus \alpha_2 \cdot q \oplus \alpha_3 \cdot r$, where $(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}_+$ such that $\alpha_1 + \alpha_2 + \alpha_3 = 1$. Together with the conclusions of the previous two paragraphs, we obtain: Δ_0 is a 2–simplex.

For any $s \in \Delta_0$, let $I(s) := \{t \in \Delta_0 : t \sim s\}$. Further, let F_0, F_1 and F_2 be the "faces" of Δ_0 defined by the respective pairs (r,q), (r,p) and (p,q). Formally, F_0, F_1 and F_2 are defined as follows:

$$\begin{aligned} F_R &:= \{\alpha \cdot r \oplus (1-\alpha) \cdot q : \alpha \in (0,1)\}, \\ F_1 &:= \{\alpha \cdot r \oplus (1-\alpha) \cdot p : \alpha \in (0,1)\}, \\ F_2 &:= \{\alpha \cdot p \oplus (1-\alpha) \cdot q : \alpha \in (0,1]\}. \end{aligned}$$

They are pairwise disjoint. Let $F_L := F_1 \cup F_2$. Fix any $s \in \Delta_0 \setminus \{q, r\}$. Continuity and steps 1–2 imply: there exists a unique $(s_L, s_R) \in F_L \times F_R$ such that $I(s) = \{\alpha \cdot s_R \oplus (1-\alpha) \cdot s_L : \alpha \in [0,1]\}$.

Let $\mathbf{d}$ and $\mathbf{d}_*$ be the "direction vectors" of $I(p)$ and $I(p_*)$. Since $r \succ p \succ q$, Continuity implies $p \sim s$ for some $s \in F_R$. Thus, $\mathbf{d} \neq \mathbf{d}_*$. Pick $(\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}_+$ such that $\alpha_1 + \alpha_2 + \alpha_3 = 1$, and let $r_* := \alpha_1 \cdot p_* \oplus \alpha_2 \cdot q_* \oplus \alpha_3 \cdot r$. Let $\mathbf{d}^\dagger$ be the "direction vector" of $I(r_*)$. By Weak Independence–0 and Continuity $\mathbf{d}^\dagger = \mathbf{d}$. Similarly, $\mathbf{d}^\dagger = \mathbf{d}_*$. Thus, $\mathbf{d} = \mathbf{d}_*$ which contradicts $\mathbf{d} \neq \mathbf{d}_*$. ■.

We close this subsection with one final remark. All testability claims made in this subsection can be formalised in the sense of Chambers et al. (2017). In particular, they outline all formal statements of a particular form as testable by using the model theoretic framework as proposed in Chambers et al. (2014).

First, we "geometrize" the problem. Let $\phi : Z \to N := \{1, \ldots, n\}$ be a bijection, where $n := |Z|$. Let $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be the standard inner–product over $\mathbb{R}^n$, $\mathbb{1}$ be the "all ones" vector and $\mathbf{e}_i$ be the ith standard basis vector of $\mathbb{R}^n$. Then, $\Delta := \{\mathbf{x} \in \mathbb{R}^n_+ : \langle \mathbf{x}, \mathbb{1} \rangle = 1\}$ is the $(n-1)$–dimensional unit simplex.

The enumeration ϕ induces the bijection $p \in \mathscr{L}(Z) \mapsto \mathbf{p} \in \Delta$, where $\mathbf{p} = \sum_{i=1}^{n} \langle \mathbf{e}_i, \mathbf{p} \rangle \mathbf{e}_i$ with $\langle \mathbf{e}_i, \mathbf{p} \rangle := [p \circ \phi^{-1}](i)$ for all $i \in N$. Since the inner–product is bilinear, observe that the compound lottery $\alpha \cdot p \oplus (1-\alpha) \cdot q$ is mapped to the vector $\alpha \mathbf{p} + (1-\alpha)\mathbf{q}$. The preference $\succsim$ over $\mathscr{L}(Z)$ induces a preference $\succsim^*$ on Δ as: $p \succsim q \iff \mathbf{p} \succsim^* \mathbf{q}$. Then, Independence–3 of $\succsim$ translates to that of $\succsim^*$ as follows:

$$\big[\mathbf{p} \succ^* \mathbf{q}\big] \quad \text{iff} \quad \big(\, \forall \alpha \in (0,1) \,\big)\big[\, \alpha \mathbf{p} + (1-\alpha)\mathbf{r} \succ^* \alpha \mathbf{q} + (1-\alpha)\mathbf{r} \,\big]$$

The enumeration ϕ also induces a bijection of EUs to vectors in $\mathbb{R}^n$ as $u \in \mathscr{E}(Z) \mapsto \mathbf{u} \in \mathbb{R}^n$, where $\mathbf{u} = \sum_{i=1}^{n} \langle \mathbf{e}_i, \mathbf{u} \rangle \mathbf{e}_i$ with $\langle \mathbf{e}_i, \mathbf{u} \rangle := [u \circ \phi^{-1}](i)$ for every $i \in N$. The bijections imply the crucial property:

$$u(p) = \langle \mathbf{u}, \mathbf{p} \rangle \text{ for every } u \in \mathscr{E}(Z) \text{ and any } p \in \mathscr{L}(Z).$$

Let $\mathbf{a} := \mathbb{1}/n$ be the *centroid* of Δ, and $O_{\mathbb{1}} := \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbb{1}, \mathbf{x} \rangle = 0\}$ be the orthogonal subspace in $\mathbb{R}^n$ to the vector $\mathbb{1}$. Let $\mathbf{u}_\perp$ be the orthogonal projection of $\mathbf{u} \in \mathbb{R}^n$ onto $O_{\mathbb{1}}$. Thus, $\langle \mathbf{p}-\mathbf{q}, \mathbf{u} \rangle = \langle \mathbf{p}-\mathbf{q}, \mathbf{u}_\perp \rangle$ if $\mathbf{p}, \mathbf{q} \in \Delta$. Further, $\mathbf{p}_\perp := \mathbf{p} - \mathbf{a}$ is the orthogonal projection of $\mathbf{p}$ onto $O_{\mathbb{1}}$ because $\langle \mathbf{p}, \mathbb{1} \rangle = 1 = \langle \mathbf{a}, \mathbb{1} \rangle$. Then, $\langle \mathbf{p} - \mathbf{q}, \mathbf{u}_\perp \rangle = \langle \mathbf{p}_\perp - \mathbf{q}_\perp, \mathbf{u}_\perp \rangle$. Thus, for any $\mathbf{p}, \mathbf{q} \in \Delta$ and $\mathbf{u} \in \mathbb{R}^n$, the following holds:

$$\langle \mathbf{p}, \mathbf{u} \rangle \geq \langle \mathbf{q}, \mathbf{u} \rangle \quad \text{iff} \quad \langle \mathbf{p}_\perp - \mathbf{q}_\perp, \mathbf{u}_\perp \rangle \geq 0.$$

The statement " there exists an LEU representation for $\succsim$" can then be rephrased as follows: there exist a K–tuple $\langle \mathbf{u}_k \in O_{\mathbb{1}} : k = 1, \ldots, K \rangle$ of *orthonormal* vectors in $O_{\mathbb{1}}$ such that,

$$\mathbf{p} \succ^* \mathbf{q} \quad \text{iff} \quad [\langle \mathbf{p}_\perp - \mathbf{q}_\perp, \mathbf{u}_1 \rangle, \ldots, \langle \mathbf{p}_\perp - \mathbf{q}_\perp, \mathbf{u}_K \rangle] >_L \mathbf{0}_K,$$

where $\mathbf{0}_K$ is the *origin* of $\mathbb{R}^K$ and $>_L$ is the asymmetric component of the lexicographic order $\geq_L$ over $\mathbb{R}^K$. Note, the subscript $\perp$ has been dropped as each $\mathbf{u}_k$ is assumed to be in $O_{\mathbb{1}}$ from the outset.

Moreover, the $\mathbf{u}_k$'s are assumed to be orthonormal. To see why, write $\mathbf{u}_k = \mathbf{u}_k^\perp + \mathbf{u}_k^\parallel$ where $\mathbf{u}_k^\perp$ and $\mathbf{u}_k^\parallel$, respectively, are the components of $\mathbf{u}_k$ perpendicular and parallel to the span of $\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}$. Assume, $\langle \mathbf{p}^\perp - \mathbf{q}^\perp, \mathbf{u}_l \rangle = 0$ for each $1 \leq l \leq k-1$. Then, $\langle \mathbf{p}_\perp - \mathbf{q}_\perp, \mathbf{u}_k^\parallel \rangle = 0$, and $\langle \mathbf{p}_\perp - \mathbf{q}_\perp, \mathbf{u}_k \rangle = \langle \mathbf{p}_\perp - \mathbf{q}_\perp, \mathbf{u}_k^\perp \rangle$. Hence, we may assume $\mathbf{u}_k^\parallel = \mathbf{0}$.

Let the list $\langle \mathbf{u}_k : k = 1, \ldots, K \rangle$ of orthonormal vectors in $O_{\mathbb{1}}$ be denoted by $\mathbf{U}$. Recall, from subsection 2.1, the *graded halfspace* induced by $\mathbf{U}$ is $H_{\mathbf{U}} \coloneqq \bigcup_{k=1}^{K} \mathbf{U}^k$, where $\mathbf{U}^k$ is the kth *slice* defined as:

$$\mathbf{U}^k \coloneqq \left\{ \mathbf{w} \in O_{\mathbb{1}} : \langle \mathbf{w}, \mathbf{u}_l \rangle = 0 \text{ for all } l < k, \text{ and } \langle \mathbf{w}, \mathbf{u}_k \rangle > 0 \right\}$$

for all $k = 1, \ldots, K$. Also, recall that the reflection of $H_{\mathbf{U}}$ through the origin, $-H_{\mathbf{U}}$, is the graded halfspace $H_{-\mathbf{U}}$. Also, let

$$O_{\mathbf{U}} \coloneqq \{ \mathbf{w} \in O_{\mathbb{1}} : \langle \mathbf{w}, \mathbf{u}_k \rangle = 0 \text{ for all } k = 1, \ldots, K \}$$

be the orthogonal subspace of $\mathbf{U}$ in $O_{\mathbb{1}}$. Observe, if $\mathbf{x} \in O_{\mathbb{1}}$ then:

$$[\langle \mathbf{x}, \mathbf{u}_1 \rangle, \ldots, \langle \mathbf{x}, \mathbf{u}_K \rangle] >_L \mathbf{0} \text{ iff } \mathbf{x} \in H_{\mathbf{U}},$$

by the definition of $\geq_L$. For any $\mathbf{p} \in \Delta$, define the sets:

$$\begin{aligned} U(\mathbf{p}) &\coloneqq \{ \mathbf{q} \in \Delta : \mathbf{q} \succ^* \mathbf{p} \}, && (\text{“strict upper contour set of } \mathbf{p}\text{”}) \\ I(\mathbf{p}) &\coloneqq \{ \mathbf{q} \in \Delta : \mathbf{q} \sim^* \mathbf{p} \}, && (\text{“indifference set of } \mathbf{p}\text{”}) \\ L(\mathbf{p}) &\coloneqq \{ \mathbf{q} \in \Delta : \mathbf{p} \succ^* \mathbf{q} \}. && (\text{“strict lower contour set of } \mathbf{p}\text{”}) \end{aligned}$$

Then, $\succsim^*$ admits an LEU representation via the vectors in $\mathbf{U}$ iff:

$$U(\mathbf{p}) = \Delta \cap (\mathbf{p} + H_{\mathbf{U}}),\ I(\mathbf{p}) = \Delta \cap (\mathbf{p} + O_{\mathbf{U}}),\ L(\mathbf{p}) = \Delta \cap (\mathbf{p} + H_{-\mathbf{U}}).$$

Now, let $W_* \coloneqq O_{\mathbb{1}}$ and consider the following sets:

$$\begin{aligned} U_* &\coloneqq \left\{ \mathbf{w} \in W_* : \mathbf{a} + t\mathbf{w} \succ^* \mathbf{a} \quad \text{for some } t > 0 \right\}, \\ V_* &\coloneqq \left\{ \mathbf{w} \in W_* : \mathbf{a} \succ^* \mathbf{a} + t\mathbf{w} \quad \text{for some } t > 0 \right\}, \\ S_* &\coloneqq \left\{ \mathbf{w} \in W_* : \mathbf{a} + t\mathbf{w} \sim^* \mathbf{a} \quad \text{for some } t > 0 \right\}. \end{aligned}$$

Note, $\succsim^*$ admits an LEU representation via $\mathbf{U}$ iff: $U_* = H_{\mathbf{U}}$, $V_* = H_{-\mathbf{U}}$ and $S_* = O_{\mathbf{U}}$. Moreover, then the structure of the graded halfspaces $H_{\mathbf{U}}$ and $H_{-\mathbf{U}}$ imply: (U_*, V_*, S_*) is a partition of W_* where U_*, V_* are cones satisfying $V_* = -U_*$ and S_* is a subspace. Hence, to prove Theorem 2, from Theorem 1 it is enough to establish,

LEMMA 3: *Suppose $\succsim^*$ satisfies Independence–3. Then, (U_*, V_*, S_*) is a partition of W_* where U_*, V_* are cones that satisfy $V_* = -U_*$, and S_* is a subspace. Further, $U(\mathbf{p}) = \Delta \cap (\mathbf{p} + U_*)$, $I(\mathbf{p}) = \Delta \cap (\mathbf{p} + S_*)$ and $L(\mathbf{p}) = \Delta \cap (\mathbf{p} + V_*)$ for all $\mathbf{p} \in \Delta$.*

The formal proof is in subsection A.II.2 of the Appendix. However, a sketch is provided in the following subsection.

We now present a geometric outline of the proof of Lemma 3. To begin, consider Figure 4 which shows an embedding of the set of lotteries $\mathscr{L}(Z)$, over a set Z of three basic prizes, in the three dimensional Eucidean space. Thus, each point on the simplex Δ corresponds to a lottery. A typical lottery is $\mathbf{p}$ whereas $\mathbf{a}$ is the centroid of the simplex. It corresponds to that lottery which randomly selects any basic prize with the same probability for every prize to be selected.

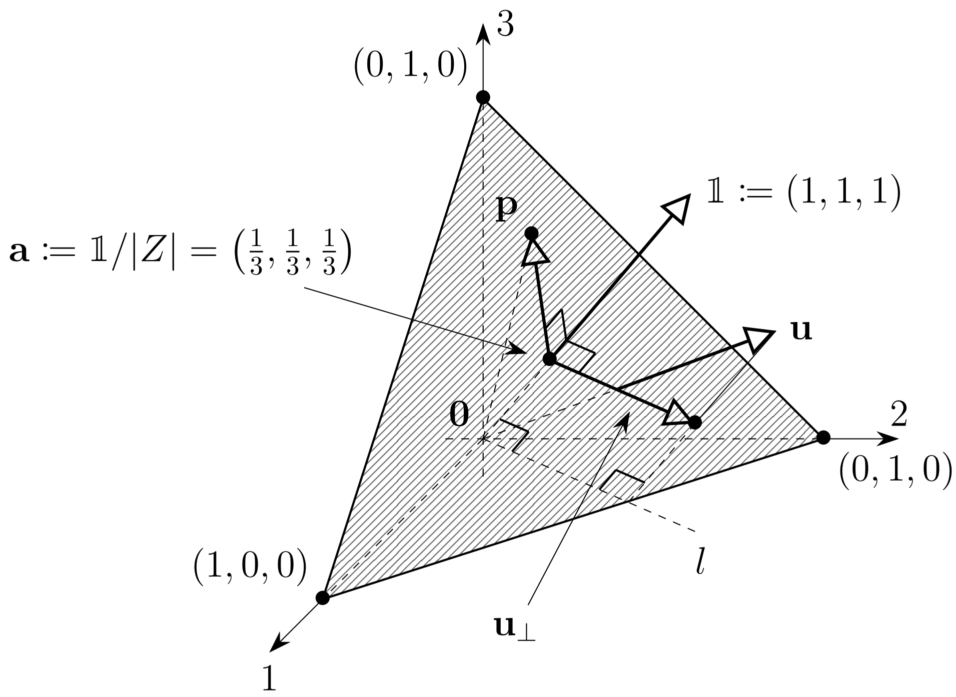


FIGURE 4: The simplex Δ and an expected utility vector $\mathbf{u} \in \mathbb{R}^{|Z|}$.

Next, the vector $\mathbf{u}$ is a Bernoullian to be used in ascribing expected utilities to various lotteries. For instance, the expected utility of the lottery $\mathbf{p}$ according to the Bernoullian $\mathbf{u}$ is the inner product $\langle \mathbf{u}, \mathbf{p} \rangle$. The collection of all vectors in $\mathbb{R}^{|Z|}$ which are perpendicular to the vector of "all ones" $\mathbb{1}$ form a subspace denoted by $O_\mathbb{1}$.

Of interest shall be the orthogonal projections, of the lotteries and the Bernoullians, onto $O_\mathbb{1}$ because: $\langle \mathbf{u}, \mathbf{p} \rangle \geq \langle \mathbf{u}, \mathbf{q} \rangle$ iff $\langle \mathbf{u}, \mathbf{p} - \mathbf{q} \rangle \geq 0$. Note, $\mathbf{p} = \mathbf{a} + \mathbf{p}_\perp$ and $\mathbf{q} = \mathbf{a} + \mathbf{q}_\perp$ where $\mathbf{p}_\perp$ and $\mathbf{q}_\perp$ are the orthogonal projections onto $O_\mathbb{1}$ of $\mathbf{p}$ and $\mathbf{q}$. Thus, $\langle \mathbf{u}, \mathbf{p} - \mathbf{q} \rangle = \langle \mathbf{u}, \mathbf{p}_\perp - \mathbf{q}_\perp \rangle$. Moreover, with $\mathbf{u}_\perp$ as the orthogonal projection of $\mathbf{u}$ onto $O_\mathbb{1}$, $\mathbf{u} - \mathbf{u}_\perp$ is perpendicular to $\mathbf{p}_\perp - \mathbf{q}_\perp$. Thus, $\langle \mathbf{u}, \mathbf{p} - \mathbf{q} \rangle = \langle \mathbf{u}_\perp, \mathbf{p}_\perp - \mathbf{q}_\perp \rangle$. Further, note that the orthogonal projection of $\mathbf{a}$ onto $O_\mathbb{1}$ is the origin $\mathbf{0}$. Hence, all the action essentially takes place in the translation by $-\mathbf{a}$ of the simplex Δ which is part of the hyperplane $O_\mathbb{1}$.

Henceforth, the perspective is such that the eye is located at $\mathbb{1}$ and looks in the direction of $\mathbf{0}$. Thus, the simplex Δ appears as shown in Figure 5. To illustrate Independence–3, let $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{r}$ be arbitrary points in Δ. If $\alpha \in (0,1)$ then the line segment joining $\mathbf{s}_\alpha := \alpha\mathbf{p} + (1-\alpha)\mathbf{r}$ and $\mathbf{t}_\alpha := \alpha\mathbf{q} + (1-\alpha)\mathbf{r}$ is *parallel* to the line segment joining $\mathbf{p}$ and $\mathbf{q}$. This is because $\mathbf{s}_\alpha$ divides the line segment joining $\mathbf{r}$ to $\mathbf{p}$ in the ratio $\alpha : 1-\alpha$ which is the same ratio in which $\mathbf{t}_\alpha$ divides the line segment joining the point $\mathbf{r}$ to $\mathbf{q}$. Then, Independence–3 places two requirements on the binary relation $\succsim^*$ defined over Δ. First, if $\mathbf{p} \succ^* \mathbf{q}$ then $\mathbf{s}_\alpha \succ^* \mathbf{t}_\alpha$ for each $\alpha \in (0,1)$. Moreover, if $\mathbf{s}_\alpha \succ^* \mathbf{t}_\alpha$ for *every* $\alpha \in (0,1)$ then $\mathbf{p} \succ^* \mathbf{q}$. Note, when using Independence–3 to conclude $\mathbf{p} \succ^* \mathbf{q}$, it is not enough that $\mathbf{s}_\alpha \succ^* \mathbf{t}_\alpha$ for *some* $\alpha \in (0,1)$.

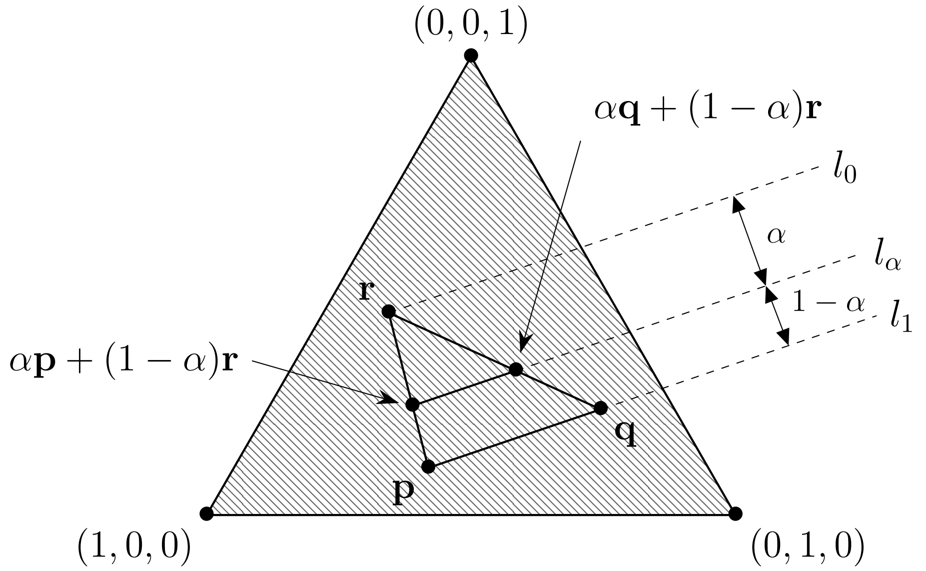


FIGURE 5: The Independence axiom and "Similar Triangles".

First, we shall argue that $\succsim^*$ is "consistent along any ray". For this, consider Figure 6 which shows two lotteries $\mathbf{p}$ and $\mathbf{q}$ defining the ray l which emanates from $\mathbf{q}$ and passes through $\mathbf{p}$. First, assume $\mathbf{p} \succ^* \mathbf{q}$. By Independence–3, every point on the "open line segment" with end points as $\mathbf{p}$ and $\mathbf{q}$ is strictly preferred to $\mathbf{q}$.

Let $\mathbf{r}$ be on the ray l is such that $\mathbf{p}$ lies on the "open line segment" whose end points are $\mathbf{q}$ and $\mathbf{r}$. Suppose $\mathbf{r} \sim^* \mathbf{q}$. Then, $\mathbf{p} \succ^* \mathbf{q}$ implies $\mathbf{p} \succ^* \mathbf{r}$. By Independence–3, every point on the "open line segment" with end points as $\mathbf{p}$ and $\mathbf{r}$ is strictly preferred to $\mathbf{r}$. Since $\mathbf{r} \sim^* \mathbf{q}$, every such point is strictly preferred to $\mathbf{q}$. Then, every point on the "open line segment" with $\mathbf{q}$ and $\mathbf{r}$ is strictly preferred to $\mathbf{q}$. Thus, $\mathbf{r} \succ^* \mathbf{q}$ by Independence–3 which contradicts $\mathbf{r} \sim^* \mathbf{q}$. Thus, $\mathbf{r} \sim^* \mathbf{q}$ is not possible. Further, by the argument in the previous paragraph, $\mathbf{q} \succ^* \mathbf{r}$ implies $\mathbf{q} \succ^* \mathbf{p}$. However, $\mathbf{p} \succ^* \mathbf{q}$ by assumption.

Thus, if some lottery on the ray l is strictly preferred to $\mathbf{q}$ then each lottery on the ray, which is distinct from $\mathbf{q}$, is strictly preferred to $\mathbf{q}$. A similar argument shows, if $\mathbf{q}$ is strictly preferred to some lottery on the ray l then $\mathbf{q}$ is strictly preferred to every lottery on the ray provided it is distinct from $\mathbf{q}$. Now, assume that some lottery, say $\mathbf{p}$, on the ray l is such that $\mathbf{p} \sim^* \mathbf{q}$. Thus, neither $\mathbf{p} \succ^* \mathbf{q}$ nor $\mathbf{q} \succ^* \mathbf{p}$ holds. Then, for any lottery $\mathbf{r}$ on the ray l, it must be the case that neither $\mathbf{r} \succ^* \mathbf{q}$ nor $\mathbf{q} \succ^* \mathbf{r}$ holds. For instance, note that $\mathbf{r} \succ^* \mathbf{q}$ would imply $\mathbf{p} \succ^* \mathbf{q}$ which is a contradiction. That is, if the decision maker is indifferent between $\mathbf{q}$ and some lottery on the ray l which is distinct from q then he is indifferent between $\mathbf{q}$ and every lottery on the ray. In other words, any two lottery on the ray l which are distinct from $\mathbf{q}$ must be ranked *consistently* with respect to $\mathbf{q}$.

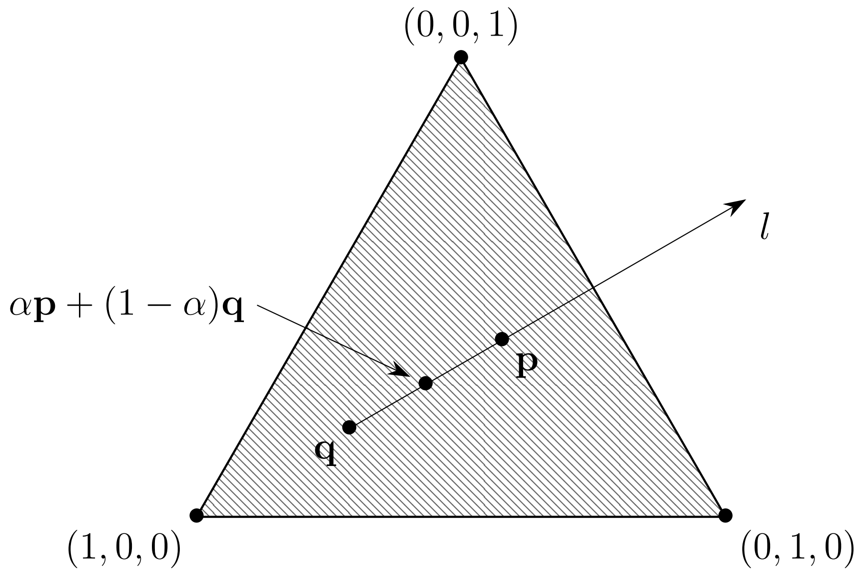


FIGURE 6: Consistency of $\succsim^*$ along a ray.

We demonstrate "anti–consistency along reflected rays". Consider Figure 7 which shows $\mathbf{p}$ in Δ and two rays l_1 and l_2 emanating from $\mathbf{p}$ which contain the points $\mathbf{q}$ and $\mathbf{r}$, respectively. First, assume that $\mathbf{q} \succ^* \mathbf{p}$. Let $\alpha \in (0, 1)$ be such that $\alpha\mathbf{q} + (1-\alpha)\mathbf{r} = \mathbf{p}$ and define $\mathbf{s} := \alpha\mathbf{p} + (1-\alpha)\mathbf{r}$. Thus, $\mathbf{q} \succ^* \mathbf{p}$ implies $\mathbf{p} \succ^* \mathbf{s}$. Then, "consistency along a ray" requires that if every point on the "open ray" l_1 is strictly preferred to $\mathbf{p}$ then $\mathbf{p}$ is strictly preferred to every point on the "open ray" l_2. The converse also holds by a similar argument.

Now, assume that $\mathbf{q} \sim^* \mathbf{p}$. Suppose $\mathbf{r} \succ^* \mathbf{p}$. Then, $\mathbf{p} \succ^* \mathbf{q}$ by the previous paragraph. Thus, both $\mathbf{p} \sim^* \mathbf{q}$ and $\mathbf{p} \succ^* \mathbf{q}$ hold which is impossible because $\sim^*$ and $\succ^*$ are disjoint. Thus, $\mathbf{r} \succ^* \mathbf{p}$ fails. Similarly, $\mathbf{p} \succ^* \mathbf{r}$ fails. Hence, $\mathbf{q} \sim^* \mathbf{p}$ implies $\mathbf{r} \sim^* \mathbf{p}$. We say, the rays l_1 and l_2 are ranked *anti–consistently* with respect to $\mathbf{p}$.

Recall, $U(\mathbf{p})$ and $L(\mathbf{p})$ are the strict upper and lower contour sets of any $\mathbf{p}$ in the simplex. By "consistency along a ray", it follows that $U(\mathbf{p})$ and $L(\mathbf{p})$ are made up of "open rays" which emanate from $\mathbf{p}$ as the "origin". Moreover, $L(\mathbf{p}) = -U(\mathbf{p})$ by "anti–consistency along reflected rays". Also, recall that $I(\mathbf{p})$ is the indifference set of $\mathbf{p}$. Then, $I(\mathbf{p})$ too is made up of rays emanating from $\mathbf{p}$ as the origin. However, $I(\mathbf{p}) = -I(\mathbf{p})$ by "anti–consistency of reflected rays". Because $\succ^*$ is asymmetric and $\sim^*$ is symmetric, the sets $U(\mathbf{p})$, $L(\mathbf{p})$ and $I(\mathbf{p})$ are pairwise disjoint. Moreover, since the union of $\succ^*$ and $\sim^*$ is complete, they form a partition of the simplex.

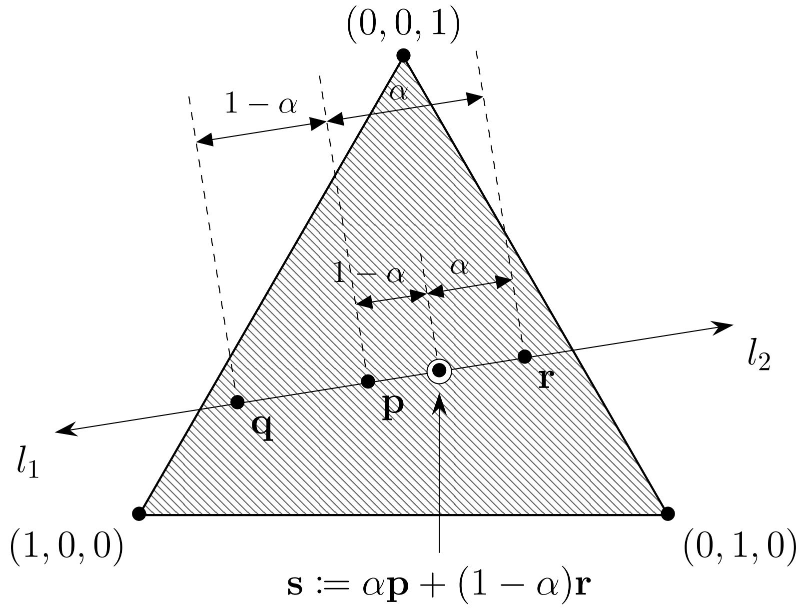
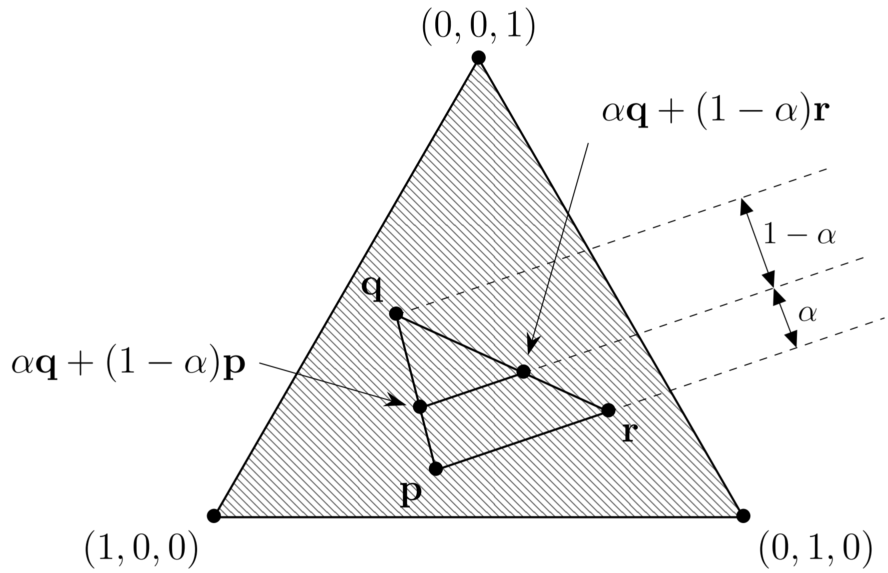


FIGURE 7: Anti–consistency of $\succsim^*$ along reflected rays.

Since we wish to invoke the Decomposition Theorem, we shall now proceed to argue that each of the sets $U(\mathbf{p})$, $L(\mathbf{p})$ and $I(\mathbf{p})$ is convex. Coupled with the observations as in the previous paragraph, this will imply that $U(\mathbf{p})$ and $L(\mathbf{p})$ are a pair of mutually reflecting (convex) cones while $I(\mathbf{p})$ is a subspace. Further, they partition the simplex which, essentially, can be thought of as the hyperplane $O_{\mathbb{1}}$.

First, we argue that $U(\mathbf{p})$ and $L(\mathbf{p})$ are convex. Consider Figure 8 which shows a point $\mathbf{p}$ of the simplex. Also, let $\mathbf{q}$ and $\mathbf{r}$ be two arbitrary points in $U(\mathbf{p})$. That is, both $\mathbf{q} \succ^* \mathbf{p}$ and $\mathbf{r} \succ^* \mathbf{p}$ hold. Fix any $\alpha \in (0, 1)$. Let $\mathbf{s} := \alpha\mathbf{q} + (1 - \alpha)\mathbf{r}$ and $\mathbf{t} := \alpha\mathbf{q} + (1 - \alpha)\mathbf{p}$. Then, $\mathbf{r} \succ^* \mathbf{p}$ implies $\mathbf{s} \succ^* \mathbf{t}$. Also, $\mathbf{q} \succ^* \mathbf{p}$ implies $\mathbf{t} \succ^* \mathbf{p}$ because $\mathbf{p} = \alpha\mathbf{p} + (1 - \alpha)\mathbf{p}$. Since $\succ^*$ is transitive, $\mathbf{s} \succ^* \mathbf{t}$ and $\mathbf{t} \succ^* \mathbf{p}$ imply $\mathbf{s} \succ^* \mathbf{p}$. Since $\mathbf{s} = \alpha\mathbf{q} + (1 - \alpha)\mathbf{r}$ where $\mathbf{q}, \mathbf{r} \in U(\mathbf{p})$ and $\alpha \in (0, 1)$ are arbitrary, we have: $U(\mathbf{p})$ is convex. By a similar argument, $L(\mathbf{p})$ is convex. Thus, $U(\mathbf{p})$ and $L(\mathbf{p})$ are (convex) cones.

It remains to argue that $I(\mathbf{p})$ is convex. For this, let $\mathbf{q}$ and $\mathbf{r}$ in $L(\mathbf{p})$. That is, both $\mathbf{q} \succ^* \mathbf{p}$ and $\mathbf{r} \sim^* \mathbf{p}$ hold. Fix any $\alpha \in (0,1)$ and let $\mathbf{s} := \alpha\mathbf{q} + (1-\alpha)\mathbf{r}$. Since $\mathbf{q} \sim^* \mathbf{p}$ and $\mathbf{r} \sim^* \mathbf{p}$, the symmetry and transitivity of $\sim^*$ implies $\mathbf{q} \sim^* \mathbf{r}$. Then, $\mathbf{r}$ and $\mathbf{s}$ are two points on the "open ray" emanating from $\mathbf{q}$ which passes through $\mathbf{r}$. By "consistency along a ray", $\mathbf{r} \sim^* \mathbf{q}$ implies $\mathbf{s} \sim^* \mathbf{q}$. Again, $\mathbf{s} \sim^* \mathbf{q}$ and $\mathbf{q} \sim^* \mathbf{p}$ imply $\mathbf{s} \sim^* \mathbf{p}$ because $\sim^*$ is transitive. Since $\mathbf{s} = \alpha\mathbf{q} + (1-\alpha)\mathbf{r}$ where $\mathbf{q}, \mathbf{r} \in I(\mathbf{p})$ and $\alpha \in (0,1)$ are arbitrary, we have: $I(\mathbf{p})$ is convex. Recall, $I(\mathbf{p}) = -I(\mathbf{p})$ by "anti–consistency along reflected rays". Hence, "consistency along a ray" and convexity of $I(\mathbf{p})$ imply: $I(\mathbf{p})$ is a subspace.



$\mathbf{q} \succ^* \mathbf{p}$; $\mathbf{r} \succ^* \mathbf{p}$

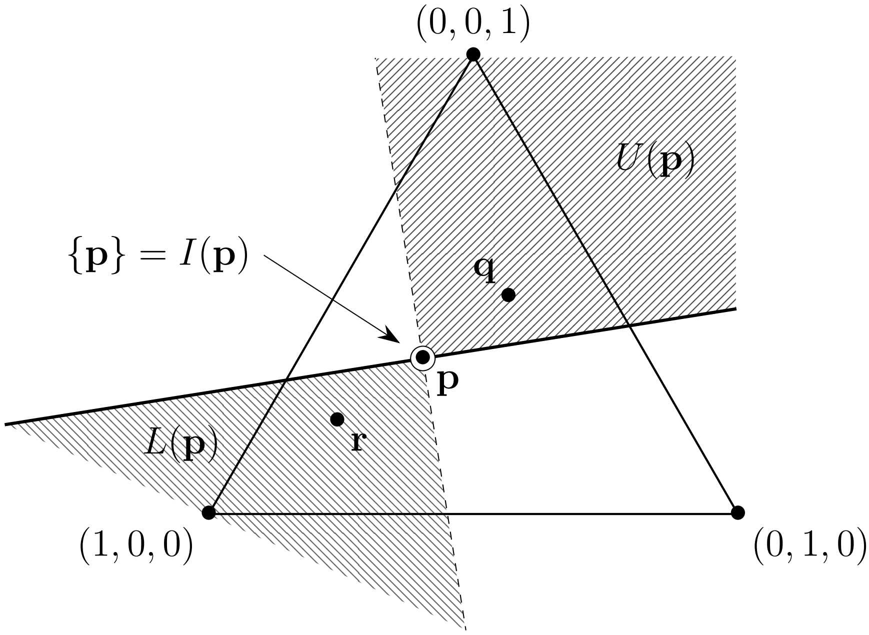
FIGURE 8: Convexity of $U(\mathbf{p}) := \{\mathbf{q} \in \Delta : \mathbf{q} \succ^* \mathbf{p}\}$.

The conclusions thus far can be represented in a drawing such as Figure 9. With the point $\mathbf{p}$ in the simplex as the "vertex", two cones have been drawn to represent the strict upper contour set $U(\mathbf{p})$ and the strict lower contour set $L(\mathbf{p})$. Then, if a coordinate system is so chosen that the vertex $\mathbf{p}$ becomes the "origin" then, the cones $U(\mathbf{p})$ and $V(\mathbf{p})$ must satisfy $L(\mathbf{p}) = -U(\mathbf{p})$ which is the algebraic experssion of the geometric fact that $U(\mathbf{p})$ and $L(\mathbf{p})$ are "reflections" of each other through the origin (or, the vertex). Hence, to any ray emanating from $\mathbf{p}$ which passes through a typical point $\mathbf{q}$ in $U(\mathbf{p})$, the reflected ray through $\mathbf{p}$ is part of $L(\mathbf{p})$. Likewise, the converse holds. In particular, consider the two rays shown in "bold" which are part of the boundaries of $U(\mathbf{p})$ and $L(\mathbf{p})$, respectively. Then, the former belongs to $U(\mathbf{p})$, if and only if, the latter belongs to $L(\mathbf{p})$. However, these rays may not belong to $U(\mathbf{p})$ and $L(\mathbf{p})$. Then, *both* rays are part of $I(\mathbf{p})$.

Notice, the subspace $I(\mathbf{p})$ is shown to be the singleton $\{\mathbf{p}\}$. Observe, the "white spaces" in the simplex. This is indicative of the possibiity that there are lotteries which are not comparable to $\mathbf{p}$ according to either $\succ^*$ or $\sim^*$. However, this is *not* possible because the relation $\sim^*$ satisfies the following:

$$\mathbf{p} \sim^* \mathbf{q} \quad \text{iff} \quad \big(\text{ not } \mathbf{p} \succ^* \mathbf{q} \text{ ; not } \mathbf{q} \succ^* \mathbf{p} \big).$$

That is, for any $\mathbf{q}$ in the simplex, if neither $\mathbf{q} \in U(\mathbf{p})$ nor $\mathbf{q} \in L(\mathbf{p})$ hold then $\mathbf{q} \in I(\mathbf{p})$. Moreover, $I(\mathbf{p})$ must be disjoint from the union of $U(\mathbf{p})$ and $L(\mathbf{p})$. The asymmetry of $\succ^*$ forces the two cones $U(\mathbf{p})$ and $L(\mathbf{p})$ to be disjoint. Thus, the claim that "$U(\mathbf{p})$, $L(\mathbf{p})$ and $I(\mathbf{p})$ partition the simplex" has been established.



$\mathbf{q} \succ^* \mathbf{p} \succ^* \mathbf{r}$

FIGURE 9: The pair of cones and the subspace for $\mathbf{p}$.

Thus, the cones $U(\mathbf{p})$, $L(\mathbf{p})$ and the subspace $I(\mathbf{p})$ must "fan out", while maintaining $L(\mathbf{p}) = -U(\mathbf{p})$, to cover the whole simplex and must do so without any "overlaps". At this stage, recall Figure 2 which was presented in section 2. Just as in Figure 9, the setting shown in Figure 2 involves two mutually reflecting cones and a subspace each pair of which is disjoint. However, for them to "fan out" so as to cover the whole plane implied that the cones must be graded halfspaces of the form illustrated, for instance, in Figure 1. Moreover, the structure of graded halfspaces then imply that the ranking of lotteries with respect to $\mathbf{p}$ is according to lexicographic expected utilities.

However, for this strategy to be complete, it must be ensured that the expected utility maps that define the lexicographic expected utility representations must *not* depend on the lottery $\mathbf{p}$. That is, if $\mathbf{p}$ and $\mathbf{p}'$ are arbitrary lotteries then, the sets $U(\mathbf{p}')$, $L(\mathbf{p}')$ and $I(\mathbf{p}')$ must be translations of $U(\mathbf{p})$, $L(\mathbf{p})$ and $I(\mathbf{p})$, respectively. Our strategy to show this will be as follows. Recall, $\mathbf{a}$ is the centroid of the simplex. For any arbitrary point $\mathbf{p}$ of the simplex, we shall argue that the sets $U(\mathbf{p})$, $L(\mathbf{p})$ and $I(\mathbf{p})$ are translations by the vector $\mathbf{p} - \mathbf{a}$ of the sets $U(\mathbf{a})$, $L(\mathbf{a})$ and $I(\mathbf{a})$, respectively.

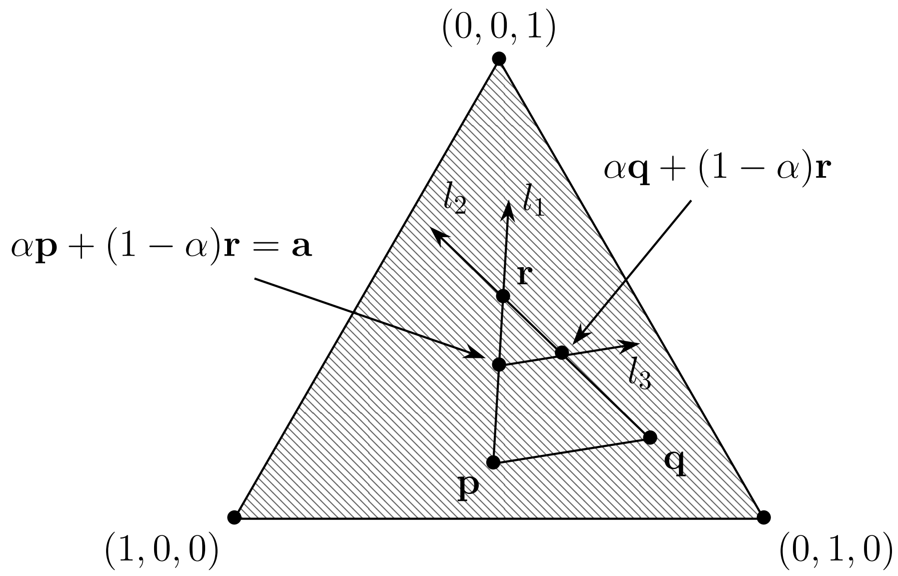


FIGURE 10: $U(\mathbf{p})$ is a *subset* of the translation of $U(\mathbf{a})$.

For this, consider Figure 10 which shows the arbitrary lottery $\mathbf{p}$ and the centroid $\mathbf{a}$ of the simplex. Also, let $\mathbf{q}$ be a lottery distinct from $\mathbf{p}$. Let l_1 be a ray emanating from $\mathbf{p}$ and passing through $\mathbf{a}$. Pick any point in the simplex on the ray l_1 such that $\mathbf{a}$ lies on the "open segment" whose end points are $\mathbf{p}$ and $\mathbf{r}$. That is, there exists an $\alpha \in (0,1)$ such that $\mathbf{a} = \alpha\mathbf{p} + (1-\alpha)\mathbf{r}$. Now, draw the ray l_2 which emanates from $\mathbf{q}$ and passes through $\mathbf{r}$. Further, draw the ray l_3 emanating from $\mathbf{a}$ that is *parallel* to the segment with end points $\mathbf{p}$ and $\mathbf{q}$. Observe, the intersection of l_3 with l_2 is $\alpha\mathbf{q} + (1-\alpha)\mathbf{r}$ by construction.

First, assume $\mathbf{q} \succ^* \mathbf{p}$. Then, $\alpha\mathbf{q} + (1-\alpha)\mathbf{r} \succ^* \alpha\mathbf{p} + (1-\alpha)\mathbf{r}$. That is, $\alpha\mathbf{q} + (1-\alpha)\mathbf{r} \succ^* \mathbf{a}$. Since $\mathbf{q} \succ^* \mathbf{p}$, "consistency along a ray" ensures that all points on the ray emanating from $\mathbf{p}$ and passing through $\mathbf{q}$ must be strictly preferred to $\mathbf{p}$. Also, since $\alpha\mathbf{q} + (1-\alpha)\mathbf{r} \succ^* \mathbf{a}$, "consistency along a ray" ensures that all points on the ray emanating from $\mathbf{a}$ and is parallel to former. The argument thus far has shown that $U(\mathbf{p})$ is a *subset* of the translation, by the vector $\mathbf{p} - \mathbf{a}$, of $U(\mathbf{a})$. To show equality, we argue: $\alpha\mathbf{q} + (1-\alpha)\mathbf{r} \succ^* \mathbf{a}$ implies $\mathbf{q} \succ^* \mathbf{p}$.

For this, consider Figure 11. Assume $\mathbf{s} \coloneqq \alpha\mathbf{q} + (1-\alpha)\mathbf{r} \succ^* \mathbf{a}$. For any arbitrary $\beta \in (0,1)$, $\mathbf{t}_\beta \coloneqq \beta\mathbf{a} + (1-\beta)\mathbf{p}$ is on the "open segment" with end points $\mathbf{p}$ and $\mathbf{a}$. The ray l_4 emanates from $\mathbf{t}_\beta$ and is parallel to l_3. The intersection of l_4 with the segment joining $\mathbf{s}$ and $\mathbf{p}$ is $\mathbf{v}_\beta \coloneqq \beta\mathbf{s} + (1-\beta)\mathbf{p}$. Then, $\mathbf{s} \succ^* \mathbf{a}$ implies $\mathbf{v}_\beta \succ^* \mathbf{t}_\beta$. Let l_4 intersect l_2 at $\mathbf{w}_\beta$. Since l_4 is parallel to the segment joining $\mathbf{p}$ and $\mathbf{q}$, we have $\mathbf{w}_\beta = \beta\mathbf{s} + (1-\beta)\mathbf{q}$. Since $\mathbf{v}_\beta$ and $\mathbf{w}_\beta$ are on l_4, "consistency along a ray" forces $\mathbf{v}_\beta \succ^* \mathbf{t}_\beta$ to imply $\mathbf{w}_\beta \succ^* \mathbf{t}_\beta$.

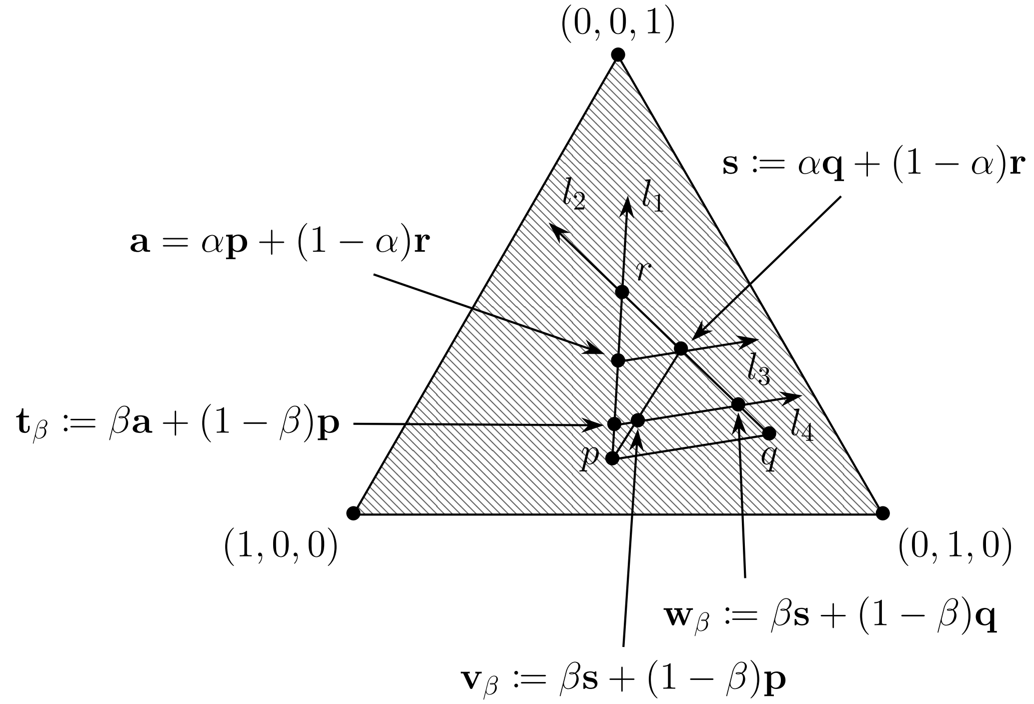


FIGURE 11: $U(\mathbf{p})$ is *equal* to the translation of $U(\mathbf{a})$.

Observe, $\mathbf{t}_\beta = [1-\beta(1-\alpha)]\mathbf{p} + [\beta(1-\alpha)]\mathbf{r}$ because $\mathbf{a} = \alpha\mathbf{p} + (1-\alpha)\mathbf{r}$ and $\mathbf{t}_\beta = \beta\mathbf{a} + (1-\beta)\mathbf{p}$. Also, $\mathbf{w}_\beta = [1-\beta(1-\alpha)]\mathbf{q} + [\beta(1-\alpha)\mathbf{r}]$ because $\mathbf{s} = \alpha\mathbf{q} + (1-\alpha)\mathbf{r}$ and $\mathbf{w}_\beta = \beta\mathbf{s} + (1-\beta)\mathbf{q}$. That $\mathbf{w}_\beta \succ^* \mathbf{t}_\beta$ holds for any arbitrary $\beta \in (0,1)$ is equivalent to:

$$\gamma\mathbf{q} + (1-\gamma)\mathbf{r} \succ^* \gamma\mathbf{p} + (1-\gamma)\mathbf{r} \quad \text{for every } \gamma \in (\alpha, 1).$$

Of course, the above holds at $\gamma = \alpha$ because $\mathbf{s} \succ^* \mathbf{a}$. To see why it also holds for any $\gamma \in (0,\alpha)$, let $\mathbf{m}_\beta \coloneqq \beta\mathbf{a} + (1-\beta)\mathbf{r}$ and $\mathbf{n}_\beta \coloneqq \beta\mathbf{s} + (1-\beta)\mathbf{r}$. Thus, $\mathbf{s} \succ^* \mathbf{a}$ implies $\mathbf{m}_\beta \succ^* \mathbf{n}_\beta$ for any $\beta \in (0,1)$. Also, $\mathbf{m}_\beta = (\alpha\beta)\mathbf{p} + (1-\alpha\beta)\mathbf{r}$ and $\mathbf{n}_\beta = (\alpha\beta)\mathbf{q} + (1-\alpha\beta)\mathbf{r}$ because $\mathbf{a} = \alpha\mathbf{p} + (1-\alpha)\mathbf{r}$ and $\mathbf{s} = \alpha\mathbf{q} + (1-\alpha)\mathbf{r}$. Since $\alpha\beta$ increases from 0 to α as β increases from 0 to 1, the above relation holds for *every* $\gamma \in (0,1)$. Then, Indpendence–3 implies $\mathbf{q} \succ^* \mathbf{p}$ as was required. Thus, $U(\mathbf{p})$ is *equal* to the translation, by $\mathbf{p} - \mathbf{a}$, of the set $U(\mathbf{a})$.

A similar argument shows that $L(\mathbf{p})$ is the translation, by the vector $\mathbf{p} - \mathbf{a}$, of the set $L(\mathbf{a})$. Let us reconsider Figure 10. We already have (1) $\mathbf{q} \succ^* \mathbf{p}$ iff $\mathbf{s} \succ^* \mathbf{a}$, and (2) $\mathbf{p} \succ^* \mathbf{q}$ iff $\mathbf{a} \succ^* \mathbf{s}$. Thus, we must have: $\mathbf{q} \sim^* \mathbf{p}$ iff $\mathbf{s} \sim^* \mathbf{a}$. To see why, assume $\mathbf{q} \sim^* \mathbf{p}$. Suppose $\mathbf{s} \succ^* \mathbf{a}$. Then, $\mathbf{q} \succ^* \mathbf{p}$ by (1) which is a contradiction. Thus, $\mathbf{s} \succ^* \mathbf{a}$ does not hold. Similarly, (2) implies $\mathbf{a} \succ^* \mathbf{s}$ does not hold. Thus, $\mathbf{s} \sim^* \mathbf{a}$ must hold. Hence, $\mathbf{q} \sim^* \mathbf{p}$ implies $\mathbf{s} \sim^* \mathbf{a}$. A similar argument implies the converse. Hence, $I(\mathbf{p})$ is the translation, by $\mathbf{p} - \mathbf{a}$, of the set $I(\mathbf{a})$.

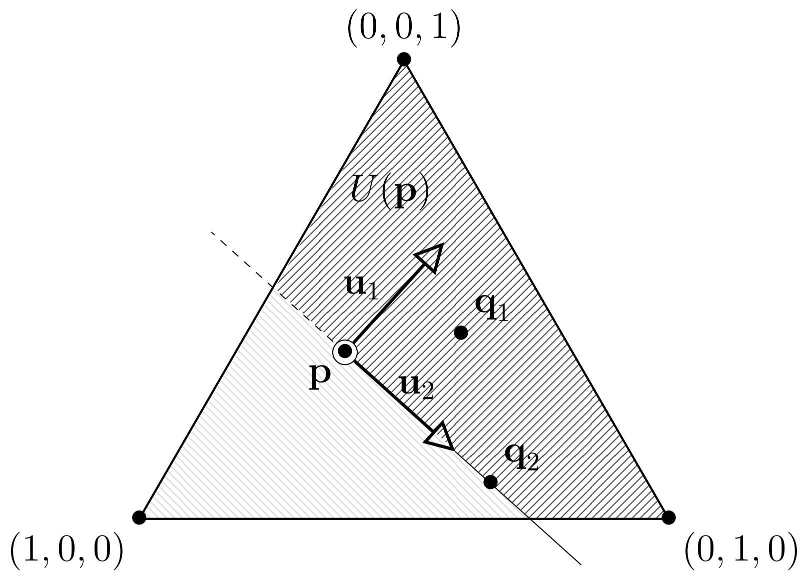


FIGURE 12: Lexicographic expected utilities for $\succsim^*$.

Thus, for any arbitrary lottery $\mathbf{p}$, the sets $U(\mathbf{p})$, $L(\mathbf{p})$ and $I(\mathbf{p})$ are translations by $\mathbf{p} - \mathbf{a}$ of the sets $U(\mathbf{a})$, $L(\mathbf{a})$ and $I(\mathbf{a})$, respectively. This is equivalent to asserting that there exists a pair of cones U_*, V_* satisfying $V_* = -U_*$, and a subspace S_*, where (U_*, V_*, S_*) partitions $O_{\mathbb{1}}$ such that, for any $\mathbf{p} \in \Delta$, $U(\mathbf{p}) = \Delta \cap (\mathbf{p} + U_*)$, $L(\mathbf{p}) = \Delta \cap (\mathbf{p} + V_*)$ and $I(\mathbf{p}) = \Delta \cap (\mathbf{p} + S_*)$. This proves Lemma 3.

To complete the picture, observe that the Decomposition Theorem applies on the parition (U_*, V_*, S_*). That is, there exists a list $\mathbf{U}$ of orthonormal vectors in $O_{\mathbb{1}}$ such that U_* is the graded halfspace that is generated by $\mathbf{U}$. Recall, a typical graded halfspace appears as is shown in Figure 1 of section 2. Importing such a structure for U_*, which generates $U(\mathbf{p})$ through translations by $\mathbf{p}$, we obtain Figure 12 which illustrates the strict upper contour set of the lottery $\mathbf{p}$.

The list, shown here, consists of two orthonormal vectors $\mathbf{u}_1$ and $\mathbf{u}_2$. Thus, lottery $\mathbf{q}_1$ satisfies $\mathbf{q}_1 \succ^* \mathbf{p}$ because $\langle \mathbf{u}_1, \mathbf{q}_1 - \mathbf{p} \rangle > 0$. However, note that $\langle \mathbf{u}_1, \mathbf{q}_2 - \mathbf{p} \rangle = 0$. But, observe that $\langle \mathbf{u}_2, \mathbf{q}_2 - \mathbf{p} \rangle > 0$. Thus, $\mathbf{q}_2 \succ^* \mathbf{p}$. Equivalently, the maps $\mathbf{p} \in \Delta \mapsto \langle \mathbf{u}_1, \mathbf{p} \rangle$ and $\mathbf{p} \in \Delta \mapsto \langle \mathbf{u}_2, \mathbf{p} \rangle$ specify a lexicographic expected utility representation for $\succsim^*$.

4. SOCIAL CHOICE THEORY

The second application is to the aggregation of individual preferences into a social preference. The framework is as follows. Let A be the set of alternatives; A is non–empty and $|A| \geq 3$. Also, let $N = \{1, 2, \ldots, n\}$ be the set of individuals. A utility profile u is a $\langle u_i \in \mathbb{R}^A : 1, \ldots, n\rangle$ where u_i is the utility function representing individual i's ranking over A. Let $\mathscr{U}$ be the class of all utility functions for an individual. Let $\mathscr{R}$ be the class of all preferences[30] over A. A *social welfare functional* is a map $F : \mathscr{U}^n \to \mathscr{R}$. For any $u \in \mathscr{U}^n$, let $\hat{F}(u)$ and $\bar{F}(u)$ be respectively the strict and indifference components of $F(u)$. For $u \in \mathscr{U}^n$ and $a \in A$, let $u(a) := \langle u_i(a) : i = 1, \ldots, n\rangle$. Also, for $u \in \mathscr{U}^n$ and $a, b \in A$, let $F(u)|_{\{a,b\}}$ be the restriction of $F(u)$ to the set $\{a, b\}$.

DEFINITION 5: *A* lexicographic generalized utilitarianism *is a social welfare functional F which admits some $\lambda = \langle \lambda^k \in \mathbb{R}^n : k = 1, \ldots, K\rangle$ such that $\lambda^k \neq \mathbf{0}$ and, for any $u \in \mathscr{U}^n$ and $a, b \in A$:*

$$aF(u)b \iff [\lambda^1 \cdot u(a), \ldots, \lambda^K \cdot u(a)] \geq_L [\lambda^1 \cdot u(b), \ldots, \lambda^K \cdot u(b)],$$

where $\lambda^k \cdot u(a) := \sum_{i=1}^n \lambda_i^k u_i(a)$ and $\geq_L$ is the lexicographic order on R^K.

Additionally, if K is 1 and $\lambda \in \mathbb{R}^n_+$, F is a *generalised utilitarianism.* Consider the following two axioms that F may satisfy.

BINARY INDEPENDENCE OF IRRELEVANT ALTERNATIVES (BIIA):

$$[\, u(a) = u'(a) \ ; \ u(b) = u'(b) \,] \implies [\, F(u)|_{\{a,b\}} = F(u')|_{\{a,b\}} \,].$$

PARETO INDIFFERENCE (PI): $[\, u(a) = u(b) \,] \implies [\, a\bar{F}(u)b \,]$.

Any social welfare functional that satisfies each of the above two axioms is a *welfarism.* Another property is as follows.

STRONG NEUTRALITY (SN): If $u, u' \in \mathscr{U}^n$ and $a, b, c, d \in A$ then:

$$[\, u(a) = u'(c) \ ; \ u(b) = u'(d) \,] \implies [\, aF(u)b \iff cF(u')d \,].$$

The key result characterizing strong neutrality is the following.

THEOREM OF WELFARISM: *A social choice functional satisfies strong neutrality, if and only if, it is a welfarism.*

[30] A binary relation over A is a *preference* if it is complete and transitive.

This result is well–known in the literature. It appears as Theorem 2.1 in Blackorby et al. (1984) for instance. Strong Neutrality of a social welfare function implies that it admits a description through a single complete and transitive binary relation over the space $\mathbb{R}^n$ of all utility n–tuples under any utility profile and any alternative. Thus, information apart from individuals' utility values to alternatives is not relevant. The result appears in Blackorby et al. (1984) as Theorem 2.2 and one formulation of this result is as follows.

Representation Lemma: *Let F be a social welfare functional that satisfies strong neutrality. Then, there exists a complete and transitive binary relation $\succsim$ (an "ordering") over $\mathbb{R}^n$ such that:*

$$aF(u)b \iff u(a) \succsim u(b).$$

for any $a, b \in A$ and any $u \in \mathscr{U}^n$.

At this stage, we point out a matter regarding the terminology. Both the terms "preference" and "ordering" refer to complete and transitive binary relations. However, the term "preference" shall apply when the binary relation is defined over the set A of alternatives. On the other hand, the term "ordering" shall be invoked when the binary relation is defined over the space $\mathbb{R}^n$ of utility n–tuples.

We proceed to state some normative axioms that a given social welfare functional may satisfy. For this, the notation for the standard partial orders on n–vectors will be useful. Denote by $\geq$, $>$ and $\gg$ the binary relations over $\mathbb{R}^n$ which are defined by:

$$\begin{aligned} \mathbf{x} \geq \mathbf{y} &\iff (\forall i \in N)[\, x_i \geq y_i \,], \\ \mathbf{x} > \mathbf{y} &\iff (\, \mathbf{x} \geq \mathbf{y} \,;\, \mathbf{x} \neq \mathbf{y} \,), \\ \mathbf{x} \gg \mathbf{y} &\iff (\forall i \in N)[\, x_i > y_i \,], \end{aligned}$$

where $\mathbf{x} \equiv (x_1, \ldots, x_n)$ and $\mathbf{y} \equiv (y_1, \ldots, y_n)$ are arbitrary vectors in $\mathbb{R}^n$. Then, the axioms can be stated as follows.

Weak Pareto (WP): $[\, u(a) \gg u(b) \,] \implies [\, a\hat{F}(u)b \,]$.

Strong Pareto (SP): $[\, u(a) > u(b) \,] \implies [\, a\hat{F}(u)b \,]$.

Continuity: *Suppose $\{u^k\}_{k\in\mathbb{N}}$ is $\mathscr{U}^n$–valued and $u^* \in \mathscr{U}^n$ such that $\lim_{k\to\infty} u^k(a) = u^*(a)$ for all $a \in A$. Then, for any $a, b \in A$,*

$$(\forall k \in \mathbb{N})[\, aF(u^k)b \,] \implies [\, aF(u^*)b \,].$$

We now come to the question: how "sensitive" is a social welfare functional to the "informational content" of utility profiles? Formally, we are interested in specifying the finest partition, given some social welfare functional F, of the space of all $\mathscr{U}^n$ of utility profiles such that F is constant over partition elements. Since any such partition is equivalently described by an equivalence relation over $\mathscr{U}^n$, we must specify the nature of the equivalence relation given the question. Since elements of an utility profile are utility representations of individual preferences, the equivalence relation over $\mathscr{U}^n$ will be defined through classes of "monotone tranformations" of utility profiles.

Let Φ_* the class of all n–tuples $\phi := (\phi_1, \dots, \phi_n)$, where each ϕ_i is a strictly increasing map on $\mathbb{R}$. For any $\phi \in \Phi_*$ and $u \in \mathscr{U}^n$, let $\phi \circ u$ be the utility profile $u' \in \mathscr{U}^n$, where $u'_i = \phi_i \circ u_i$ for every $i \in N$. The equivalence relations of interest are described as follows.

Definition 6: *Suppose $\Phi \subseteq \Phi_*$ is a subclass of transformations and F is a social welfare functional. Then, F is* Φ–invariant *if,*

$$F(\phi \circ u) = F(u) \quad \text{for all } u \in \mathscr{U}^n \text{ and } \phi \in \Phi.$$

Suitable choices for Φ in the above definition allow formalization of different notions of "comparability" of utility levels across individuals and of "measurability" of utility levels for each individual. For instance, let $\Phi_{\text{CMUC}} \subseteq \Phi$ consist of all $\phi = (\phi_1, \dots, \phi_n) \in \Phi$ corresponding to which there exists $\alpha > 0$ and $(\beta_1, \dots, \beta_n) \in \mathbb{R}^n$ such that,

$$\phi_i(t) = \alpha t + \beta_i \text{ for all } t \in \mathbb{R},$$

for every $i \in N$. Now, consider the following definition.

Definition 7: *A social welfare functional F is* cardinally measurable unit–comparable *if, F is Φ_{CMUC}–invariant.*

Observe, each ϕ_i of ϕ in Φ_{CMUC} is a positive affine transformation. Further, across individuals, ϕ_i's have a common "scale" α but possibly differing "offsets" β_i's. Thus, Φ_{CMUC}–invariance of F means that F processes, at most, the "cardinal information" in each utility profile. Moreover, the utility differences across individuals matter.

A social welfare functional F is *null* if $aF(u)b$ for any $a, b \in A$ and every $u \in \mathscr{U}^n$. That is, F ranks every pair of alteratives indifferently under every utility profile. In the rest of this section, we assume social welfare functionals to be not–null. Then, our first main result regarding social welfare functionals is the following.

Theorem 7: *A social welfare functional is a lexicographic generalized utilitarianism, if and only if, it is a welfarism that satisfies Cardinal Measurability Unit–Comparability and is non–null.*

Proof: For "sufficiency", let F be a Φ_{CMUC}–invariant welfarism. The Theorem of Welfarism and the Representation Lemma imply existence of an ordering $\succsim$ such that, for any $a, b \in A$ and any $u \in \mathscr{U}^n$:

$$aF(u)b \iff u(a) \succsim u(b).$$

We argue: there exists $\langle \lambda^k \in \mathbb{R}^n \setminus \{\mathbf{0}\} : k = 1, \ldots, K\rangle$ such that

$$\mathbf{x} \succsim \mathbf{y} \iff [\lambda^1 \cdot \mathbf{x}, \ldots, \lambda^K \cdot \mathbf{x}] \geq_L [\lambda^1 \cdot \mathbf{y}, \ldots, \lambda^K \cdot \mathbf{y}],$$

where $\geq_L$ is the lexicographic order over $\mathbb{R}^K$, and $\lambda^k \cdot \mathbf{x}$ denotes the standard inner product of the vectors λ^k and $\mathbf{x}$ in $\mathbb{R}^n$. Then, substituting $u(a)$ and $u(b)$ for $\mathbf{x}$ and $\mathbf{y}$, respectively, shows that F satisfies definition 5 as is required.

Let us "translate" the Φ_{CMUC}–invariance of F to $\succsim$. For any ϕ in Φ_{CMUC}, $u \in \mathscr{U}^n$ and $a \in A$, recall that $\phi \circ u = (\phi_1 \circ u_1, \ldots, \phi_n \circ u_n)$ and $u(a) = (u_1(a), \ldots, u_n(a))$. Thus, we shall write:

$$[\phi \circ u](a) := \big([\phi_1 \circ u_1](a), \ldots, [\phi_n \circ u_n](a)\big).$$

Then, $aF(\phi \circ u)b$ iff $[\phi \circ u](a) \succsim [\phi \circ u](b)$. Also, $aF(u)b$ iff $u(a) \succsim u(b)$. By Φ_{CMUC}–invariance of F, $aF(\phi \circ u)b$ iff $aF(u)b$. Thus:

$$u(a) \succsim u(b) \iff [\phi \circ u](a) \succsim [\phi \circ u](b).$$

Now, pick any $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ in $\mathbb{R}^n$. Also, let $\alpha > 0$ be arbitrary. Define $\beta_i := z_i$ for every $i \in N$, where $\mathbf{z} = (z_1, \ldots, z_n)$. Fix distinct $a, b \in A$ and construct an utility profile $u \in \mathscr{U}^n$ as follows. Let $u(a) := \mathbf{x}$, $u(b) := \mathbf{y}$, and $u(c) := \mathbf{0}$ for every $c \in A \setminus \{a, b\}$. Also, for each $i \in N$, let $\phi_i(t) := \alpha t + \beta_i$ for every $t \in \mathbb{R}$. Then, $\phi := (\phi_1, \ldots, \phi_n)$ is in Φ_{CMUC}. Observe, $[\phi \circ u](a) = \alpha\mathbf{x} + \mathbf{z}$ and $[\phi \circ u](b) = \alpha\mathbf{y} + \mathbf{z}$. Then, because ϕ belongs to Φ_{CMUC}, for any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ and $\alpha > 0$:

$$\mathbf{x} \succsim \mathbf{y} \iff \alpha\mathbf{x} + \mathbf{z} \succsim \alpha\mathbf{y} + \mathbf{z}. \tag{1}$$

Consider, for any $\mathbf{x} \in \mathbb{R}^n$, the three sets $U(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^n : \mathbf{y} \succ \mathbf{x}\}$, $L(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^n : \mathbf{x} \succ \mathbf{y}\}$ and $I(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^n : \mathbf{y} \sim \mathbf{x}\}$. By (1), for any $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, we have: $\mathbf{y} \succsim \mathbf{0}$ iff $\mathbf{x} + \mathbf{y} \succsim \mathbf{x}$. Thus, $\mathbf{y} \in U(\mathbf{0})$ iff $\mathbf{x} + \mathbf{y} \in U(\mathbf{x})$. That is, $U(\mathbf{x}) = \mathbf{x} + U(\mathbf{0})$. Similarly, $L(\mathbf{x}) = \mathbf{x} + L(\mathbf{0})$ and $I(\mathbf{x}) = \mathbf{x} + I(\mathbf{0})$. That is, for any $\mathbf{x} \in \mathbb{R}^n$, $U(\mathbf{x})$, $L(\mathbf{x})$ and $I(\mathbf{x})$ are translations by $\mathbf{x}$ of $U(\mathbf{0})$, $L(\mathbf{0})$ and $I(\mathbf{0})$, respectively.

By (1), if $\alpha > 0$ and $\mathbf{y} \in U(\mathbf{0})$ then $\alpha\mathbf{y} \in U(\mathbf{0})$. Also, $\mathbf{y}_1, \mathbf{y}_2 \in U(\mathbf{0})$ implies $\mathbf{y}_1 + \mathbf{y}_2 \in U(\mathbf{0})$. Thus, $U(\mathbf{0})$ is a (convex) cone. Similarly, $L(\mathbf{0})$ and $I(\mathbf{0})$ are cones. Moreover, $\mathbf{y} \succ \mathbf{0}$ iff $\mathbf{0} \succ -\mathbf{y}$. Thus, $L(\mathbf{0}) = -U(\mathbf{0})$. Also, $\mathbf{y} \sim \mathbf{0}$ iff $\mathbf{0} \sim -\mathbf{y}$. Thus, $I(\mathbf{0}) = -I(\mathbf{0})$. Since $I(\mathbf{0})$ is a cone and $I(\mathbf{0}) = -I(\mathbf{0})$, $I(\mathbf{0})$ is a subspace. Finally, note that $\big(U(\mathbf{0}), L(\mathbf{0}), I(\mathbf{0})\big)$ paritions $\mathbb{R}^n$ because $(\succ, \sim)$ partition $\succsim$ which is complete.

Thus, the Decomposition Theorem (Theorem 1 of section 2) applies. Hence, there exists a list $\mathbf{U} \equiv (\mathbf{u}_1, \ldots, \mathbf{u}_K)$ of some K orthonormal vectors such that $U(\mathbf{0}) = H_{\mathbf{U}}$, $L(\mathbf{0}) = -H_{\mathbf{U}}$ and $I(\mathbf{0}) = O_{\mathbf{U}}$, where $H_{\mathbf{U}}$ is the graded halfspace generated by $\mathbf{U}$ and $O_{\mathbf{U}}$ is the subspace of $\mathbb{R}^n$ orthogonal to $\mathbf{U}$. Let $\mathbf{y} \in \mathbb{R}^n$ be arbitrary. Since $U(\mathbf{y}) = \mathbf{y} + U(\mathbf{0})$, we have: $U(\mathbf{y}) = \mathbf{y} + H_{\mathbf{U}}$. Thus, by the definition of the graded halfspace $H_{\mathbf{U}}$ (that is, definition 1 of section 2) and because $U(\mathbf{y})$ is the strict upper contour set according to $\succsim$ of $\mathbf{y}$, we have:

$$\mathbf{x} \succ \mathbf{y} \iff [\mathbf{u}_1 \cdot (\mathbf{x} - \mathbf{y}), \ldots, \mathbf{u}_K \cdot (\mathbf{x} - \mathbf{y})] >_L \mathbf{0}_K,$$

where $\mathbf{u}^k \cdot \mathbf{x}$ is the standard inner product of vectors in $\mathbb{R}^n$, $\geq_L$ is the lexicographic order over $\mathbb{R}^K$ and $\mathbf{0}_K$ is the origin of $\mathbb{R}^K$. This is equivalent to the following:

$$\mathbf{x} \succsim \mathbf{y} \iff [\mathbf{u}_1 \cdot \mathbf{x}, \ldots, \mathbf{u}_K \cdot \mathbf{x}] \geq_L [\mathbf{u}_1 \cdot \mathbf{y}, \ldots, \mathbf{u}_K \cdot \mathbf{y}]. \tag{2}$$

because: $\mathbf{u}_k \cdot (\mathbf{x} - \mathbf{y}) > 0$ iff $\mathbf{u}_k \cdot \mathbf{x} > \mathbf{u}_k \cdot \mathbf{y}$. Then, defining $\lambda_k := \mathbf{u}_k$ for every $k = 1, \ldots, K$ completes the proof. ■

Then, generalized utilitarianism admits a characterization, which appears as Theorem 7.1 in Blackorby et al. (1984), which is seen to be an immediate consequence of the above theorem.

Theorem 8: *A social welfare functional is a generalized utilitarianism, if and only if, it is a welfarism that satisfies Weak Pareto, Continuity and Cardinal Measurability Unit–Comparability.*

Proof: We build on the proof of Theorem 7. In particular, recall that (2) holds where $\mathbf{u}_1, \ldots, \mathbf{u}_K$ are orthogonal. Thus, the ordering $\succsim$ over $\mathbb{R}^n$ is continuous iff $K = 1$. Further, the social welfare functional F satisfies Continuity iff $\succsim$ is continuous. Hence, we must have $K = 1$. Thus, it remains to argue that $\mathbf{u}_1 \in \mathbb{R}^n_+$. Suppose $\mathbf{u}_1 \notin \mathbb{R}^n_+$. Let $i_* \in N$ satisfy $\mathbf{u}_1 \cdot \mathbf{e}_{i_*} < 0$ where $\mathbf{e}_{i_*}$ be the i_*th standard basis vector. For any $\varepsilon \in (0, 1)$, let $\mathbf{x}_\varepsilon := (1 - \varepsilon)\mathbf{e}_{i_*} + \sum_{i \in N \setminus \{i_*\}} \varepsilon \mathbf{e}_i$. Thus, $\mathbf{u}_1 \cdot \mathbf{x}_\varepsilon < 0$ for all small enough $\varepsilon > 0$. Then, $K = 1$ and (2) implies $\mathbf{0} \succ \mathbf{x}_\varepsilon$ which contradicts Weak Pareto. Thus, $\mathbf{u}_1 \in \mathbb{R}^n_+$. ■

As in the above proof, axioms on a welfarism F such as Weak Pareto and Strong Pareto "translate" to properties of the ordering $\succsim$ which represents F in the following manner:

$$\mathbf{x} \gg \mathbf{y} \implies \mathbf{x} \succ \mathbf{y} \qquad \text{(Weak Pareto for } \succsim\text{)}$$
$$\mathbf{x} > \mathbf{y} \implies \mathbf{x} \succ \mathbf{y} \qquad \text{(Strong Pareto for } \succsim\text{)}$$

To see why, let us assume F satisfies Weak Pareto. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ be arbitrary such that $\mathbf{x} \gg \mathbf{y}$. Fix distinct a and b in A. Construct a utility profile $u \in \mathscr{U}^n$ as follows. Let $u(a) := \mathbf{x}$, $u(b) := \mathbf{y}$ and $u(c) := \mathbf{0}$ for all $c \in A \setminus \{a, b\}$. Thus, $u(a) \gg u(b)$. Since F satisfies Weak Pareto, we have $a\hat{F}(u)b$. Also, $aF(u)b$ iff $u(a) \succsim u(b)$. Thus, $a\hat{F}(u)b$ implies $u(a) \succ u(b)$. That is, $\mathbf{x} \succ \mathbf{y}$. This proves the "Weak Pareto for $\succsim$". Similar considerations establish "Strong Pareto for $\succsim$".

Now, we consider the following strengthening of the "invariance" requirement on F. Let $\Phi_{\text{CMNC}} \subseteq \Phi$ consist of all $\phi = (\phi_1, \ldots, \phi_n) \in \Phi_*$ corresponding to which there exists $(\alpha_i, \beta_i) \in \mathbb{R}_{++} \times \mathbb{R}$ for each $i \in N$ such that, for every $i \in N$,

$$\phi_i(t) = \alpha_i t + \beta_i \text{ for all } t \in \mathbb{R},$$

Observe, in contrast to elements of Φ_{CMUC}, now even the α_i's may depend on the individuals' identity. In fact, the subscript "CMNC" (instead of the earlier "CMUC") reflects "non–comparability" across individuals. The following definition is in order.

Definition 8: *A social welfare functional F is* cardinally measurable non–comparable *if, F is Φ_{CMNC}–invariant.*

Note, $\Phi_{\text{CMUC}} \subsetneq \Phi_{\text{CMNC}}$. Thus, any F which is Φ_{CMNC}–invariant must also be Φ_{CMUC}–invariant. Thus, our Theorem 7 will be useful in investigating the effect of Φ_{CMNC}–invariance on F. To that end, we must introduce the following two definitions.

Definition 9: *A social welfare functional F is a* dictatorship *if, there exists $i_* \in N$ such that, for any $a, b \in A$ and $u \in \mathscr{U}^n$,*

$$u_{i_*}(a) > u_{i_*}(b) \implies a\hat{F}(u)b.$$

Definition 10: *A social welfare functional F is a* serial dictatorship *if, there exists a permutation $i_1, \ldots, i_n$ of the individuals N such that, for any $a, b \in A$ and $u \in \mathscr{U}^n$,*

$$\big(\exists k \in N\big)\big[\, u_{i_l}(a) = u_{i_l}(b) \text{ if } l < k \;;\; u_{i_k}(a) > u_{i_k}(b) \,\big] \iff a\hat{F}(u)b.$$

Two remarks are in order. "Dictatorship" as in definition 9 is a weak notion: if the "dictator" i_* exhibits a strict preference for an alternative over another then the two alternatives are socially ranked according to his preference. However, definition 9 does not require the converse. Second, the idea underlying definition 10 is that individuals have been prioritized such that i_1 gets to be the "dictator" first but if i_1 exhibits indifference then i_2 gets to be the "dictator" ... and so on. Moreover, note that definition 10 requires a "two–way implication" in contrast to just the "one–way implication" as in definition 9. We provide a characterization of serial dictatorships as follows.

An inspection of definitions 5 and 10 reveals that serial dictatorships form a specifc subclass of lexicographic generalized utilitarianisms. As already noted, Φ_{CMNC}–invariance is stronger than Φ_{CMUC}–invariance with the latter characterizing lexicographic generalized utilitarianisms. Serial dictatorships are characterized by Φ_{CMNC}–invariance.

Theorem 9: *A social welfare functional is a serial dictatorship, if and only if, it is a welfarism which satisfies Strong Pareto and Cardinal Measurability Non–Comparability.*

Proof: For "sufficiency", let F be a welfarism that is Φ_{CMNC}–invariant and satisfies Strong Pareto. Then, the Theorem of Welfarism and the Representation Lemma imply the existence of an ordering $\succsim$ over $\mathbb{R}^n$ such that, for any $a, b \in A$ and $u \in \mathscr{U}^n$,

$$aF(u)b \iff u(a) \succsim u(b).$$

As $\Phi_{\mathrm{CMUC}} \subseteq \Phi_{\mathrm{CMNC}}$, note $\succsim$ satisfies (2) as in the proof of Theorem 7. Thus, there exists K orthonormal $\mathbf{u}_1, \ldots, \mathbf{u}_K \in \mathbb{R}^n$ such that:

$$\mathbf{x} \succsim \mathbf{y} \iff [\mathbf{u}_1 \cdot \mathbf{x}, \ldots, \mathbf{u}_K \cdot \mathbf{x}] \geq_L [\mathbf{u}_1 \cdot \mathbf{y}, \ldots, \mathbf{u}_K \cdot \mathbf{y}], \tag{3}$$

where $\geq_L$ is lexicographic order over $\mathbb{R}^K$. Also, note that because F satisfies Strong Pareto, it satisfies Weak Pareto. Observe, to show that F is a serial dictatorship, it is enough to show: $K = n$, and there exists a bijection $\sigma : N \to N$ such that $\mathbf{u}_k = \mathbf{e}_{\sigma(k)}$ for all $k = 1, \ldots, n$.

Step 1: We argue: $\mathbf{u}_1 \in \mathbb{R}^n_+$. Suppose $i_* \in N$ satisfies $\mathbf{u}_1 \cdot \mathbf{e}_{i_*} < 0$. Let $\varepsilon \in (0,1)$ and define $\mathbf{x}_\varepsilon := (1-\varepsilon)\mathbf{e}_{i_*} + \sum_{i \in N \setminus \{i_*\}} \varepsilon \mathbf{e}_i$. Note, $\mathbf{x}_\varepsilon \gg \mathbf{0}$. Then, as F satisfies Weak Pareto, we have: $\mathbf{x}_\varepsilon \succ \mathbf{0}$. Also, $\mathbf{u}_1 \cdot \mathbf{e}_{i_*} < 0$ implies that $\mathbf{u}_1 \cdot \mathbf{x}_\varepsilon < 0$ for all small enough $\varepsilon > 0$. Then, $\mathbf{0} \succ \mathbf{x}_\varepsilon$ by (3). However, this is a contradiction to the asymmetry of $\succ$. Hence, $\mathbf{u}_1 \cdot \mathbf{e}_i \geq 0$ for all $i = 1, \ldots, n$. That is, $\mathbf{u}_1 \in \mathbb{R}^n_+$.

Step 2: We argue: $\mathbf{u}_1 = \mathbf{e}_{i_*}$ for some $i_* \in N$. By normality of $\mathbf{u}_1$ and step 1, observe that it is enough to show: there does not exist distinct i and j in N such that $\mathbf{u}_1 \cdot \mathbf{e}_i > 0$ and $\mathbf{u}_1 \cdot \mathbf{e}_j > 0$ hold. Thus, suppose that i_* and j_* are distinct elements in N such that $\mathbf{u}_1 \cdot \mathbf{e}_{i_*} > 0$ and $\mathbf{u}_1 \cdot \mathbf{e}_{j_*} > 0$. For $\varepsilon > 0$, let $\mathbf{x}_\varepsilon \coloneqq -\varepsilon \mathbf{e}_{i_*} + \mathbf{e}_{j_*}$ and $\mathbf{y}_\varepsilon \coloneqq -\mathbf{e}_{i_*} + \varepsilon \mathbf{e}_{j_*}$. Then, $\mathbf{u}_1 \cdot \mathbf{e}_{j_*} > 0$ and $\mathbf{u}_1 \cdot \mathbf{e}_{i_*} > 0$ imply, respectively, $\mathbf{u}_1 \cdot \mathbf{x}_\varepsilon > 0$ and $\mathbf{u}_1 \cdot \mathbf{y}_\varepsilon < 0$ for all small enough $\varepsilon > 0$. Thus, (3) implies:

$$\mathbf{x}_\varepsilon \succ \mathbf{0} \quad \text{and} \quad \mathbf{0} \succ \mathbf{y}_\varepsilon. \tag{4}$$

Thus far, we have not appealed to the fact that F is Φ_{CMNC}–invariant. Now, we proceed to do so as follows. For each $i \in N$, define the map $\phi_i^* : \mathbb{R} \to \mathbb{R}$ by the following rule:

$$\phi_i^*(t) \coloneqq \kappa_i t \quad \text{for all } t \in \mathbb{R},$$

where κ_i is $1/\varepsilon$ or ε or 1 according as i is i_* or j_* or belongs to $N \backslash \{i_*, j_*\}$. Let $\phi^* \coloneqq (\phi_1^*, \dots, \phi_n^*)$. Since ϕ_i^* is a positive affine transformation for every $i \in N$, we have: $\phi^* \in \Phi_{\text{CMNC}}$. Recall, Φ_{CMNC}–invariance of F "translates" to an invariance property of $\succsim$ as follows:

$$\mathbf{x} \succsim \mathbf{y} \quad \text{iff} \quad \phi(\mathbf{x}) \succsim \phi(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and any $\phi \in \Phi_{\text{CMNC}}$. In particular, since $\phi^* \in \Phi_{\text{CMNC}}$ with $\mathbf{x}_\varepsilon$ and $\mathbf{0}$ in $\mathbb{R}^n$, we obtain:

$$\mathbf{x}_\varepsilon \succsim \mathbf{0} \quad \text{iff} \quad \phi^*(\mathbf{x}_\varepsilon) \succsim \phi^*(\mathbf{0}). \tag{5}$$

Now, observe that the definition of $\mathbf{x}_\varepsilon$, $\mathbf{y}_\varepsilon$ and ϕ^* imply: $\phi^*(\mathbf{0}) = \mathbf{0}$ and $\phi^*(\mathbf{x}_\varepsilon) = \mathbf{y}_\varepsilon$. Then, (5) implies the following:

$$\mathbf{x}_\varepsilon \succsim \mathbf{0} \quad \text{iff} \quad \mathbf{y}_\varepsilon \succsim \mathbf{0}. \tag{6}$$

However, (4) and (6) constitute a contradiction. Thus, there does not exist distinct i and j in N such that $\mathbf{u}_1 \cdot \mathbf{e}_i > 0$ and $\mathbf{u}_1 \cdot \mathbf{e}_j > 0$. Therefore, $\mathbf{u}_1 = \mathbf{e}_{i_*}$ for some $i_* \in N$.

Step 3: We argue: there is an injection $\sigma : \{1, \dots, K\} \to \{1, \dots, n\}$ such that $\mathbf{u}_k = \mathbf{e}_{\sigma(k)}$ for all $k \in \{1, \dots, K\}$. For this claim to hold, let $\sigma(1) \coloneqq i_*$ where i_* is as in the claim proven in step 2. Let $N^* \coloneqq N \backslash \{i_*\}$. Thus, $|N^*| = n - 1$. Recall, F maps any $u \in \mathscr{U}^n$ to some elment in $\mathscr{R}$. Define $F^* : \mathscr{U}^{n-1} \to \mathscr{R}$ as follows. For any $u^* \in \mathscr{U}^{n-1}$, let $u \in \mathscr{U}^n$ be that utility profile where $u_i = u_i^*$ for all $i \in N^*$, and $u_{i_*} : A \to \mathbb{R}$ be defined as $u_{i_*}(a) \coloneqq 0$ for all $a \in A$. Define $F^*(u^*) \coloneqq F(u)$. We next show that F^* satisfies the axioms required in Theorem 9 of F.

For Binary Independence of Irrelevant Alternatives of F^*, let $a, b \in A$ and $u^*, v^* \in \mathscr{U}^{n-1}$ satisfy $u^*(a) = v^*(a)$ and $u^*(b) = v^*(b)$. Then, by definition of the mapping $u^* \in \mathscr{U}^{n-1} \mapsto u \in \mathscr{U}^n$, we have $u(a) = v(a)$ and $u(b) = v(b)$ because $u_{i_*}(c) = 0 = v_{i_*}(c)$ for any $c \in \{a, b\}$. Since F satisfies this axiom, we obtain: $aF(u)b$ iff $aF(v)b$. However, $F^*(u^*) = F(u)$ and $F^*(v^*) = F(v)$ by definition. Hence, $aF^*(u^*)b$ iff $aF(v^*)b$. That is, F^* satisfies this axiom.

For Pareto Indifference, let $a, b \in A$ and $u^* \in \mathscr{U}^{n-1}$ be such that $u^*(a) = u^*(b)$. Then, by definition of the map $u^* \in \mathscr{U}^{n-1} \mapsto u \in \mathscr{U}^n$, we have $u(a) = u(b)$ as $u_{i_*}(a) = 0 = u_{i_*}(b)$. Since F satisfies Pareto Indifference, we have $a\bar{F}(u)b$. That is, $aF(u)b$ and $bF(u)a$. Further, $F^*(u^*)$ is $F(u)$ by definition. Thus, $aF^*(u^*)b$ and $bF^*(u^*)a$. That is, $a\bar{F}^*(u^*)b$. Hence, F^* satisfies Pareto Indifference.

For Strong Pareto, let $a, b \in A$ and $u^* \in \mathscr{U}^{n-1}$ satisfy $u^*(a) > u^*(b)$. That is, (i) $u_i^*(a) \geq u_i^*(b)$ for all $i \in N^*$, and (ii) $u_{i_{**}}^*(a) > u_{i_{**}}^*(b)$ for some $i_{**} \in N^*$, where $N^* = N \setminus \{i_*\}$. Now, by definition of the map $u^* \in \mathscr{U}^{n-1} \mapsto u \in \mathscr{U}^n$, we have $u_{i_*}(a) = 0 = u_{i_*}(b)$. Thus, both (i) $u_i(a) \geq u_i(b)$ for all $i \in N$, and (ii) $u_{i_{**}}(a) > u_{i_{**}}(b)$ for some $i_{**} \in N$, hold. That is, $u(a) > u(b)$ holds. Since F satisfies Strong Pareto, we have $a\hat{F}(u)b$. That is, $aF(u)b$ holds but $bF(u)a$ does not. As $F^*(u^*)$ is $F(u)$ by definition, it follows: $aF^*(u^*)b$ holds but $bF^*(u^*)a$ does not. That is, $a\hat{F}^*(u^*)b$ holds. Hence, F^* satisfies Strong Pareto.

For Cardinal Measurability Non–Comparability, let $a, b \in A$ and $u^* \in \mathscr{U}^{n-1}$. Denote by Φ_{CMNC}^{n-1} the collection of all $(n-1)$–tuples $\phi^* \equiv \langle \phi_i^* : i \in N^* \rangle$ such that, for each $i \in N^*$, $\phi_i(t) = \alpha_i t + \beta_i$ for all $t \in \mathbb{R}$ with $\alpha_i > 0$ and $\beta_i \in \mathbb{R}$. Fix an arbitrary $\phi^* \in \Phi_{\text{CMNC}}^{n-1}$. Let $\phi^* \circ u^* := \langle \phi_i^* \circ u_i^* : i \in N^* \rangle$. We must argue: $aF^*(u^*)b$ iff $aF^*(\phi^* \circ u^*)b$. With $u \in \mathscr{U}^n$ corresponding to u^*, let the n–tuple $\phi := \langle \phi_i : i \in N \rangle$ be defined by $\phi_i := \phi_i^*$ if $i \in N^*$, and $\phi_{i_*}(t) := t$ for all $t \in \mathbb{R}$. Now, observe that $\phi_i \circ u_i = \phi_i^* \circ u_i^*$ for all $i \in N^*$ as $u_i = u_i^*$ and $\phi_i = \phi_i^*$ for all $i \in n^*$. Moreover, $\phi_{i_*} \circ u_{i_*} = u_{i_*}$ as ϕ_{i_*} is the identity map on $\mathbb{R}$. Since u_{i_*} is the zero map on A, we have: $\phi \circ u$ corresponds to $\phi^* \circ u^*$. Thus, $aF^*(\phi^* \circ u^*)b$ iff $aF(\phi \circ u)$. Also, $aF^*(u^*)b$ iff $aF(u)b$. Note, $\phi \in \Phi_{\text{CMNC}}$. Since F is Φ_{CMNC}–invariant, we have: $aF(u)b$ iff $aF(\phi \circ u)b$. Thus, $aF^*(u^*)b$ iff $aF^*(\phi^* \circ u^*)b$. That is, F^* is Φ_{CMNC}^{n-1}–invariant.

Since F^* is a welfarism, let $\succsim^*$ be an ordering over $\mathbb{R}^{n-1}$ such that: $aF^*(u^*)b$ iff $u^*(a) \succsim^* u^*(b)$. With $u \in \mathscr{U}^n$ corresponding to $u^* \in \mathscr{U}^n$, note that $F^*(u^*)$ is $F(u)$ by definition. Also, $aF(u)b$ iff $u(a) \succsim u(b)$. Further, $u_{i_*}(a) = 0 = u_{i_*}(b)$ and $\mathbf{u}_1 = \mathbf{e}_{i_*}$. Then, (3) implies:

$$\mathbf{x} \succsim^* \mathbf{y} \iff [\mathbf{u}_2 \cdot \mathbf{x}, \ldots, \mathbf{u}_2 \cdot \mathbf{x}] \geq_L^* [\mathbf{u}_2 \cdot \mathbf{y}, \ldots, \mathbf{u}_2 \cdot \mathbf{y}],$$

for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{n-1}$ with $\geq_L^*$ as the lexicographic order on $\mathbb{R}^{K-1}$.

Hence, the arguments in steps 1 and 2 imply: $\mathbf{u}_2 = \mathbf{e}_{j_*}$ for some $j_* \in N^* = N \setminus \{i_*\}$. Again, we let $\sigma(2) := j_*$. Then, we iteratively repeat the generation of $(N^*, F^*, \succsim^*, \geq_L^*)$ from $(N, F, \succsim, \geq_L)$, thereby, assigning *distinct* values to $\sigma(1)$, $\sigma(2)$, ... up to $\sigma(K)$. This results in an injection $\sigma : \{1, \ldots, K\} \to \{1, \ldots, n\}$ with the property that $\mathbf{u}_k = \mathbf{e}_{\sigma(k)}$ for all $k = 1, \ldots, K$.

Step 4: We argue: $K = n$ and σ is a bijection from N to N. Since σ was already constructed to be an injection from $\{1, \ldots, K\}$ to $N = \{1, \ldots, n\}$, it is enough to show: $K = n$. Also, note that $K \in \mathbb{N}$ is such that $u_1, \ldots, \mathbf{u}_K$ are orthonormal vectors in $\mathbb{R}^n$. Since any system of orthonormal vectors must be linearly independent and the dimension of $\mathbb{R}^n$ is n, it follows that $K \leq n$.

Suppose $K < n$. Consider any $i_{**} \in N \setminus \sigma\big(\{1, \ldots, K\}\big)$ and define $\mathbf{x}_* := \mathbf{e}_{i_{**}}$. Clearly, $\mathbf{x}_* \neq \mathbf{0}$ and $\mathbf{x}_* \geq \mathbf{0}$. That is, $\mathbf{x}_* > \mathbf{0}$ holds. Since F satisfies Strong Pareto, $\mathbf{x}_* > \mathbf{0}$ implies: $\mathbf{x}_* \succ \mathbf{0}$. Also, since $\mathbf{u}_k = \mathbf{e}_{\sigma(k)}$ for all $k \in \{1, \ldots, K\}$, the fact that $i_{**} \in N \setminus \sigma\big(\{1, \ldots, K\}\big)$ implies: $\mathbf{u}_k \cdot \mathbf{x}_* = 0$ for all $k = 1, \ldots, K$. Then, $\mathbf{x}_* \sim \mathbf{0}$ by (3). However, $\mathbf{x}_* \succ \mathbf{0}$ and $\mathbf{x}_* \sim \mathbf{0}$ constitute a contradiction. Hence, $K = n$.

The proof of the theorem is complete. ■

Observe, steps 1 and 2 in the proof of Theorem 9 do *not* require the full force of the Strong Pareto axiom; Weak Pareto suffices. Also, steps 1 and 2 imply: $\mathbf{u}_1 = \mathbf{e}_{i_*}$ for some $i_* \in N$. Then, representation (3) and definition 9 immediately lead us to the following conclusion.

COROLLARY 1: *Suppose a social welfare functional is a welfarism that satisfies Weak Pareto and Cardinal Measurability Non–Comparability. Then, it must be a dictatorship.*

Obviously, the above conclusion continues to hold with any stronger invariance requirement on F. Let $\Phi_{\text{OMNC}} := \Phi_*$. Thus, $\phi \equiv (\phi_1, \ldots, \phi_n)$ is in Φ_{OMNC} iff ϕ_i is a strictly increasing for each $i \in N$ (the ϕ_i's may differ across i's). Consider the following definition.

DEFINITION 11: *A social welfare functional F is* ordinally measurable non–comparable if, *F is Φ_{OMNC}–invariant.*

Strengthening the invariance requirement in corollary 1 according to definition 11 implies Arrow's Impossibility Theorem as a further corollary to our Theorem 9. Arrow's result appears in this form as Theorem 4.1 in BLACKORBY ET AL. (1984).

5. BLACKWELL–GIRSHICK THEOREM

We consider a non–trivial binary relation $\succsim$ over a given non–empty convex subset C of an Euclidean space $\mathbb{R}^n$. We shall say that $\succsim$ admits a *linear representation* if, there exists[31] $\lambda \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ such that,

$$x \succsim y \iff \lambda \cdot x \geq \lambda \cdot y \quad \text{for all } x, y \in \mathbb{R}^n. \tag{7}$$

The binary relation $\succsim$ is a *preference* if it is complete and transitive. Observe, if a linear representation exists then the binary relation must be a preference. The asymmetric and symmetric components of $\succsim$ are denoted by $\succ$ and $\sim$. Existence of linear representations is the focus of this section. We begin with some axioms on $\succsim$.

INVARIANCE–1: *If $x, y \in C$ and $z \in \mathbb{R}^n$ satisfy $x+z, y+z \in C$ then,*

$$x \succsim y \iff x+z \succsim y+z.$$

INVARIANCE–2: *If $x, y \in C$ and $z \in \mathbb{R}^n$ satisfy $x+z, y+z \in C$ then,*

$$x \succ y \implies x+z \succ y+z, \text{ and}$$
$$x \sim y \implies x+z \sim y+z.$$

Observe, if a linear representation exists then Invariance–1 holds. Further, Invariance–1 and Invariance–2 are equivalent for a preference. More specifically, observe the following.

PROPOSITION 4: *Let $\succsim$ be a binary relation over any $C \subseteq \mathbb{R}^n$. Then,*

1. *Invariance–1 implies Invariance–2.*

2. *Invariance–2 and completeness imply Invariance–1.*

PROOF: Let $\succsim$ be a binary relation defined over a set $C \subseteq \mathbb{R}^n$. Fix any $x, y \in C$ and $z \in \mathbb{R}^n$ such that $x+z, y+z \in C$. First, assume $\succsim$ satisfies Invariance–1. Let $x \succ y$. Then, $x \succsim y$ by definition[32] of $\succ$. Thus, $x+z \succsim y+z$ by Invariance–1. Further, suppose $y+z \succsim x+z$. Then, $y \succsim x$ by Invariance–1. However, this is a contradiction to $x \succ y$ by the definition of $\succ$. Thus, $y+z \succsim x+z$ does not hold. Hence, $x+z \succ y+z$ by definition of $\succ$. That is, $(x \succ y \implies x+z \succ y+z)$ holds if $\succsim$ satisfies Invariance–1.

[31] Throughtout this section, by $\lambda \cdot x$ we shall denote $\lambda^1 x^1 + \ldots + \lambda^n x^n$ which is the standard inner product of the vectors $\lambda \equiv (\lambda^1, \ldots, \lambda^n)$ and $x \equiv (x^1, \ldots, x^n)$ in $\mathbb{R}^n$.

[32] The *asymmetric component* $\succ$ of $\succsim$ is defined as: $x \succ y \iff (\, x \succsim y$; not $y \succsim x\,)$. Further, the *symmetric component* $\sim$ of $\succsim$ is defined as: $x \sim y \iff (\, x \succsim y$; $y \succsim x\,)$.

Now, let $x \sim y$. Then, $x \succsim y$ by definition of $\sim$. Thus, $x+z \succsim y+z$ by Invariance–1. Also, $x \sim y$ implies $y \succsim x$ by definition of $\sim$. Then, $y+z \succsim x+z$ by definition of Invariance–1. Hence, $x+z \sim y+z$ by definition of $\sim$. That is, $(x \sim y \implies x+z \sim y+z)$ holds if $\succsim$ satisfies Invariance–1. Hence, Invariance–1 implies Invariance–2.

Now, assume $\succsim$ satisfies Invariance–2 and completeness. Let $x \succsim y$. Then, completeness of $\succsim$ implies either $x \succ y$ or $x \sim y$ holds by the definitions of $\succ$ and $\sim$. By Invariance–2, $x \succ y$ implies $x+z \succ y+z$. Further, $x+z \succ y+z$ implies $x+z \succsim y+z$ by definition of $\succ$. Thus, $x \succ y$ would imply $x+z \succsim y+z$. Again, by Invariance–2, $x \sim y$ implies $x+z \sim y+z$. Further, $x+z \sim y+z$ implies $x+z \succsim y+z$ by definition of $\sim$. Thus, $x \sim y$ would also imply $x+z \succsim y+z$. Hence, $x+z \succsim y+z$ holds. That is, $(x \succsim y \implies x+z \succsim y+z)$ holds.

Finally, let $x+z \succsim y+z$. Suppose $x \succsim y$ does not hold. Then, $y \succsim x$ by completeness of $\succsim$. Hence, $y \succ x$ by definition of $\succ$. By Invariance–2, $y \succ x$ implies $y+z \succ x+z$. Thus, $x+z \succsim y+z$ does not hold by definition of $\succ$. Since this is a contradiction, we must conclude that $x \succsim y$ holds. That is, $(x+z \succsim y+z \implies x \succsim y)$ holds. Hence, Invariance–1 follows from Invariance–2 and completeness. ■

Since $\succsim$ shall be a preference throughout this section, we shall not make any distinction between the two versions. Henceforth, we shall simply refer to either statement as "Invariance".

Resuming the discussion on necessary conditions, note $u(x) := \lambda \cdot x$ defines a $\mathbb{R}$–valued continuous utility representation $\succsim$ if (7) holds. Let C be endowed with the restriction of the standard topology of $\mathbb{R}^n$. Then, the following axiom is also necessary.

Continuity: *The sets $\{y \in C : y \succ x\}$ and $\{y \in C : x \succ y\}$ are open subsets of C for every $x \in C$.*

Invariance and Continuity are clearly necessary for the existence of a linear representation of the preference $\succsim$ over C. It was shown in Blackwell & Girshick (1954), which is their Theorem 4.3.1, if the set C is equal to $\mathbb{R}^n$ then these axioms are also sufficient.

Blackwell–Girshick Theorem: *Suppose $\succsim$ is a binary relation on $C = \mathbb{R}^n$. Then, $\succsim$ admits a linear representation, if and only if, $\succsim$ is a preference that satisfies Invariance and Continuity.*

It has been used extensively in microeconomic theory, for instance, in the minimax theory of games, foundations of utilitarianism in social choice and Roberts' type characterizations in mechanism design.

A closer inspection of its proof, which is based on the "Separating Hyperplane Theorem" of convex sets, has allowed adaptation when the set C is any *open* convex subset of $\mathbb{R}^n$ instead of the entire Euclidean space $\mathbb{R}^n$. It is the ability to adapt this result to more general domains which has in large measure made the result ubiquitous in applications. We shall generalize the Blackwell–Girshick Theorem to *arbitrary* convex subsets C of $\mathbb{R}^n$. However, to state our result, we need to formalize the intuitive idea of "a vector in $\mathbb{R}^n$ whose direction is *along* a convex set". To that end, some preliminaries are in order.

Definition 12: *Let $C \subseteq \mathbb{R}^n$ be non–empty. A* subspace generated by *C is a linear subspace S_0 of $\mathbb{R}^n$ that satisfies*:

1. *There exists $x_0 \in \mathbb{R}^n$ such that $C \subseteq x_0 + S_0$, and*
2. *If $x \in \mathbb{R}^n$ and S is a linear subspace of $\mathbb{R}^n$ such that $C \subseteq x + S$ then S_0 is a linear subspace of S.*

Some remarks are in order. Given any non–empty $C \subseteq \mathbb{R}^n$, there is the question of whether a subspace generated by C exists? Moreover, if it exists then is it unique? The answers to both these questions is in the affirmative which is formally stated as follows.

Proposition 5: *Let $C \subseteq \mathbb{R}^n$ be non–empty. Then, there exists a unique $S_C \subseteq \mathbb{R}^n$ such that S_C is the subspace generated by C. Moreover, if $x \in \mathbb{R}^n$ and $x_0 \in C$ then the following holds*:

$$C \subseteq x + S_C \iff x - x_0 \in S.$$

Note, all translations of S_C which contain C has been characterized. In terms of geometric intuition, S_C is the linear span of all vectors in C relative to some point in C chosen to be the origin. The proof is supplied in section A.III.1 of the Appendix.

Definition 13: *If $x_0 \in \mathbb{R}^n$ and $C \subseteq \mathbb{R}^n$ non–empty, x_0* is along *C if $x_0 \in S_C$ where S_C is the subspace generated by C.*

When C as in the above definition is an abstract subset of $\mathbb{R}^n$, the notion of a vector x_0 being "along" the set C is harder to intuitively justify. However, if the set C is convex then the above notion makes geometric sense. Moreover, observe that if $C \subseteq \mathbb{R}^n$ is non–empty and *open* then $S_C = \mathbb{R}^n$. Then, the phrase "λ is along S_C" is equivalent to "λ is in $\mathbb{R}^n$". Our generalization of the Blackwell–Girshick Theorem to *arbitrary* convex sets is as follows.

Theorem 10: *Suppose* $\succsim$ *is a binary relation over a convex* $C \subseteq \mathbb{R}^n$. *Then, there exists a unique* $\lambda \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ *along* C *such that*:

$$x \succsim y \iff \lambda \cdot x \geq \lambda \cdot y \quad \text{for all } x, y \in C,$$

if and only if, $\succsim$ *is a non–trivial preference that satisfies Continuity and Invariance.*

The proof of the above result is presented in section A.III.1 of the Appendix. We introduce the following axiom, on the binary relation, for the existence of lexicographic linear representations.

Convexity: *Let* C *be convex. If* $x, y \in C$ *and* $\alpha \in (0, 1)$ *then,*

$$x \succ y \implies \alpha x + (1 - \alpha) y \succ y.$$

This axiom is one of the standard assumptions on a preference in many settings. We now relax Continuity by replacing it in Theorem 10 with Convexity. This guarantees the existence of a unique lexicographic linear representation for an *arbitrary* convex set. Our characterization, proven in section A.III.2 of the Appendix, is as follows.

Theorem 11: *Suppose* $\succsim$ *is a binary relation over a convex* $C \subseteq \mathbb{R}^n$. *Then, there exists a unique list* $(\lambda_1, \ldots, \lambda_K)$ *of orthonormal vectors along* C *such that, for any* $x, y \in C$,

$$x \succsim y \iff [\lambda_1 \cdot x, \ldots, \lambda_K \cdot x] \geq_L [\lambda_1 \cdot y, \ldots, \lambda_K \cdot y],$$

where $\geq_L$ *is the lexicographic order over* $\mathbb{R}^K$, *if and only if,* $\succsim$ *is a non–trivial preference that satisfies Invariance and Convexity.*

The above axiom may be reminiscent of the "Independence" from section 3. Then, why is Invariance *additionally* required for existence of lexicographic linear representations? To see why, note that the clause "$\alpha \cdot p \oplus (1 - \alpha) \cdot r \succ \alpha \cdot q \oplus (1 - \alpha) \cdot r$", in "Independence", allows r to be *arbitrary.* However, in "Convexity", r is *equal* to q which is more restrictive—that is, *weaker*—than "Independence".

To place Theorems 8 and 9 in context, three remarks are in order. First, standard versions of the Blackwell–Girshick theorem also require "monotonicity" to substantially simplify proofs. Second, the existence of (lexicographic) linear representations is *not* assured under the said axioms for non–convex domains. Third, convex subsets exist which are both nowhere dense and are not Lebesgue measurable. Thus, additional assumptions such as "open subset" are restrictive.

6. ORDERED VECTOR SPACES

Our last application of the Decomposition Theorem is a characterization of ordered (real) vector spaces. The landmark result in this direction is by HAUSNER & WENDEL (1952) who considered arbitrary vector spaces over the real numbers. We shall only consider finite dimensional vector spaces. So, let V be an n–dimensional vector space over $\mathbb{R}$ and $\succ$ be a linear order[33] on V. Elements of V are denoted by x, y, ... and so on but the origin is denoted by $\mathbf{0}$. Scalars are denoted by α, β, ... and so on. Then, the pair $(V, \succ)$ is an *ordered vector space* if, the following properties are hold:

1. If $x \succ \mathbf{0}$ and $\lambda > 0$ then $\lambda x \succ \mathbf{0}$.
2. If $x \succ \mathbf{0}$ and $y \succ \mathbf{0}$ then $x + y \succ \mathbf{0}$.
3. $x \succ y$ if and only if $x - y \succ \mathbf{0}$.

In addition to the above defining properties, three simple but useful consequences are now stated and proved as follows.

PROPOSITION 6: *Let $(V, \succ)$ be an ordered vector space over $\mathbb{R}$. Then,*

1. *If $x \succ y$ then $x + z \succ y + z$.*
2. *If $x \succ y$ and $\lambda > 0$ then $\lambda x \succ \lambda y$.*
3. *$x \succ \mathbf{0}$ if and only if $\mathbf{0} \succ -x$.*

PROOF: For the first property, observe that $x \succ y$ iff $x - y \succ \mathbf{0}$ iff $(x + z) - (y + z) \succ \mathbf{0}$ iff $x + z \succ y + z$. Now, for the second property, suppose $\lambda > 0$ and $x \succ y$. Then, $x \succ y$ implies $x - y \succ \mathbf{0}$. Also, $x - y \succ \mathbf{0}$ and $\lambda > 0$ implies $\lambda(x - y) \succ \mathbf{0}$. That is, $\lambda x - \lambda y \succ \mathbf{0}$. Thus, we have $\lambda x \succ \lambda y$. Hence, $x \succ y$ and $\lambda > 0$ implies $\lambda x \succ \lambda y$.

Next, suppose $x \succ \mathbf{0}$. Then, $x + (-x) \succ \mathbf{0} + (-x)$. Since $x + (-x) = \mathbf{0}$ and $\mathbf{0} + (-x) = -x$, we have $\mathbf{0} \succ -x$. That is, $x \succ \mathbf{0}$ implies $\mathbf{0} \succ -x$. For the converse, suppose $\mathbf{0} \succ -x$. Then, $\mathbf{0} + x \succ -x + x$. Since $\mathbf{0} + x = x$ and $-x + x = \mathbf{0}$, it follows that $x \succ \mathbf{0}$. That is, $\mathbf{0} \succ -x$ implies $x \succ \mathbf{0}$. Hence, $x \succ \mathbf{0}$ if and only if $\mathbf{0} \succ -x$. ∎

In the rest of this section, we shall denote by $[n]$ the set $\{1, \ldots, n\}$. Note that $[n]$ is well–ordered by the restriction to $[n]$ of the standard order over $\mathbb{R}$. The following definition is critical.

[33] A *linear order* on a set X is a binary relation over X which is weakly connected, asymmetric and transitive. The standard order $>$ on $\mathbb{R}$ is an example.

Definition 14: *The* lexicographic function space *on* $[n]$ *is the ordered vector space* $(\mathscr{L}_n, \succ_n)$*, where* $\mathscr{L}_n$ *is the vector space of all* $\mathbb{R}$*–valued maps on* $[n]$ *and* $\succ_n$ *is the linear order on* $\mathscr{L}_n$ *that satisfies:*[34]

$$f \succ_n 0_n \text{ if and only if } f(k_f) > 0,$$

where $k_f := \min\{k \in [n] : f(k) \neq 0\}$ *for every* $f \in \mathscr{L}_n$*.*

Since any $f \in \mathscr{L}_n$ is naturally identifiable with a corresponding unique n–tuple of real numbers, it is clear that $\mathscr{L}_n$ is an n–dimensional vector space over $\mathbb{R}$. That is, $f \in \mathscr{L}_n \mapsto x_f \in \mathbb{R}^n$ is the linear bijection such that $\langle x_f, e_k \rangle = f(k)$ for all $k \in [n]$, where e_k is the kth standard basis vector in $\mathbb{R}^n$. Observe, the linear order $\succ_n$ satisfies:

$$f \succ_n g \quad \text{if and only if} \quad x_f >_L x_g,$$

where $>_L$ is the strict component of the standard lexicographic order $\geq_L$ on $\mathbb{R}^n$. Thus, definition 14 is justified.

For each $k \in [n]$, let $f_{k,n}$ be the $\mathbb{R}$–valued map over $[n]$ defined by: $f_{k,n}(i) := 1$ if $i = k$; otherwise, 0. Clearly, the n–tuple of maps $(f_{1,n}, \ldots, f_{n,n})$ is an ordered basis of $\mathscr{L}_n$. For an arbitrary n–dimensional ordered vector space $(V, \succ)$ and an ordered basis $\mathscr{B} \equiv (v_1, \ldots, v_n)$ of V, the linear bijection $\phi_{\mathscr{B}} : V \to \mathscr{L}_n$ such that:

$$\phi_{\mathscr{B}}(v_k) := f_{k,n} \quad \text{for every } k \in [n],$$

induces the linear order $\succ_{\mathscr{B}}$ on $\mathscr{L}_n$ defined by:

$$x \succ y \quad \text{iff} \quad \phi_{\mathscr{B}}(x) \succ_{\mathscr{B}} \phi_{\mathscr{B}}(y).$$

Thus, the moment an ordered basis $\mathscr{B}$ of V is chosen, the map $\phi_{\mathscr{B}}$ implements a linear embedding of the vector space V into the vector space $\mathscr{L}_n$. Recall, $\mathscr{L}_n$ already has the linear order $\succ_n$ defined over it which makes it an ordered vector space. Additionally, the linear order $\succ_{\mathscr{B}}$ induced by the embedding $\phi_{\mathscr{B}}$ also makes $\mathscr{L}_n$ a (possibly different) ordered vector space. A definition is in order.

Definition 15: *Let* $(V, \succ)$ *be an* n*–dimensional ordered vector space and* $\mathscr{B}$ *be an ordered basis of* V*. Then,* $(V, \succ)$ *is* isomorphic *to* $(\mathscr{L}_n, \succ_n)$ *via the ordered basis* $\mathscr{B}$ *if* $\succ_{\mathscr{B}} = \succ_n$*.*

Then, the fundamental result can be stated as follows.

[34] We denote by 0_n the map on the set $[n]$ which takes the value 0 for all $k \in [n]$.

Hausner–Wendel Theorem: *Suppose $(V, \succ)$ is an n–dimensional order vector space. Then, there exists an ordered basis $\mathscr{B}$ of V such that $(V, \succ)$ is isomorphic to $(\mathscr{L}_n, \succ_n)$ via $\mathscr{B}$.*

The lexicographic function space $(\mathscr{L}_n, \succ_n)$ on $[n]$ is *one* example of an n–dimensional ordered vector space. It is a basic fact in linear algebra that any n–dimensional vector space over $\mathbb{R}$, by the choice of an *arbitrary* ordered basis, is essentially $\mathbb{R}^n$. The above theorem claims that *every* n–dimensional ordered vector space over $\mathbb{R}$ is essentially the lexicographic function space on $[n]$ through the choice of *some* ordered basis. The objective of this section is to show that the above theorem is a consequence of our Decomposition Theorem.

Proof: Let $(V, \succ)$ be an n–dimensional ordered vector space over $\mathbb{R}$. Fix an arbitrary ordered basis $\mathscr{B}_0 \equiv (v_1, \ldots, v_n)$ of V. Let $\phi_{\mathscr{B}_0}$ be the linear bijection from V to $\mathbb{R}^n$ that satisfies the following:

$$\phi_{\mathscr{B}_0}(v_k) = e_k \quad \text{for all } k = 1, \ldots, n$$

with e_k being the kth standard basis vector of $W_* := \mathbb{R}^n$. Let $\succ^*$ be the linear order on $\mathbb{R}^n$ induced by $\succ$ under $\phi_{\mathscr{B}}$. That is,

$$x \succ y \quad \text{if and only if} \quad \phi_{\mathscr{B}_0}(x) \succ^* \phi_{\mathscr{B}_0}(y).$$

Observe, $(\mathbb{R}^n, \succ^*)$ is an n–dimensional ordered vector space. Define $U_* := \{x \in W_* : x \succ^* \mathbf{0}\}$, $V_* := \{x \in W_* : \mathbf{0} \succ^* x\}$ and $S_* := \{\mathbf{0}\}$. By the definition and properties of ordered vector spaces, U_* and V_* are cones with $V_* = -U_*$. Clearly, S_* is a 0–dimensional subspace of W_*. Moreover, (U_*, V_*, S_*) partition W_*. Then, by the Decomposition Theorem (Theorem 1 in section 2), there exists a unique orthonormal basis $\mathbf{U} \equiv (u_1, \ldots, u_n)$ of W_* such that $U_* = H_{\mathbf{U}}$ and $V_* = -H_{\mathbf{U}}$, where $H_{\mathbf{U}}$ is the graded halfspace generated by $\mathbf{U}$. Now, define the ordered basis $\mathscr{B} \equiv (w_1, \ldots, w_n)$ of V as follows:

$$w_k := \phi_{\mathscr{B}_0}^{-1}(u_k) \quad \text{for all } k = 1, \ldots, n.$$

Also, define the map $\psi : \mathbb{R}^n \to \mathscr{L}_n$ as follows. For each $k \in [n]$ let $\psi(u_k)$ be the function from $[n]$ to $\mathbb{R}$ defined by: $[\psi(u_k)](i) := 1$ if $i = k$; otherwise, 0. Moreover, uniquely extend ψ linearly to all of $\mathbb{R}^n$. Thus, ψ is linear bijection from $\mathbb{R}^n$ to $\mathscr{L}_n$. Hence, $\psi \circ \phi_{\mathscr{B}_0}$ is a linear bijection from V to $\mathscr{L}_n$. Let $\succ_{\mathscr{B}}$ be the linear order indcued by $\succ$ under $\psi \circ \phi_{\mathscr{B}_0}$. Then, the definitions of a graded halfspace (definition 1 in section 2) implies that $\succ_{\mathscr{B}} = \succ_n$. That is, $(V, \succ)$ is isomorphic to $(\mathscr{L}_n, \succ_n)$ via the ordered basis $\mathscr{B}$. This completes the proof. ∎

APPENDIX

*A.*I.1 *The Decomposition Theorem*

PROOF OF THEOREM 1: We first prove "existence". For any subspace $W_* \subseteq \mathbb{R}^m$, with U_*, V_* as cones in W_* and $S_* \subseteq W_*$ as a subspace, such that (U_*, V_*, S_*) is a partition of W_* and $V_* = -U_*$, let K be the codimension of S_* in W_*. That is, $K \coloneqq \dim(W_*) - \dim(S_*)$. Let $\text{St}[K]$ be the name of the following statement:

If U_* is nonempty then there exists a list of K orthonormal vectors in W_*, $\mathbf{U}_* \equiv \langle \mathbf{u}_*^k : k = 1, \ldots, K \rangle$, such that $U_* = \bigcup_{k=1}^{K} U_*^k$, where:

$$U_*^k \coloneqq \{\mathbf{w} \in W_* : \langle \mathbf{u}_*^l, \mathbf{w} \rangle = 0 \text{ if } l < k, \text{ and } \langle \mathbf{u}_*^k, \mathbf{w} \rangle > 0\}.$$

The proof that $\text{St}[K]$ holds is by mathematical induction on K.

Basis — We argue that $\text{St}[0]$ is true. Observe, with $K = 0$, we have $\dim(S_*) = \dim(W_*)$. This is because S_* is a *subspace* of W_* and has codimension K which is 0. Thus, $U_* = \varnothing = V_*$ as (U_*, V_*, S_*) is a partition of W_*. Thus, $\text{St}[0]$ is vacuously true.

Induction Step — Assume $\text{St}[K-1]$ holds. If $U_* \cup V_* = \varnothing$ then $\text{St}[K]$ holds vacuously. So, we assume $U_* \cup V_* \neq \varnothing$. Since $V_* = -U_*$, both $U_* \neq \varnothing$ and $V_* \neq \varnothing$. Hence, by Lemma 2, there exists a unique $\mathbf{u}_{**} \in W_* \setminus \{\mathbf{0}\}$ such that each of the following holds:

1. $S_* \subseteq \partial U_* = \partial V_* = \bar{U}_* \cap \bar{V}_* = \{\mathbf{w} \in W_* : \langle \mathbf{u}_{**}, \mathbf{w} \rangle = 0\}$.
2. $U_*^\circ = \{\mathbf{w} \in W_* : \langle \mathbf{u}_{**}, \mathbf{w} \rangle > 0\}$.
3. $V_*^\circ = \{\mathbf{w} \in W_* : \langle \mathbf{u}_{**}, \mathbf{w} \rangle < 0\}$.

Define $\mathbf{u}_*^1 \coloneqq \mathbf{u}_{**}/\|\mathbf{u}_{**}\|$, $S_{**} \coloneqq S_*$ and $W_{**} \coloneqq \{\mathbf{w} \in W_* : \langle \mathbf{u}_*^1, \mathbf{w} \rangle = 0\}$. Also, let $U_{**} \coloneqq U_* \cap W_{**}$ and $V_{**} \coloneqq V_* \cap W_{**}$. Clearly, U_{**} and V_{**} are (convex) cones in W_{**} and $S_{**} \subseteq W_{**}$ is a subspace such that $V_{**} = -U_{**}$. Moreover, (U_{**}, V_{**}, S_{**}) is a partition of W_{**}.

If $U_{**} \cup V_{**} = \varnothing$ then $W_{**} = S_*$. Also, $W_{**} = \partial U_*$ implies $\partial U_* = S_*$. Since $U_*^\circ \subseteq U_* \subseteq \bar{U}_*$ and $\partial U_* = \bar{U}_* \setminus U_*^\circ$, $U_* \setminus U_*^\circ \subseteq \partial U_*$. Also, $S_* = \partial U_*$ and $S_* \cap U_* = \varnothing$ implies $U_* = U_*^\circ$. Thus, $U_* = \{\mathbf{w} \in W_* : \langle \mathbf{u}_*^1, \mathbf{w} \rangle > 0\}$. Since $V_* = -U_*$, $V_* = \{\mathbf{w} \in W_* : \langle \mathbf{u}_*^1, \mathbf{w} \rangle < 0\}$. Moreover, $S_* = W_{**}$ implies $S_* = \{\mathbf{w} \in W_* : \langle \mathbf{u}_*^1, \mathbf{w} \rangle = 0\}$. Thus, $K = 1$ and $\text{St}[K]$ holds. That is, if $U_{**} \cup V_{**} = \varnothing$ then: $\text{St}[K-1]$ implies $\text{St}[K]$. Henceforth, we shall assume that $U_{**} \cup V_{**} \neq \varnothing$.

As $V_{**} = -U_{**}$, $U_{**} \neq \varnothing$ and $V_{**} \neq \varnothing$. Observe, W_{**} is a subspace of W_* with codimension 1. Then, $K' := \dim(W_{**}) - \dim(S_{**}) = K - 1$ as $S_* = S_{**}$. By $\textsc{St}[K-1]$, there exists a list of K_{**} orthonormal vectors in W_{**}, $\mathbf{U}_{**} \equiv \langle \mathbf{u}_{**}^k : k = 1, \ldots, K' \rangle$ such that $U_{**} = \bigcup_{k=1}^{K'} U_{**}^k$, where:

$$U_{**}^k := \{\mathbf{w} \in W_{**} : \langle \mathbf{u}_{**}^l, \mathbf{w} \rangle = 0 \text{ if } 1 \leq l < k, \text{ and } \langle \mathbf{u}_{**}^k, \mathbf{w} \rangle > 0\}$$

for $1 \leq k \leq K'$. Let $\mathbf{u}_*^k := \mathbf{u}_{**}^{k-1}$ for $2 \leq k \leq K$. Thus, $U_{**}^k = U_*^{k+1}$ for $1 \leq k \leq K'$ because $W_{**} = \{\mathbf{w} \in W_* : \langle \mathbf{u}_*^1, \mathbf{w} \rangle = 0\}$, where:

$$U_*^k := \{\mathbf{w} \in W_* : \langle \mathbf{u}_*^l, \mathbf{w} \rangle = 0 \text{ if } 1 \leq l < k, \text{ and } \langle \mathbf{u}_*^k, \mathbf{w} \rangle > 0\}$$

if $2 \leq k \leq K$. Since $U_{**} = \bigcup_{k=1}^{K'} U_{**}^k$, and $U_{**}^k = U_*^{k+1}$ if $1 \leq k \leq K'$, from $K = K' + 1$ we obtain: $U_{**} = \bigcup_{k=2}^{K} U_*^k$. Let $U_*^1 := U_*^\circ$. Recall, $U_*^\circ = \{\mathbf{w} \in W_* : \langle \mathbf{u}_*^1, \mathbf{w} \rangle > 0\}$. Thus, if we show $U_* = U_*^\circ \cup U_{**}$, then $U_* = \bigcup_{k=1}^{K} U_*^k$ follows. Now, $U_{**} = W_{**} \cap U_*$ and $W_{**} = \partial U_*$ imply $U_{**} = (\partial U_*) \cap U_*$. Also, $U_* \subseteq \bar{U}_*$ and $\partial U_* = \bar{U}_* \setminus U_*^\circ$ imply $U_* \setminus U_*^\circ = (\partial U_*) \cap U_*$. Thus, $U_{**} = U_* \setminus U_*^\circ$. Then, $U_* = U_*^\circ \cup (U_* \setminus U_*^\circ)$ as $U_*^\circ \subseteq U_*$. Hence, $U_* = U_*^\circ \cup U_{**}$ as required. That is, $\textsc{St}[K]$ holds. This completes the induction step and the proof of "existence".

We now prove "uniqueness". Let $\mathbf{U}_*^1 = \langle \mathbf{u}_*^{1,k} \in W_* : k = 1, \ldots, K_1 \rangle$ and $\mathbf{U}_*^2 = \langle \mathbf{u}_*^{2,k} \in W_* : k = 1, \ldots, K_2 \rangle$ be two lists of orthonormal vectors. For each $l \in \{1, 2\}$ and $1 \leq k \leq K_l$, define:

$$U_*^{l,k} := \{\mathbf{w} \in W_* : \langle \mathbf{u}_*^{l,j}, \mathbf{w} \rangle = 0 \text{ if } 1 \leq j < k, \text{ and } \langle \mathbf{u}_*^{l,k}, \mathbf{w} \rangle > 0\}.$$

Let $U_*^l := \bigcup_{k=1}^{K_1} U_*^{l,k}$ for $l = 1, 2$. We argue: $U_*^1 = U_*^2$ implies $\mathbf{U}_*^1 = \mathbf{U}_*^2$. Suppose, $U_*^1 = U_*^2$ and $\mathbf{U}_*^1 \neq \mathbf{U}_*^2$. Assume $K_1 \leq K_2$. Since $\mathbf{U}_*^1 \neq \mathbf{U}_*^2$, exactly one of the following cases must hold:

1. For some $K \leq K_1$, $\mathbf{u}_*^{1,K} \neq \mathbf{u}_*^{2,K}$ and $\mathbf{u}_*^{1,k} = \mathbf{u}_*^{2,k}$ if $k \leq K - 1$.

2. $\mathbf{u}_*^{1,k} = \mathbf{u}_*^{2,k}$ for each $k \in \{1, \ldots, K_*^1\}$ and $K_1 < K_2$.

In case (1), $\langle \mathbf{u}_*^{1,K}, \mathbf{u}_*^{2,K} \rangle < 1$; else, $\mathbf{u}_*^{1,K} = \mathbf{u}_*^{2,K}$ by Cauchy–Schwarz. Also, $\mathbf{w} := \mathbf{u}_*^{1,K} - \mathbf{u}_*^{2,K}$ implies $\langle \mathbf{u}_*^{1,K}, \mathbf{w} \rangle = 1 - \langle \mathbf{u}_*^{1,K}, \mathbf{u}_*^{2,K} \rangle = -\langle \mathbf{u}_*^{2,K}, \mathbf{w} \rangle$. Thus, $\langle \mathbf{u}_*^{1,K}, \mathbf{w} \rangle > 0$ and $\langle \mathbf{u}_*^{2,K}, \mathbf{w} \rangle < 0$. As $\langle \mathbf{u}_*^{2,K}, \mathbf{w} \rangle \neq 0$, $\mathbf{w} \notin U_*^{2,k}$ if $K + 1 \leq k \leq K_2$. Also, $\langle \mathbf{u}_*^{2,k}, \mathbf{w} \rangle \leq 0$ if $1 \leq k \leq K$ implies $\mathbf{w} \notin U_*^{2,k}$ if $1 \leq k \leq K$. Thus, $\mathbf{w} \notin U_*^2$. By orthonormality of the vectors in $\mathbf{U}_*^1$, $\mathbf{U}_*^2$ and that $\mathbf{u}_*^{1,k} = \mathbf{u}_*^{2,k}$ for $1 \leq k < K$, we have: $\langle \mathbf{u}_*^{1,k}, \mathbf{w} \rangle = 0$ if $1 \leq k < K$. Thus, $\langle \mathbf{u}_*^{1,K}, \mathbf{w} \rangle > 0$ implies $\mathbf{w} \in U_*^{1,K} \subseteq U_*^1$. That is, $\mathbf{w} \in U_*^1 \setminus U_*^2$ which is a contradiction to $U_*^1 = U_*^2$. Thus, the first of the two cases is ruled out.

In case (2), let $\mathbf{w} \coloneqq \mathbf{u}^{2,K_1+1}$. By orthonormality of $\mathbf{U}_*^1$ and that $\mathbf{u}_*^{1,k} = \mathbf{u}_*^{2,k}$ if $1 \leq k \leq K_1$, $\langle \mathbf{u}^{l,k}, \mathbf{w} \rangle = 0$ for $1 \leq k \leq K_1$ and $l \in \{1, 2\}$. Thus, $\mathbf{w} \notin U_*^{l,k}$ for $1 \leq k \leq K_1$ and $1 \leq l \leq 2$. In particular, $\mathbf{w} \notin U_*^1$. However, $\langle \mathbf{u}^{2,K_1+1}, \mathbf{w} \rangle = 1 > 0$. Hence, together with $\langle \mathbf{u}_*^{2,k}, \mathbf{w} \rangle = 0$ for $1 \leq k \leq K_1$, we have: $\mathbf{w} \in U_*^{2,K_1+1}$. That is, $\mathbf{w} \in U_*^2$ which contradicts $U_*^1 = U_*^2$. Thus, the second case is also ruled out. This completes the proof of "uniqueness". ■

A.I.2 Proof of Lemma 1

PROOF: Let $T_* \subseteq W_*$ be a subspace of codimension at least 2. We shall argue: $W_* \setminus T_*$ is path connected. Let $R_* \coloneqq O_{T_*} \subseteq W_*$ be the subspace orthogonal to T_*. Fix two arbitrary points $\mathbf{x}, \mathbf{y} \in W_* \setminus T_*$. Let $P^{R_*}(\mathbf{x})$ and $P^{R_*}(\mathbf{y})$ be the orthogonal projections of $\mathbf{x}$ and $\mathbf{y}$ onto R_*, respectively. Define $\pi_1 : [0, 1] \to W_*$ as: $\pi_1(t) \coloneqq \mathbf{x} + t(P^{R_*}(\mathbf{x}) - \mathbf{x})$ for every $t \in [0, 1]$. Since $\mathbf{x} \in W_* \setminus T_*$, $\alpha\mathbf{x} + (1 - \alpha)P^{R_*}(\mathbf{x}) \in W_* \setminus T_*$ for every $\alpha \in [0, 1]$. Thus, $\pi_1([0, 1]) \subseteq W_* \setminus T_*$. Further, π_1 is continuous with $\pi_1(0) = \mathbf{x}$ and $\pi_1(1) = P^{R_*}(\mathbf{x})$. Likewise, $\pi_2 : [0, 1] \to W_*$ defined by, $\pi_2(t) \coloneqq \mathbf{y} + t(P^{R_*}(\mathbf{y}) - \mathbf{y})$ for all $t \in [0, 1]$, is continuous, satisfies $\pi_2([0, 1]) \subseteq W_* \setminus T_*$. Further, $\pi_2(0) = \mathbf{y}$ and $\pi_2(1) = P^{R_*}(\mathbf{y})$.

Let $\mathbf{w}_1, \mathbf{w}_2 \in R_*$ be linearly independent and define $\psi : \mathbb{R}^2 \to R_*$ by: $\psi(\alpha_1, \alpha_2) \coloneqq \alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2$ for every $(\alpha_1, \alpha_2) \in \mathbb{R}^2$. Thus, ψ is linear homeomorphism. Then, the path connectedness of $\mathbb{R}^2 \setminus \{\mathbf{0}\}$ implies: there exists a continuous map $\pi_0 : [0, 1] \to R_* \setminus \{\mathbf{0}\}$ such that $\pi_0(0) = P^{R_*}(\mathbf{x})$ and $\pi_0(1) = P^{R_*}(\mathbf{y})$. Since $T_* \cap R_* = \mathbf{0}$, $\pi_0([0, 1]) \subseteq W_* \setminus T_*$. Now, consider the map $\pi_* : [0, 1] \to W_*$ defined as follows:

$$\pi_*(t) = \begin{cases} \pi_1(3t) & ; \quad \text{if } 0 \leq t < 1/3. \\ \pi_0(3t - 1) & ; \quad \text{if } 1/3 \leq t < 2/3. \\ \pi_2(3 - 3t) & ; \quad \text{if } 2/3 \leq t \leq 1. \end{cases}$$

Clearly, π_* is continuous, $\pi_*([0, 1]) \subseteq W_* \setminus T_*$ with $\pi_*(0) = \mathbf{x}$ and $\pi_*(1) = \mathbf{y}$. Thus, $W_* \setminus T_*$ is path connected.

Now, let $T_* \subsetneq W_*$ be a subspace with $W_* \setminus T_*$ path connected. By $W_* \setminus T_* \neq \varnothing$, the codimension of T_* in W_* is at least 1. Suppose the codimension is 1. Thus, T_* is a hyperplane in W_*. Let $\mathbf{w}_*$ be satisfy $\|\mathbf{w}_*\| = 1$ and $\langle \mathbf{w}_*, \mathbf{w} \rangle = 0$ for all $\mathbf{w} \in T_*$. Pick $\mathbf{x}, \mathbf{y} \in W_*$ such that $\langle \mathbf{x}, \mathbf{w}_* \rangle > 0$ and $\langle \mathbf{y}, \mathbf{w}_* \rangle < 0$. Let $\pi : [0, 1] \to W_* \setminus T_*$ be continuous with $\pi(0) = \mathbf{x}$ and $\pi(1) = \mathbf{y}$. Then, $f : [0, 1] \to \mathbb{R}$, defined by $f(t) \coloneqq \langle \pi(t), \mathbf{w}_* \rangle$ for all $t \in [0, 1]$, is continuous with $f(0) > 0$ and $f(1) < 0$. By continuity of f, $\langle \pi(t_*), \mathbf{w}_* \rangle = 0$ for some $t_* \in (0, 1)$. Thus, $\pi(t_*) \in T_*$ which contradicts $\pi([0, 1]) \subseteq W_* \setminus T_*$. ■

PROOF: We write $B(\mathbf{w},\varepsilon)$ for $B^{W_*}_{\|\cdot\|}(\mathbf{w},\varepsilon)$. First, we show "existence".

Step 1 — We claim: $(U_*^\circ)^c = \overline{V}_*$ and $(V_*^\circ)^c = \overline{U}_*$. For $(U_*^\circ)^c \subseteq \overline{V}_*$, suppose $\mathbf{w} \in (U_*^\circ)^c$ and $\mathbf{w} \notin \overline{V}_*$. As $\mathbf{w} \in (U_*^\circ)^c$, for some $\varepsilon_1 > 0$, $B(\mathbf{w},\varepsilon) \not\subseteq U_*$ if $\varepsilon \in (0,\varepsilon_1)$. By $U_* \cap (V_* \cup S_*) = \varnothing$, $B(\mathbf{w},\varepsilon) \cap (S_* \cup V_*) \neq \varnothing$ if $\varepsilon \in (0,\varepsilon_1)$. By $\mathbf{w} \in W_* \setminus \overline{V}_*$, for some $\varepsilon_2 > 0$, $B(\mathbf{w},\varepsilon) \subseteq W_* \setminus \overline{V}_*$ if $\varepsilon \in (0,\varepsilon_2)$. As $W_* \setminus \overline{V}_* \subseteq V_*^c$, $B(\mathbf{w},\varepsilon) \subseteq V_*^c$ if $\varepsilon \in (0,\varepsilon_2)$. As (U_*, V_*, S_*) partitions W_*, $B(\mathbf{w},\varepsilon) \subseteq U_* \cup S_*$ if $\varepsilon \in (0,\varepsilon_2)$.

Suppose $\mathbf{w} \notin U_*$. Let $\varepsilon \in (0,\varepsilon_2)$. Since $\mathbf{w} \in B(\mathbf{w},\varepsilon) \subseteq U_* \cup S_*$, $\mathbf{w} \notin U_*$ implies $\mathbf{w} \in S_*$. Also, $B(\mathbf{w},\varepsilon) \not\subseteq S_*$ because S_* is a proper subspace of W_* as $S_*^c = U_* \cup V_*$ is non–empty. Then, $B(\mathbf{w},\varepsilon) \cap U_* \neq \varnothing$. Let $\mathbf{w}_1 \in B(\mathbf{w},\varepsilon) \cap U_*$, $\delta\mathbf{w} := \mathbf{w}_1 - \mathbf{w}$ and $\mathbf{w}_2 := \mathbf{w} - \delta\mathbf{w}$. Note, $\mathbf{w}_2 \in B(\mathbf{w},\varepsilon)$. Observe, $\mathbf{w}_2 \notin S_*$. Else, $\mathbf{w}, \mathbf{w}_2 \in S_*$ implies $\delta\mathbf{w} \in S_*$. Then, $\mathbf{w}, \delta\mathbf{w} \in S_*$ implies $\mathbf{w}_1 \in S_*$ contradicting $U_* \cap S_* = \varnothing$. Thus, $\mathbf{w}_2 \in B(\mathbf{w},\varepsilon) \setminus S_*$. As $\mathbf{w}_2 \notin S_*$ and $\mathbf{w}_2 \in B(\mathbf{w},\varepsilon) \subseteq U_* \cup S_*$, $\mathbf{w}_2 \in U_*$. By $\mathbf{w}_1, \mathbf{w}_2 \in U_*$, $\mathbf{w} = (1/2)[\mathbf{w}_1 + \mathbf{w}_2] \in U_*$. Thus, $\mathbf{w} \in U_*$.

Let $\varepsilon_3 := \inf\{\|\mathbf{w}' - \mathbf{w}\| : \mathbf{w}' \in S_*\}$. Suppose, $\varepsilon_3 = 0$. Let $\{\mathbf{w}'_m\}_{m\in\mathbb{N}}$ be S_*–valued with $\lim_{m\to\infty} \|\mathbf{w}'_m - \mathbf{w}\| = 0$. As S_* is closed, $\mathbf{w} \in S_*$. But, $\mathbf{w} \in U_*$ and $U_* \cap S_* = \varnothing$ imply $\mathbf{w} \notin S_*$. Thus, $\varepsilon_* := \min\{\varepsilon_1, \varepsilon_2, \varepsilon_3\} > 0$. As $\varepsilon_* < \varepsilon_3$, $B(\mathbf{w},\varepsilon_*) \cap S_* = \varnothing$. By $\varepsilon_* < \varepsilon_1$ and $B(\mathbf{w},\varepsilon_*) \cap S_* = \varnothing$, $B(\mathbf{w},\varepsilon_*) \cap V_* \neq \varnothing$. By $\varepsilon_* < \varepsilon_2$, $B(\mathbf{w},\varepsilon_*) \cap S_* = \varnothing$ implies $B(\mathbf{w},\varepsilon_*) \subseteq U_*$. Thus, $U_* \cap V_* \neq \varnothing$—a contradiction. Hence, $(U_*^\circ)^c \subseteq \overline{V}_*$.

For $\overline{V}_* \subseteq (U_*^\circ)^c$, let $\mathbf{w} \in \overline{V}_*$ and suppose $\mathbf{w} \in U_*^\circ$. Then, $-\mathbf{w} \in V_*^\circ$ as $U_*^\circ = -V_*^\circ$ by $U_* = -V_*$. As $\mathbf{w} \in \overline{V}_*$, for some V_*–valued $\{\mathbf{w}_k\}_{k\in\mathbb{N}}$ with $\lim_{k\to\infty} \|\mathbf{w}_k - \mathbf{w}\| = 0$. As $U_* = -V_*$, $\{-\mathbf{w}_k\}_{k\in\mathbb{N}}$ is U_*–valued and converges to $-\mathbf{w}$. Since $-\mathbf{w} \in V_*^\circ$, there exists $k_* \in \mathbb{N}$ such that $-\mathbf{w}_k \in V_*^\circ$ for all $k \geq k_*$. As $V_*^\circ \subseteq V_*$, $-\mathbf{w}_k \in V_*$ if $k \geq k_*$. Thus, $\mathbf{w}_{k_*} \in V_*$ and $-\mathbf{w}_{k_*} \in V_*$. Then, $\mathbf{0} = \mathbf{w}_{k_*} + (-\mathbf{w}_{k_*}) \in V_*$ contradicting $V_* \cap S_* \neq \varnothing$ as $\mathbf{0} \in S_*$. Thus, $\overline{V}_* \subseteq (V_*^\circ)^c$. Hence, $(U_*^\circ)^c = \overline{V}_*$.

Step 2 — We claim: $T_* := \overline{U}_* \cap \overline{V}_*$ is a subspace. Since U_* and V_* are cones, $\overline{U}_*$ and $\overline{V}_*$ are closed cones. Then, $\alpha\mathbf{w} \in T_*$ if $\alpha \geq 0$ and $\mathbf{w} \in T_*$. Also, $T_* = -T_*$ as $V_* = -U_*$ implies $\overline{V}_* = -\overline{U}_*$. Thus, $\alpha\mathbf{w} \in T_*$ if $\alpha \in \mathbb{R}$ and $\mathbf{w} \in T_*$. Now, let $\mathbf{w}_1, \mathbf{w}_2 \in T_* \subseteq \overline{U}_*$. Get $\overline{U}_*$–valued $\{\mathbf{w}_k^1\}_{k\in\mathbb{N}}$ and $\{\mathbf{w}_k^2\}_{k\in\mathbb{N}}$ with $\lim_{k\to\infty} \|\mathbf{w}_k^1 - \mathbf{w}_1\| = 0$ and $\lim_{k\to\infty} \|\mathbf{w}_k^2 - \mathbf{w}_2\| = 0$. As $\overline{U}_*$ is a cone, $\mathbf{w}_k^1 + \mathbf{w}_k^2 \in \overline{U}_*$ if $k \in \mathbb{N}$. Also, $\lim_{k\to\infty} \|(\mathbf{w}_k^1 + \mathbf{w}_k^2) - (\mathbf{w}_1 + \mathbf{w}_2)\| = 0$. As $\overline{U}_*$ is closed, $\mathbf{w}_1 + \mathbf{w}_2 \in \overline{U}_*$. Similarly, $\mathbf{w}_1 + \mathbf{w}_2 \in \overline{V}_*$. Thus, $\mathbf{w}_1 + \mathbf{w}_2 \in T_*$ if $\mathbf{w}_1, \mathbf{w}_2 \in T_*$.

Step 3 — We claim: $\partial U_* = \partial V_* = T_*$. As $\partial U_* = \overline{U} \setminus U_*^\circ = \overline{U} \cap (U_*^\circ)^c$ and $(U_*^\circ)^c = \overline{V}_*$ by step 1, $\partial U_* = T_*$. Similarly, $\partial V_* = T_*$.

Step 4 — We claim: S_* is a subspace of T_*. As $U_*^\circ \subseteq U_*$ and $V_*^\circ \subseteq V_*$, we obtain $U_*^\circ \cup V_*^\circ \subseteq U_* \cup V_*$. Thus, $(U_* \cup V_*)^c \subseteq (U_*^\circ \cup V_*^\circ)^c$. As (U_*, V_*, S_*) partitions W_*, $S_* = (U_* \cup V_*)^c$. Hence, $S_* \subseteq (U_*^\circ \cup V_*^\circ)^c$. As $(U_*^\circ \cup V_*^\circ)^c = (U_*^\circ)^c \cap (V_*^\circ)^c$, step 1 implies $S_* \subseteq \overline{U}_* \cap \overline{V}_* = T_*$.

Step 5 — We claim: $U_*^\circ \neq \varnothing$ and $V_*^\circ \neq \varnothing$. Let the intersection of all subspaces of W_*, which contain U_*, be Z_*. Clearly, Z_* is the smallest subspace of W_* containing U_*. As $V_* = -U_*$, $V_* \subseteq Z_*$. Since W_* is finite dimensional, Z_* is a closed subset of W_*. Thus, $\overline{U}_*$ and $\overline{V}_*$ are contained in Z_*. Hence, $T_* = \overline{U}_* \cap \overline{V}_* \subseteq Z_*$. By step 4, $S_* \subseteq Z_*$. Since (U_*, V_*, S_*) partitions W_*, $W_* \subseteq Z_*$. Hence, $Z_* = W_*$. That is, W_* is the *minimal* subspace of W_* which contains U_*.

Let $P \coloneqq \{\mathbf{w}_k \in W_* : k = 1, \ldots, K\}$ be a set of (distinct) vectors in U_* which is maximally linearly independent . Thus, U_* is contained in the linear span of P. However, W_* is the minimal subspace of W_* that contains U_*. Hence, $K = \dim(W_*)$. Moreover, U_* is a cone containing P. Thus, the open set $\left\{ \sum_{k=1}^{K} \alpha_k \mathbf{w}_k : (\alpha_1, \ldots, \alpha_k) \in \mathbb{R}_{++}^K \right\}$ is contained in U_*. Hence, $U_*^\circ \neq \varnothing$. Similarly, $V_*^\circ \neq \varnothing$.

Step 6 — We claim: $\partial U_* = T_*$ has codimension 1 in W_*. Observe, $U_* \subseteq \overline{U}_* = \partial U_* \cup U_*^\circ = T_* \cup U_*^\circ$. Similarly, $V_* \subseteq T_* \cup V_*^\circ$. Also, $S_* \subseteq T_*$. As (U_*, V_*, S_*) partitions W_*, $W_* = T_* \cup (U_*^\circ \cup V_*^\circ)$. Also, $T_* = \partial U_*$ implies $T_* \cap U_*^\circ = \varnothing$. Similarly, $T_* \cap V_*^\circ = \varnothing$. Thus, $T_* \cap (U_*^\circ \cup V_*^\circ) = \varnothing$. Hence, $U_*^\circ \cup V_*^\circ = W_* \setminus T_*$. Now, $U_*^\circ \cap V_*^\circ = \varnothing$ as $U_* \cap V_* = \varnothing$. Thus, $W_* \setminus T_*$ is *not* connected. Hence, $W_* \setminus T_*$ is not path–connected. Also, if $\partial U_* = W_*$ then $\partial U_* = \overline{U}_* \setminus U_*^\circ$ implies $U_*^\circ = \varnothing$. However, step 5 implies $U_*^\circ \neq \varnothing$. Thus, ∂U_* is a *proper* subspace of W_*. Then, lemma 1 implies that T_* has codimension 1 in W_*.

Step 7 — We claim: there exists $\mathbf{u} \in W_*$ with $\|\mathbf{u}\| = 1$ such that $\partial U_* = \{\mathbf{w} \in W_* : \langle \mathbf{u}, \mathbf{w} \rangle = 0\}$ and $U_*^\circ = \{\mathbf{w} \in W_* : \langle \mathbf{u}, \mathbf{w} \rangle > 0\}$. As $U_*^\circ \neq \varnothing$, pick $\mathbf{w}_0 \in U_*^\circ$. Let $\mathbf{w}_1 \in \partial U_*$ be *the* orthogonal projection of $\mathbf{w}$ onto the subspace ∂U_*. Note, $\mathbf{w}_0 \neq \mathbf{w}_1$ as $\partial U_* \cap U_*^\circ = \varnothing$. Let $\mathbf{u} \coloneqq (\mathbf{w}_0 - \mathbf{w}_1)/\|\mathbf{w}_0 - \mathbf{w}_1\|$. Then, $T_* = \partial U_* = I_* \coloneqq \{\mathbf{w} \in W_* : \langle \mathbf{u}, \mathbf{w} \rangle = 0\}$ by step 6. Consider the cones, $P_* \coloneqq \{\mathbf{w} \in W_* : \langle \mathbf{u}, \mathbf{w} \rangle > 0\}$ and $N_* \coloneqq \{\mathbf{w} \in W_* : \langle \mathbf{u}, \mathbf{w} \rangle < 0\}$. As $(U_*^\circ, V_*^\circ, T_*)$ and (P_*, N_*, I_*) partition W_*, $U_*^\circ \cup V_*^\circ = P_* \cup N_*$ with $U_*^\circ \cap V_*^\circ = \varnothing$ and $P_* \cap N_* = \varnothing$. Also, U_*°, V_*°, P_* and N_* are each connected being cones. Observe, $\mathbf{w}_0 \in U_*^\circ \cap P_*$. Thus, $U_*^\circ = P_*$ and $V_*^\circ = N_*$. This proves "existence".

For "uniqueness", observe: $\mathbf{u}_1, \mathbf{u}_2 \in W_*$ with $\|\mathbf{u}_1\| = 1 = \|\mathbf{u}_2\|$ and $\{\mathbf{w} \in W_* : \langle \mathbf{u}_1, \mathbf{w} \rangle > 0\} = \{\mathbf{w} \in W_* : \langle \mathbf{u}_2, \mathbf{w} \rangle > 0\}$ implies $\mathbf{u}_1 = \mathbf{u}_2$. ■

Our objective is to prove Proposition 3. However, we first "geometrize" the problem as follows. Let $n := |Z|$ and $\phi : Z \to N := \{1, \ldots, n\}$ be an enumeration (*i.e.*, a bijection with N) of the set of basic prizes. Let $\mathbf{e}_i$ be the ith standard basis vector of $\mathbb{R}^n$. Then, $\mathscr{L}(Z)$ is in a bijection with the $(n-1)$–dimensional unit simplex $\Delta := \{\mathbf{x} \in \mathbb{R}^n_+ : \langle \mathbf{x}, \mathbb{1} \rangle = 1\}$, where $p \in \mathscr{L}(Z)$ is mapped to $\mathbf{p} \in \Delta$ such that:

$$\langle \mathbf{p}, \mathbf{e}_i \rangle = [p \circ \phi^{-1}](i) \text{ for all } i \in N.$$

Let $O_{\mathbb{1}} := \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x}, \mathbb{1} \rangle = 0\}$ be the subspace of $\mathbb{R}^n$ orthogonal to the "all–ones" vector $\mathbb{1}$. Also, define $\mathbf{a} := \mathbb{1}/n$. Note, $\mathbf{a} \in \Delta$ and $\Delta \subseteq \mathbf{a} + O_{\mathbb{1}}$. To each affine screening criterion $f \in \mathscr{F}$ associate the corresponding vector $\mathbf{f} \in \mathbb{R}^n$ such that:

$$\langle \mathbf{f}, \mathbf{e}_i \rangle = [f \circ \phi^{-1}](i) \text{ for all } i \in N.$$

Now, the definition of "affine screening criterion" requires that, if $p, q \in \mathscr{L}(Z)$ and $\alpha \in [0,1]$ then $f(\alpha \cdot p \oplus [1-\alpha] \cdot q)$ is must equal $\alpha f(p) + [1-\alpha] f(q)$. Thus, by the definition of $p \in \mathscr{L}(Z) \mapsto \mathbf{p} \in \Delta$ and $f \in \mathscr{F} \mapsto \mathbf{f} \in \mathbb{R}^n$, the bilinearity of the standard inner product on $\mathbb{R}^n$ implies the following the crucial property:

$$f(p) = \langle \mathbf{f}, \mathbf{p} \rangle \text{ for all } f \in \mathscr{F} \text{ and } p \in \mathscr{L}(Z).$$

We begin by translation of the structure of an affine local order to the "embedding space" $\mathbb{R}^n$. For this, consider any filter ϑ. Define the subset $S \subseteq O_{\mathbb{1}}$ corresponding to ϑ as:

$$S := \bigcap_{f \in \mathscr{F}} \big\{ \mathbf{x} \in O_{\mathbb{1}} : \langle \mathbf{f}, \mathbf{x} \rangle \leq \vartheta(f) \big\}.$$

Being the intersection of closed halfspaces, S is a closed convex subset of $O_{\mathbb{1}}$. Further, $\mathbf{0} \in S$ because $\vartheta(f) > 0$ for every $f \in \mathscr{F}$. However, observe that S may *fail* to be compact.

Now, we characterize the relation R_ϑ. First, let $p, q \in \mathscr{L}(Z)$ satisfy $pR_\vartheta q$. That is, $f(p) \leq f(q) + \vartheta(f)$ for all $f \in \mathscr{F}$ by the definition of R_ϑ. Hence, $\mathbf{p} \in \Delta \cap (\mathbf{q} + S)$. Second, assume $p, q \in \mathscr{L}(Z)$ satisfy $\mathbf{p} \in \Delta \cap (\mathbf{q} + S)$. Fix any $f \in \mathscr{F}$. Then, $\langle \mathbf{f}, \mathbf{p} - \mathbf{q} \rangle \leq \vartheta(f)$. Thus, $f(p) \leq f(q) + \vartheta(f)$ for each $f \in \mathscr{F}$. Hence, $pR_\vartheta q$. Thus, $pR_\vartheta q$ iff $\mathbf{p} \in \Delta \cap (\mathbf{q} + S)$. Then, the definition of S_ϑ implies:

$$pS_\vartheta q \quad \text{iff} \quad \big(\exists \mathbf{x} \in \Delta\big)\big[\, \mathbf{p}, \mathbf{q} \in \Delta \cap (\mathbf{x} + S) \,\big].$$

Let $\succ_0$ be a total order on $\mathscr{L}(Z)$. Let $\succ_0^*$ over Δ be defined as: $\mathbf{p} \succ_0^* \mathbf{q}$ iff $p \succ_0 q$. Define $\sim_0^*$ over Δ as: $\mathbf{p} \sim_0^* \mathbf{q}$ iff $p \sim_0 q$. Further, assume $\succsim_0$ satisfies Independence–3. This is equivalent to:

$$\mathbf{p} \succ_0^* \mathbf{q} \quad \text{iff} \quad \big(\, \forall \alpha \in (0,1) \,\big)\big[\, \alpha\mathbf{p} + (1-\alpha)\mathbf{r} \succ_0^* \alpha\mathbf{q} + (1-\alpha)\mathbf{r} \,\big].$$

Now, let $\succ_\vartheta$ be the affine local preorder on $\mathscr{L}(Z)$ induced by ϑ and $\succ_0$. Also, define $\succ_\vartheta^*$ over Δ as: $\mathbf{p} \succ_\vartheta^* \mathbf{q}$ iff $p \succ_\vartheta q$. Further, define $\sim_\vartheta^*$ over Δ as: $\mathbf{p} \sim_\vartheta^* \mathbf{q}$ iff $p \sim_\vartheta q$. Then, we have:

$$\mathbf{p} \succ_\vartheta^* \mathbf{q} \quad \text{iff} \quad \big(\exists \mathbf{x} \in \Delta\big)\big[\, \mathbf{p} \neq \mathbf{q}\ ;\ \mathbf{p}, \mathbf{q} \in \Delta \cap (\mathbf{x} + S)\ ;\ \mathbf{p} \succ_0^* \mathbf{q} \,\big].$$

Observe, $\mathbf{p} \sim_\vartheta^* \mathbf{q}$ iff (not $\mathbf{p} \succ_\vartheta^* \mathbf{q}$; not $\mathbf{q} \succ_\vartheta^* \mathbf{p}$). Let $\succsim_\vartheta^*$ be defined as $\succ_\vartheta^* \cup \sim_\vartheta^*$. Note, $\succ_\vartheta^*$ and $\sim_\vartheta^*$ are, respectively, the asymmetric and symmetric components of $\succsim_\vartheta^*$. Also, $p \succsim_\vartheta q$ iff $\mathbf{p} \succsim_\vartheta^* \mathbf{q}$. Now, we present a set of basic lemmas as follows.

Lemma A.II.1(a): *The relation $\succsim_\vartheta^*$ satisfies Independence–3.*

Proof: Assume $\mathbf{p}, \mathbf{q}, \mathbf{r} \in \Delta$ and $\mathbf{p} \succ_\vartheta^* \mathbf{q}$. Pick an arbitrary $\alpha \in (0,1)$. Let $\mathbf{s}_\alpha := \alpha\mathbf{p} + (1-\alpha)\mathbf{r}$ and $\mathbf{t}_\alpha := \alpha\mathbf{q} + (1-\alpha)\mathbf{r}$. Note, $\mathbf{s}_\alpha \neq \mathbf{t}_\alpha$ as $\mathbf{p} \neq \mathbf{q}$ because $\mathbf{p} \succ_\vartheta^* \mathbf{q}$. Also, $\mathbf{p} \succ_\vartheta^* \mathbf{q}$ implies $\mathbf{p} \succ_0^* \mathbf{q}$. Further, $\mathbf{p} \succ_0^* \mathbf{q}$ implies $\mathbf{s}_\alpha \succ_0^* \mathbf{t}_\alpha$. Now, $\mathbf{s}_\alpha$ and $\mathbf{t}_\alpha$ are in Δ because Δ is convex. Note, $\mathbf{p} \succ_\vartheta^* \mathbf{q}$ requires, there exists $\mathbf{x} \in \Delta$ such that $\mathbf{p}$ and $\mathbf{q}$ are in $\Delta \cap (\mathbf{x} + S)$. Let $\mathbf{x}_\alpha := \alpha\mathbf{x} + (1-\alpha)\mathbf{r}$. Convexity of Δ implies $\mathbf{x}_\alpha \in \Delta$. Since $\mathbf{p} \in \mathbf{x} + S$, let $\mathbf{y} \in S$ such that $\mathbf{p} = \mathbf{x} + \mathbf{y}$. Recall, $\mathbf{0} \in S$ and S is convex. Thus, $\mathbf{y}_\alpha := \alpha\mathbf{y} \in S$. Since $\mathbf{s}_\alpha = \mathbf{x}_\alpha + \mathbf{y}_\alpha$, we have: $\mathbf{s}_\alpha \in \Delta \cap (\mathbf{x}_\alpha + S)$. Similarly, $\mathbf{t}_\alpha \in \Delta \cap (\mathbf{x}_\alpha + S)$. As $\mathbf{x}_\alpha \in \Delta$, we obtain: $\mathbf{s}_\alpha \succ_\vartheta^* \mathbf{t}_\alpha$. Since $\alpha \in (0,1)$ was arbitrary, we conclude:

$$\mathbf{p} \succ_\vartheta^* \mathbf{q} \text{ implies } \big(\, \forall \alpha \in (0,1) \,\big)\big[\, \alpha\mathbf{p} + (1-\alpha)\mathbf{r} \succ_\vartheta^* \alpha\mathbf{q} + (1-\alpha)\mathbf{r} \,\big].$$

For the converse, let $\mathbf{s}_\alpha := \alpha\mathbf{p} + (1-\alpha)\mathbf{r}$ and $\mathbf{t}_\alpha := \alpha\mathbf{q} + (1-\alpha)\mathbf{r}$ for each $\alpha \in (0,1)$. Assume, $\mathbf{s}_\alpha \succ_\vartheta^* \mathbf{t}_\alpha$ for every $\alpha \in (0,1)$. Then, $\mathbf{s}_\alpha \succ_0^* \mathbf{t}_\alpha$ for every $\alpha \in (0,1)$. Hence, $\mathbf{p} \succ_0^* \mathbf{q}$ obtains. Further, $\mathbf{p} \neq \mathbf{q}$ because $\mathbf{s}_\alpha \neq \mathbf{t}_\alpha$ as required by $\mathbf{s}_\alpha \succ_\vartheta^* \mathbf{t}_\alpha$. Fix a $(0,1)$–valued sequence (α_n) such that $\lim_{n\to\infty} \alpha_n = 1$. Define $\mathbf{s}_n := \mathbf{s}_{\alpha_n}$ and $\mathbf{t}_n := \mathbf{t}_{\alpha_n}$ for all $n \in \mathbb{N}$. Then, for any $n \in \mathbb{N}$, there exists $\mathbf{x}_n \in \Delta$ such that $\mathbf{s}_n$ and $\mathbf{t}_n$ belong to $\Delta \cap (\mathbf{x}_n + S)$. Since Δ is compact and each sequence $(\mathbf{x}_n)$, $(\mathbf{s}_n)$ and $(\mathbf{t}_n)$ is Δ–valued, it is without loss of generality to assume that there exists $\mathbf{x}_*$, $\mathbf{s}_*$ and $\mathbf{t}_*$ in Δ such that $\lim_{n\to\infty}\|\mathbf{x}_n - \mathbf{x}_*\|_2 = 0$, $\lim_{n\to\infty}\|\mathbf{s}_n - \mathbf{s}_*\|_2 = 0$ and $\lim_{n\to\infty}\|\mathbf{t}_n - \mathbf{t}_*\|_2 = 0$. Since $\lim_{n\to\infty} \alpha_n = 1$ and $\lim_{n\to\infty}\|\mathbf{s}_n - \mathbf{s}_*\|_2 = 0$, the definition of $\mathbf{s}_\alpha$ implies $\mathbf{s}_* = \mathbf{p}$. By a similar argument, we obtain $\mathbf{t}_* = \mathbf{q}$.

We argue: $\mathbf{s}_* \in \Delta \cap (\mathbf{x}_* + S)$. Define $\mathbf{y}_* := \mathbf{s}_* - \mathbf{x}_*$, and $\mathbf{y}_n := \mathbf{s}_n - \mathbf{x}_n$ for each $n \in \mathbb{N}$. Then, $\lim_{n\to\infty}\|\mathbf{s}_n - \mathbf{s}_*\|_2 = 0$ and $\lim_{n\to\infty}\|\mathbf{x}_n - \mathbf{x}_*\|_2 = 0$ imply $\lim_{n\to\infty}\|\mathbf{y}_n - \mathbf{y}_*\|_2 = 0$. Now, the sequence $(\mathbf{y}_n)$ is S–valued as $\mathbf{s}_n \in \mathbf{x}_n + S$ and $\mathbf{y}_n = \mathbf{s}_n - \mathbf{x}_n$ for all $n \in \mathbb{N}$. Then, $\mathbf{y}_* \in S$ because S is a closed set. As $\mathbf{s}_* = \mathbf{x}_* + \mathbf{y}_*$ by definition of $\mathbf{y}_*$, we obtain: $\mathbf{s}_* \in \mathbf{x}_* + S$. Since $\mathbf{s}_* \in \Delta$, we have $\mathbf{s}_* \in \Delta \cap (\mathbf{x}_* + S)$. Likewise, $\mathbf{t}_* \in \Delta \cap (\mathbf{x}_* + S)$. Thus, $\mathbf{p}, \mathbf{q} \in \Delta \cap (\mathbf{x}_* + S)$ because $\mathbf{s}_* = \mathbf{p}$ and $\mathbf{t}_* = \mathbf{q}$. Then, $\mathbf{p} \neq \mathbf{q}$ and $\mathbf{p} \succ_0^* \mathbf{q}$ imply $\mathbf{p} \succ_\vartheta^* \mathbf{q}$. The converse has been proven. ■

Lemma A.II.1(b): *$\succ_\vartheta^*$ is acyclic.*

Proof: Suppose $\mathbf{p}, \mathbf{q}, \mathbf{r} \in \Delta$ satisfy $\mathbf{p} \succ_\vartheta^* \mathbf{q}$, $\mathbf{q} \succ_\vartheta^* \mathbf{r}$ and $\mathbf{r} \succ_\vartheta^* \mathbf{p}$. Then, the definition of $\succ_\vartheta^*$ and $\mathbf{p} \succ_\vartheta^* \mathbf{q}$ imply $\mathbf{p} \succ_0^* \mathbf{q}$. Similarly, we obtain $\mathbf{q} \succ_0^* \mathbf{r}$ and $\mathbf{r} \succ_0^* \mathbf{p}$. However, $\mathbf{p} \succ_0^* \mathbf{q}$ and $\mathbf{q} \succ_0^* \mathbf{r}$ imply $\mathbf{p} \succ_0^* \mathbf{r}$ because $\succ_0^*$ being a total order, over Δ, is transitive. Thus, both $\mathbf{p} \succ_0^* \mathbf{r}$ and $\mathbf{r} \succ_0^* \mathbf{p}$ hold. However, this contradicts the asymmetry of $\succ_0^*$. Hence, ($\mathbf{p} \succ_\vartheta^* \mathbf{q}$; $\mathbf{q} \succ_\vartheta^* \mathbf{r}$) implies (not $\mathbf{r} \succ_\vartheta^* \mathbf{p}$). ■

Let Θ be the class of all filters and Θ_c be the subclass of continuous filters. Also, let $F := \{\mathbf{x} \in O_{\mathbb{1}} : \|\mathbf{x}\|_2 = 1\}$. Recall, a filter $\vartheta \in \Theta$ is a map $\vartheta : \mathscr{F} \to \mathbb{R}_{++}$ that satisfies:

$$f' = \alpha f + \beta \quad \text{implies} \quad \vartheta(f') = \alpha\vartheta(f)$$

if $f, f' \in \mathscr{F}$ and $(\alpha, \beta) \in \mathbb{R}_{++} \times \mathbb{R}$. Thus, the map $f \in \mathscr{F} \mapsto \mathbf{f} \in \mathbb{R}^n$ induces a correspondence $\vartheta \in \Theta \mapsto \theta$, where θ corresponding to $\vartheta \in \Theta$ is a map from F to $\mathbb{R}_{++}$ defined as follows:

$$\theta(\mathbf{f}) := \vartheta(f) \quad \text{for every } \mathbf{f} \in F.$$

Note, $\vartheta \in \Theta_c$ iff θ is a continuous map, where F inherits from the standard topology on $\mathbb{R}^n$. Observe, for any $\vartheta \in \Theta$, the set S defined as $\bigcap_{f\in\mathscr{F}}\{\mathbf{x} \in O_{\mathbb{1}} : \langle \mathbf{f}, \mathbf{x}\rangle \leq \vartheta(f)\}$ can also be expressed as:

$$S = \bigcap_{\mathbf{f}\in F} \big\{\mathbf{x} \in O_{\mathbb{1}} : \langle \mathbf{f}, \mathbf{x}\rangle \leq \theta(\mathbf{f})\big\}.$$

Since $\|\cdot\|_2$ is continuous, the set F is closed; it is obviously bounded. Thus, F is compact. Then, θ achieves both a minimum and a maximum over F by continuity. For any $\vartheta \in \Theta_c$, define κ_ϑ as:

$$\kappa_\vartheta := \big[\sqrt{n} \cdot \max_{\mathbf{f}\in F} \theta(\mathbf{f})\big]^{-1}.$$

Define $\kappa \cdot S \coloneqq \{\kappa\mathbf{x} : \mathbf{x} \in S\}$. Then, observe that:

$$\kappa \cdot S = \bigcap_{\mathbf{f} \in F} \big\{\mathbf{x} \in O_{\mathbb{1}} : \langle \mathbf{f}, \mathbf{x} \rangle \leq [\kappa \cdot \theta](\mathbf{f})\big\}.$$

LEMMA *A*.II.1(*c*): *If $\kappa \in (0, \kappa_\vartheta)$ then $\succsim^*_{\kappa \cdot \vartheta}$ violates Independence–2.*

PROOF: Let M be a non–empty proper subset of N and set $m \coloneqq |M|$. Consider the following two vectors in $\mathbb{R}^n$.

$$\mathbf{p} \coloneqq \frac{1}{m} \sum_{j \in M} \mathbf{e}_j \quad \text{and} \quad \mathbf{q} \coloneqq \frac{1}{n-m} \sum_{j \in N \setminus M} \mathbf{e}_j.$$

Note, $\langle \mathbf{p}, \mathbb{1} \rangle = 1$ and $\langle \mathbf{p}, \mathbf{e}_i \rangle \geq 0$ for every $i \in N$. That is, $\mathbf{p} \in \Delta$. Likewise, $\mathbf{q} \in \Delta$. Fix any $\kappa \in (0, \kappa_\vartheta)$. Suppose, there exists $\mathbf{x}_0 \in \Delta$ such that $\mathbf{p}$ and $\mathbf{q}$ belong to $\Delta \cap (\mathbf{x}_0 + \kappa \cdot S)$. Let $\mathbf{f}_0 \coloneqq (\mathbf{p} - \mathbf{q})/\|\mathbf{p} - \mathbf{q}\|_2$. Thus, $\mathbf{f}_0 \in F$. Further, $\|\mathbf{p} - \mathbf{q}\|_2^2 = m \cdot [1/m^2] + (n-m) \cdot [1/(n-m)^2]$. That is, $\|\mathbf{p} - \mathbf{q}\|_2 = n^{1/2}/[m(n-m)]^{1/2}$. Since $\mathbf{f}_0 \in F$ and $\mathbf{p} \in \mathbf{x}_0 + \kappa \cdot S$, we have: $\langle \mathbf{f}_0, \mathbf{p} - \mathbf{x}_0 \rangle \leq [\kappa \cdot \theta](\mathbf{f}_0)$. Also, $\mathbf{f}_0 \in F$ and the definition of F implies $-\mathbf{f}_0 \in F$. Then, $\mathbf{q} \in \mathbf{x}_0 + \kappa \cdot S$ implies $\langle -\mathbf{f}_0, \mathbf{q} - \mathbf{x}_0 \rangle \leq [\kappa \cdot \theta](-\mathbf{f}_0)$. That is, $\langle \mathbf{f}_0, \mathbf{x}_0 - \mathbf{q} \rangle \leq [\kappa \cdot \theta](-\mathbf{f}_0)$. Note, $\langle \mathbf{f}_0, \mathbf{p} - \mathbf{x}_0 \rangle + \langle \mathbf{f}_0, \mathbf{x}_0 - \mathbf{q} \rangle = \langle \mathbf{f}_0, \mathbf{p} - \mathbf{q} \rangle$. Thus, $\langle \mathbf{f}_0, \mathbf{p} - \mathbf{q} \rangle \leq [\kappa \cdot \theta](\mathbf{f}_0) + [\kappa \cdot \theta](-\mathbf{f}_0) \leq 2\kappa \cdot \max_{\mathbf{f} \in F} \theta(\mathbf{f})$. Then, $\kappa \in (0, \kappa_\vartheta)$ implies $\langle \mathbf{f}_0, \mathbf{p} - \mathbf{q} \rangle < 2\kappa_\vartheta \cdot \max_{\mathbf{f} \in F} \theta(\mathbf{f}) \leq 2/n^{1/2}$ by definition of κ_ϑ. Observe, $\langle \mathbf{f}_0, \mathbf{p} - \mathbf{q} \rangle = \|\mathbf{p} - \mathbf{q}\|_2$. Hence, $n^{1/2}/[m(n-m)]^{1/2} < 2/n^{1/2}$. That is, $[m(n-m)]^{1/2} > (n/2)$. However, by the Arithmetic–Geometric Mean Inequality, we have: $(n/2) = [m + (n-m)]/2 \geq [m(n-m)]^{1/2}$. Thus, we have a contradiction. Hence, there does *not* exist $\mathbf{x}_0 \in \Delta$ such that $\mathbf{p}$ and $\mathbf{q}$ belong to $\Delta \cap (\mathbf{x}_0 + \kappa \cdot S)$. Thus, neither $\mathbf{p} \succ^*_{\kappa \cdot \vartheta} \mathbf{q}$ nor $\mathbf{q} \succ^*_{\kappa \cdot \vartheta} \mathbf{p}$ holds. That is, $\mathbf{p} \sim^*_{\kappa \cdot \vartheta} \mathbf{q}$ holds.

Recall, $\mathbf{a} \coloneqq \mathbb{1}/n$. Now, $m \in \{1, \ldots, n-1\}$ implies $\|\mathbf{p} - \mathbf{a}\|_2 > 0$ and $\|\mathbf{q} - \mathbf{a}\|_2 > 0$. Note, θ achieves a minimum over F by continuity as F is compact. Clearly, $\min_{\mathbf{f} \in F} \theta(\mathbf{f}) > 0$. Let $\varepsilon_* \coloneqq \min\{1, \mu_{\mathbf{p}}, \mu_{\mathbf{q}}\}$ where $\mu_{\mathbf{p}} \coloneqq \kappa \cdot \min_{\mathbf{f} \in F} \theta(\mathbf{f})/\|\mathbf{p} - \mathbf{a}\|_2$ and $\mu_{\mathbf{q}} \coloneqq \kappa \cdot \min_{\mathbf{f} \in F} \theta(\mathbf{f})/\|\mathbf{q} - \mathbf{a}\|_2$. Thus, $\varepsilon \in (0, 1]$. Consider any $\varepsilon \in (0, \varepsilon_*)$. Let $\mathbf{p}_\varepsilon \coloneqq \varepsilon\mathbf{p} + (1-\varepsilon)\mathbf{a}$ and fix an arbitrary $\mathbf{f} \in F$. Since $\|\mathbf{f}\|_2 = 1$ and $\mathbf{p}_\varepsilon - \mathbf{a} = \varepsilon(\mathbf{p} - \mathbf{a})$, Cauchy–Schwarz Inequality implies $|\langle \mathbf{f}, \mathbf{p}_\varepsilon - \mathbf{a} \rangle| \leq \varepsilon \|\mathbf{f}\|_2 \cdot \|\mathbf{p} - \mathbf{a}\|_2$. Then, from $\varepsilon \in (0, \varepsilon_*)$ we obtain: $\langle \mathbf{f}, \mathbf{p}_\varepsilon - \mathbf{a} \rangle \leq [\kappa \cdot \theta](\mathbf{f})$. Thus, $\mathbf{p}_\varepsilon \in \Delta \cap (\mathbf{a} + \kappa \cdot S)$ as $\mathbf{f} \in F$ is arbitrary. Similarly, $\mathbf{q}_\varepsilon \coloneqq \varepsilon\mathbf{q} + (1-\varepsilon)\mathbf{a}$ satisfies $\mathbf{q}_\varepsilon \in \Delta \cap (\mathbf{a} + \kappa \cdot S)$. Further, $\mathbf{p} \neq \mathbf{q}$ implies $\mathbf{p}_\varepsilon \neq \mathbf{q}_\varepsilon$ as $\varepsilon > 0$. Moreover, $\mathbf{p}_\varepsilon \succ^*_0 \mathbf{q}_\varepsilon$ or $\mathbf{q}_\varepsilon \succ^*_0 \mathbf{p}_\varepsilon$ as $\succ^*_0$ is a total order over Δ. Hence, $\mathbf{p}_\varepsilon \succ^*_{\kappa \cdot \vartheta} \mathbf{q}_\varepsilon$ or $\mathbf{q}_\varepsilon \succ^*_{\kappa \cdot \vartheta} \mathbf{p}_\varepsilon$ holds. Since $\mathbf{p} \sim^*_{\kappa \cdot \vartheta} \mathbf{q}$, Independence–2 is violated. ■

Thus, each claim in proposition 3 has been established.

A.II.2 *Proof of Lemma 3*

Let us recall from subsection 3.4, $\succsim^*$ is a complete and transitive binary relation over the $(n-1)$–dimensional unit simplex in $\mathbb{R}^n$ satisfying Independence–3 (henceforth, simply "Independence"):

$$\big[\mathbf{p} \succ^* \mathbf{q}\big] \quad \text{iff} \quad \big(\,\forall\alpha \in (0,1)\,\big)\big[\,\alpha\mathbf{p} + (1-\alpha)\mathbf{r} \succ^* \alpha\mathbf{q} + (1-\alpha)\mathbf{r}\,\big]$$

Moreover, $W_* := O_{\mathbb{1}}$ is the subspace of $\mathbb{R}^n$ orthogonal to $\mathbb{1}$ and we have its subsets U_*, V_* and S_* whose definitions are as follows:

$$\begin{aligned} U_* &:= \big\{\mathbf{w} \in W_* : \mathbf{a} + t\mathbf{w} \succ^* \mathbf{a} \quad \text{for some } t > 0\big\}, \\ V_* &:= \big\{\mathbf{w} \in W_* : \mathbf{a} \succ^* \mathbf{a} + t\mathbf{w} \quad \text{for some } t > 0\big\}, \\ S_* &:= \big\{\mathbf{w} \in W_* : \mathbf{a} + t\mathbf{w} \sim^* \mathbf{a} \quad \text{for some } t > 0\big\}. \end{aligned}$$

Further, $U(\mathbf{p})$, $L(\mathbf{p})$ and $I(\mathbf{p})$ are, respectively, the strict upper contour set, the strict lower contour set and the indifference set of an arbitrary $\mathbf{p} \in \Delta$. Now, we proceed to establish Lemma 3.

Proof: The argument involves the following steps.

Step 1 — We claim: for any $\mathbf{p} \in \Delta$, if $t_1, t_2 > 0$ and $\mathbf{w} \in W_*$ satisfy $\mathbf{p} + t_1\mathbf{w} \in \Delta$ and $\mathbf{p} + t_2\mathbf{w} \in \Delta$, then $\big(\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p}$ iff $\mathbf{p} + t_2\mathbf{w} \succ^* \mathbf{p}\big)$. Let $\mathbf{w} \in W_*$ and assume $0 < t_1 < t_2$ such that $\mathbf{p} + t_1\mathbf{w} \in \Delta$ and $\mathbf{p} + t_2\mathbf{w} \in \Delta$. First, assume $\mathbf{p} + t_2\mathbf{w} \succ^* \mathbf{p}$. Define $\alpha := t_1/t_2$. By Independence, $\mathbf{p} + t_1\mathbf{w} = \alpha(\mathbf{p} + t_2\mathbf{w}) + (1-\alpha)\mathbf{p} \succ^* \alpha\mathbf{p} + (1-\alpha)\mathbf{p} = \mathbf{p}$. That is, $\mathbf{p} + t_2\mathbf{w} \succ^* \mathbf{p}$ implies $\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p}$.

Assume $\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p}$. Suppose $\mathbf{p} \succ^* \mathbf{p} + t_2\mathbf{w}$. Let $\alpha := t_1/t_2$. By Independence, $\mathbf{p} = \alpha\mathbf{p} + (1-\alpha)\mathbf{p} \succ^* \alpha(\mathbf{p} + t_2\mathbf{w}) + (1-\alpha)\mathbf{p} = \mathbf{a} + t_1\mathbf{w}$ contradicting $\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p}$. Thus, $\mathbf{p} \succ^* \mathbf{p} + t_2\mathbf{w}$ is *not* possible.

Suppose $\mathbf{p} + t_2\mathbf{w} \sim^* \mathbf{p}$. Then, $\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p}$ implies $\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p} + t_2\mathbf{w}$. Fix an arbitrary $t \in (t_1, t_2)$ and let $\alpha := (t - t_1)/(t_2 - t_1)$. Note, $\alpha \in (0,1)$. Further, $\alpha t_2 + (1-\alpha)t_1 = t$. Thus, $\alpha(\mathbf{p} + t_2\mathbf{w}) + (1-\alpha)(\mathbf{p} + t_1\mathbf{w}) = \mathbf{p} + t\mathbf{w}$. By Independence, $\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p} + t_2\mathbf{w}$ implies $\mathbf{p} + t\mathbf{w} \succ^* \mathbf{p} + t_2\mathbf{w}$. Then, $\mathbf{p} + t_2\mathbf{w} \sim^* \mathbf{p}$ implies $\mathbf{p} + t\mathbf{w} \succ^* \mathbf{p}$. As $t \in (t_1, t_2)$ was arbitrary, $\mathbf{p} + t\mathbf{w} \succ^* \mathbf{p}$ for all $t \in (t_1, t_2)$. Further, $\mathbf{p} + t_1\mathbf{w} \succ^*$ implies: $\mathbf{p} + t\mathbf{w} \succ^* \mathbf{w}$ for all $t \in (0, t_1]$. Thus, $\mathbf{p} + t\mathbf{w} \succ^* \mathbf{p}$ for all $t \in (0, t_2)$. That is, $\alpha(\mathbf{p} + t_2\mathbf{w}) + (1-\alpha)\mathbf{p} \succ^* \alpha\mathbf{p} + (1-\alpha)\mathbf{p}$ for all $\alpha \in (0,1)$. By Independence, $\mathbf{p} + t_2\mathbf{w} \succ^* \mathbf{p}$ which contradicts $\mathbf{p} + t_2\mathbf{w} \sim^* \mathbf{p}$. Thus, $\mathbf{p} + t_2\mathbf{w} \sim^* \mathbf{p}$ is also *not* possible. Since $\succsim^*$ is complete, we have $\mathbf{p} + t_2\mathbf{w} \succ^* \mathbf{p}$. That is, $\mathbf{p} + t_1\mathbf{w}$ implies $\mathbf{p} + t_2\mathbf{w}$. The converse was already established. Thus, $\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p}$ iff $\mathbf{p} + t_2\mathbf{w} \succ^* \mathbf{p}$. Note, the assumption that $t_1 < t_2$ is without loss of generality.

Step 2 — We claim: for any $\mathbf{p} \in \Delta$, if $t_1, t_2 > 0$ and $\mathbf{w} \in W_*$ satisfy $\mathbf{p} + t_1\mathbf{w} \in \Delta$ and $\mathbf{p} + t_2\mathbf{w} \in \Delta$, then $\big(\mathbf{p} \succ^* \mathbf{p} + t_1\mathbf{w}$ iff $\mathbf{p} \succ^* \mathbf{p} + t_2\mathbf{w}\big)$. Define the binary relation $\succsim^{**}$ over Δ as follows: $\mathbf{q} \succsim^{**} \mathbf{r}$ iff $\mathbf{r} \succsim^{**} \mathbf{q}$. Observe that $\succsim^{**}$ is complete, transitive and satisfies Independence. Moreover, its asymmetric component $\succ^{**}$ satisfies: $\mathbf{q} \succ^{**} \mathbf{r}$ iff $\mathbf{r} \succ^* \mathbf{q}$. Thus, the argument in step 1 implies the claim.

Step 3 — We claim: for any $\mathbf{p} \in \Delta$, if $t_1, t_2 > 0$ and $\mathbf{w} \in W_*$ satisfy $\mathbf{p} + t_1\mathbf{w} \in \Delta$ and $\mathbf{p} + t_2\mathbf{w} \in \Delta$, then $\big(\mathbf{p} + t_1\mathbf{w} \sim^* \mathbf{p}$ iff $\mathbf{p} + t_2\mathbf{w} \sim^* \mathbf{p}\big)$. Let $t_1, t_2 > 0$ and $\mathbf{w} \in W_*$ satisfy $\mathbf{p} + t_1\mathbf{w} \in \Delta$ and $\mathbf{p} + t_2\mathbf{w} \in \Delta$. Assume, $\mathbf{p} + t_1\mathbf{w} \sim^* \mathbf{p}$. Suppose $\mathbf{p} + t_2\mathbf{w} \succ^* \mathbf{p}$. By step 1, $\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p}$ which is a contradiction. Now, suppose $\mathbf{p} \succ^* \mathbf{p} + t_2\mathbf{w}$. By step 2, $\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p}$ which is also a contradiction. Then, the completeness of $\succsim^*$ implies $\mathbf{p} + t_2\mathbf{w} \sim^* \mathbf{p}$. That is, $\mathbf{p} + t_1\mathbf{w} \sim^* \mathbf{p}$ implies $\mathbf{p} + t_2\mathbf{w} \sim^* \mathbf{p}$. Interchanging the roles of t_1 and t_2 implies the converse.

Step 4 — We claim: for any $\mathbf{w} \in W_*$, there exists $\varepsilon > 0$ such that $t \in (0, \varepsilon)$ implies $\mathbf{a} + t\mathbf{w} \in \Delta^\circ$. Note, if $\mathbf{w} = \mathbf{0}$ then every $\varepsilon > 0$ works. So, assume $\mathbf{w} \neq \mathbf{0}$. Let $\varepsilon \coloneqq \big(n \cdot \max\big\{|\langle \mathbf{e}_i, \mathbf{w}\rangle| : i = 1, \ldots, n\big\}\big)^{-1}$, where $n = |Z|$. Thus, $\varepsilon > 0$. Pick an arbitrary $t \in (0, \varepsilon)$ and let $\mathbf{p} \coloneqq \mathbf{a} + t\mathbf{w}$. Since $\mathbf{a} = \mathbb{1}/n$, $\langle \mathbf{e}_i, \mathbf{p}\rangle > 0$ for all $i = 1, \ldots, n$. That is, $\mathbf{p} \in \mathbb{R}^n_{++}$. Also, $\langle \mathbf{p}, \mathbb{1}\rangle = 1$ as $\mathbf{a} \in \Delta$ and $\mathbf{w} \in W_* = O_{\mathbb{1}}$. Thus, $\mathbf{p} \in \Delta^\circ$.

Step 5 — We claim: (U_*, V_*, S_*) partitions W_*. Note, each of U_*, V_* and S_* are subsets of W_* by their definitions. Let $\mathbf{w} \in W_*$. By step 4, $\mathbf{a} + t\mathbf{w} \in \Delta$ for some $t > 0$. Since $\succsim^*$ is complete, exactly one of $\mathbf{a} + t\mathbf{w} \succ^* \mathbf{a}$, $\mathbf{a} \succ^* \mathbf{a} + t\mathbf{w}$ or $\mathbf{a} + t\mathbf{w} \sim^* \mathbf{a}$ holds. Accordingly, $\mathbf{w}$ belongs to exactly one of U_*, V_* or S_*. Thus, (U_*, V_*, S_*) partitions W_*.

Step 6 — We claim: if $\mathbf{w} \in \mathbb{R}^n$ and $t > 0$ such that $\mathbf{a} + t\mathbf{w} \in \Delta$ then $\mathbf{w} \in W_*$. Note, $\langle \mathbf{a} + t\mathbf{w}, \mathbb{1}\rangle = 1 = \langle \mathbf{a}, \mathbb{1}\rangle$ as $\mathbf{a} + t\mathbf{w}$ and $\mathbf{a}$ are in Δ. Since $t \neq 0$, $\langle \mathbf{w}, \mathbb{1}\rangle = 0$. That is, $\mathbf{w} \in O_{\mathbb{1}}$. Recall, $W_* = O_{\mathbb{1}}$.

Step 7 — We claim: U_* and V_* are (convex) cones. Let $\mathbf{w} \in U_*$ and $\alpha > 0$. Since $\mathbf{w} \in U_*$, there exists $t > 0$ such that $\mathbf{a} + t\mathbf{w} \succ^* \mathbf{a}$. Define $t_* \coloneqq t/\alpha$. Then, $\mathbf{a} + t_*(\alpha\mathbf{w}) = \mathbf{a} + t\mathbf{w} \succ^* \mathbf{a}$. Thus, $\alpha\mathbf{w} \in U_*$. Hence, if $\mathbf{w} \in U_*$ and $\alpha > 0$ then $\alpha\mathbf{w} \in U_*$. Now, assume $\mathbf{w}_1, \mathbf{w}_2 \in U_*$. Then, $\mathbf{a} + t_1\mathbf{w}_1 \succ^* \mathbf{a}$ and $\mathbf{a} + t_2\mathbf{w}_2 \succ^* \mathbf{a}$ for some $t_1, t_2 > 0$. Let $\alpha \coloneqq t_2/(t_1 + t_2)$. Thus, $\alpha \in (0, 1)$ and $\alpha t_1 = (1 - \alpha)t_2$. Let $t_* \coloneqq \alpha t_1$ and note $t_* > 0$. By Independence, $\alpha(\mathbf{a} + t_1\mathbf{w}_1) + (1 - \alpha)(\mathbf{a} + t_2\mathbf{w}_2) \succ^* \mathbf{a}$; that is, $\mathbf{a} + t_*(\mathbf{w}_1 + \mathbf{w}_2) \succ^* \alpha\mathbf{a} + (1 - \alpha)\mathbf{a} = \mathbf{a}$. That is, $\mathbf{w}_1 + \mathbf{w}_2 \in U_*$. Thus, $\mathbf{w}_1, \mathbf{w}_2 \in U_*$ implies $\mathbf{w}_1 + \mathbf{w}_2 \in U_*$. Hence, U_* is a cone.

Step 9 — We claim: for any $\mathbf{p} \in \Delta$ and $\mathbf{w} \in W_*$, if $t_1, t_2 > 0$ satisfy $\mathbf{p} + t_1\mathbf{w} \in \Delta$ and $\mathbf{a} + t_2\mathbf{w} \in \Delta$ then each of the following hold.

$$\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p} \quad \text{iff} \quad \mathbf{a} + t_2\mathbf{w} \succ^* \mathbf{a}, \tag{8}$$
$$\mathbf{p} \succ^* \mathbf{p} + t_1\mathbf{w} \quad \text{iff} \quad \mathbf{a} \succ^* \mathbf{a} + t_2\mathbf{w}, \tag{9}$$
$$\mathbf{p} + t_1\mathbf{w} \sim^* \mathbf{p} \quad \text{iff} \quad \mathbf{a} + t_2\mathbf{w} \sim^* \mathbf{a}. \tag{10}$$

For a proof, assume throughout this step that $\mathbf{p} \in \Delta$, $\mathbf{w} \in W_*$ and $t_1, t_2 > 0$ satisfy $\mathbf{p} + t_1\mathbf{w} \in \Delta$ and $\mathbf{a} + t_2\mathbf{w} \in \Delta$. Moreover, assume $\mathbf{w} \neq \mathbf{0}$ and $\mathbf{p} \neq \mathbf{a}$. Otherwise, steps 1–3 imply the claim.

To show (8), let $\mathbf{q} \in \Delta$ satisfy $\mathbf{a} = \theta_1\mathbf{p} + (1-\theta_1)\mathbf{q}$ for some $\theta_1 \in (0,1)$. This is possible by step 4 and because Δ is convex. Let $t_3 \coloneqq \theta_1 t_1$ and note $t_3 > 0$. Then, $\mathbf{a} + t_3\mathbf{w} = \theta_1(\mathbf{p} + t_1\mathbf{w}) + (1-\theta_1)\mathbf{q}$.

Assume $\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p}$. Since $\theta_1 \in (0,1)$, Independence implies $\theta_1(\mathbf{p} + t_1\mathbf{w}) + (1-\theta_1)\mathbf{q} \succ^* \theta_1\mathbf{p} + (1-\theta_1)\mathbf{q}$. As $\mathbf{a} = \theta_1\mathbf{p} + (1-\theta_1)\mathbf{q}$ and $\mathbf{a} + t_3\mathbf{w} = \theta_1(\mathbf{p} + t_1\mathbf{w}) + (1-\theta_1)\mathbf{q}$, we have: $\mathbf{a} + t_3\mathbf{w} \succ^* \mathbf{a}$. By step 1, $\mathbf{a} + t_2\mathbf{w} \succ^* \mathbf{a}$. Since t_1 and t_2 are arbitrary, we obtain:

$$\mathbf{p} + t_1\mathbf{p} \succ^* \mathbf{p} \quad \text{implies} \quad \mathbf{a} + t_2\mathbf{w} \succ^* \mathbf{a}.$$

Assume $\mathbf{a} + t_2\mathbf{w} \succ^* \mathbf{a}$. Then, $\mathbf{a} + t_3\mathbf{w} \succ^* \mathbf{a}$. Consider an arbitrary $\theta \in (0,1)$ such that $\theta(\mathbf{p}+t_1\mathbf{w})+(1-\theta)\mathbf{q} \succ^* \theta\mathbf{p}+(1-\theta)\mathbf{q}$. Then, for any $\theta' \in (0,\theta)$, Independence implies $\theta'(\mathbf{p}+t_1\mathbf{w})+(1-\theta')\mathbf{q} \succ^* \theta'\mathbf{p}+(1-\theta')\mathbf{q}$. Also, by Independence: if $\theta'(\mathbf{p} + t_1\mathbf{w}) + (1-\theta')\mathbf{q} \succ^* \theta'\mathbf{p} + (1-\theta')\mathbf{q}$ for every $\theta' \in (0,\theta)$, then $\theta(\mathbf{p} + t_1\mathbf{w}) + (1-\theta)\mathbf{q} \succ^* \theta\mathbf{p} + (1-\theta)\mathbf{q}$. That is, if the set $\Theta \subseteq (0,1]$, defined as follows:

$$\Theta \coloneqq \big\{\theta \in (0,1] : \theta(\mathbf{p} + t_1\mathbf{w}) + (1-\theta)\mathbf{q} \succ^* \theta\mathbf{p} + (1-\theta)\mathbf{q}\big\},$$

is non–empty then: $\Theta = (0,\theta_*]$ for some unique $\theta_* \in (0,1]$. Observe, $\theta_1 \in \Theta \cap (0,1)$ because $\mathbf{a} + t_3\mathbf{w} \succ^* \mathbf{a}$. Hence, there exists a unique $\theta_* \in (0,1]$ such that $\Theta = (0,\theta_*]$.

Suppose $\theta_* \neq 1$. That is, $(\theta_*,1) \neq \varnothing$ and $\Theta \cap (\theta_*,1) = \varnothing$. Pick an arbitrary $\theta \in (\theta_*,1)$. Let $\mathbf{s} \coloneqq \alpha_*[\theta_*(\mathbf{p} + t_1\mathbf{w}) + (1-\theta_*)\mathbf{q}] + (1-\alpha_*)\mathbf{p}$ and $\mathbf{r} \coloneqq \theta\mathbf{p} + (1-\theta)\mathbf{q}$, where $\alpha_* \coloneqq (1-\theta)/(1-\theta_*)$. As $\theta \in (\theta_*,1)$, note $\alpha_* \in (0,1)$. Further, $\mathbf{s} = \mathbf{r} + t_4\mathbf{w}$ where $t_4 = \alpha_* t_1$. Note, $\mathbf{r} = \alpha_*[\theta_*\mathbf{p} + (1-\theta_*)\mathbf{q}] + (1-\alpha_*)\mathbf{p}$ by the definition of α_*. As $\theta_* \in \Theta$ and $\alpha_* \in (0,1)$, Independence implies $\mathbf{s} \succ^* \mathbf{r}$. That is, $\mathbf{r} + t_4\mathbf{w} \succ^* \mathbf{r}$. Let $t_5 \coloneqq \theta t_1$. Then, $\mathbf{r} + t_5\mathbf{w} = \theta(\mathbf{p} + t_1\mathbf{w}) + (1-\theta)\mathbf{q}$ as $\mathbf{r} = \theta\mathbf{p} + (1-\theta)\mathbf{q}$. Since $t_4, t_5 > 0$ and $\mathbf{r} + t_4\mathbf{w} \succ^* \mathbf{r}$, step 1 implies $\mathbf{r} + t_5\mathbf{w} \succ^* \mathbf{r}$. That is, $\theta(\mathbf{p}+t_1\mathbf{w})+(1-\theta)\mathbf{q} \succ^* \theta\mathbf{p}+(1-\theta)\mathbf{q}$. Hence, $\theta \in \Theta$ by the definition of Θ. This contradicts $\Theta \cap (\theta_*,1) = \varnothing$. Thus, $\theta_* = 1$. That is, $\Theta = (0,1]$. Hence, $\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p}$. Since t_1 and t_2 are arbitrary, the converse is established. This completes the proof of (8).

To show (9), define $\succsim^{**}$ over Δ by: $\mathbf{q} \succsim^{**} \mathbf{r}$ iff $\mathbf{r} \succsim^{**} \mathbf{q}$. Observe, $\succsim^{**}$ is complete, transitive and satisfies Independence. Moreover, its asymmetric component $\succ^{**}$ satisfies: $\mathbf{q} \succ^{**} \mathbf{r}$ iff $\mathbf{r} \succ^{*} \mathbf{q}$. Thus, the argument for (8) establishes (9).

To show (10), first assume $\mathbf{p} + t_1\mathbf{w} \sim^* \mathbf{p}$. Suppose $\mathbf{a} + t_2\mathbf{w} \succ^* \mathbf{a}$. Then, (8) implies $\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p}$. However, $\sim^*$ and $\succ^*$ are disjoint. This contradicts $\mathbf{p} + t_1\mathbf{w} \sim^* \mathbf{p}$. Thus, $\mathbf{a} + t_2\mathbf{w} \succ^* \mathbf{a}$ is *not* possible. Similarly, (9) implies $\mathbf{a} \succ^* \mathbf{a} + t_2\mathbf{w}$ is *not* possible. However, the union of $\succ^*$ and $\sim^*$ is $\succsim^*$. Moreover, $\succsim^*$ is a complete binary relation over Δ. Thus, $\mathbf{a} + t_2\mathbf{w} \sim^* \mathbf{a}$. Since t_1 and t_2 are arbitrary, we have:

$$\mathbf{p} + t_1\mathbf{w} \sim^* \mathbf{p} \quad \text{implies} \quad \mathbf{a} + t_2\mathbf{w} \sim^* \mathbf{a}.$$

For the converse, interchange the role of $\mathbf{p}$ with $\mathbf{a}$, and t_1 with t_2. This completes the proof of (10) and the step.

Step 10 — We claim: S_* is a cone. Pick an arbitrary $\mathbf{w} \in S_*$ and any $\alpha > 0$. Since $\mathbf{w} \in U_*$, there exists $t > 0$ such that $\mathbf{a} + t\mathbf{w} \sim^* \mathbf{a}$. Define $t_* := t/\alpha$. Then, $\mathbf{a} + t_*(\alpha\mathbf{w}) = \mathbf{a} + t\mathbf{w} \sim^* \mathbf{a}$. Thus, $\alpha\mathbf{w} \in S_*$. Hence, if $\mathbf{w} \in S_*$ and $\alpha > 0$ then $\alpha\mathbf{w} \in S_*$.

Now, assume $\mathbf{w}_1, \mathbf{w}_2 \in S_*$. Then, there exists $t_1, t_2 > 0$ such that $\mathbf{p}_1 := \mathbf{a} + t_1\mathbf{w}_1 \sim^* \mathbf{a}$ and $\mathbf{p}_2 := \mathbf{a} + t_2\mathbf{w}_2 \sim^* \mathbf{a}$. Let $\alpha_* := t_2/(t_1 + t_2)$ and $t_* := 2t_1t_2/(t_1 + t_2)$. Note, $\alpha_* \in (0, 1)$, $t_* > 0$ and $\alpha_* t_1 = (1 - \alpha_*)t_2 = t_*/2$. Define $\mathbf{p} := \mathbf{a} + t_*(\mathbf{w}_1 + \mathbf{w}_2)$. Observe, $\alpha_*\mathbf{p}_1 + (1 - \alpha_*)\mathbf{p}_2 = \mathbf{p}$. Define $t_{**} := \min\{t_1, t_2, t_*\}$. Clearly, $t_{**} > 0$. Define $\mathbf{q}_1 := \mathbf{a} + t_{**}\mathbf{w}_1$, $\mathbf{q}_2 := \mathbf{a} + t_{**}\mathbf{w}_2$ and $\mathbf{q} := \mathbf{a} + t_{**}(\mathbf{w}_1 + \mathbf{w}_2)$.

Suppose $\mathbf{w}_1 + \mathbf{w}_2 \in U_*$. Then, $\mathbf{q} \succ^* \mathbf{a}$ by step 1. As $\mathbf{p}_1 \sim^* \mathbf{a}$, step 9 implies $\mathbf{q}_1 \sim^* \mathbf{a}$. Thus, $\mathbf{q}_1 + t_{**}\mathbf{w}_2 = \mathbf{q} \succ^* \mathbf{q}_1$. Then, $\mathbf{a} + t_{**}\mathbf{w}_2 \succ^* \mathbf{a}$ by step 9. By step 1, $\mathbf{p}_2 \succ^* \mathbf{a}$. However, $\mathbf{p}_2 \sim^* \mathbf{a}$. This contradicts the fact that $\succ^*$ and $\sim^*$ are disjoint. Hence, $\mathbf{w}_1 + \mathbf{w}_2 \notin U_*$. Similarly, $\mathbf{w}_1 + \mathbf{w}_2 \notin V_*$. That is, $\mathbf{w}_1 + \mathbf{w}_2 \notin U_* \cup V_*$.

Since $\mathbf{w}_1, \mathbf{w}_2 \in S_*$ and $S_* \subseteq W_*$, that W_* is a subspace implies $\mathbf{w}_1 + \mathbf{w}_2 \in W_*$. Moreover, (U_*, V_*, S_*) is a partition of W_* by step 5. However, $\mathbf{w}_1 + \mathbf{w}_2 \notin U_* \cup V_*$. Thus, $\mathbf{w}_1 + \mathbf{w}_2 \in S_*$. Hence, if $\mathbf{w}_1, \mathbf{w}_2 \in S_*$ then $\mathbf{w}_1 + \mathbf{w}_2 \in S_*$. Moreover, we have already shown: if $\alpha > 0$ and $\mathbf{w} \in S_*$ then $\alpha\mathbf{w} \in S_*$. Hence, S_* is a cone.

Step 11 — We claim: S_* is a subspace. Since S_* has been shown to be cone, it is enough to argue: $S_* = -S_*$. Assume $\mathbf{w} \in S_*$. Thus, $\mathbf{a} + t_1\mathbf{w} \sim^* \mathbf{a}$ for some $t_1 > 0$. Also, by step 4, let $t_2 > 0$ be such that $\mathbf{a} - t_2\mathbf{w} \in \Delta$. Let $t := \min\{t_1, t_2\}$ and note that $t > 0$. Further, $\mathbf{a} + t\mathbf{w} \in \Delta$ and $\mathbf{a} - t\mathbf{w} \in \Delta$. Since $\mathbf{a} + t_1\mathbf{w} \sim^* \mathbf{a}$ and $t > 0$, step 3 implies: $\mathbf{a} + t\mathbf{w} \sim^* \mathbf{a}$. We shall now argue: $-\mathbf{w} \in S_*$.

Suppose $\mathbf{a} - t\mathbf{w} \succ^* \mathbf{a}$. Then, $\mathbf{a} + t\mathbf{w} \sim^* \mathbf{a}$ implies $\mathbf{a} - t\mathbf{w} \succ^* \mathbf{a} + t\mathbf{w}$. As $\alpha \in (0,1)$, Independence implies $\alpha(\mathbf{a} - t\mathbf{w}) + (1-\alpha)(\mathbf{a} + t\mathbf{w}) \succ^* \alpha(\mathbf{a} + t\mathbf{w}) + (1-\alpha)(\mathbf{a} + t\mathbf{w})$. As $\alpha = 1/2$, we have: $\mathbf{a} \succ^* \mathbf{a}$. This is a contradiction to the asymmetry of $\succ^*$. Thus, $\mathbf{a} - t\mathbf{w} \succ^* \mathbf{a}$ is *not* possible. Similarly, $\mathbf{a} \succ^* \mathbf{a} - t\mathbf{w}$ is *not* possible. However, $\succsim^*$ is a complete binary relation. Then, $\mathbf{a} - t\mathbf{w} \sim^* \mathbf{a}$. Thus, $-\mathbf{w} \in S_*$. Hence, $\mathbf{w} \in S_*$ implies $-\mathbf{w} \in S_*$. Note that $-(-\mathbf{w}) = \mathbf{w}$ for any $\mathbf{w} \in W_*$. Thus, $-\mathbf{w} \in S_*$ implies $\mathbf{w} \in S_*$. Hence, $\mathbf{w} \in S_*$ iff $-\mathbf{w} \in S_*$. Then, $-S_* = \{\mathbf{w} \in W_* : -\mathbf{w} \in S_*\}$ implies: $S_* = -S_*$.

Step 12 — We claim: $V_* = -U_*$. First, we argue: $-U_* \subseteq V_*$. Let $\mathbf{w} \in -U_*$. That is, $-\mathbf{w} \in U_*$. Thus, $\mathbf{a} + t_1(-\mathbf{w}) \succ^* \mathbf{a}$ for some $t_1 > 0$. By step 4, pick $t_2 > 0$ such that $\mathbf{a} + t_2\mathbf{w} \in \Delta$. Let $t \coloneqq \min\{t_1, t_2\}$ and note that $t > 0$. By step 1, $\mathbf{a} + t_1(-\mathbf{w}) \succ^* \mathbf{a}$ implies $\mathbf{a} + t(-\mathbf{w}) \succ^* \mathbf{a}$. That is, $\mathbf{a} - t\mathbf{w} \succ^* \mathbf{a}$. Also, $\mathbf{a} + t_2\mathbf{w} \in \Delta$ implies $\mathbf{a} + t\mathbf{w} \in \Delta$. With $\alpha \coloneqq 1/2$, Independence implies $\alpha(\mathbf{a} - t\mathbf{w}) + (1-\alpha)(\mathbf{a} + t\mathbf{w}) \succ^* \alpha\mathbf{a} + (1-\alpha)(\mathbf{a} + t\mathbf{w})$. That is, $\mathbf{a} \succ^* \mathbf{a} + t_*\mathbf{w}$ where $t_* \coloneqq (1-\alpha)t > 0$. Thus, $\mathbf{w} \in V_*$. Hence, we have: $-U_* \subseteq V_*$.

Second, we argue: $V_* \subseteq -U_*$. Let $\mathbf{w} \in V_*$. Thus, $\mathbf{a} \succ^* \mathbf{a} + t_1\mathbf{w}$ for some $t_1 > 0$. By step 4, pick $t_2 > 0$ such that $\mathbf{a} + t_2(-\mathbf{w}) \in \Delta$. Let $t \coloneqq \min\{t_1, t_2\}$. and note that $t > 0$. Then, $\mathbf{a} + t(-\mathbf{w}) \in \Delta$. That is, $\mathbf{a} - t\mathbf{w} \in \Delta$. Also, $\mathbf{a} \succ^* \mathbf{a} + t_1\mathbf{w}$ implies $\mathbf{a} \succ^* \mathbf{a} + t\mathbf{w}$ by step 2. Let $\alpha \coloneqq 1/2$. By Independence, $\mathbf{a} \succ^* \mathbf{a} + t\mathbf{w}$ implies $\alpha\mathbf{a} + (1-\alpha)(\mathbf{a} - t\mathbf{w}) \succ^* \alpha(\mathbf{a} + t\mathbf{w}) + (1-\alpha)(\mathbf{a} - t\mathbf{w})$. That is, $\mathbf{a} + t_*(-\mathbf{w}) \succ^* \mathbf{a}$ where $t_* \coloneqq (1-\alpha)t > 0$. Thus, $-\mathbf{w} \in U_*$. That is, $\mathbf{w} \in -U_*$. Hence, $V_* \subseteq -U_*$ holds. Thus, $V_* = -U_*$.

Step 13 — We claim: $U(\mathbf{p}) = \Delta \cap (\mathbf{p} + U_*)$, $L(\mathbf{p}) = \Delta \cap (\mathbf{p} + V_*)$ and $I(\mathbf{p}) = \Delta \cap (\mathbf{p} + S_*)$ for any $\mathbf{p} \in \Delta$. First, assume $\mathbf{q} \in U(\mathbf{p})$. That is, $\mathbf{q} \in \Delta$ and $\mathbf{q} \succ^* \mathbf{p}$. Let $\mathbf{w} \coloneqq \mathbf{q} - \mathbf{p}$ and $t \coloneqq 1$. Clearly, $\mathbf{q} = \mathbf{p} + t\mathbf{p}$ where $t > 0$. Also, $\mathbf{w} \in W_* = O_{\mathbb{1}}$ as $\langle \mathbf{p}, \mathbb{1} \rangle = 1 = \langle \mathbf{q}, \mathbb{1} \rangle$. Thus, $\mathbf{w} \in U_*$ by definition of U_*. Further, $\mathbf{q} = \mathbf{p} + \mathbf{w}$ by definition of $\mathbf{w}$. Hence, $\mathbf{p} \in \mathbf{p} + U_*$. Since $\mathbf{q} \in \Delta$, we have $\mathbf{q} \in \Delta \cap (\mathbf{p} + U_*)$. Hence, $\mathbf{q} \in U(\mathbf{p})$ implies $\mathbf{q} \in \Delta \cap (\mathbf{p} + U_*)$. That is, $U(\mathbf{p}) \subseteq \Delta \cap (\mathbf{p} + U_*)$.

Now, assume $\mathbf{q} \in \Delta \cap (\mathbf{p} + U_*)$. Thus, $\mathbf{q} = \mathbf{p} + t_1\mathbf{w}$ where $t_1 \coloneqq 1 > 0$ and $\mathbf{w} \in U_*$. Since $\mathbf{w} \in U_*$, $\mathbf{a} + t_2\mathbf{w} \succ^* \mathbf{a}$ for some $t_2 > 0$. By (8) of step 9, we have: $\mathbf{p} + t_1\mathbf{w} \succ^* \mathbf{p}$. That is, $\mathbf{q} \succ^* \mathbf{p}$. Thus, $\mathbf{q} \in U(\mathbf{p})$. Since $\mathbf{q} \in \Delta \cap (\mathbf{p} + U_*)$ is arbitrary, we have: $\mathbf{q} \in \Delta \cap (\mathbf{p} + U_*)$ implies $\mathbf{q} \in U(\mathbf{p})$. Hence, $\Delta \cap (\mathbf{p} + U_*) \subseteq U(\mathbf{p})$. Since $U(\mathbf{p}) \subseteq \Delta \cap (\mathbf{p} + U_*)$ also holds, we obtain: $U(\mathbf{p}) = \Delta \cap (\mathbf{p} + U_*)$. The remaining two equalities follow by similar arguments using (9) and (10) of step 9.

This completes the proof of the lemma. ■

We prove Theorem 10 from section 5, and associated results, generalizing the Blackwell–Girshick Theorem to *arbitrary* convex domains.

PROOF OF PROPOSITION 5: Let C be a non–empty subset of $\mathbb{R}^n$. Let $\mathscr{S}$ be the collection of every linear subspace S of $\mathbb{R}^n$ for which there exists a corresponding $x \in \mathbb{R}^n$ such that $C \subseteq x + S$. Note, $\mathbb{R}^n \in \mathscr{S}$ as $C \subseteq \mathbf{0} + \mathbb{R}^n$. Also, if $S \in \mathscr{S}$ then $\dim(S) \leq n$.

Fix $x_0 \in C$ and $S_* \in \mathscr{S}$. Also, let $x_* \in \mathbb{R}^n$ be such that $C \subseteq x_* + S_*$. Then, $C \subseteq x_0 + S_*$. To see why, let $x \in C$ be arbitrary. Define $y_0 := x_0 - x_*$. Also, let $y := x - x_*$. Since $C \subseteq x_* + S_*$, both $y_0 \in S_*$ and $y \in S_*$. Then, $y - y_0 \in S_*$ because S_* is a subspace. Since $y - y_0 = x - x_0$, we have $x - x_0 \in S_*$. Thus, $x \in x_0 + S_*$. Since $x \in C$ is arbitrary, we have: $C \subseteq x_0 + S_*$.

Let S_C be the intersection of all elements in $\mathscr{S}$. Since each element of $\mathscr{S}$ is a linear subspace of $\mathbb{R}^n$, so must be S_C. Further, fix any $x_0 \in C$. Then, $C \subseteq x_0 + S$ for all $S \in \mathscr{S}$. Then, $C \subseteq x_0 + S_C$ as well. Thus, $S_C \in \mathscr{S}$. Of course, $S_C \subseteq S$ for any $S \in \mathscr{S}$ by definition of S_C. That is, S_C is the unique subspace generated by C.

Now, let $x_0 \in C$ and $x_* \in \mathbb{R}^n$ such that $C \subseteq x_* + S_C$. Then, $x_0 = x_* + y_*$ for some $y_* \in S_C$. That is, $x_0 - x_* \in S_C$. Since S_C is a subspace, we have $x_* - x_0 \in S_C$. Finally, assume $x_0 \in C$ and $x_* \in \mathbb{R}^n$ such that $x_* - x_0 \in S_C$. Let $y_* := x_0 - x_*$. Then, $x_0 = x_* + y_*$. Further, since $S_C \in \mathscr{S}$, we know: $C \subseteq x_0 + S_C$. That is, $C \subseteq (x_* + y_*) + S_C$. Since $y_* \in S_C$ and S_C is a linear subspace, it follows that $y_* + S_C = S_C$. Thus, $(x_* + y_*) + S_C = x_* + S_C$. Hence, $C \subseteq x_* + S_C$. ■

For Theorem 10, we begin with some preliminaries. Fix a non–empty $C \subseteq \mathbb{R}^n$. For any $(m+1)$–tuple $(x_1, \ldots, x_{m+1})$ of vectors in C, define $x_0 := \sum_{k=1}^{m+1} x_k/(m+1)$ to be the *centroid* and the vectors $(p_1, \ldots, p_{m+1})$, where $p_k := x_k - x_0$, to be the *vertices*.

LEMMA *A.III.1(a)*: *Let x_0 be the centroid and $(p_1, \ldots, p_{m+1})$ be the vertices defined by any $(m+1)$–tuple $(x_1, \ldots, x_{m+1})$ of points in C. If some m of the vertices are linearly independent, then every collection of m vertices is linearly independent. Moreover, the collection of $(m+1)$ vertices is linearly dependent.*

PROOF: Fix any $(m+1)$–tuple $(x_1, \ldots, x_{m+1})$ of vectors in C. Let x_0 be the centroid and the $(m+1)$ vertices be $(p_1, \ldots, p_{m+1})$. Without any loss of generality, we assume that $(p_1, \ldots, p_m)$ are linearly independent and argue: $(p_2, \ldots, p_{m+1})$ are linearly independent.

First, note that $\sum_{k=1}^{m+1} p_k = \mathbf{0}$ by the definition of x_0 and the p_k's. In particular, the $(m+1)$ vertices are linearly independent. Moreover, $p_{m+1} = -\sum_{k=1}^{m} p_k$. Suppose there exists $\alpha_2, \ldots, \alpha_{m+1}$ in $\mathbb{R}$, not all equal to 0, such that $\sum_{k=2}^{m+1} \alpha_k p_k = \mathbf{0}$. Let $\beta_1 \coloneqq -\alpha_{m+1}$, and $\beta_k \coloneqq \alpha_k - \alpha_{m+1}$ for all $k = 2, \ldots, m$. Then, $\sum_{k=1}^{m} \beta_k p_k = \mathbf{0}$. Since $(p_1, \ldots, p_m)$ are linearly independent, we have: $\beta_k = 0$ for all $k = 1, \ldots, m$. That is, $\alpha_{m+1} = 0$ and $\alpha_k = \alpha_{m+1}$ for all $k = 2, \ldots, m$. This contradicts the supposition that not all α_k's are 0. Hence, $(p_2, \ldots, p_{m+1})$ is linearly independent. This completes the proof. ■

For the set C, let M_* denote the set of all $m \in \mathbb{N}$ for which some $(x_1, \ldots, x_{m+1})$ in C induces a centroid and $m+1$ vertices such that any m of the vertices are linearly independent but all the $m+1$ vertices are linearly dependent. Since $C \subseteq \mathbb{R}^n$, the set $M_* \subseteq \mathbb{N}$ is non–empty and bounded above by n. Define $m_* \coloneqq \max M_*$.

Definition A.III.1(b): *Let $C \subseteq \mathbb{R}^n$ be non–empty. Any (m_*+1)–tuple $(x_1, \ldots, x_{m_*+1})$ of vectors in C is a* coordinate system *for C if, every collection of its m_* vertices is linearly independent.*

By definition, m_* is the largest m such that any $(m+1)$–tuple in C induces vertices such that any proper subcollection, but not the whole, of it can be linearly independent. The above definition calls any such (m_*+1)–tuple a "coordinate system" for C. The reason for this choice of terminology is the following basic result about the representability of any arbitrary element x of the set C.

Lemma A.III.1(c): *Let $\mathscr{X} \equiv (x_1, \ldots, x_{m_*+1})$ be a coordinate system for C. Suppose x_0 is the centroid and $(p_1, \ldots, p_{m_*+1})$ are the m_*+1 vertices induced by $\mathscr{X}$. Then, for any $x \in C$, there exists $\nu_1, \ldots, \nu_{m_*+1}$ in $\mathbb{R}$ such that $x = x_0 + \sum_{k=1}^{m_*+1} \nu_k p_k$.*

Proof: Fix a coordinate system $\mathscr{X} \equiv (x_1, \ldots, x_{m_*+1})$ for C. Let $x \in C$ be arbitrary. Let $\mathscr{Y} \equiv (y_1, \ldots, y_{m_*+2})$ be defined as (1) $y_{m_*+2} \coloneqq x$, and (2) $y_k \coloneqq x_k$ for every $k = 1, \ldots, m_*+1$. Let $y_0 \coloneqq \sum_{k=1}^{m_*+2} y_k/(m_*+2)$ be the centroid and $(q_1, \ldots, q_{m_*+2})$, with $q_k \coloneqq y_k - y_0$ for $k = 1, \ldots, m_*+2$, be the vertices induced by $\mathscr{Y}$. Then, the m_*+1 vertices $(q_1, \ldots, q_{m_*+1})$ are linearly dependent. For otherwise, lemma A.III.1(a) would imply that every m_*+1 of the vertices are linearly independent with all the m_*+2 being linearly dependent. This would contradict the maximality of m_* in in the set M_*. That is, there exists $\alpha_1, \ldots, \alpha_{m_*+1}$ in $\mathbb{R}$, not all equal to 0, such that $\sum_{k=1}^{m_*+1} \alpha_k q_k = \mathbf{0}$.

Let $\mathscr{X}$ induce the centroid $x_0 \coloneqq \sum_{k=1}^{m_*+1} x_k/(m_*+1)$ and vertices $(p_1, \ldots, p_{m_*+1})$, where $p_k \coloneqq x_k - x_0$ for every $k = 1, \ldots, m_*+1$. Then, the following algebraic equality holds:

$$\sum_{k=1}^{m_*+1} \alpha_k q_k = \sum_{k=1}^{m_*+1} \alpha_k p_k - \frac{\theta}{m_*+2}(x - x_0), \tag{11}$$

where $\theta \coloneqq \sum_{k=1}^{m_*+1} \alpha_k$. Suppose $\theta = 0$. Then, $\sum_{k=1}^{m_*+1} \alpha_k q_k = \mathbf{0}$ and (11) imply $\sum_{k=1}^{m_*+1} \alpha_k p_k = \mathbf{0}$. Observe, $p_{m_*+1} = -\sum_{k=1}^{m_*} p_k$ by definition of x_0 and the p_k's. Thus, we obtain: $\sum_{k=1}^{m_*}(\alpha_k - \alpha_{m_*+1})p_k = \mathbf{0}$. Since $\mathscr{X}$ is a coordinate system, the vectors in $(p_1, \ldots, p_{m_*+1})$ are linearly independent. Hence, $\alpha_1 = \ldots = \alpha_{m_*+1}$. Since $\theta = 0$, we obtain: $\alpha_k = 0$ for all $k = 1, \ldots, m_*+1$. However, recall that not all of $\alpha_1, \ldots, \alpha_{m_*+1}$ are 0. Hence, we have a contradiction. Thus, $\theta \neq 0$. Then, (11) and $\sum_{k=1}^{m_*+1} \alpha_k q_k = \mathbf{0}$ imply the following:

$$x = x_0 + \sum_{k=1}^{m_*+1} \frac{(m_*+2)\alpha_k}{\theta} p_k.$$

Define $\nu_k \coloneqq (m_*+2)\alpha_k/\theta$ for every $k = 1, \ldots, m_*+1$ to complete the proof of the Lemma. ■

Henceforth, we fix a coordinate system $\mathscr{X} \equiv (x_1, \ldots, x_{m_*+1})$ for the set C. Also, let x_0 be the centroid and $(p_1, \ldots, p_{m_*+1})$ be the vertices induced by $\mathscr{X}$. Denote by W_* the linear span of $(p_1, \ldots, p_{m_*+1})$. Also, recall that S_C is the subspace generated by C.

Lemma A.III.1(d): $W_* = S_C$.

Proof: Since W_* is the linear span of $(p_1, \ldots, p_{m_*+1})$, Lemma A.III.1(c) implies that $x \in x_0 + W_*$ for every $x \in C$. That is, $C \subseteq x_0 + S_*$. Hence, $S_C \subseteq S_*$ because S_C is the subspace generated by C (see Definition 12 in section 5). Also, note that the dimension of W_* is m_*. This is because any m_* elements from $\{p_1, \ldots, p_{m_*+1}\}$ are linearly independent but the set of all the m_*+1 elements is linearly dependent.

Now, consider any p_k where $k \in \{1, \ldots, m_*\}$. Since $x_0 + p_k = x_k$ and $x_k \in C$, we have $p_k \in S_C$ because $C \subseteq x_0 + S_C$. Thus, S_C is a linear subspace of $\mathbb{R}^n$ containing the m_* linearly independent vectors $(p_1, \ldots, p_{m_*})$. Hence, dimension of S_C is *at least* m_*. That is, S_C is a linear subspace of the linear subspace W_* with the dimension of S_C at least as much as the dimension of W_*. Thus, $W_* = S_C$. ■

Recall, the notion of "subspace generated by C" was defined as the intersection of those linear subspaces S such that $C \subseteq x + S$ for some x in $\mathbb{R}^n$. This is an "extrinsic" description. The above lemma provides an "intrinsic" description of the same concept in terms of the (arbitrary) coordinate system $\mathscr{X}$ for the set C. The following lemma, which builds on the previous ones, shall be critical in the proof of Theorem 10.

LEMMA A.III.1(e): *Let $\mathscr{X} \equiv (x_1, \ldots, x_{m_*+1})$ be a coordinate system for C and x_0 be the centroid induced by $\mathscr{X}$. Then, for any $x \in C$, there exists $\lambda \in (0,1)$ and $\lambda_1, \ldots, \lambda_{m_*+1}$* in $\mathbb{R}_{++}$ *which satisfy $\sum_{k=1}^{m_*+1} \lambda_k = 1$ such that the following holds*:

$$x_0 = \lambda x + (1-\lambda) \sum_{k=1}^{m_*+1} \lambda_k x_k.$$

PROOF: Let the centroid and the vertices induced by the coordinate system $\mathscr{X}$ be x_0 and $(p_1, \ldots, p_{m_*+1})$, respectively. Fix an arbitrary $x \in C$. Then, by Lemma A.III.1(c), there exists $\nu_1, \ldots, \nu_{m_*+1}$ in $\mathbb{R}$ such that $x = x_0 + \sum_{k=1}^{m_*+1} \nu_k p_k$. For any $\lambda \in (0,1)$, define:

$$\mu_k(\lambda) := \frac{1}{1-\lambda}\Big(\frac{1}{m_*+1}\Big[1 + \lambda\Big(\sum_{k=1}^{m_*+1} \nu_k - 1\Big)\Big] - \lambda\nu_k\Big) \tag{12}$$

for every $k = 1, \ldots, m_* + 1$. Note, $\lim_{k\to 0} \mu_k(\lambda) = 1/(m_* + 1) > 0$. Further, the map $\lambda \in [0,1) \mapsto \mu_k(\lambda) \in \mathbb{R}$ is continuous. Thus, there exists $\lambda^* \in (0,1)$ such that, for any $\lambda \in (0, \lambda^*]$, $\mu_k(\lambda) > 0$ for all $k = 1, \ldots, m_* + 1$. Define $\lambda_k^* := \mu_k(\lambda_*)$ for all $k = 1, \ldots, m_* + 1$. Since $x_0 = \sum_{k=1}^{m_*+1} x_k/(m_* + 1)$ and $p_k = x_k - x_0$, from $x = x_0 + \sum_{k=1}^{m_*+1} \nu_k p_k$ and (12) we find that the following equality holds:

$$x_0 = \lambda^* x + (1-\lambda^*) \sum_{k=1}^{m_*+1} \lambda_k^* x_k.$$

Moreover, from (12) we obtain: $\sum_{k=1}^{m_*+1} \mu_k(\lambda) = 1$ for any $\lambda \in (0,1)$. In particular, $\sum_{k=1}^{m_*+1} \lambda_k^* = 1$ holds. ■

Geometrically, for every $x \in C$, there exists a "weighted average" $y_x := \sum_{k=1}^{m_*+1} \lambda_k x_k$ of the coordinate system $\mathscr{X}$ such that the centroid x_0 is some "weighted average" $\lambda x + (1-\lambda) y_x$ of the points x and y_x. Finally, we shall also need the following technical result.

Lemma A.III.1(f): *Let* $L \subseteq \mathbb{R}$ *be an interval of the form* $(0.\theta)$ *or* $(0, \theta]$ *for some* $\theta > 0$. *Suppose* $\tau \subseteq L$ *satisfies the following*:

1. $\big(q \in \mathbb{Q}_{++} \, ; \, t \in \tau \, ; \, qt \in L \big) \implies qt \in \tau$, *and*
2. $\big(\exists \, t_* > 0 \big) \big(\exists \, \varepsilon > 0 \big) \big[\, (t_* - \varepsilon, t_* + \varepsilon) \subseteq \tau \, \big]$.

Then, $\tau = L$.

Proof: Let L be the interval $(0, \theta)$ or $(0, \theta]$ for some $\theta > 0$. It will be enough to argue, $L \subseteq \tau$. Assume $t_* > 0$ and $\varepsilon > 0$ are such that $(t_* - \varepsilon, t_* + \varepsilon) \subseteq \tau$. First, let $s \in (0, \theta)$ be arbitrary.

Define $\alpha_* := s/(t_* + \varepsilon)$ and $\beta_* := \min\{\beta_1, \beta_2\}$, where $\beta_1 := s/(t_* - \varepsilon)$ and $\beta_2 := \theta/(t_* + \varepsilon)$. Note, $\alpha_* < \beta_1$ as $s > 0$ and $t_* + \varepsilon > t_* - \varepsilon > 0$. Also, $s < \theta$ and $t_* + \varepsilon > 0$ imply $\alpha < \beta_2$. Thus, $\alpha_* < \beta_*$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists $q \in \mathbb{Q}$ such that $\alpha_* < q_* < \beta_*$. Note, $\alpha_* > 0$ by definition. Thus, $\alpha_* < q_* < \beta_*$ implies $q_* \in \mathbb{Q}_{++}$.

Let $\gamma_* := q_*(t_* - \varepsilon)$ and $\delta_* := q_*(t_* + \varepsilon)$. Pick an arbitrary $t_0 \in L$ such that $\gamma_* < t_0 < \delta_*$. Define $t_1 := t_0/q_*$. Thus, $t_1 \in \tau$ because $(t_* - \varepsilon, t_* + \varepsilon) \subseteq \tau$. Note, $q_* t_1 = t_0 \in L$. Since $q_* \in \mathbb{Q}_{++}$, $t_1 \in \tau$ and $q_* t_1 \in L$, we obtain $q_* t_1 \in \tau$. Then, $t_0 = q_* t_1$ implies $t_0 \in \tau$. Since $t_0 \in L \cap (\gamma_*, \delta_*)$ was arbitrary, we obtain: $L \cap (\gamma_*, \delta_*) \subseteq \tau$.

Note that $\beta_* \leq \beta_1$ by definition of β_*. Then, $\alpha_* < q_* < \beta_*$ implies $\alpha_* < q_* < \beta_1$. Since $\alpha_* = s/(t_* + \varepsilon)$, $\beta_1 = s/(t_* - \varepsilon)$ and $\alpha_* < q_* < \beta_1$, it follows that $q_*(t_* - \varepsilon) < s < q_*(t_* + \varepsilon)$. That is, $s \in (\gamma_*, \delta_*)$. Also, $s \in L$ as $s \in (0, \theta)$ and $(0, \theta) \subseteq L$. Thus, $s \in L \cap (\gamma_*, \delta_*)$. As we have already shown that $L \cap (\gamma_*, \delta_*) \subseteq \tau$, it follows that $s \in \tau$. Since $s \in (0, \theta)$ was arbitrary, we have: $(0, \theta) \subseteq \tau$.

Recall, L is either $(0, \theta)$ or $(0, \theta]$. If L is indeed the interval $(0, \theta)$ then we already have $L \subseteq \tau$. So, we assume that L is the interval $(0, \theta]$. Of course, since we have already established $(0, \theta) \subseteq \tau$, it remains to show that $\theta \in \tau$. Let $t_2 := \theta/2$. Since $(0, \theta) \subseteq \tau$, we have $t_2 \in \tau$. Also, let $q := 2$. Thus, $q \in \mathbb{Q}_{++}$ and $qt_2 = \theta$. Then, $\theta \in L$ implies $qt_2 \in L$. Since $q \in \mathbb{Q}_{++}$, $t_2 \in \tau$ and $qt_2 \in L$, it follows that $qt_2 \in \tau$. As $qt_2 = \theta$, we obtain $\theta \in \tau$. Thus, $L \subseteq \tau$ if L is the interval $(0, \theta]$. ■

Roughly, the import of the above lemma can be described as follows. The ambient space L is the interval $(0, \theta)$ or $(0, \theta]$. Now, depending on the problem at hand, suppose that a particular subset $\tau \subseteq L$ has been defined. If τ has a non–empty interior then, for any arbitrary $x \in \tau$, there exists a neighborhood of x which is contained in τ. This is because τ is closed under the "multiplication from left" action of the subgroup $\mathbb{Q}_{++}$ which is dense in the group $\mathbb{R}_{++}$. With these lemmas stated and established, we are now ready to prove Theorem 10.

PROOF OF THEOREM 10: We establish "sufficiency". Let C be a non–empty *convex* subset of $\mathbb{R}^n$ and $\succsim$ be a non–trivial preference on C. Since $\succsim$ is a non–trivial, it follows that $\succ$ is non–empty. Further, $\sim$ is non–empty by Reflexivity of $\succsim$. Fix $\mathscr{X} \equiv (x_1, \ldots, x_{m_*+1})$ as the coordinate system for C. Then, $x_0 \coloneqq \sum_{k=1}^{m_*+1} x_k/(m_*+1)$ is the centroid and $(p_1, \ldots, p_{m_*+1})$, where $p_k \coloneqq x_k - x_0$ for all $k = 1, \ldots, m_* + 1$, are the vertices induced by $\mathscr{X}$. For $x \in C$, let $U(x) \coloneqq \{y \in C : y \succ x\}$, $L(x) \coloneqq \{y \in C : x \succ y\}$ and $I(x) \coloneqq \{y \in C : y \sim x\}$. Recall, W_* is the m_*–dimensional linear span of the vectors in $(p_1, \ldots, p_{m_*+1})$. Consider the subsets U_*, V_* and S_* of W_* defined as follows:

$$\begin{aligned} U_* &\coloneqq \big\{w \in W_* : x_0 + tw \succ x_0 \quad \text{for some } t > 0\big\}, \\ V_* &\coloneqq \big\{w \in W_* : x_0 \succ x_0 + tw \quad \text{for some } t > 0\big\}, \\ S_* &\coloneqq \big\{w \in W_* : x_0 + tw \sim x_0 \quad \text{for some } t > 0\big\}. \end{aligned}$$

We assume that $\succsim$ satisfies Continuity and Invariance. The argument proceeds through the following steps.

Step 1 — We argue: if $x \in C$, $w \in W_*$ and $t_1, t_2 > 0$ are such that $x + t_1 w$ and $x + t_2 w$ are in C then the following hold:

$$x + t_1 w \succ x \quad \text{iff} \quad x + t_2 w \succ x \tag{13}$$

$$x \succ x + t_1 w \quad \text{iff} \quad x \succ x + t_2 w \tag{14}$$

$$x + t_1 w \sim x \quad \text{iff} \quad x + t_2 w \sim x \tag{15}$$

To prove (13), fix $x \in C$, $w \in W_*$ and $t_1 > 0$ such that $x + t_1 w \in C$. Consider an arbitrary $t_2 > 0$ such that $x + t_2 w \in C$. Observe, it is enough to show: $x + t_1 w \succ x$ implies $x + t_2 w \succ x$. So, let $x + t_1 w \succ x$. First, assume $t_2 = a t_1$ for some $a \in \mathbb{N}$. We have nothing to argue if $a = 1$. So, assume $a > 1$. By convexity of C, $x + b t_1 w \in C$ for every $b = 1, \ldots, a$. By Invariance, $x + t_1 w \succ x$ implies $x + 2t_1 w \succ x + t_1 w$. Similarly, $x + b t_1 w \succ x + (b-1) t_1 w$ for all $b = 1, \ldots, a$. Transitivity of $\succ$ implies $x + a t_1 w \succ x$. That is, $x + t_2 w \succ x$ holds.

Now, assume $t_2 = t_1/a$ for some $a \in \mathbb{N}$. By an argument as above, if $x \succ x + t_2 w$ then $x \succ x + t_1 w$. However, this contradicts $x + t_1 w \succ x$. Thus, $x \succ x + t_2 w$ is not possible. Similarly, $x + t_2 w \sim x$ is not possible. However, $\succsim$ is complete. Thus, $x + t_2 w \succ x$ holds.

Next, assume $t_2 = b t_1/a$ for some $a, b \in \mathbb{N}$ with $a \neq 1$. Let $t_3 \coloneqq t_1/a$. Then, $x + t_1 w \succ x$ implies $x + t_3 w \succ x$. Also, $t_2 = b t_3$ where $b \in \mathbb{N}$. Then, $x + t_3 w \succ x$ implies $x + t_2 w \succ x$. Thus, $x + t_2 w \succ x$ holds. Let $\tau \coloneqq \{t > 0 : x + tw \succ x\}$ and $L_{x,w} \coloneqq \{t > 0 : x + tw \in C\}$. Thus, we have shown: $(\ q \in \mathbb{Q}_{++}\ ;\ t \in \tau\ ;\ qt \in L_{x,w}\ ) \implies qt \in \tau$. Also, note that $L_{x,w}$ is $(0, \theta)$ or $(0, \theta]$ for some $\theta > 0$ as C is convex.

Now, we *also* show: $(t_* - \varepsilon, t_* + \varepsilon) \subseteq \tau$ for some $t_* > 0$ and $\varepsilon > 0$. Then, Lemma A.III.1(f) will imply $\tau = L_{x,w}$ which is equivalent to: $x + t_2 w \succ x$ for all $t_2 > 0$ such that $x + t_2 w \in C$. Let $t_* := t_1/2$. Thus, $y_* := x + t_* w \succ x$. By Continuity of $\succsim$, let $\varepsilon > 0$ be such that the ε–ball $B_\varepsilon(y_*)$ in $\mathbb{R}^n$ satisfies: $z \in C \cap B_\varepsilon(y_*) \implies z \succ x$. Since $x, x + t_1 w \in C$ and C is convex, $(t_* - \varepsilon, t_* + \varepsilon) \subseteq \tau$. This proves (13).

To prove (14), define $\succsim^*$ over C by: $u \succsim^* v$ iff $v \succsim u$. Observe, $\succsim^*$ saisfies the axioms on $\succsim$. Further, the strict component $\succ^*$ of $\succsim^*$ satisfies: $u \succ^* v$ iff $v \succ u$. Moreover, by the argument for (13), we have the equivalence: $x + t_1 w \succ^* x$ iff $x + t_2 w \succ^* x$. Thus, we obtain: $x \succ x + t_1 w$ iff $x \succ x + t_2 w$. This proves (14).

To prove (15), assume $x + t_1 w \sim x$. Suppose $x + t_2 w \succ x$. Then, $x + t_2 w \succ x$ by (13) which is a contradiction. Thus, $x + t_2 w \succ x$ is not possible. Similarly, (14) implies that $x \succ x + t_2 w$ is not possible. However, $\succsim$ is complete. Thus, $x + t_2 w \sim x$ holds. That is, $x + t_1 w \sim x$ implies $x + t_2 w \sim x$. The converse also holds because t_1 and t_2 are arbitrary. This proves (15). The step is complete.

Step 2 — We argue: if $x \in C$, $w \in W_*$ and $t_1, t_2 > 0$ are such that $x + t_1 w$ and $x_0 + t_2 w$ are in C then the following hold:

$$x + t_1 w \succ x \quad \text{iff} \quad x_0 + t_2 w \succ x_0 \tag{16}$$

$$x \succ x + t_1 w \quad \text{iff} \quad x_0 \succ x_0 + t_2 w \tag{17}$$

$$x + t_1 w \sim x \quad \text{iff} \quad x_0 + t_2 w \sim x_0 \tag{18}$$

Let $x \in C$, $w \in W_*$ and $t_1, t_2 > 0$ be such that $x + t_1 w$ and $x_0 + t_2 w$ are in C. If $w = \mathbf{0}$ then the claim is trivial. If $x = x_0$ then step 1 implies the claim. Thus, we assume $w \neq \mathbf{0}$ and $x \neq x_0$.

To prove (16), note that since $x \in C$, Lemma A.III.1(e) implies the existence of $\lambda \in (0,1)$ and $\lambda_1, \ldots, \lambda_{m_*+1}$ in $\mathbb{R}_{++}$ such that $\sum_{k=1}^{m_*+1} \lambda_k = 1$ and $x_0 = \lambda x + (1-\lambda) y_1$, where $y_1 := \sum_{k=1}^{m_*+1} \lambda_k x_k$. As $x_1, \ldots, x_{m_*+1} \in C$, the convexity of C implies $y_1 \in C$. Let $y_2 := x_0 + \lambda t_1 w$. Thus, $y_2 = \lambda(x + t_1 w) + (1-\lambda) y_1$. Since $x + t_1 w, y_1 \in C$ and $\lambda \in (0,1)$, the convexity of C implies $y_2 \in C$. Also, let $z_1 := x_0 - x$ and $y_3 := y_2 - z_1$. Thus, $y_3 = x + \lambda t_1 w$. Since $x, x + t_1 w \in C$ and $\lambda \in (0,1)$, the convexity of C implies $y_3 \in C$. Since $x + t_1 w \succ x$ implies $x + \lambda t_1 w \succ x$ by step 1, we have: $y_3 \succ x$. Moreover, $x + z_1 = x_0$ and $y_3 + z_1 = y_2$. Then, $y_3 \succ x$ implies $y_2 \succ x_0$ by Invariance. That is, $x_0 + \lambda t_1 w \succ x_0$. Hence, $x_0 + t_2 w \succ x_0$ by step 1. This proves the forward implication claimed in (16). For the reverse implication, let $z_2 := -z_1$. Then, observe that $x_0 + z_2 = x$ and $y_2 + z_2 = y_3$. By step 1, $x_0 + t_2 w \succ x_0$ implies $x_0 \succ y_2$ because $y_2 = x_0 + \lambda t_1 w$. Then, $y_3 \succ x$ by Invariance. By step 1, $x + t_1 w \succ x$ because $y_3 = x + \lambda t_1 w$. This proves (16).

To prove (17), define $\succsim^*$ over C by: $u \succsim^* v$ iff $v \succsim u$. Observe, $\succsim^*$ satisfies the axioms on $\succsim$. Further, the strict component $\succ^*$ of $\succsim^*$ satisfies: $u \succ^* v$ iff $v \succ u$. Moreover, by the argument for (16), we have the equivalence: $x + t_1 w \succ^* x$ iff $x_0 + t_2 w \succ^* x_0$. Thus, we obtain: $x \succ x + t_1 w$ iff $x_0 \succ x_0 + t_2 w$. This proves (17).

To prove (18), assume $x + t_1 w \sim x$. Suppose $x_0 + t_2 w \succ x_0$. Then, $x_0 + t_2 w \succ x_0$ by (16) which is a contradiction. Thus, $x_0 + t_2 w \succ x_0$ is not possible. Similarly, $x_0 \succ x_0 + t_2 w$ is not possible by (17). However, $\succsim$ is complete. Thus, $x_0 + t_2 w \sim x_0$ holds. That is, $x + t_1 w \sim x$ implies $x_0 + t_2 w \sim x_0$. Interchanging the role of x with x_0 and t_1 with t_2, in this argument, implies the converse. This proves (18).

Step 3 — We claim: $U(x) = C \cap (x + U_*)$, $L(x) = C \cap (x + V_*)$ and $I(x) = C \cap (x + S_*)$ for every $x \in C$. We shall only argue that $U(x) = C \cap (x + U_*)$. To show $U(x) \subseteq C \cap (x) + U_*$, let $y_0 \in U(x)$ be arbitrary. That is, $y_0 \in C$ and $y \succ x$. Let $w := y_0 - x$ and $t_1 := 1$. By Lemma A.III.1(e), there exists $\lambda \in (0, 1)$ and $\lambda_1, \ldots, \lambda_{m_*+1}$ in $\mathbb{R}_{++}$ such that $\sum_{k=1}^{m_*+1} \lambda_k = 1$ and $x_0 = \lambda x + (1 - \lambda) y_1$, where $y_1 := \sum_{k=1}^{m_*+1} \lambda_k x_k$. As $x_1, \ldots, x_{m_*+1} \in C$, the convexity of C implies $y_1 \in C$. Let $y_2 := \lambda y_0 + (1 - \lambda) y_1$. Thus, $y_2 \in C$ by convexity of C. Also, $y_2 = x_0 + \lambda t_1 w$. Note, $y_0 \succ x$ is equivalent to $x + t_1 w \succ x$ by the definition of w and t_1. Also, $x + t_1 w \succ x$ implies $x_0 + \lambda t_1 w \succ x_0$ by step 2. Then, if we show that $w \in W_*$ then $w \in U_*$. By Lemma A.III.1(c), $y_2 \in C$ implies there exists $\nu_1, \ldots, \nu_{m_*+1}$ in $\mathbb{R}$ such that $y_2 = x_0 + \sum_{k=1}^{m_*+1} \nu_k p_k$. Thus, $w \in W_*$ because $w = (y_2 - x_0)/(\lambda t_1)$ and W_* is the linear span of $(p_1, \ldots, p_{m_*+1})$. Hence, $w \in U_*$ and $y_0 = x + w$. Since $y_0 \in C$ already, we obtain: $y_0 \in C \cap (x + U_*)$. As $y_0 \in U(x)$ was arbitrary, it follows: $U(x) \subseteq C \cap (x + U_*)$.

For the converse, let $y_0 \in C \cap (x + U_*)$ be arbitrary. Then, $y_0 \in C$ and there exists $w \in U_*$ such that $y_0 = x + w$. Let $t_1 := 1$. By lemma A.III.1(e), there exists $\lambda \in (0, 1)$ and $\lambda_1, \ldots, \lambda_{m_*+1}$ in $\mathbb{R}_{++}$ such that $\sum_{k=1}^{m_*+1} \lambda_k = 1$ and $x_0 = \lambda x + (1 - \lambda) y_1$, where $y_1 := \sum_{k=1}^{m_*+1} \lambda_k x_k$. As $x_1, \ldots, x_{m_*+1} \in C$, $y_1 \in C$ by convexity of C. Let $y_2 := \lambda y_0 + (1 - \lambda) y_1$. Thus, $y_2 \in C$ by convexity of C. Also, $y_2 = x_0 + \lambda t_1 w$. Since $w \in W_*$ and $x_0 + \lambda t_1 w \in C$, we have $x_0 + \lambda t_1 w \succ x_0$. By step 2, $x + t_1 w \succ x$ follows. Since $t_1 = 1$ and $y_0 = x + w$, we obtain $y_0 \in U(x)$. As $y_0 \in C \cap (x + U_*)$ was arbitrary, we have: $C \cap (x + U_*) \subseteq U(x)$. Thus, we have shown: $U(x) = C \cap (x + U_*)$ for every $x \in C$. The arguments for $L(x) = C \cap (x + V_*)$ and $I(x) = C \cap (x + S_*)$ are similar.

Step 4 — We claim: (U_*, V_*, S_*) is a partition of W_*. First, we shall argue, if $w \in W_*$, there exists $t > 0$ such that $x_0 + tw \in C$. Because $\succsim$ is complete, this will imply $W_* = U_* \cup V_* \cup S_*$.

Fix an arbitrary $w \in W_*$. Since W_* is the linear span of the vertices $(p_1, \ldots, p_{m_*+1})$ induced by $\mathscr{X}$, there exists $\mu_1, \ldots, \mu_{m_*+1}$ in $\mathbb{R}$ such that $w = \sum_{k=1}^{m_*+1} \mu_k p_k$. For every $k \in \{1, \ldots, m_*+1\}$, consider the $\mathbb{R}$–valued map ψ_k on $\mathbb{R}_+$ which is defined as follows:

$$\psi_k(t) \coloneqq \mu_k t + \frac{1}{m_*+1}\Big(1 - t\sum_{l=1}^{m_*+1} \mu_l\Big) \quad \text{for all } t \in \mathbb{R}_+.$$

Since each ψ_k is continuous and $\lim_{t\to 0} \psi_k(t) = 1/(m_*+1) > 0$, there exists $t_* > 0$ such that $\psi_k(t_*) > 0$ for all k. Let $\lambda_k \coloneqq \psi_k(t_*)$ for every k. Thus, $\lambda_k > 0$ for every k. Note, $\sum_{k=1}^{m_*+1} \psi_k(t) = 1$ for any $t \in \mathbb{R}_+$. Thus, $\sum_{k=1}^{m_*+1} \lambda_k = 1$. Recall, $x_0 = \sum_{k=1}^{m_*+1} x_k/(m_*+1)$ and $p_k = x_k - x_0$ for every $k = 1, \ldots, m_*+1$. Then, by definition of the ψ_k's:

$$x_0 + tw = \sum_{k=1}^{m_*+1} \psi_k(t) x_k \quad \text{for any } t \in \mathbb{R}_+.$$

In particular, $x_0 + t_* w = \sum_{k=1}^{m_*+1} \lambda_k x_k$. Since $\mathscr{X} \equiv (x_1, \ldots, x_{m_*+1})$ is a coordinate system for C, the points $x_1, \ldots, x_{m_*+1}$ are in C. Then, because $\lambda_1, \ldots, \lambda_{m_*+1}$ are in $\mathbb{R}_+$ and $\sum_{k=1}^{m_*+1} \lambda_k = 1$, the convexity of C implies that $\sum_{k=1}^{m_*+1} \lambda_k x_k \in C$. That is, $x_0 + t_* w \in C$. Since $x_0 + t_* w \in C$ and $\succsim$ is complete, at least one of $x_0 + t_* w \succ x_0$ or $x_0 \succ x_0 + t_* w$ or $x_0 + t_* w \sim x_0$ must hold. Then, $w \in W_*$ and $t_* > 0$ imply that w belongs to at least of U_*, V_* or S_*. Since $w \in W_*$ was arbitrary, we have: $W_* \subseteq U_* \cup V_* \cup S_*$. Moreover, each of U_*, V_* and S_* is a subset of W_* by definition. Thus, $W_* = U_* \cup V_* \cup S_*$.

We now argue: U_*, V_* and S_* are pairwise disjoint. First, suppose $w \in U_* \cap V_*$. Since $w \in U_*$, there exists $t_1 > 0$ such that $x_0 + t_1 w \succ x_0$. Since $w \in V_*$, there exists $t_2 > 0$ such that $x_0 \succ x_0 + t_2 w$. As t_1 and t_2 are positive, $x_0 + t_1 w \succ x_0$ implies $x_0 + t_2 w \succ x_0$ by step 1. That is, both $x_0 + t_2 w \succ x_0$ and $x_0 \succ x_0 + t_2 w$ hold. This contradicts the asymmetry of $\succ$. Hence, $U_* \cap V_* = \varnothing$.

Now, suppose $w \in U_* \cap S_*$. As $w \in U_*$, there exists $t_1 > 0$ such that $x_0 + t_1 w \succ x_0$. As $w \in S_*$, there exists $t_2 > 0$ such that $x_0 + t_2 \sim x_0$. As t_1 and t_2 are positive, $x_0 + t_1 w \succ x_0$ implies $x_0 + t_2 \succ x_0$ by step 1. That is, both $x_0 + t_2 w \succ x_0$ and $x_0 + t_2 w \sim x_0$ hold. However, $\succ$ and $\sim$ are disjoint. Thus, $U_* \cap S_* = \varnothing$. Similarly, $V_* \cap S_* = \varnothing$.

As $\succsim$ is non–trivial, let $y_0, y_1 \in C$ satisfy $y_1 \succ y_0$ and set $w \coloneqq y_1 - y_0$. Thus, $y_0 + w \in U(y_0)$. Then, $y_0 + w \in C \cap (y_0 + U_*)$ by step 3. Hence, $w \in U_*$. Thus, $U_* \neq \varnothing$. Similarly, $V_* \neq \varnothing$. Observe, $\mathbf{0} \in S_*$.

Step 5 — We claim: U_*, V_* and S_* are (convex) cones. We only argue: U_* is a cone. First, let $w \in U_*$ and $\lambda > 0$. Since $w \in U_* \subseteq W_*$ and W_* is a linear subspace, we have $w' := \lambda w \in W_*$. Also, there exists $t > 0$ such that $x_0 + tw \succ x_0$ as $w \in U_*$. Let $t' := t/\lambda$. Thus, $x_0 + t'w' \succ x_0$ as $t'w' = tw$. Hence, $w' \in U_*$. Since $w \in U_*$ and $\lambda > 0$ are arbitrary, we have: $\big(w \in U_* \;;\; \lambda > 0 \big) \implies \lambda w \in U_*$.

Now, fix any $w_1, w_2 \in U_*$ and let $w := w_1 + w_2$. Since $U_* \subseteq W_*$ and W_* is a linear subspace, we have $w \in W_*$. Also, $w_1, w_2 \in U_*$ imply the existence of $t_1, t_2 > 0$ such that $x_0 + t_1 w_1 \succ x_0$ and $x_0 + t_2 w_2 \succ x_0$. Let $t_* := \min\{t_1, t_2\}$ and $t_{**} := t_*/2$. Note, $x_0 + t_1 w_1 \succ x_0$ implies $x_0 + t_1 w_1 \in C$. Further, $x_0 \in C$ and $x_0 + t_1 w_1 \in C$ imply $x_0 + t_{**} w_1 \in C$ because C is convex. By step 1, $x_0 + t_1 w_1 \succ x_0$ implies $x_0 + t_{**} w_1 \succ x_0$. Similarly, $x_0 + t_{**} w_2 \in C$ and $x_0 + t_{**} w_2 \succ x_0$. Moreover, $x_0 + t_* w_1 \in C$ and $x_0 + t_* w_2 \in C$ by convexity of C. Observe,

$$x_0 + t_{**} w = \frac{1}{2}(x_0 + t_* w_1) + \frac{1}{2}(x_0 + t_* w_2)$$

because $t_{**} = t_*/2$ and $w = w_1 + w_2$ by definition. Hence, $x_0 + t_{**} w \in C$ by convexity of C. Also, note that $x_0 + t_{**} w = (x_0 + t_{**} w_1) + t_{**} w_2$. Then, $x_0 + t_{**} w_1 \succ x_0$ implies $x_0 + t_{**} w \succ x_0 + t_{**} w_2$ by Invariance. Recall, $x_0 + t_{**} w_2 \succ x_0$. Transitivity of $\succ$ implies $x_0 + t_{**} w \succ x_0$. Then, $w \in W_*$ and $t_{**} > 0$ imply $w \in U_*$. As $w = w_1 + w_2$ where $w_1, w_2 \in U_*$ are arbitrary, we have: $\big(w_1 \in U_* \;;\; w_2 \in U_* \big) \implies w_1 + w_2 \in U_*$. Thus, U_* is a cone. Similarly, V_* and S_* are cones.

Step 6 — We argue: $V_* = -U_*$ and S_* is a subspace. First, let us show that $V_* = -U_*$. Let $w \in U_*$. Thus, $w \in W_*$ and there exists $t_1 > 0$ such that $x_0 + t_1 w \succ x_0$. Let $x := x_0 + t_1 w$. Note, $x_0 + t_1 w \succ x_0$ implies $x \in C$ in particular. Then, by Lemma A.III.1(e), there exists $\lambda \in (0, 1)$ and $\lambda_1, \ldots, \lambda_{m_*+1}$ such that $\sum_{k=1}^{m_*+1} \lambda_k = 1$ and $x_0 = \lambda x + (1 - \lambda) y$, where $y := \sum_{k=1}^{m_*+1} \lambda_k x_k$. Since $x_1, \ldots, x_{m_*+1}$ are in C, $y \in C$ as C is convex. Observe, $y = x_0 + t_2(-w)$, where $t_2 := \lambda t_1/(1 - \lambda)$, because $x = x_0 + t_1 w$ and $x_0 = \lambda x + (1 - \lambda) y$. Note, $t_2 > 0$.

Let $t_* := \min\{t_1, t_2\}$. Then, $x_0 + t_* w$ and $x_0 + t_*(-w)$ are in C because $x_0, x, y, y \in C$ and C is convex. Also, $x_0 + t_* w \succ x_0$ as $x_0 + t_1 w \succ x_0$ by step 1. Since $(x_0 + t_* w) + t_*(-w) = x_0$ and $x_0 + t_* w \succ x_0$, Invariance implies $x_0 \succ x_0 + t_*(-w)$. Also, $w \in W_*$ implies $-w \in W_*$ as W_* is a linear subspace. Hence, $-w \in V_*$. That is, $-U_* \subseteq V_*$. Similarly, $-V_* \subseteq U_*$. Hence, $V_* \subseteq -U_*$. Thus, $V_* = -U_*$. By a similar argument, $S_* = -S_*$. Moreover, S_* is a cone by step 5. Hence, S_* must be a subspace.

Step 7 — We argue: there exists $K \in \mathbb{N}$ and a list $(\lambda_1, \dots, \lambda_K)$ of K orthonormal vectors in W_* such that:

$$x \succ y \iff [\lambda_1 \cdot x, \dots, \lambda_K \cdot x] >_L [\lambda_1 \cdot y, \dots, \lambda_K \cdot y], \tag{19}$$

for any $x, y \in C$, where $>_L$ is the strict component of the standard lexicographic order $\geq_L$ over $\mathbb{R}^K$.

By steps 4–6, (U_*, V_*, S_*) is a partition of the linear space W_* such that U_*, V_* are cones satisfying $V_* = -U_*$ and S_* is a subspace. Then, by the Decomposition Theorem (Theorem 1 of section 2), there exists K and a list $\mathbf{U} := (\lambda_1, \dots, \lambda_K)$ of K orthonormal vectors in W_* such that $U_* = H_{\mathbf{U}}$, $V_* = -H_{\mathbf{U}}$ and $S_* = O_{\mathbf{U}}$, where $H_{\mathbf{U}}$ is the graded halfspace (Definition 1 of section 2) generated by $\mathbf{U}$ and $O_{\mathbf{U}}$ is the subspace of W_* which is orthogonal to the vectors in the list $\mathbf{U}$.

Fix an arbitrary $x, y \in C$. Then, $x \succ y$ iff $x \in U(y)$. By step 3, $U(y) = C \cap (y + U_*)$. Then, $x \succ y$ iff, $x = y + w$ for some $w \in H_{\mathbf{U}}$. By definition of $H_{\mathbf{U}}$, $w \in W_*$ is equivalent to:

$$[\lambda_1 \cdot (x - y), \dots, \lambda_K \cdot (x - y)] >_L \mathbf{0}_K,$$

where $\mathbf{0}_K$ is the origin of $\mathbb{R}^K$. Note, $\lambda_k \cdot (x - y) > 0$ iff $\lambda_k \cdot x > \lambda_k \cdot y$ for every $k = 1, \dots, K$. Hence, (19) follows from the definition of $\geq_L$. Since $x, y \in C$ are arbitrary, the step is complete.

Step 8 — We claim: if $(\lambda_1, \dots, \lambda_{K_1})$ and $(\mu_1, \dots, \mu_{K_2})$ are two lists of K_1 and K_2 orthonormal vectors in W_* such that:

$$x \succ y \iff [\lambda_1 \cdot x, \dots, \lambda_{K_1} \cdot x] >_L^1 [\lambda_1 \cdot y, \dots, \lambda_{K_1} \cdot y], \text{ and} \tag{20}$$

$$x \succ y \iff [\mu_1 \cdot x, \dots, \mu_{K_2} \cdot x] >_L^2 [\mu_1 \cdot y, \dots, \mu_{K_2} \cdot y], \tag{21}$$

for every $x, y \in C$, where $\succ_L^1$ and $\succ_L^2$ denote the strict components of the standard lexicographic orders over $\mathbb{R}^{K_1}$ and $\mathbb{R}^{K_2}$, then it must be that $K_1 = K_2 =: K_0$ and $\lambda_k = \mu_k$ for all $k = 1, \dots, K_0$.

Let $K_0 := \min\{K_1, K_2\}$. Denote the set $\{1, \dots, K_0\}$ by $[K_0]$. Now, suppose $\lambda_k \neq \mu_k$ for some $k \in [K_0]$. Then, define:

$$k_* := \min\big\{k \in [K_0] : \lambda_k \neq \mu_k\big\}.$$

We claim the existence of $w_1, w_2 \in W_*$ with the following properties:

(a) $\lambda_{k_*} \cdot w_1 > 0$ and $\lambda_{k_*} \cdot w_2 < 0$.

(b) $\mu_{k_*} \cdot w_1 < 0$ and $\mu_{k_*} \cdot w_2 > 0$.

(c) For any $j \in \{1, 2\}$, $\lambda_k \cdot w_j = 0$ and $\mu_k \cdot w_j = 0$ if $1 \leq k < k_*$.

By the Cauchy–Schwarz inequality, $|\lambda_{k_*} \cdot \mu_{k_*}| \leq \|\lambda_{k_*}\|_2 \cdot \|\mu_{k_*}\|_2 = 1$ with equality iff $\lambda_{k_*} = \pm\mu_{k_*}$. First, consider the case when $|\lambda_{k_*} \cdot \mu_{k_*}| = 1$. Since $\lambda_{k_*} \neq \mu_{k_*}$, we have $\lambda_{k_*} = -\mu_{k_*}$. Define $w_1 := \lambda_{k_*}$ and $w_2 := \mu_{k_*}$. Clearly, properties (a) and (b) hold. Moreover, (c) holds as the vectors $\lambda_1, \ldots, \lambda_{k_*}$ are orthogonal and $\lambda_k = \mu_k$ if $1 \leq k < k_*$.

Now, we assume $|\lambda_{k_*} \cdot \mu_{k_*}| < 1$. That is, $1 - (\lambda_{k_*} \cdot \mu_{k_*})^2 > 0$. Then, fix any θ and ψ in $\mathbb{R}_{++}$. Also, let $w_1 := \alpha\lambda_{k_*} + \beta\mu_{k_*}$ and $w_2 := -w_1$, where $\alpha, \beta \in \mathbb{R}$ are defined as follows:

$$\begin{aligned} \alpha &:= \quad \big[\theta + \psi(\lambda_{k_*} \cdot \mu_{k_*})\big] / \big[1 - (\lambda_{k_*} \cdot \mu_{k_*})^2\big], \text{ and} \\ \beta &:= -\big[\psi + \theta(\lambda_{k_*} \cdot \mu_{k_*})\big] / \big[1 - (\lambda_{k_*} \cdot \mu_{k_*})^2\big] \end{aligned}$$

Observe, $\lambda_{k_*} \cdot w_1 = \theta$ and $\mu_{k_*} \cdot w_1 = -\psi$. As θ and ψ are in $\mathbb{R}_{++}$ and $w_2 = -w_1$, properties (a) and (b) obtain. Since $\lambda_k = \mu_k$ if $1 \leq k < k_*$ and $\lambda_1, \ldots, \lambda_{k_*}$ are orthogonal, property (c) obtains because $w_2 = -w_1$ where w_1 is a linear combination of *only* λ_{k_*} and μ_*. Thus, we have demonstrated the existence of $w_1, w_2 \in W_*$ as claimed.

In step 4, recall that we argued: if $w \in W_*$ then there exists $t > 0$ such that $x_0 + tw \in C$. Then, as $w_1, w_2 \in C$, there exists $t_1, t_2 > 0$ such that $x_0 + t_1w_1$ and $x_0 + t_2w_2$ are in C. Thus, properties (a)–(c) imply $x_0 + t_1w_1 \succ x_0 + t_2w_2$ and $x_0 + t_2w_2 \succ x_0 + t_1w_1$ by representations (20) and (21), respectively. However, the relation $\succ$ is asymmetric. Hence, our supposition that there exists $k \in [K_0]$ such that $\lambda_k \neq \mu_k$ must be wrong. That is, $\lambda_k = \mu_k$ for all $k \in [K_0]$.

It remains to argue: $K_1 = K_2$. Suppose $K_1 < K_2$. Let $w := \mu_{K_1+1}$. Clearly, $w \in W_*$ as the vectors $\mu_1, \ldots, \mu_{K_2}$ are in W_*. Thus, there exists $t > 0$ such that $x_0 + tw \in C$. Note, since $(\mu_1, \ldots, \mu_{K_1+1})$ are orthogonal and $\lambda_k = \mu_k$ for all $k \in [K_0]$, representations (20) and (21) imply $x_0 + tw \sim x_0$ and $x_0 + tw \succ x_0$, respectively. This contradicts the fact that $\succ$ and $\sim$ are disjoint. Hence, the supposition that $K_1 < K_2$ must be wrong. That is, $K_1 < K_2$ is not possible. Similarly, $K_2 < K_1$ is not possible. Thus, $K_1 = K_2$. This step is complete.

Step 9 — By steps 7 and 8, there exists a unique $K \in \mathbb{N}$ and a unique list $(\lambda_1, \ldots, \lambda_K)$ of orthonormal vectors in W_* such that:

$$x \succsim y \iff [\lambda_1 \cdot x, \ldots, \lambda_K \cdot x] \geq_L [\lambda_1 \cdot y, \ldots, \lambda_K \cdot y],$$

for any $x, y \in C$, where $\geq_L$ is the standard lexicographic order over $\mathbb{R}^K$. By Lemma A.III.1(d), $W_* = S_C$ where S_C is the subspace generated by the set C. Thus, the vectors $\lambda_1, \ldots, \lambda_K$ are in S_C.

To complete the proof, observe that $K = 1$ by Continuity of $\succsim$. ■

Proof of Theorem 11: Consider the existence claim in the statement of Theorem 11 and step 9 in the proof of Theorem 10 (see section *A*.III.1). Note, both Theorems 10 and 11 assume the Invariance axiom. Thus, if step 9 continues to hold under the additional assumption of Convexity of the binary relation, instead of Continuity as in the proof of Theorem 10, then Theorem 11 is proven.

Observe, Continuity was referred to only in step 1 of the proof of Theorem 10, in paragraph 4, to establish *only* the following:

$$\big(\exists\, t_* > 0\big)\big(\exists\, \varepsilon > 0\big)\big[\ (t_* - \varepsilon, t_* + \varepsilon) \subseteq \tau\ \big]. \tag{22}$$

Of course, in addition to Continuity, (22) was established under the additional assumption that $C \subseteq \mathbb{R}^n$ is convex and $t_1 > 0$ exists such that $x + t_1 \succ x$. We shall now argue that (22) continues to hold when Continuity is replaced by Convexity.

Recall, $\tau = \{t > 0 : x + tw \succ x\}$. Clearly, $t_1 \in \tau$ and $x + t_1 w \in C$. Moreover, since C is a convex subset of $\mathbb{R}^n$, it follows that $x + tw \in C$ for every $t \in (0, t_1)$. Let $y \coloneqq x + t_1 w$. Since $y \succ x$, by Convexity of $\succsim$ we have the following:

$$\alpha x + (1 - \alpha) y \succ x \quad \text{for every } \alpha \in (0, 1).$$

Since $y = x + t_1 w$, note that $\alpha x + (1 - \alpha) y = x + [(1 - \alpha) t_1] w$ for every $\alpha \in (0, 1)$. Thus, $\{t \in (0, t_1) : x + tw \succ x\} = (0, t_1)$. Let $t_* \coloneqq t_1/2$ and $\varepsilon \coloneqq t_*$. Hence, $(t_* - \varepsilon, t_* + \varepsilon) = (0, t_1)$. That is,

$$(t_* - \varepsilon, t_* + \varepsilon) = \{t \in (0, t_1) : x + tw \succ x\}.$$

Therefore, $(t_* - \varepsilon, t_* + \varepsilon) \subseteq \tau$ which proves (22).

Having established (22), we note that the other clause that had to be established in step 1 of Theorem 10 was the following:

$$\big(\ q \in \mathbb{Q}_{++}\ ;\ t \in \tau\ ;\ qt \in L_{x,w}\ \big) \implies qt \in \tau, \tag{23}$$

where $L_{x,w} = \{t > 0 : x + tw \in C\}$. However, observe that the proof of (23) relied only the convexity of C and the axiom of Invariance. Since Invariance has been assumed in Theorem 11 as well, while maintaining that C is convex, (23) also continues to hold.

We now come to one final observation. To complete the proof of step 1 (and Theorem 11), it is enough to show that, for any $x, y \in C$ and $\alpha \in (0, 1)$, $x \succ y$ implies $x \succ \alpha x + (1 - \alpha) y$. For this, let $z \coloneqq \alpha(x - y)$. Note, Convexity implies $x' \coloneqq (1 - \alpha) x + \alpha y \succ y$. As $x' + z = x$ and $y + z = \alpha x + (1 - \alpha) y$, Invariance implies $x \succ \alpha x + (1 - \alpha) y$. ■

REFERENCES

ALLAIS, M. (1953): "Le comportement de l'homme rationnel devant de le risque", *Econometrica*, vol. 21, pp. 503–546.

ANSCOMBE, F. J. & R. J. AUMANN (1963): "A Definition of Subjective Probabilities", *Annals of Mathematical Statistics*, vol. 34, no. 1, pp. 199–205.

ARROW, K. J. (1963): *Social Choice and Individual Values* New York: John Wiley and Sons, Inc.

AUMANN, R. J. (1962): "Utility Theory without the Completeness Axiom", *Econometrica*, vol. 30, no. 3, pp. 445–462.

AZRIELI, Y., C. P. CHAMBERS & P. J. HEALY (2018): "Incentives in Experiments: A Theoretical Analysis", *Journal of Political Economy*, vol. 126, no. 4, pp. 1472–1503.

BLACKORBY, C., D. DONALDSON & J. A. WEYMARK (1984): "Social Choice with Interpersonal Utility Comparisons: A Diagrammatic Introduction", *International Economic Review*, vol. 25(2), pp. 327–356.

BLACKWELL, D. & M. A. GIRSHICK (1954): *Theory of Games and Statistical Decisions* New York: John Wiley and Sons, Inc.

BLUME, L., A. BRANDENBURGER & E. DEKEL (1989): "An Overview of Lexicographic Choice under Uncertainty", *Annals of Operations Research*, vol. 19, pp. 231–246.

BLUME, L., A. BRANDENBURGER & E. DEKEL (1991*a*): "Lexicographic Probabilities and Choice under Uncertainty", *Econometrica*, vol. 59, no. 1, pp. 61–79.

BLUME, L., A. BRANDENBURGER & E. DEKEL (1991*b*): "Lexicographic Probabilities and Equilibrium Refinements", *Econometrica*, vol. 59, no. 1, pp. 81–98.

CHAMBERS, C. P., F. ECHENIQUE & E. SHMAYA (2017): "General Revealed Preference Theory", *Theoretical Economics*, vol. 12, pp. 493–511.

CHAMBERS, C. P., F. ECHENIQUE & E. SHMAYA (2014): "The Axiomatic Structure of Empirical Content", *American Economic Review*, vol. 104, pp. 2303–19.

CHEW, S. H. (1983): "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory resolving the Allais Paradox", *Econometrica*, vol. 51, no. 4, pp. 1065–92.

CHEW, S. H., L. G. EPSTEIN & U. SEGAL: (1991): "Mixture Symmetry and Quadratic Utility", *Econometrica*, vol. 59, no. 1, pp. 139–163.

Chew, S. H., E. Karni & Z. Safra (1987): "Risk Aversion in the Theory of Expected Utility with Rank Dependent Probabilities", *Journal of Economic Theory*, vol. 42, no. 2, pp. 370–381.

d'Aspremont, C., & L. Gevers (1977): "Equity and te Informational Basis of Collective Choice", *Review of Economic Studies*, vol. 44, no. 2. pp. 199–209.

d'Aspremont, C., & L. Gevers (2002): "Social Welfare Functionals and Interpersonal Comparability", in K. J. Arrow, A. K. Sen and K. Suzumura (editors) *Handbook of Social Choice and Welfare: Volume 1* Amsterdam: North–Holland Publishing.

Dekel, E. (1986): "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom", *Journal of Economic Theory*, vol. 40, no. 2, pp. 304–18.

Dubra, J., F. Maccheroni & E. A. Ok (2004): "Expected Utility Theory without the Completeness Axiom", *Journal of Economic Theory*, vol. 115, no. 1, pp. 118–133.

Fishburn, P. C. (1969): "A Study of Independence in Multivariate Utility Theory", *Econometrica*, vol. 37, no. 1, pp. 107–121.

Fishburn, P. C. (1970): *Utility Theory for Decision Making* New York: John Wiley and Sons, Inc.

Fishburn, P. C. (1974): "Lexicographic Orders, Utilities and Decision Rules: A Survey", *Management Science*, vol. 20, no. 11, pp. 1442–1471.

Grether, D. & C. Plott (1979): "Economic Theory of Choice and the Preference Reversal Phenomenon", *American Economic Review*, vol. 69, no. 4, pp. 623–638.

Hara, K., E. A. Ok & G. Riella (2019): "Coalitional Expected Multi–Utility Theory", *Econometrica*, vol. 87, no. 3, pp. 933–980.

Harsanyi, J. C. (1955): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility", *Journal of Political Economy*, vol. 63, no. 4, pp. 309–321.

Hausner, M. (1954): "Multidimensional Utilities", in R. M. Thrall, C. H. Coombs and R. L. Davis (editors) *Decision Processes*, New York: John Wiley & Sons Inc., 1954.

Hausner, M. & J. G. Wendel (1952): "Ordered Vector Spaces", *Proceedings of the American Mathematical Society*, vol. 3, no. 6, pp.

977–982.

Herstein, I. N. & J. Milnor (1953): "An Axiomatic Approach to Measurable Utility", *Econometrica*, vol. 21, No. 2, pp. 291–297.

Holt, C. (1986): "Preference Reversals and the Independence Axiom", *American Economic Review*, vol. 76, no. 3, pp. 508–515.

Karni, E. & Z. Safra (1987): "'Preference Reversal' and the Observability of Preferences by Experimental Methods", *Econometrica*, vol. 55, no. 3, pp. 675–685.

Krantz, D. H., R. D. Luce, P. Suppes & A. Tversky (1971): *Foundations of Measurement* (Volume 1), New York: Academic Press, Inc.

Loomes, G., C. Starmer & R. Sugden (1991): "Observing Violations of Transitivity by Expermental Methods", *Econometrica*, vol. 59, no. 2, pp. 425–439.

Machina, M. J. (1982): ""Expected Utility" Analysis without the Independence Axiom", *Econometrica*, vol. 50, no. 2, pp. 277–323.

Manzini, P. & M. Mariotti (2007): "Sequentially Rationalizable Choice", *American Economic Review*, vol. 97, no. 5, pp. 1824–39.

Manzini, P. & M. Mariotti (2014): "Stochastic Choice and Consideration Sets", *Econometrica*, vol. 82, no. 3, pp. 1153–76.

Marschak, J. (1950): "Rational Behavior, Uncertain Prospects, and Measurable Utility", *Econometrica*, vol. 18, no. 2, pp. 277–323.

Masatlioglu, Y., D. Nakajima & E. Y. Ozbay (2012): "Revealed Attention", *American Economic Review*, vol. 102, no. 5, pp. 2183–2205.

Mishra, D. & A. Sen (2012): "Roberts' Theorem with Neutrality: A Social Welfare Ordering Approach", *Games and Economic Behavior*, vol. 75, no. 1, pp. 283–298.

Nielsen, K. & J. Rehbeck (2022): "When Choices Are Mistakes", *American Economic Review*, vol. 112, no. 7, pp. 2237–68.

Pommerehne, W., W. F. Schneider & P. Zweifel (1982): "Economic Theory of Choice and the Preference Reversal Phenomenon: A Re–examination", *American Economic Review*, vol. 72, no. 3, pp. 569–574.

Quiggin, J. (1982): "A Theory of Anticipated Utility", *Journal of Economic Behavior and Organization*, vol. 3, no. 4, pp. 323–43.

Regenwetter, M., J. Dana & C. P. Davis–Stober (2011): "Transitivity of Preferences", *Psychological Review*, vol. 118, no. 1, pp.

42–56.

Roberts, K. W. S. (1979): "The Characterization of Implementable Choice Rules", in J–. J. Laffont (editor) *Aggregation and Revelation of Preferences*, Amsterdam: North–Holland Publishing.

Roberts, K. W. S. (1980*a*): "Possibility Theorems with Interpersonally Comparable Welfare Levels", *The Review of Economic Studies*, vol. 47, no. 2, pp. 409–420.

Roberts, K. W. S. (1980*b*): "Interpersonal Comparability and Social Choice Theory", *The Review of Economic Studies*, vol. 47, no. 2, pp. 421–439.

Roberts, K. W. S. (1980*c*): "Social Choice Theory: The Single–profile and Multi–profile Approaches", *The Review of Economic Studies*, vol. 47, no. 2, pp. 441–450.

Samuelson, P. A. (1952): "Probability, Utility and the Independence Axiom", *Econometrica*, vol. 20, no. 4, pp. 670–678.

Segal, U. (1988): "Does the Preference Reversal Phenomenon necessarily contradict the Independence Axiom?", *American Economic Review*, vol. 78, no. 1, pp. 233–236.

Segal, U. (1990): "Two–Stage Lotteries without the Reduction Axiom", *Econometrica*, vol. 58, no. 2, pp. 349–77.

Segal, U. (2023): "$\forall$ or $\exists$?", *Theoretical Economics*, vol. 18, no. 1, pp. 1–13.

Lichtenstein, S. & P. Slovic (1983): "Preference Reversals: A Broader Perspective", *American Economic Review*, vol. 73, no. 4, pp. 596–605.

Tversky, A. (1969): "Intransitivity of Preferences", *Psychological Review*, vol. 76, no. 1, pp. 31–48.

Tversky, A., P. Slovic & D. Kahneman (1990): "The Causes of Preference Reversal", *American Economic Review*, vol. 80, no. 1, pp. 204–217.

Young, H. P. (1975): "Social Choice Scoring Functions", *SIAM Journal of Applied Mathematics*, vol. 28, no. 4, pp. 824–838.

von Neumann, J. & O. Morgenstern (1944) *Theory of Games and Economic Behavior* New Jersey: Princeton University Press.

Chapter 3

Preferences with Norms as Representations

1. introduction

1.1 *An Overview*

Norms over Euclidean spaces define natural weak orderings over the vectors. We consider continuous weak orders on any given Euclidean space and ask the following question: what axioms characterize those weak orders which admit some norm as a representation? Of particular interest is the subclass of p–norms. To place our question in proper perspective, we now describe the background comprising of applications involving preferences which admit norms as representations in various aspects of economic theory.

Over a span of several decades, such preferences have been assumed as the model of individuals comprising of the society in the theory of strategic voting in spatial models or multiple issues. For instance, McKelvey & Wendell (1976) generalize the results, on the majority rule admitting voting equilibria, due to Plott (1967) and Davis et al. (1972) by assuming individuals to have arbitrary "quadratic" preferences which subsume Euclidean preferences.

However, WENDELL & THORSON (1974) already recognized that preferences other than the "quadratic" preferences are at least as important. They assume individual preferences admit some norm as a representation and proceed to analyse the consequences in voting and its equilibria. Similarly, BORDER & JORDAN (1983) recognize the need to allow into consideration preferences that are more general than the Euclidean preferences. They show that under strategy–proofness considertions, voting rules in spatial models must be driven by only "ideal points" of individuals whose preferences are "star–shaped and separable". As we shall observe at the end of section 4, these preferences admit representations which are essentialy the p–norms except that their "balls" may not be convex.

Further, ZHOU (1991) showed that Gibbard's theorem on dictatorships holds in public goods problem for multidimensional Euclidean spaces with quasi–concave preferences. More recently, GERSHKOV ET AL. (2019, 2022) have considered the problem of voting on multiple issues. They emphasize the need to consider general norms as preferences and show that dominant strategy incentive compatibility is equivalent to the geometric property of "orthant monotonicity".

Moreover, ENELOW & HINCH (1982, 1984) and ENELOW ET AL. (1986, 1988) show that empirical testing, via regression analysis, of predictions made by the theory on spatial voting heavily depends on the correctness of the specification of the norm representing individual preferences. For further examples of norms considered in strategic voting for settings with spatial models, one may consider BARBERÀ ET AL. (1993) and PETERS ET AL. (1993) for instance. Also, for "quadratic" functionals that generalize the classical utilitarianism, one may consider EPSTEIN & SEGAL (1992).

More recently, applications in matching theory have considered the Euclidean norm such as the school choice functions generated by "ideal points" as in ECHENIQUE & YENMEZ (2015). Just as WENDELL & THORSON (1974), BORDER & JORDAN (1983) and ZHOU (1991) have argued—in strategic voting over multiple issues—for considering individual preferences that admit arbitrary norm like representations, a similar argument applies for matching probems as considered in ECHENIQUE & YENMEZ (2015) for instance.

Two further applications are as follows. Measurement theory concerns itself with specific functional forms as representations for weak orders. For instance, MACHINA & MÜLLER (1987) characterize weak orders that admit polynomial representations up to some moments. Second, FIELDS & OK (1996) and MITRA & OK (1996) characterize real–valued measures of income mobility as p–norms. Perhaps, such problems can be based on orders as primitives.

The many applications which assume general norms as primitives make it imperative to supply a decision theoretic foundation for preferences which admit norms as representations. However, Kanai (1977), Bogomolnaia & Laslier (2007) and Eguia (2011) show existence of some "embeddings" in normed Euclidean spaces. Similarly, characterizing those real–valued maps which are the Euclidean norm, as in D'Agostino & Dardanoni (2009) for instance, does *not* accomplish the task set forth by the applications.

For Euclidean preferences, a decision theoretic foundation has been provided in Chambers & Echenique (2020). In measurement theory, Tversky & Krantz (1970) give a foundation for the metric induced by the Euclidean norm. However, they do so by considering a weak order on $\mathbb{R}^n \times \mathbb{R}^n$ as the primitives, whereas, we must consider a weak order on $\mathbb{R}^n$ as the primitive. Thus, it is *not* possible to adapt their foundation for our primitive as otherwise the axioms would involve differences of vectors which are harder to justify normatively. Moreover, axioms must involve only the universal quantifier from both the normative and falsifiability perspectives — see Dekel & Lipman (2010) and Chambers et al. (2014) for instance.

We first generalize the notion of a norm to "pre–norm" and extend the scope of our question of existence of representations from norms to pre–norms. For concreteness, we state the definition. A *pre–norm* is any map $f : \mathbb{R}^n \to \mathbb{R}_+$ that satisfies (1) $f(x) = 0$ iff $x = \mathbf{0}$, (2) $f(\alpha \cdot x) = \alpha \cdot f(x)$ for every $\alpha > 0$ and $x \in \mathbb{R}^n$, and (3) $f(x + y) \leq f(x) + f(y)$ for all $x, y \in \mathbb{R}^n$.

Thus, a pre–norm satisfies the definition of norms except for the "symmetry" requirement: $f(x) = f(-x)$ for all $x \in \mathbb{R}^n$. Property (2) says that a pre–norm is a function which is homogenous of degree one. Property (3) is the "Triangle Inequality". One key element in our analysis is the following observation: any homogenous function satisfies the Triangle Inequality if and only if it is a convex function.

Homotheticity requires that $x \succ y$ implies $\alpha \cdot x \succ \alpha \cdot y$. Further, Convexity requires all weak lower contour sets to be convex.[35] Moreover, we introduce an axiom, which we call Scale Monotonicity, that requires the weak order to exhibit increasing returns to scale. Denote by $\mathcal{P}$ the class of all binary relations over $\mathbb{R}^n$ which are weak orders that satisfy Continuity, Homotheticity, Convexity and Scale Monotonicity. Our first result is that $\mathcal{P}$ is precisely the class of binary relations which admit some pre–norm as a representation.

[35]We observe that in many standard settings, such as consumer choice theory, the axiom of Convexity requires the weak *upper* contour sets to be convex. However, many problems in social choice and othe settings involving geospatial preferences often require the weak *lower* contour set to be convex. It is the latter axiom that we call Convexity.

By definition, f is a pre–norm if it is homogenous function of degree one which evaluates to 0 only at the origin, and satisfies the Triangle Inequality. Thus, any pre–norm uniquely identifies a compact convex subset C_f of $\mathbb{R}^n$ which contains the origin in its interior. Here, C_f is the set of all vectors whose f–value is atmost 1. Geometrically, the pre–norm generates "open balls" which are all the scalings and translations of the interior of C_f.

Any pre–norm is a continuous map. If a binary relation $\succ$ admits a pre–norm as a represeentation, then $\succ$ must be a weak order and satisfy Continuity. Further, the weak lower contour sets of $\succ$ must be convex as C_f is convex. Also, Scale Monotonicity and Homotheticity should obviously hold. Thus, binary relations which admit some pre–norm as a representation must be in the class $\mathcal{P}$. However, for "existence" of pre–norms as representations, it must be shown that the weak lower contour sets of $\succ$ satisfy (1) compactness, and (2) the origin is in the interior. Obtaining these properties from the axioms are the major challenges in establishing our main result.

As a corollary to our main representation theorem, we obtain a characterization of norms as representations. Within $\mathcal{P}$, the subclass of those binary relations which admit some norm as a representation are characterized by an additional axiom called Reflection Symmetry which requires $-x$ to be indifferent to x. This is so as a norm f is a pre–that also satisfies $f(x) = f(-x)$ for any vector x.

We then move on to the characterization of p–norms. For this, we consider any n–tuple $\theta \equiv (\theta_1, \ldots, \theta_n)$ of positive numbers and $p \geq 1$ to define a map $\|\cdot\|_{(\theta,p)} : \mathbb{R}^n \to \mathbb{R}_+$ as follows:

$$\|x\|_{(\theta,p)} := \Big(\sum_{i=1}^{n} \theta_i |x_i|^p\Big)^{1/p} \quad \text{for all } x \equiv (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

We call $\|\cdot\|_{(\theta,p)}$ the (θ,p)–norm a special case of which is the p–norm, denoted by $\|\cdot\|_p$, when all the θ_i's are equal to unity. Since the map $\xi \in \mathbb{R}_+ \mapsto \xi^p$ is monotone, a binary relation $\succ$ which admits some (θ,p)–norm as a representation must satisfy the Separability axiom(s) due to Debreu (1959). Now, consider $\succ$ which admits some norm as a representation. Our second main result is, if $\succ$ satisfies Separability then $\succ$ admits some (θ,p)–norm as a representation. Thus, we have a characterization of those binary relations which admit some (θ,p)–norm as a representation.

Note, Debreu's theorem on the existence of additive representations does not alone characterize the particular functional form as required by the definition of the (θ,p)–norm. We arrive at the suitable functional equation from the combination of the axioms.

To pin down those binary relation which are represented by some p–norm, we additionally impose Permutation Symmetry which requires an vector x to be indifferent to the vector x_σ obtained by permuting the components of x. This completes the high–level description of our results on the existence of representations.

We also develop a "duality theory" for binary relations represented via pre–norms. This is possible because maximization of the support function of a compact convex set is a homogenous functional which is convex. Essentially, this duality theory is analogous to the relationship of the Utility Maximization Problem and the Expenditure Minimization Problem as in the classical theory of consumer choice.

Suppose $\succ$ admits the pre–norm f as a representations. As we have outlined above, there is a compact convex set C_f with the origin in its interior that is naturally associated to $\succ$. All weak lower contour sets of $\succ$ are scalings of C_f. Then, the support function $T(f)$ of C_f is also pre–norm. Then, the weak order induced by $T(f)$, which we denote by $\succ^*$, is the *dual* of $\succ$.

Thus, to each $\succ$ in $\mathcal{P}$ the dual $\succ^*$ is also in $\mathcal{P}$. Hence, every $\succ$ in $\mathcal{P}$ admits a *second dual* $\succ^{**}$ which, by definition, is the dual of the dual of $\succ$. Our first main result on duality is that the second dual of any binary relation in $\mathcal{P}$ must be itself. That is, "take dual" is an idempotent operator on $\mathcal{P}$. Further, we define a binary relation to be *self–dual* if its dual is identical to itself. Our second main result is: a weak order is self–dual iff it admits the Euclidean norm as a representation — "spherical preferences". Our third result is: dual of the p–norm is the q–norm, where $1/p + 1/q = 1$.

In functional analysis, any pair (p, q) such that $1/p + 1/q = 1$ are called *conjugate indices.* They feature, for instance, in the statement of Hölder's inequality which generalizes the Cauchy–Schwarz Inequality. Hölder's inequality claims the following:

$$|x \cdot y| \leq \|x\|_p \cdot \|y\|_q,$$

where (p, q) are any pair of conjugate indices and $x \cdot y$ is the standard inner product on $\mathbb{R}^n$. Since the notion of conjugate index is seen to be intimately related to the notion of dual of a weak order, we ask: does the Hölder's inequality generalize to *arbitrary* pre–norms? We show that the answer to this question is in the affirmative.

The rest of the article is organized as follows. Section 2 presents the framework. Results for general pre–norms are presented in section 3. The theory is specialized to (θ, p)–norms in section 4 which also obtains the classical inequalities due to Minkowski and Hölder as corollaries. Proofs ommitted from the text are supplied in the Appendix.

Of interest shall be binary relations over $\mathbb{R}^n$, with $n \in \mathbb{N}$ fixed, which shall be typically denoted by $\succ$. For any given $\succ$ over $\mathbb{R}^n$, we define the corresponding binary relation $\sim$ over $\mathbb{R}^n$ as follows:

$$x \sim y \iff \big(\text{not } x \succ y \text{ ; not } y \succ x \big).$$

From the definition of $\sim$, it is clear that $\sim$ is symmetric.[36] Then, define the binary relation $\succsim$ over $\mathbb{R}^n$ as follows:

$$x \succsim y \iff \big(x \succ y \text{ or } x \sim y \big)$$

Note, if $\succ$ is asymmetric[37] then $\succsim$ admits $\succ$ and $\sim$ as its asymmetric and symmetric components, respectively. We say, $\succ$ is a *weak order* if $\succ$ is asymmetric and negatively transitive.[38] The binary relation $\succsim$ is called a *preference* if $\succsim$ is complete[39] and transitive.[40] Then, observe that $\succ$ is weak order if and only if $\succsim$ is a preference.

Let $\mathcal{U}$ be a given subclass of the collection of all maps from $\mathbb{R}^n$ to $\mathbb{R}$. Then, a *$\mathcal{U}$–representation* of $\succ$ is any $u \in \mathcal{U}$ such that

$$x \succ y \iff u(x) > u(y).$$

Our primary objective is to axiomatically characterize binary relations over $\mathbb{R}^n$ which admit a $\mathcal{U}$–representation, where $\mathcal{U}$ is the class of objects which we call "pre–norms". Let $\mathbf{0}$ denote the "origin" of $\mathbb{R}^n$.

DEFINITION 1: *Any map* $f : \mathbb{R}^n \to \mathbb{R}$ *is a* pre–norm *on* $\mathbb{R}^n$ *if* f *satisfies each of the following properties*:

1. $f(x) \geq 0$ *for all* $x \in \mathbb{R}^n$.
2. $f(x) = 0$ *iff* $x = \mathbf{0}$.
3. $f(\alpha \cdot x) = \alpha \cdot f(x)$ *for all* $\alpha > 0$ *and* $x \in \mathbb{R}^n$.
4. $f(x + y) \leq f(x) + f(y)$ *for all* $x, y \in \mathbb{R}^n$.

A pre–norm f is a *norm* on $\mathbb{R}^n$ if condition 3 is strengthened as follows:

$$f(\alpha \cdot x) = |\alpha| \cdot f(x) \quad \text{for all } \alpha \in \mathbb{R} \text{ and } x \in \mathbb{R}^n.$$

[36] A binary relation R over X is *symmetric* if: $xRy \implies yRx$.

[37] A binary relation R over X is *asymmetric* if: $xRy \implies \text{not } yRx$.

[38] A binary relation R over X is *negatively transitive* if: $\big(\text{not } xRy \text{ ; not } yRz \big) \implies \text{not } xRz$.

[39] A binary relation R over X is *complete* if: $\big(xRy \text{ or } yRx \big)$.

[40] A binary relation R over X is *transitive* if: $\big(xRy \text{ ; } yRz \big) \implies xRz$.

Let the classes of all pre–norms and norms over $\mathbb{R}^n$ be denoted by $\mathcal{N}_*$ and $\mathcal{N}$, respectively. By definition of the terms "pre–norm" and "norm", it follows that $\mathcal{N} \subseteq \mathcal{N}_*$. In fact, this set–inclusion is proper as there exists pre–norms on $\mathbb{R}^n$ which are not norms—examples are provided in section 3.

We now consider the standard notion of a "p–norm" over $\mathbb{R}^n$. For any $1 \leq p < \infty$, let the map $\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R}$ be defined as follows:

$$\|x\|_p := \big(\sum_{i=1}^{n} |x_i|^p\big)^{1/p} \quad \text{for all } x \equiv (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

It is a non–trivial result in the theory of normed linear spaces that *Minkowski's inequality* holds which states that, for any $1 \leq p < \infty$ and any $x, y \in \mathbb{R}^n$, the following holds:

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Thus, Minkowski's inequality asserts that the map $\|\cdot\|_p$, for any $1 \leq p < \infty$, satisfies condition 4 as in Definition 1. That the map $\|\cdot\|_p$ satisfies the other conditions in the definition of the term "norm" hold is easy to observe from the definition of $\|\cdot\|_p$. Thus, $\|\cdot\|_p$ is a norm over $\mathbb{R}^n$ if $1 \leq p < \infty$. The maps $\|\cdot\|_p$ are called *p–norms.* Denote by $\mathcal{N}_\pi$ the set $\big\{\|\cdot\|_p : 1 < p < \infty\big\}$. Observe, $\mathcal{N}_\pi \subseteq \mathcal{N}$. In fact, this set–inclusion is also proper. We shall demonstrate in section 4 that there exists norms over $\mathbb{R}^n$ which are not p–norms.

We must note that though here we have appealed to the fact that Minkowski's inequality holds, in order to conclude that p–norms are indeed norms, our development in sections 3 and 4 will in fact lead to Minkowski's inequality as a corollary.

Another remark is in order. The standard proof that Minkowski's inequality holds rests on another non–trivial fact from the theory of normed linear spaces which is the Hölder's inequality that generlaizes the well–known Cauchy–Schwarz Inequality. For any $1 < p < \infty$, the unique number $1 < q < \infty$ such that $1/p + 1/q = 1$ is called the *conjugate* of p. Then, *Hölder's inequality* states that, if $1 < p < \infty$ and q is the conjugate of p then, for any $x, y \in \mathbb{R}^n$:

$$|x \cdot y| \leq \|x\|_p \|y\|_q,$$

where $x \cdot y := \sum_{i=1}^{n} x_i y_i$ is the standard inner product on $\mathbb{R}^n$.

In sections 3 and 4, we generalize the Hölder's inequality to any pre–norm. Moreover, our conlcusion that the Minkowski's inequality holds will not rely on the fact that Hölder's inequality holds.

3. GENERAL THEORY

3.1 *The Basic Representation Theorem*

Our basic result is a characterization of those binary relations $\succ$ over $\mathbb{R}^n$ for which some pre–norm on $\mathbb{R}^n$ is a representation. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *homogenous* (*of degree one*) if,

$$f(\alpha \cdot x) = \alpha \cdot f(x) \quad \text{for all } \alpha > 0 \text{ and } x \in \mathbb{R}^n.$$

Any pre–norm is a homogenous function. Let $\mathcal{H}$ be the class of all homogenous functions on $\mathbb{R}^n$. Recall, $\mathcal{N}_*$ and $\mathcal{N}$ denote the class of all pre–norms and norms on $\mathbb{R}^n$, respectively. Thus, $\mathcal{N} \subseteq \mathcal{N}_* \subseteq \mathcal{H}$ holds. We begin with the following preliminary result.

PROPOSITION 1: *Suppose $f : \mathbb{R}^n \to \mathbb{R}_+$ is an $\mathcal{H}$–representation of the binary relation $\succ$ over $\mathbb{R}^n$. Then, $g : \mathbb{R}^n \to \mathbb{R}$ is an $\mathcal{H}$–representation of $\succ$ if and only if, there exists $\alpha > 0$ such that $g = \alpha \cdot f$.*

The primary content of the above proposition broadly is as follows: homogenous maps of degree one that represent a given binary relation are *unique* up to a positive multiplicative constant. However, there is one caveat. Homogenous maps of degree one can possibly have a range which includes both positive and negative real numbers. In fact, the definition of $\mathcal{H}$–representations does not exclude such possibilities. The above proposition claims that uniqueness of $\mathcal{H}$–representations up to a positive multiplicative factor holds *if* at least one of the homogenous maps has a non–negative (or, non–positive) range.

The proof of Proposition 1 is in section A.I.1 of the Appendix. We come to the question of "existence" pre–norms as representations. The following notation shall be used. For any binary relation $\succ$ and $x \in \mathbb{R}^n$, the sets $U_{\succ}(x) := \{y \in \mathbb{R}^n : y \succ x\}$ and $L_{\succ}(x) := \{y \in \mathbb{R}^n : x \succ y\}$ are the strict upper and strict lower contour sets of x.

WEAK ORDER: *$\succ$ over $\mathbb{R}^n$ is asymmetric and negatively transitive.*

CONTINUITY: *The sets $U_{\succ}(x)$ and $L_{\succ}(x)$ are open in $\mathbb{R}^n$.*

HOMOTHETICITY: $\big(x \succ y\,;\, \alpha > 0\big) \implies \alpha \cdot x \succ \alpha \cdot y.$

CONVEXITY: $\big(x \succsim y\,;\, 0 < \alpha < 1\big) \implies x \succsim \alpha \cdot x + (1-\alpha) \cdot y.$

SCALE MONOTONICITY: $\big(x \neq \mathbf{0}\,;\, \alpha > 1\big) \implies \alpha \cdot x \succ x.$

Of the five axioms stated above, the first two are standard necessary and sufficient conditions on the binary relation $\succ$ to admit a continuous $\mathbb{R}$–valued representation. Recall that $\mathcal{N}_*$ is the class of all pre–norms over $\mathbb{R}^n$ and $\mathcal{N}_* \subseteq \mathcal{H}$, where $\mathcal{H}$ is the class of all homogenous functions of degree one. Thus, the Homotheticity of $\succ$ is a necessary condition for $\succ$ to admit some pre–norm as a representation. Then, our basic representation theorem can be stated as follows.

THEOREM 1: *A binary relation $\succ$ on $\mathbb{R}^n$ admits an $\mathcal{N}_*$–representation, if and only if, $\succ$ is a weak order satisfying Continuity, Homotheticity, Convexity and Scale Monotonicity.*

Thus, the additional axioms of Convexity and Scale Monotonicity characterize those binary relations which admit some pre–norm as their representation. The "uniqueness" result is as follows.

PROPOSITION 2: *Suppose the binary relation $\succ$ over $\mathbb{R}^n$ admits the map f as an $\mathcal{N}_*$–representation and $g : \mathbb{R}^n \to \mathbb{R}$ as an $\mathcal{H}$–representation. Then, $g = \alpha \cdot f$ for some unique $\alpha > 0$.*

Note, if f is an $\mathcal{N}_*$–representation of $\succ$ then f is a pre–norm on $\mathbb{R}^n$. In particular, f must be an $\mathbb{R}_+$–valued map which is homogenous of degree one (see Definition 1 of section 2). Now, consider $g : \mathbb{R}^n \to \mathbb{R}$ to be any $\mathcal{H}$–representation of $\succ$. Thus, g is a homogenous function of degree one possibly with a range comprising of both positive and negative real numbers. However, Proposition 1 requires $g = \alpha \cdot f$ for some $\alpha > 0$. Thus, the only additional claim in Proposition 2 is that α is unique. This follows from the facts that (1) $f(x) > 0$ if $x \neq \mathbf{0}$, and (2) both f and g are homogenous maps representing the same underlying weak order. In particular, (1) is true as f is a pre–norm.

Thus, the only non–trivial claim in Theorem 2 is the "uniqueness" of the multiplicative constant. The proof of this part of Proposition 2, and the proof of Theorem 1, is in section A.I.1 of the Appendix. However, we indicate the proof strategy of Theorem 1.

Let $\succ$ over $\mathbb{R}^n$ satisfy the axioms in Theorem 1. Fix an arbitrary $x_0 \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Define C to be the closure of $L_\succ(x_0)$. The axioms imply that C is convex and compact with $\mathbf{0}$ in the interior of C. Then, the map $\|\cdot\|_\succ : \mathbb{R}^n \to \mathbb{R}_+$ is defined as follows:

$$\|x\|_\succ := \inf\big\{\kappa > 0 : x \in \kappa \cdot C\big\} \quad \text{for all } x \in \mathbb{R}^n,$$

where $\kappa \cdot C := \{\kappa \cdot y : y \in C\}$. Then, $\|\cdot\|_\succ$ is shown to be a pre–norm that represents $\succ$. That is, $\|\cdot\|_\succ$ is an $\mathcal{N}_*$–representation of $\succ$.

The foregoing discussion suggests a geometric structure induced by any arbitrary pre–norm. It is of interest to formalize this geometric interpretation for two reasons. First, it serves to provide examples of binary relations which admit pre–norms as representations. Second, it shall aid in the organization of the proof of Theorem 1. We cast this presentation as the following characterization.

THEOREM 2: *Let C be convex and compact subset of $\mathbb{R}^n$ such that $\mathbf{0}$ is in the interior of C. Then, the map $\|\cdot\|_C : \mathbb{R}^n \to \mathbb{R}_+$ defined as:*

$$\|x\|_C := \inf \big\{ \kappa > 0 : x \in \kappa \cdot C \big\} \quad \textit{for all } x \in \mathbb{R}^n,$$

where $\kappa \cdot C := \{\kappa \cdot y : y \in C\}$, is a pre–norm on $\mathbb{R}^n$ and satisfies:

$$C = \big\{ x \in \mathbb{R}^n : \|x\|_C \leq 1 \big\}.$$

Moreover, suppose $f : \mathbb{R}^n \to \mathbb{R}_+$ be any pre–norm on $\mathbb{R}^n$ and define:

$$C_f := \big\{ x \in \mathbb{R}^n : f(x) \leq 1 \big\}.$$

Then, C_f is a convex and compact subset of $\mathbb{R}^n$ with $\mathbf{0}$ in its interior. Further, the map $\|\cdot\|_{C_f}$ is identical to f.

That is, there is a one–to–one correspondence between pre–norms and convex compact sets with the origin in their interior. Next, we come to the characterization of those binary relations which admit some norm as a representation. Recall, a norm is a pre–norm f on $\mathbb{R}^n$ that satisfies the following stronger property than condition 3 in Definition 1:

$$f(\alpha \cdot x) = |\alpha| \cdot f(x) \quad \text{for all } \alpha \in \mathbb{R} \text{ and } x \in \mathbb{R}^n.$$

It turns out that the following symmetry axiom, in addition to those listed in Theorem 1, achieves the desired characterization.

REFLECTION SYMMETRY: $x \sim -x$.

Recall, the symbol $\mathcal{N}$ denotes the class of all norms over $\mathbb{R}^n$. Thus, the phrase "the norm f is a representation of $\succ$" is equivalent to the phrase "f is an $\mathcal{N}$–representation of $\succ$". The result is as follows.

PROPOSITION 3: *The binary relation $\succ$ admits an $\mathcal{N}$–representation, iff, $\succ$ admits an $\mathcal{N}_*$–representation and satisfies Reflection Symmetry.*

A "uniqueness" claim analogous to Proposition 2 clearly holds.

In this section, we shall investigate the consequence of maximization of any linear numerical objective over feasible sets which are the weak lower contour set of those binary relations on $\mathbb{R}^n$ which admit some pre–norm as a representation.

Let $\mathcal{P}$ be the class of all binary relations over $\mathbb{R}^n$ which admit some pre–norm as a representation. Fix any $x_0 \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Then, associate with any $\succ$ in $\mathcal{P}$, the map $f_\succ : \mathbb{R}^n \to \mathbb{R}$ defined as follows:

$$f_\succ(y) := \max_{x_0 \succsim x} \; x \cdot y \quad \text{for all } y \in \mathbb{R}^n.$$

We begin with the following observation.

PROPOSITION 4: *If $\succ$ is in $\mathcal{P}$ then $f_\succ$ is a pre–norm.*

The proof is almost obvious but is supplied, for completeness, in section A.I.2 of the Appendix. However, one may compare $f_\succ$ with the profit function of a price–taking competitive firm whose objective is to maximize profits. It is a standard exercise in microeconomic theory that the profit function is a non–negative homogenous map of degree one which is convex. Observe, these properties are almost equivalent to asserting that the map is a pre–norm.

Since Proposition 4 says that to each $\succ$ in $\mathcal{P}$ the corresponding map $f_\succ$ is a pre–norm, we may now define a map $(\,\cdot\,)^* : \mathcal{P} \to \mathcal{P}$ which shall assosiate to each $\succ$ in $\mathcal{P}$ a "dual" $(\succ)^*$ in $\mathcal{P}$. For notational brevity, we shall write $\succ^*$ for $(\succ)^*$. The definition of $(\,\cdot\,)^* : \mathcal{P} \to \mathcal{P}$ is as:

$$x \succ^* y \iff f_\succ(x) > f_\succ(y)$$

Note, the definition of the map $f_\succ$ rested on the choice of some x_0 in $\mathbb{R}^n \setminus \{\mathbf{0}\}$. Thus, before proceeding further, there is a need to argue that $\succ^*$ is well–defined in the sense that its definition does not depend on the choice of the x_0 from $\mathbb{R}^n \setminus \{\mathbf{0}\}$. That such is indeed the case is an immediate consequence of the Homotheticity of $\succ$ which holds as $\succ$ admits a pre–norm as a representation.

We shall call $\succ^*$ the *dual* of $\succ$. We shall also write $(\succ^*)^*$ as $\succ^{**}$. We shall call $\succ^{**}$ the *second dual* of $\succ$. With these preliminaries in place, our first key result regarding duals is as follows.

THEOREM 3: *If $\succ$ is in $\mathcal{P}$ then $\succ^{**}$ is equal to $\succ$.*

That is, $(\,\cdot\,)^*$ is an unary operator on $\mathcal{P}$ such that its composition with itself is the identity map on $\mathcal{P}$. Thus, $(\,\cdot\,)^*$ is an involution.

In words, for any binary relation on $\mathbb{R}^n$ that admits some pre–norm as a representation, its second dual is itself. We say that $\succ$ in $\mathcal{P}$ is *self–dual* if $\succ^*$ is equal to $\succ$. Our second key result is as follows.

THEOREM 4: *Let $\succ$ be a binary relation in $\mathcal{P}$. Then, $\succ^*$ equals $\succ$, if and only if, $\succ$ admits $\|\cdot\|_2$ as a representation.*

That is, among all binary relations on $\mathbb{R}^n$ that admit some pre–norm as a representation, the one which is self–dual is unique and it admits the Euclidean norm as its representation.

Theorems 3 and 4 may remind the reader of the Hölder's inequality from the theory of normed linear spaces. It generalizes the well–known Cauchy–Schwarz Inequality. It claims that the absolute value of the inner product of any two vectors is bounded above by the product of the p–norm of one vector with the q–norm of the other, where $p, q > 1$ are "conjugates" in the sense that $1/p + 1/q = 1$. Thus, the conjugate of the conjugate of p is p itself for any arbitrary $p > 1$. Moreover, the conjugate of $p > q$ is itself iff $p = 2$. We show that this "parallel" is in fact tight by generalizing the Hölder's inequality.

Recall, $\mathcal{N}_*$ is the class of all pre–norms on $\mathbb{R}^n$. To each $f : \mathbb{R}^n \to \mathbb{R}_+$ on $\mathbb{R}^n$, associate the map $g_f : \mathbb{R}^n \to \mathbb{R}_+ \cup \{\infty\}$ defined as:

$$g_f(y) := \max_{f(x)\leq 1} \; x \cdot y \quad \text{for all } y \in \mathbb{R}^n.$$

We say g_f is the *conjugate* of f. The key result is as follows.

THEOREM 5: *Suppose f is a pre–norm. Then, its conjugate g_f is also a pre–norm, and the map $T : \mathcal{N}_* \to \mathcal{N}_*$ defined as*:

$$T(f) := g_f \quad \textit{for every } f \in \mathcal{N}_*,$$

satisfies: $[T \circ T](f) = f$ *for every* $f \in \mathcal{N}_*$. *Further, for every* $f \in \mathcal{N}_*$, $T(f) = f$ *if, and only if,* $f = \|\cdot\|_2$. *Moreover, the following holds*:

$$x \cdot y \leq f(x) \cdot [T \circ f](y) \quad \textit{for all } x, y \in \mathbb{R}^n.$$

The consequence of assuming f to be a norm is as follows.

COROLLARY 1: *Suppose f is a norm on $\mathbb{R}^n$ and T is as defined in the statement of Theorem 5. Then, the following inequality holds*:

$$|x \cdot y| \leq f(x) \cdot [T \circ f](y) \quad \textit{for all } x, y \in \mathbb{R}^n.$$

Thus, Hölder's inequality is generalized to *any* norm and its conjugate.

Now that the notions of "dual" and "conjugate" for preferences in $\mathcal{P}$ and pre–norms respectively stand formulated, we are in position to explicitly describe the connection between the dual and the conjugate. This is essential as preferences in $\mathcal{P}$ are precisely those which admit some pre–norm as a representation. The result is as follows.

THEOREM 6: *Suppose $\succ$ is in $\mathcal{P}$ and $\succ^*$ is its dual. Let f be an $\mathcal{N}_*$–representation of $\succ$. Then, g is an $\mathcal{N}_*$–representation of $\succ^*$, if and only if, there exists $\alpha > 0$ such that $g = \alpha \cdot T(f)$.*

We conclude this subsection with some remarks. Three classes of objects have been under consideration. First, the class of all compact convex subsets of $\mathbb{R}^n$ which have the origin in their interior. Second, the class of all pre–norms over $\mathbb{R}^n$. Third, the class of all continuous weak orders over $\mathbb{R}^n$ which satisfy Homotheticity, Convexity and Scale Monotonicity. Theorems 1 and 2, of subsection 3.1, establish natural correspondences between objects across these classes.

Subsection 3.2 defines the notion of "dual" for such weak orders and "conjugate" for pre–norms. Theorems 3 claims that the second dual of a weak order is equal to the weak order itself. Theorem 5 claims that the conjugate of a pre–norm is equal to the pre–norm. Thus, the "dualization" operator defined over the class of such weak orders and the "conjugation" operator defined over the class of all pre–norms are involutions. Moreover, Theorem 5 also claims that a pre–norm is equal to its conjugate iff it is the Euclidean norm. Thus, Theorem 4 claims that a weak order is self–dual iff it is "spherical" — the indifference curves are spherical in shape.

The precise connection of a weak order and its dual through a pre–norm that represents the former and the conjugate of that pre–norm is formulated in Theorem 6. Finally, we note that, within the larger class of weak orders which admit some pre–norm as a representation, the additional requirement of the axiom called Reflection Symmetry pins down the class of those weak orders which admit some *norm* as a representation. This is our theory for *general* pre–norms.

4. STANDARD NORMS

The aim in the previous section was to characterize weak orders that admit pre–norms, of which norms are special case, as representations. Further, a duality theory was presented which culminated in three key results. First, the second dual of any such weak order is itself. Second, if a weak order is self–dual, it must be "spherical". Third, the Hölder's inequality generalizes to any pre–norm.

The purpose of this section is to specialize to the case of p–norms and a natural generaliztion of them. Such objects are important in the theory of normed linear spaces and its various applications. Our first set of results are characterizations of weak orders that admits such norms as representations.

As the reader may know, the definition of the p–norm does not make it immediate that they are indeed norms. In particular, it requires proof that "Triangle Inequality" holds — this is the well–known Minkowski's inequality. Moreover, the Hölder's inequality is fundamental to the theory of normed linear spaces since it generalizes the Cauchy–Schwarz Inequality for p–norms. In the two subsections that follow, we shall derive these inequalities based on the geometry of the general theory in section 3 adapted to the special case of p–norms.

With this background in place, we now proceed to define the class of objects called "p–norms" and a class of its generalization called "(θ, p)–norms". However, we first begin with some comments on the notation. Throughout this section, we shall denote vectors in $\mathbb{R}^n$ by symbols such as x, y, ... and so on. Further, we shall often write x as $(x_1, \ldots, x_n)$ to indicate the vector x as an n–tuple in $\mathbb{R}^n$, where x_i is the ith component of x. A definition[41] follows.

Definition 2: *Suppose $\theta \equiv (\theta_1, \ldots, \theta_n) \in \mathbb{R}^n_{++}$ and $p \geq 1$. Then, the* (θ, p)–norm *on $\mathbb{R}^n$ is the map $\|\cdot\|_{(\theta,p)} : \mathbb{R}^n \to \mathbb{R}_+$ defined as:*

$$\|x\|_{(\theta,p)} \coloneqq \Big(\sum_{i=1}^{n} \theta_i |x_i|^p \Big)^{1/p} \quad \text{for every } x \in \mathbb{R}^n.$$

Further, the p–norm *is the map $\|\cdot\|_p \coloneqq \|\cdot\|_{(\theta,p)}$ when $\theta = \mathbb{1}_n$.*

Since our interest is to characterize binary relations $\succ$ over $\mathbb{R}^n$ which admit some (θ, p)–norm as representation, we begin by observing that such a binary relation must be "separable" due to Debreu (1960). To see why, note that $\xi \in \mathbb{R}_+ \mapsto \xi^p \in \mathbb{R}_+$ is stricly increasing. Thus, if $\|\cdot\|_{(\theta,p)}$ represents $\succ$ then so does $\|\cdot\|^p_{(\theta,p)}$. Also, note that:

$$\|x\|^p_{(\theta,p)} \coloneqq \sum_{i=1}^{n} h_i(x_i) \quad \text{for every } x \in \mathbb{R}^n,$$

where $h_i : \mathbb{R} \to \mathbb{R}$ is defined as: $h_i(\xi) \coloneqq \theta_i |\xi|^p$ for all $\xi \in \mathbb{R}$. That is, $\succ$ admits an "additive" representation. Therefore, Hence, $\succ$ must satisfy "separability" if it admits a (θ, p)–norm as a representation.

[41] We denote by $\mathbb{1}_n$ the n–tuple in $\mathbb{R}^n$ whose every component is 1.

To state this axiom, we use the following notation. We denote by N the set $\{1, \ldots, n\}$ and indicate by I any typical subset of N. Now, any $x \equiv (x_1, \ldots, x_n)$ and $y \equiv (y_1, \ldots, y_n)$ in $\mathbb{R}^n$, we shall write $(x_I, y_{N \setminus I})$ for that vector in $\mathbb{R}^n$ whose kth component is x_k or y_k according as $k \in I$ or $k \in N \setminus I$. Then, "separability" is as follows.

Separability: $(x_I, x_{N\setminus I}) \succ (x'_I, x_{N\setminus I}) \iff (x_I, x'_{N\setminus I}) \succ (x'_I, x'_{N\setminus I})$.

All free variables are universally quantified over their respective range. For instance, the above statement must hold for *every* $I \subseteq N$. Next, observe that for $\succ$ to admit some p–norm as a representation, it is necessary that $\succ$ exhibits indifference between any vector and the one obtained by "permuting" its components. Denote by $\sigma : N \to N$ a typical bijection — that is, a *permutation* of N. Also, for any vector $x \equiv (x_1, \ldots, x_n)$ in $\mathbb{R}^n$ and any permutation of N, let x_σ denote that vector in $\mathbb{R}^n$ whose ith component is $x_{\sigma(i)}$ for every $i \in N$.

Permutation Symmetry: $x_\sigma \sim x$.

Since the key result in this section is on the existence of (θ, p)–norms as representations, it must logically be demonstrated first that any (θ, p)–norm is indeed a norm *if* $p \geq 1$. However, we defer the proof of this claim until later in order to arrive at the statement of the main result. For now, we assume that $n \geq 3$.

Theorem 7: *The binary relation $\succ$ on $\mathbb{R}^n$ admits a (θ, p)–norm as a representation, if and only if, $\succ$ satisfies separability and admits a norm as a representation. Further, a norm f represents $\succ$ iff, there exists $\alpha > 0$ such that $f = \alpha \|\cdot\|_{(\theta,p)}$.*

Some remarks are in order regarding the claim of "existence" in the above theorem. Observe, the characterization of $\succ$ which admits a norm as a representation has been provided, in subsection 3.1, via Theorem 1 and Proposition 3. Thus, the non–trivial part is to pin down those binary relations which admit a (θ, p)–norm as a representation. The point of Theorem 7 is that the only additional axiom needed to characterize such binary relations is separability.

Corollary 2: *The binary relation $\succ$ on $\mathbb{R}^n$ admits a p–norm as a representation, if and only if, $\succ$ satisfies permutation symmetry and the (θ, p)–norm represents $\succ$ for some $\theta \in \mathbb{R}^n_{++}$. Further, a norm f represents $\succ$ iff, there exists $\alpha > 0$ such that $f = \alpha \|\cdot\|_{(\theta,p)}$.*

The proof of the "existence" claim in the above corollary is easy because $\succ$ is assumed to satisfy permutation symmetry. Our proof of Theorem 7 is involves reducing the problem, via application of Theorem 1 and separability, to the problem of solving a particular functional equation. Concretely, we are interested in the characterization of all continuous functions $h : \mathbb{R}_{++} \to \mathbb{R}_{++}$ which satisfy:

$$h(\xi\eta) = h(\xi)h(\eta) \quad \text{for every } \xi, \eta > 0.$$

The complete proof of Theorem 7 is provided in section A.II.1 of the Appendix where we also the above problem. It is shown that a map $h : \mathbb{R}_{++} \to \mathbb{R}_{++}$ satisfies the above functional equation, if and only if, there exists $p > 0$ such that $h(\xi) = \xi^p$ for all $p > 0$. While it is possible to obtain this characterization from first principles, our approach is to transform this problem to one of characterzing all solutions to the well–known Cauchy functional equation.

We make two final remarks with regard to Theorem 7 and Corollary 2. In stating these two results we have relied on the assumption that $n \geq 3$. This is because in Debreu's charcaterization of weak orders that admits addtively separable representations, the separability axiom is sufficient for the case when $n \geq 3$. However, Debreu also provides a charcaterization for the case of $n = 2$ by using a stronger axiom which later authors have called "strong separability". Our proof of Theorem 7 works under the assumption of "strong separability" for existence when $n = 2$. Lastly, we point out that the only role of Convexity in the proof is to conclude that $p \geq 1$. Otherwise, the function $\|\cdot\|_p$ is a a representation of $\succ$ for some unique $p > 0$. This is precisely the class of "star–shaped preferences" as in Border & Jordan (1983).

4.1 *Minkowski's inequality*

We had deferred the proof of the claim that (θ, p)–norms are indeed norms on $\mathbb{R}^n$. The only non–trivial part of the claim is to show that the "Triangle Inequality" holds. We establish this claim in the present subsection. However, before that we prove some general elementary results which shall be of use in the final argument. For this, we need the following definition.

Definition 3: *A map* $f : \mathbb{R}^n \to \mathbb{R}$ *is* subadditive *if,*

1. $f(\alpha \cdot x) = \alpha \cdot f(x)$ *for all* $\alpha > 0$ *and* $x \in \mathbb{R}^n$,
2. $f(x + y) \leq f(x) + f(y)$ *for all* $x, y \in \mathbb{R}^n$.

Lemma 1: *A function is subadditive, if and only if, it is convex and homogenous of degree one.*

Proof: Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is convex and homogenous of degree one. We argue: f is subadditive. It is enough to show: $f(x+y) \leq f(x)+f(y)$ for all $x, y \in \mathbb{R}^n$. Fix any $x, y \in \mathbb{R}^n$. Let $\alpha := 1/2$ and $\mu := 1/\alpha$. Also, let $x_* := \mu \cdot x$ and $y_* := \mu \cdot y$. Clearly, $x + y = \alpha \cdot x_* + (1-\alpha) \cdot y_*$. Since f is homogenous of degree one, $\alpha \cdot f(x_*) = f(\alpha \cdot x_*) = f(x)$. Similarly, $(1-\alpha) \cdot f(y_*) = f(y)$. Since f is convex:

$$f(\alpha \cdot x_* + (1-\alpha) \cdot y_*) \leq \alpha \cdot f(x_*) + (1-\alpha) \cdot f(y_*).$$

That is, $f(x + y) \leq f(x) + f(y)$. Hence, f is subadditive.

For the converse, assume f is subaddtive. Then, it is homogenous of degree one by definition. To show convexity of f, let $x, y \in \mathbb{R}^n$ and $\alpha \in (0, 1)$. Define $x_* := \alpha \cdot x$ and $y_* := (1 - \alpha) \cdot y$. Because f is subadditive, $f(x_* + y_*) \leq f(x_*) + f(y_*)$. As $x_* + y_* = \alpha \cdot x + (1-\alpha) \cdot y$, it follows that f is convex. This completes the proof. ■

Lemma 2: *Suppose $f : \mathbb{R}^n \to \mathbb{R}_+$ is quasiconvex, homogenous of degree one, and $f(x) = 0$ iff $x = \mathbf{0}$. Then, f is convex function.*

Proof: Let $x, y \in \mathbb{R}^n$ and $\alpha \in (0, 1)$. Let $z := \alpha \cdot x + (1 - \alpha) \cdot y$. Assume, without loss of any generality, $f(x) \geq f(y)$. Note, if $f(y) = 0$ then $y = \mathbf{0}$ which implies $f(z) = \alpha \cdot f(x)$ as f is homogenous of degree one. Then, $f(y) = 0$ implies $f(z) \leq \alpha \cdot f(x) + (1 - \alpha) \cdot f(y)$. Further, if $f(x) = f(y)$ then $f(z) \leq \alpha f(x) + (1 - \alpha)f(y)$ as f is quasiconvex. Henceforth, we assume $0 < f(y) < f(x)$.

Observe, $f(\mu \cdot x) = f(y)$ for some unique $\mu \in (0, 1)$. To see why, note that $\alpha \in [0, 1] \mapsto \alpha \cdot f(x) \in [0, f(x)]$ is a continous bijection, and $0 < f(y) < f(x)$. Let $x_* := (1/\mu) \cdot y$ and $y_* := \mu \cdot x$. Then, $f(\mu \cdot x) = f(y)$ implies $f(y_*) = f(y)$. Moreover, $f(\mu \cdot x) = f(y)$ implies $f(x_*) = f(x)$ as f is homogenous of degree one.

Let $\lambda_1 := \alpha/[\alpha + (1 - \alpha)\mu]$ and $\lambda_2 := \mu(1 - \alpha)/[\alpha + (1 - \alpha)\mu]$. Note, $\lambda_1, \lambda_2 \in (0, 1)$ as $\alpha, \mu > 0$. Let $x_{**} := \lambda_1 \cdot x + (1 - \lambda_1) \cdot x_*$ and $y_{**} := \lambda_2 \cdot y + (1 - \lambda_2) \cdot y_*$. Then, $x_{**} = \theta_1 \cdot z$ and $y_{**} = \theta_2 \cdot z$, where $\theta_1 := 1/[\alpha + (1 - \alpha)\mu]$ and $\theta_2 := \mu/[\alpha + (1 - \alpha)\mu]$. To see this, recall that $x_* = (1/\mu) \cdot y$, $y_* = \mu \cdot x$ and $z = \alpha \cdot x + (1 - \alpha) \cdot y$. Note, $z = \alpha \cdot x_{**} + (1 - \alpha) \cdot y_{**}$. Since f is homogenous of degree one, $f(z) = \alpha \cdot f(x_{**}) + (1 - \alpha) \cdot f(y_{**})$ as (1) $x_{**} = \theta_1 \cdot z$, (2) $y_{**} = \theta_2 \cdot z$ and (3) $z = \alpha \cdot x_{**} + (1 - \alpha) \cdot y_{**}$. As f is quasiconvex, $f(x_{**}) \leq f(x)$ and $f(y_{**}) \leq f(y)$ as $f(x) = f(x_*)$ and $f(y) = f(y_*)$. ■

Theorem 8: *If* $\theta \in \mathbb{R}^n_{++}$ *and* $p \geq 1$, *then* $\|\cdot\|_{(\theta,p)}$ *is a norm on* $\mathbb{R}^n$.

Proof: Observe, it is enough to argue: $\|\cdot\|_{(\theta,p)}$ is subadditive. Note, $\|\cdot\|_{(\theta,p)}$ is clearly homogenous of degree one by definition. Thus, it is enough to show that $\|\cdot\|_{(\theta,p)}$ is convex by Lemma 1. For this, we appeal to Lemma 2. That is, we argue: $\|\cdot\|_{(\theta,p)}$ is quasiconvex.

Define the map $f : \mathbb{R}^n \to \mathbb{R}_+$ by: $f(x) := \sum_{i=1}^n \theta_i |x_i|^p$ for all $x \in \mathbb{R}^n$. Note, $|\cdot| : \mathbb{R} \to \mathbb{R}_+$ is a convex function. Further, $p \geq 1$ implies $\xi \in \mathbb{R}_+ \mapsto \xi^p$ is also a convex function. Since the composition of convex functions is convex, it follows that $\xi \in \mathbb{R} \mapsto \theta_i|\xi|^p$ is a convex function for every $i = 1, \ldots, n$.

Let $\pi_i : \mathbb{R}^n \to \mathbb{R}$ be ith projection map. Since π_i is a linear functional, we have: $x \in \mathbb{R}^n \mapsto \theta_i|x_i|^p$ is a convex function. Thus, f being the sum of convex functions is convex. Since the map $\xi \in \mathbb{R}_+ \mapsto \xi^{1/p}$ is strictly increasing, it follows that $x \in \mathbb{R}^n \mapsto [f(x)]^{1/p}$ is *quasiconvex.* Observe, $[f(x)]^{1/p} = \|x\|_{(\theta,p)}$ for every $x \in \mathbb{R}^n$. That is, the function $\|\cdot\|_{(\theta,p)}$ is quasiconvex. This completes the proof. ■

Corollary 3: *Suppose* $\theta \in \mathbb{R}^n_{++}$ *and* $p \geq 1$. *Then, for any* $x, y \in \mathbb{R}^n$,

$$\|x + y\|_{(\theta,p)} \leq \|x\|_{(\theta,p)} + \|y\|_{(\theta,p)}.$$

This completes our presentation of *Minkowski's inequality.*

4.2 *Hölder's inequality*

In the subsection on duality, we introduced the notion of "conjugate" of any arbitrary pre–norm on $\mathbb{R}^n$. To recall, let $f : \mathbb{R}^n \to \mathbb{R}_+$ be a pre–norm. Then, its dual is the function $T(f) : \mathbb{R}^n \to \mathbb{R}_+$ which is also a pre–norm and is defined as follows:

$$[T(f)](x) := \max_{f(y) \leq 1} \; x \cdot y \quad \text{for every } x \in \mathbb{R}^n.$$

Theorem 8 of the previous subsection shows that the function $\|\cdot\|_p$ is a norm if $p \geq 1$. Our immediate objective is to compute the dual of the norm $\|\cdot\|_p$ for any $p \geq 1$. We conduct this analysis into two parts. First, we analyse those p–norms where $p > 1$. Then, we consider the 1–norm. For any $p > 1$, the *conjugate index* of p is the unique q such that $1/p + 1/q = 1$. Note, $p > 1$ implies $q > 1$. The first main result in this direction is as follows.

Theorem 9: *Suppose $p > 1$ and let q be its conjugate index. Then, the conjugate of the norm $\|\cdot\|_p$ is the norm $\|\cdot\|_q$.*

Proof: Fix any $x \in \mathbb{R}^n$. We argue: $\max_{\|y\|_p \leq 1} x \cdot y = \|x\|_q$. It is trivial if $x = \mathbf{0}$. Further, observe that the norms $\|\cdot\|_p$ and $\|\cdot\|_q$ evaluate to the same value for any vector x and $(-x_I, x_{N \setminus I})$, where $(x_I, x_{N \setminus I})$ is the vector obtained from x by inverting the sign of x_k for each $k \in I$. Hence, without loss of generality, let $x \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\}$, and suppose $y^* \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\}$ satisfies (1) $\|y^*\|_p = 1$, and (2) $x \cdot y^* = \max_{\|y\|_p \leq 1} x \cdot y$. Define the Lagrangian $\mathcal{L} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ as:

$$\mathcal{L}(y\,; \lambda) \coloneqq x \cdot y + \lambda(1 - \|y\|_p) \quad \text{for all } y \in \mathbb{R}^n \text{ and } \lambda \in \mathbb{R}.$$

Note, $\|\cdot\|_p$ is smooth over $\mathbb{R}^n_+$ as $p > 1$. Thus, the "first–order necessary conditions" hold. Also, we have:

$$\frac{\partial}{\partial y_i} \mathcal{L}(y\,; \lambda) = x_i - \frac{\lambda}{\|y\|_p^{(p-1)}}\, y_i^{p-1} \quad \text{for all } i = 1, \ldots, n.$$

Thus, there exists $\lambda^* > 0$ such that the following holds:

$$\frac{\partial}{\partial y_i}\bigg|_{y = y^*\,;\, \lambda = \lambda^*} \mathcal{L}(y\,; \lambda) = 0 \quad \text{for all } i = 1, \ldots, n.$$

Thus, $\mu^* \coloneqq \lambda^* / \|y^*\|_p^{(p-1)}$ implies: $x_i = \mu^* (y_i^*)^{p-1}$ for all $i = 1, \ldots, n$. Fix any $i \geq 2$. Then, $y_i^* = y_1^* (x_i / x_1)^{1/(p-1)}$. Thus, $(y_i^*)^p = (y_1^*)^p (x_i/x_1)^q$ as $q = p/(p-1)$. Hence, $\|y^*\|_p^p = \sum_{i=1}^n (y_i^*)^p$ implies:

$$\|y^*\|_p^p = (y_1^*)^p + \sum_{i=2}^n (y_i^*)^p (x_i/x_1)^q.$$

That is, $\|y^*\|_p^p = (y_1^*)^p \|x\|_q^q / x_1^q$. By $\|y^*\|_p = 1$, $y_1^* = (x_1 / \|x\|_q)^{q/p}$. Note, $q/p = 1/(p-1)$ as $q = p/(p-1)$. Recall, $y_i^* = y_1^*(x_i/x_1)^{1/(p-1)}$ for all $i \geq 2$. Thus, $y_i^* = (x_i / \|x\|_q)^{1/(p-1)}$ for every $i = 1, \ldots, n$. Then, $x \cdot y^* = \sum_{i=1}^n x_i y_i^*$ implies the following:

$$x \cdot y^* = (1/\|x\|_q)^{1/(p-1)} \sum_{i=1}^n x_i^{1+1/(p-1)} = (1/\|x\|_q)^{1/(p-1)} \|x\|_q^q.$$

As $q = p/(p-1)$, $(1/\|x\|_q)^{1/(p-1)} \|x\|_q^q = \|x\|_q^{p/(p-1) - 1/(p-1)} = \|x\|_q$. Thus, $x \cdot y^* = \|x\|_q$. Recall, $x \cdot y^* = \max_{\|y\|_p \leq 1} x \cdot y$. ■

We remark that the concept of conjugate of a pre–norm, as defined in the previous section, is rooted in the geometry as represented by duality. However, the the definition of the conjugate index lacks any motivation. Theorem 9 shows that the q–norm is the conjugate of the p–norm, if and only if, q is the conjugate index of p. However, this result comes with one caveat that $p > 1$. Notwithstanding this caveat, we note that the generalized Hölder inequality, presented as Corollary 1 in subsection 3.2, now implies the classical version.

Corollary 4: *Suppose $p, q > 1$ satisfy $1/p + 1/q = 1$. Then,*

$$|x \cdot y| \leq \|x\|_p \cdot \|x\|_q \quad \textit{for every } x, y \in \mathbb{R}^n.$$

For the case when $p = 1$, the standard approach is to show that (a) $\lim_{p \downarrow 1} \|x\|_p = \|x\|_1$ and (b) $\|x\|_\infty := \lim_{q \uparrow \infty} \|x\|_q$ exists. Thus, (b) and Minkowski's inequality imply that $\|\cdot\|_\infty$ is a norm. Further, Hölder's inequality then implies the following:

$$|x \cdot y| \leq \|x\|_1 \cdot \|y\|_\infty \quad \text{for every } x, y \in \mathbb{R}^n. \tag{1}$$

Moreover, (b) implies that $\|\cdot\|_\infty : \mathbb{R}^n \to \mathbb{R}_+$ satisfies:

$$\|x\|_\infty = \max\left\{|x_i| : i = 1, \ldots, n\right\} \quad \text{for every } x \in \mathbb{R}^n. \tag{2}$$

Our approach will be to establish (1) directly. The strategy will be to take (2) as the definition of $\|\cdot\|_\infty$. We shall demonstrate, by direct computation, that $\|\cdot\|_\infty$ is the conjugate of $\|\cdot\|_1$. Then, the generalized Hölder's inequality (Corollary 1) will deliver (1).

Theorem 10: *The norms $\|\cdot\|_1$ are $\|\cdot\|_\infty$ are conjugates of each other.*

Proof: We will argue: $\max_{\|y\|_1 \leq 1} x \cdot y = \|x\|_\infty$ for any $x \in \mathbb{R}^n$. It is trivial if $x = \mathbf{0}$. Consider any $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Let $y^* \in \operatorname{argmax}_{\|y\|_1 \leq 1} x \cdot y$. Define $I_+ := \{i \in N : x_i > 0\}$ and $I_- := \{i \in N : x_i < 0\}$. Then, we have: $\|\cdot\|_1$ implies (i) $y_i^* \geq 0$ if $i \in I_+$, and (ii) $y_i^* \leq 0$ if $i \in I_-$. To see why, assume $i_0 \in I_+$ and suppose $y_{i_0}^* < 0$. Let $y^+ := (y_1^+, \ldots, y_n^+)$ such that $y_{i_0}^+ := -y_{i_o}^*$ and $y_i^+ := y_i^*$ otherwise. Note, $\|y^+\|_1 = \|y^*\|_1$ implying $\|y^+\|_1 \leq 1$. Also, $x \cdot y^+ > x \cdot y^*$ as $x_{i_0} y_{i_0}^+ > 0 > x_{i_0} y_{i_0}^*$. This contradicts $y^* \in \operatorname{argmax}_{\|y\|_1} x \cdot y$. Thus, $y_i^* \geq 0$ if $i \in I_+$. A similar argument proves $y_i^* \leq 0$ if $i \in I_-$. Thus, (i) and (ii) hold.

Now, consider $x^* := (|x_1|, \ldots, |x_n|)$ and $y^{**} := (|y_1^*|, \ldots, |y_n^*|)$. Thus, $x^* \cdot y^{**} = x \cdot y^*$ by (i) and (ii). That is, $x^* \cdot y^{**} = \max_{\|y\|_1 \leq 1} x \cdot y$. Note, $\|y^*\|_1 = \|y^{**}\|_1$ implying $\|y^{**}\|_1 \leq 1$. Thus, $x^* \cdot y^{**} \leq \max_{\|y\|_1 \leq 1} x^* \cdot y$. Hence, we have: $\max_{\|y\|_1 \leq 1} x \cdot y \leq \max_{\|y\|_1 \leq 1} x^* \cdot y$.

Let $y^{++} \in \text{argmax}_{\|y\|_1 \leq 1} x^* \cdot y$. As $x^* \in \mathbb{R}^n_+$, $y_i^{++} \geq 0$ for all $i \in N$. Define $y^- := (y_1^-, \ldots, y_n^-)$, where $y_i^- := -y_i^{++}$ if $i \in I_-$ and $y_i^- := y_i^{++}$ otherwise. Thus, $x \cdot y^- = x^* \cdot y^{++} = \max_{\|y\|_1 \leq 1} x^* \cdot y$. Also, $\|y^-\|_1 = \|y^{++}\|_1$ implying $\|y^-\|_1 \leq 1$. Thus, $\max_{\|y\|_1 \leq 1} x \cdot y \geq x \cdot y^-$. Hence, $\max_{\|y\|_1 \leq 1} x \cdot y \geq \max_{\|y\|_1 \leq 1} x^* \cdot y$. That is, we have:

$$\max_{\|y\|_1 \leq 1} x \cdot y = \max_{\|y\|_1 \leq 1} x^* \cdot y.$$

Clearly, $\|x\|_\infty = \|x^*\|_\infty$. Hence, if $\max_{\|y\|_1 \leq 1} x^* \cdot y = \|x^*\|_\infty$ then $\max_{\|y\|_1 \leq 1} x \cdot y = \|x\|_\infty$. Thus, we may assume $x \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\}$ without any loss of generality. Henceforth, let $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$.

Let $y^* \in \text{argmax}_{\|y\|_1 \leq 1} x \cdot y$. Then, without loss of generality, we may assume that $y^* \in \mathbb{R}^n_+$. To see why, define $y^+ := (y_1^+, \ldots, y_n^+)$ as: $y_i^+ := y_i^*$ if $y_i^* \geq 0$, and $y_i^+ := -y_i^*$ otherwise. Clearly, $\|y^+\|_1 = \|y^*\|_1$ implying $\|y^*\|_1 \leq 1$. Further, $x \cdot y^+ \geq x \cdot y^*$ as $x \in \mathbb{R}^n_+$. Then, $x \cdot y^* = \max_{\|y\|_1} x \cdot y$ implies: $y^+ \in \text{argmax}_{\|y\|_1 \leq 1} x \cdot y$. Also, note that $y^+ \in \mathbb{R}^n_+$.

Henceforth, we assume $y^* \in \text{argmax}_{\|y\|_1 \leq 1} x \cdot y$ such that $y^* \in \mathbb{R}^n_+$. Now, $x \neq \mathbf{0}$ implies $\|x\|_1 > 0$. Let $\alpha := 1/\|x\|_1$. Thus, $y_\alpha := \alpha \cdot x$ implies $\|y_\alpha\|_1 = 1$ and $x \cdot y_\alpha = (x \cdot x)/\|x\|_1 = \|x\|_2^2/\|x\|_1$. Thus, $x \cdot y_\alpha > x \cdot \mathbf{0}$ holds. Then, $\|y_\alpha\|_1 = 1$ implies $y^* \neq \mathbf{0}$. Further, $\max_{\|y\|_1 \leq 1} x \cdot y \geq x \cdot y_\alpha$. Thus, $y^* \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\}$ and $\max_{\|y\|_1 \leq 1} x \cdot y > 0$.

Observe, $\|y^*\|_1 = 1$. To see why, suppose $\|y^*\|_1 < 1$. Let $y_\alpha := \alpha \cdot y^*/\|y^*\|_1$. Then, $\|y^\alpha\|_1 = 1$ and $x \cdot y_\alpha = (x \cdot y^*)/\|y^*\|_1$. Since $x \cdot y^* = \max_{\|y\|_1 \leq 1} x \cdot y > 0$ and $\|y^*\|_1 \in (0, 1)$, we have: $(x \cdot y^*)/\|y^*\|_1 > x \cdot y^*$. That is, $x \cdot y_\alpha > \max_{\|y\|_1 \leq 1} x \cdot y$, where $\|y_\alpha\|_1 = 1$. Clearly, this is a contradiction. Thus, we have: $\|y^*\|_1 = 1$.

Note, $y^* \in \mathbb{R}^n_+$ implies $\|y^*\|_1 = \sum_{i=1}^n y_i^*$. Then, $\|y^*\|_1 = 1$ implies $\sum_{i=1}^n y_i^* = 1$. Clearly, $y^* = \sum_{i=1}^n y_i^* \cdot e_i$, where e_i is the ith standard basis vector of $\mathbb{R}^n$. Thus, $y^* \in V$, where $V \subseteq \mathbb{R}^n$ is the convex hull of $\{e_1, \ldots, e_n\}$. Let $\theta := \max\{x \cdot e_i : i = 1, \ldots, n\}$. Thus, $x \cdot y \leq \theta$ for all $y \in V$. Hence, $x \cdot y^* \leq \theta$ implying: $\max_{\|y\|_1} x \cdot y \leq \theta$.

Also, note that $\|e_i\|_1 = 1$ for every $i = 1, \ldots, n$. Thus, $x \cdot e_i \leq \max_{\|y\|_1 \leq 1} x \cdot y$ for all $i = 1, \ldots, n$. Hence, $\theta \leq \max_{\|y\|_1 \leq 1} x \cdot y$ because $\theta = \max\{x \cdot e_i : i = 1, \ldots, n\}$. Thus, $\theta = \max_{\|y\|_1 \leq 1} x \cdot y$. Observe, $\theta = \|x\|_\infty$. Hence, $\max_{\|y\|_1 \leq 1} x \cdot y = \|x\|_\infty$ if $x \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\}$. Thus, $\|\cdot\|_\infty$ is the conjugate of $\|\|_1$. By Theorem 5, $\|\cdot\|_\infty$ is a norm and it is the conjugate of $\|\cdot\|_1$. This completes the proof. ■.

Then, the generalized Hölder inequality (Corollary 1) implies:

$$|x \cdot y| \leq \|x\|_1 \cdot \|y\|_\infty \quad \text{for all } x, y \in \mathbb{R}^n.$$

APPENDIX

A.I.1 The Basic Representation Theorem

LEMMA *A.I.1(a)*: *Let the map $f : \mathbb{R}^n \to \mathbb{R}$ be an $\mathcal{H}$–representation of the binary relation $\succ$ over $\mathbb{R}^n$. Then, $f(\mathbf{0}) = 0$.*

PROOF: Let $f : \mathbb{R}^n \to \mathbb{R}$ be an $\mathcal{H}$–representation of $\succ$ on $\mathbb{R}^n$. Pick an $\alpha > 0$ such that $\alpha \neq 1$. Since f is a homogenous map and $\mathbf{0} = \alpha \cdot \mathbf{0}$, we have: $f(\mathbf{0}) = \alpha \cdot f(\mathbf{0})$. Then, $f(\mathbf{0}) = 0$ because $\alpha \neq 1$. ■

PROOF OF PROPOSITION 1: Suppose that the map $f : \mathbb{R}^n \to \mathbb{R}_+$ is an $\mathcal{H}$–representation of the binary relation $\succ$ over $\mathbb{R}^n$. Consider the map $g : \mathbb{R}^n \to \mathbb{R}$. Then, if there exists $\alpha > 0$ such that $g = \alpha \cdot f$ then g is also an $\mathcal{H}$–representation of $\succ$. This is because as (1) $\alpha \cdot f(x) > \alpha \cdot f(y)$ iff $f(x) > f(y)$, (2) $f(x) > f(y)$ iff $x \succ y$, and (3) g is homogenous of degree one. To see this, note that (1) holds because $\alpha > 0$, (2) holds because f is an $\mathcal{H}$–representation of $\succ$, and (3) holds because $g = \alpha \cdot f$ and f is homogenous of degree one.

For the converse, assume that $g : \mathbb{R}^n \to \mathbb{R}$ is an $\mathcal{H}$–representation of $\succ$. There are two cases. First, assume $f(x) = 0$ for all $x \in \mathbb{R}^n$. Then, $x \sim y$ for all $x, y \in \mathbb{R}^n$ as f is an $\mathcal{H}$–representation of $\succ$. As g is an $\mathcal{H}$–representation of $\succ$, there exists $\theta \in \mathbb{R}$ such that $g(x) = \theta$ for all $x \in \mathbb{R}^n$. Fix any $x_0 \in \mathbb{R}^n$ and $\alpha_0 > 1$. Let $x_1 \coloneqq \alpha_0 \cdot x_0$. Then, $g(x_1) = \alpha_0 \cdot g(x_0)$ as g is homogenous of degree one because g is an $\mathcal{H}$–representation. That is, $\theta = \alpha_0 \cdot \theta$ which is equivalent to $(1 - \alpha_0)\theta = 0$. Since $\alpha_0 \neq 1$, we have $\theta = 0$. Hence, $g(x) = 0$ for all $x \in \mathbb{R}^n$. Let $\alpha \coloneqq 1$. Thus, $\alpha > 0$ and $g = \alpha \cdot f$ as required.

Note, $f(\mathbf{0}) = 0 = g(\mathbf{0})$ by Lemma *A.I.1(a)*. Now, assume $x_0 \in \mathbb{R}^n$ satisfies $f(x_0) > 0$. Note, $f(x_0) > 0 = f(\mathbf{0})$ implies $x_0 \succ \mathbf{0}$. Since g is an $\mathcal{H}$–representation of $\succ$, we have $g(x_0) > g(\mathbf{0})$. Then, $g(\mathbf{0}) = 0$ implies $g(x_0) > 0$. Observe that by this argument, for any $x \in \mathbb{R}^n$, we have: $f(x) > 0$ iff $g(x) > 0$. In particular, the map f being $\mathbb{R}_+$–valued implies that the map g is $\mathbb{R}_+$–valued.

Let $\alpha_0 \coloneqq g(x_0)/f(x_0)$. Then, $\alpha_0 > 0$ since $f(x_0) > 0$. Now, we argue: $g(x) = \alpha_0 \cdot f(x)$ for all $x \in \mathbb{R}^n$. Fix an arbitrary $x \in \mathbb{R}^n$. Then, $f(x) > 0$ iff $g(x) > 0$ implies: $g(x) = \alpha_0 \cdot f(x)$ if $f(x) = 0$. Henceforth, we assume $f(x) > 0$. Note, $\mathbb{R}_{++} = \{\alpha \cdot f(x_0) : \alpha > 0\}$ as $f(x_0) > 0$. Hence, $f(x) \in \mathbb{R}_{++}$ implies, there exists $\alpha_x > 0$ such that $f(x) = \alpha_x \cdot f(x_0)$. Then, $x \sim \alpha_x \cdot x_0$ because f is an $\mathcal{H}$–representation of $\succ$. Hence, $g(x) = \alpha_x \cdot g(x_0)$ as g is an $\mathcal{H}$–representation of $\succ$. Thus, $g(x)/g(x_0) = \alpha_x = f(x)/f(x_0)$. That is, $g(x) = \alpha_0 \cdot f(x)$. The proof is complete as $x \in \mathbb{R}^n$ is arbitrary. ■

In the rest of the development, we adopt two standard notational devices. First, if A and B are subsets of $\mathbb{R}^n$ then $A + B$ is the name of the set $\{x + y : x \in A \text{ and } y \in B\}$. Second, if A is a subset of $\mathbb{R}^n$ and $\alpha \in \mathbb{R}$ then $\alpha \cdot A$ is the name of the set $\{\alpha \cdot x : x \in A\}$. Thus, if $A, B \subseteq \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ then $\alpha \cdot A + \beta \cdot B$ is the name of the set $\{\alpha \cdot x + \beta \cdot y : x \in A \text{ and } y \in B\}$. Before proceeding to the proofs of Theorems 1 and 2, we establish some preliminary results.

LEMMA *A*.I.1(*b*): *Let $\kappa_1, \kappa_2 > 0$ and $C \subseteq \mathbb{R}^n$ be convex. Then,*

$$\kappa_1 \cdot C + \kappa_2 \cdot C = (\kappa_1 + \kappa_2) \cdot C.$$

PROOF: First, we argue: $(\kappa_1 + \kappa_2) \cdot C \subseteq \kappa_1 \cdot C + \kappa_2 \cdot C$. For this, pick an arbitrary $x \in C$. Then, $\kappa_1 \cdot x \in \kappa_1 \cdot C$ and $\kappa_2 \cdot x \in \kappa_2 \cdot C$. Then, $\kappa_1 \cdot x + \kappa_2 \cdot x \in \kappa_1 \cdot C + \kappa_2 \cdot C$. Since $\kappa_1 \cdot x + \kappa_2 \cdot x = (\kappa_1 + \kappa_2) \cdot x$, it follows that $(\kappa_1 + \kappa_2) \cdot x \in \kappa_1 \cdot C + \kappa_2 \cdot C$. Since $x \in C$ is arbitrary, we have: if $z \in (\kappa_1 + \kappa_2) \cdot C$ then $z \in \kappa_1 \cdot C + \kappa_2 \cdot C$. That is,

$$(\kappa_1 + \kappa_2) \cdot C \subseteq \kappa_1 \cdot C + \kappa_2 \cdot C.$$

Thus far, we have not appealed to the fact that $\kappa_1, \kappa_2 > 0$ or the convexity of C. Now, we proceed to establish the reverse set–inclusion. So, let $z_1 \in \kappa_1 \cdot C$ and $z_2 \in \kappa_2 \cdot C$ be arbitrary. Thus, $z_1 = \kappa_1 \cdot x_1$ and $z_2 = \kappa_2 \cdot x_2$ for some $x_1, x_2 \in C$. Let $\alpha := \kappa_1/(\kappa_1 + \kappa_2)$. Since $\kappa_1, \kappa_2 > 0$, we have $\alpha \in (0, 1)$. Define $x_* := \alpha \cdot x_1 + (1 - \alpha) \cdot x_2$. Then, $x_* \in C$ because $x_1, x_2 \in C$ and the set C is convex. Observe, the definition of α and that $x_* \in C$ implies: $(\kappa_1 \cdot x_1 + \kappa_2 \cdot x_2)/(\kappa_1 + \kappa_2) \in C$. Thus, $(\kappa_1 + \kappa_2) \cdot x_* = \kappa_1 \cdot x_1 + \kappa_2 \cdot x_2$. Also, since $x_* \in C$, it follows that $(\kappa_1 + \kappa_2) \cdot x_* \in (\kappa_1 + \kappa_2) \cdot C$. That is, $\kappa_1 \cdot x_1 + \kappa_2 \cdot x_2 \in (\kappa_1 + \kappa_2) \cdot C$. Hence, $z_1 + z_2 \in (\kappa_1 + \kappa_2) \cdot C$. Since $z_1 \in \kappa_1 \cdot C$ and $z_2 \in \kappa_2 \cdot C$ are arbitrary, $z \in \kappa_1 \cdot C + \kappa_2 \cdot C$ implies $z \in (\kappa_1 + \kappa_2) \cdot C$. That is,

$$\kappa_1 \cdot C + \kappa_2 \cdot C \subseteq (\kappa_1 + \kappa_2) \cdot C.$$

This completes the proof of the lemma. ■

PROOF OF THEOREM 2: Let C be a convex and compact subset of $\mathbb{R}^n$ with $\mathbf{0}$ in the interior of C. Also, define $\|\cdot\|_C : \mathbb{R}^n \to \mathbb{R}_+$ as:

$$\|x\|_C := \inf\left\{\kappa > 0 : x \in \kappa \cdot C\right\} \quad \text{for all } x \in \mathbb{R}^n,$$

where $\kappa \cdot C := \{\kappa \cdot y : y \in C\}$. First, we show that $\|\cdot\|_C$ is a pre–norm on $\mathbb{R}^n$. For this, we must argue that conditions 1 to 4 in Definition 1 (of section 2) hold for the map $\|\cdot\|_C$.

To see why condition 1 holds, let $x \in \mathbb{R}^n$. Since $\mathbf{0}$ is in the interior of C, there exists $\kappa > 0$ such that $x \in \kappa \cdot C$. Thus, $\|x\|_C \geq 0$. To see why condition 2 holds, we begin by observing that $\|\mathbf{0}\|_C = 0$ because $\mathbf{0} \in \kappa \cdot C$ for every $\kappa > 0$ as $\mathbf{0} \in C$. Moreover, if $x \neq \mathbf{0}$ then, there exists a corresponding $\kappa_x > 0$ such that $x \notin \kappa \cdot C$ for any $0 < \kappa < \kappa_x$. This is so as C being a compact subset of $\mathbb{R}^n$ must be bounded, say, with respect to the norm $\|\cdot\|_1$. Thus, $x \neq \mathbf{0}$ implies $\|x\|_C > 0$. This shows that condition 2 holds. Condition 3 follows from the following observation. If $x \in \mathbb{R}^n$ and $\alpha > 0$ then,

$$x \in \kappa \cdot C \iff \alpha \cdot x \in (\alpha\kappa) \cdot C \quad \text{for every } \kappa > 0.$$

Thus, to establish the claim that $\|\cdot\|_C$ is a pre–norm, it remains to argue that $\|\cdot\|_C$ satisfies condition 4. Consider any x_1 and x_2 in $\mathbb{R}^n$. Pick any arbitrary $\kappa_1, \kappa_2 > 0$ such that $x_1 \in \kappa_1 \cdot C$ and $x_2 \in \kappa_2 \cdot C$. Let $x_* := x_1 + x_2$. Thus, $x_* \in \kappa_1 \cdot C + \kappa_2 \cdot C$. Since C is convex, lemma A.I.1(b) implies: $\kappa_1 \cdot C + \kappa_2 \cdot C = (\kappa_1 + \kappa_2) \cdot C$. Hence, $x_* \in (\kappa_1 + \kappa_2) \cdot C$. Thus, $\inf\{\kappa > 0 : x_* \in \kappa \cdot C\} \leq \kappa_1 + \kappa_2$. That is, $\|x_*\|_C \leq \kappa_1 + \kappa_2$ holds. Since $\kappa_1, \kappa_2 > 0$ are arbitrary subject to satisfying $x_1 \in \kappa_1 \cdot C$ and $x_2 \in \kappa_2 \cdot C$, the following obtains:

$$\|x_*\|_C \leq \inf\big\{\kappa_1 > 0 : x_1 \in \kappa_1 \cdot C\big\} + \inf\big\{\kappa_2 > 0 : x_2 \in \kappa_2 \cdot C\big\}.$$

That is, $\|x_1 + x_2\|_C = \|x_*\|_C \leq \|x_1\|_C + \|x_2\|_C$ (recall, $x_* = x_1 + x_2$). Hence, condition 4 of Definition 1 is established. Therefore, we have shown: $\|\cdot\|_C$ is a pre–norm on $\mathbb{R}^n$.

Now, we argue: $C = \{x \in \mathbb{R}^n : \|x\|_C \leq 1\}$. First, assume $x \in C$. Then, $1 \in \{\kappa > 0 : x \in \kappa \cdot C\}$. Hence, it follows that $\|x\|_C \leq 1$ as $\|x\|_C = \inf\{\kappa > 0 : x \in \kappa \cdot C\}$ by definition. That is,

$$C \subseteq \{x \in \mathbb{R}^n : \|x\|_C \leq 1\}.$$

For the reverse set–inclusion, assume $x_0 \in \mathbb{R}^n$ satisfies $\|x_0\|_C \leq 1$. Suppose $x_0 \notin C$. As $\{x_0\}$ and C are disjoint and convex *compact* sets, the Separating Hyperplane Theorem implies that there exists $p \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ such that the following hold:

1. $p \cdot x_0 > \alpha$, and
2. $p \cdot x < \alpha$ for all $x \in C$.

First, note that $\alpha > 0$. To see why, observe that $\mathbf{0} \in C$. Thus, $p \cdot x_0 < \alpha$ must hold by (2). Since $p \cdot \mathbf{0} = 0$, it follows that $\alpha > 0$. Moreover, $p \neq \mathbf{0}$. To see why, suppose $p = \mathbf{0}$. Then, $p \cdot x_0 = 0$ implying $\alpha < 0$ by (1). Hence, $p \neq \mathbf{0}$ and $\alpha > 0$.

Consider an arbitrary $\kappa > 0$ such that $x_0 \in \kappa \cdot C$. Thus, there exists $y_0 \in C$ such that $x_0 = \kappa \cdot y_0$. Then, (2) implies $p \cdot y_0 < \alpha$ because $y_0 \in C$. Since $x_0 = \kappa \cdot y_0$, it follows that $p \cdot x_0 = p \cdot (\kappa \cdot y_0) < \kappa\alpha$. That is, $\kappa > \kappa_* := (p \cdot x_0)/\alpha$. Also, (1) implies $\kappa_* > 1$. Hence, we have shown that the set $\{\kappa > 0 : x_0 \in \kappa \cdot C\}$ is bounded below by κ_* with κ_* being *strictly* greater than 1. Since $\|x_0\|_C = \inf\{\kappa > 0 : x_0 \in \kappa \cdot C\}$, it follows that $\|x_0\|_C > 1$ which is a contradiction. Thus, if $x \in \mathbb{R}^n$ satisfies $\|x\|_C \leq 1$ then $x \in C$. That is,

$$\big\{x \in \mathbb{R}^n : \|x\|_C \leq 1\big\} \subseteq C.$$

This proves the reverse set–inclusion. Thus, we have established the first of the two parts of Theorem 2. Now, we proceed to prove the second part of Theorem 2. For this, consider a pre–norm $f : \mathbb{R}^n \to \mathbb{R}_+$ and let $C_f \subseteq \mathbb{R}^n$ be defined as follows:

$$C_f := \{x \in \mathbb{R}^n : f(x) \leq 1\}.$$

First, we show that C_f is a convex and compact set with $\mathbf{0}$ in the interior of C_f. To show that C_f is convex, let $x_1, x_2 \in C_f$ and $\alpha \in (0, 1)$. Let $y_1 := \alpha \cdot x_1$ and $y_2 := (1-\alpha) \cdot x_2$. Since f is a pre–norm, $f(\alpha \cdot x_1) = \alpha \cdot f(x_1)$ by condition 3 in Definition 1. Also, $f(x_1) \leq 1$ as $x_1 \in C_f$. Thus, $f(\alpha \cdot x_1) \leq \alpha$. That is, $f(y_1) \leq \alpha$. Similarly, $f(y_2) \leq 1 - \alpha$. Let $x_* := \alpha \cdot x_1 + (1-\alpha) \cdot x_2$ and note that $x_* = y_1 + y_2$. Then, as f is a pre–norm, we have $f(x_*) \leq f(y_1) + f(y_2)$ by condition 4 in Definition 1. Since $f(y_1) \leq \alpha$ and $f(y_2) \leq 1 - \alpha$, we have $f(x_*) \leq 1$. That is, $x_* \in C_f$. Recall, $x_* = \alpha \cdot x_1 + (1-\alpha) \cdot x_2$, where $x_1, x_2 \in C_f$ and $\alpha \in (0, 1)$ are arbitrary. Thus, C_f is convex.

To show that C_f is compact and has $\mathbf{0}$ is in the interior of C, we shall use proposition 0 of subsection 5.2 according to which f being a pre–norm on $\mathbb{R}^n$ is continuous with respect to $\|\cdot\|_1$. Then, since C_f is the pullback under f of the closed set $(-\infty, 1]$, it follows that C_f is a closed subset of $\mathbb{R}^n$. Further, C_f is bounded according to $\|\cdot\|_1$ as it is clearly bounded according to f with f being equivalent to $\|\cdot\|_1$. Thus, C_f is compact by the Heine–Borel Theorem. Moreover, note that $B_f(\mathbf{0}, 1) := \{x \in \mathbb{R}^n : f(x) < 1\}$ is an open set with $\mathbf{0} \in B_f(\mathbf{0}, 1) \subseteq C_f$. Since the topologies generated by f and $\|\cdot\|_1$ are identical, it follows that $\mathbf{0}$ is in the interior of C_f.

To complete the proof of the theorem, it remains to argue that the function $\|\cdot\|_{C_f} : \mathbb{R}^n \to \mathbb{R}_+$ defined as:

$$\|x\|_C := \inf\big\{\kappa > 0 : x \in \kappa \cdot C_f\big\} \quad \text{for all } x \in \mathbb{R}^n,$$

satisfies: $\|x\|_C = f(x)$ for all $x \in \mathbb{R}^n$. We proceed as follows.

Observe, by the previous part of the claim of this theorem, the map $\|\cdot\|_C$ is a pre–norm on $\mathbb{R}^n$. This is because C_f is a convex compact set with $\mathbf{0}$ in the interior of C_f. Moreover, $\|\cdot\|_C$ satisfies:

$$C_f = \{x \in \mathbb{R}^n : \|x\|_C \leq 1\}.$$

Define, for any $\theta > 0$, the sets $D_f(\mathbf{0}, \theta) \coloneqq \{x \in \mathbb{R}^n : f(x) \leq \theta\}$ and $D_{\|\cdot\|_{C_f}}(\mathbf{0}, \theta) \coloneqq \{x \in \mathbb{R}^n : \|x\|_{C_f} \leq \theta\}$. Thus, $D_f(\mathbf{0}, 1) = D_{\|\cdot\|_{C_f}}(\mathbf{0}, 1)$ as each is equal to C_f. Since both $\|\cdot\|_{C_f}$ and f are pre–norms,

$$D_{\|\cdot\|_{C_f}}(\mathbf{0}, \theta) = D_f(\mathbf{0}, \theta) \quad \text{for all } \theta > 0.$$

To see why, fix any $\theta > 0$ and let $x \in D_{\|\cdot\|_{C_f}}(\mathbf{0}, \theta)$ be arbitrary. Thus, $\|x\|_{C_f} \leq \theta$. Let $x_\theta \coloneqq (1/\theta) \cdot x$. Since $\|\cdot\|_{C_f}$ is a pre–norm, we have $\|x_\theta\|_{C_f} = (1/\theta) \cdot \|x\|_{C_f} \leq 1$. That is, $x_\theta \in D_{\|\cdot\|_{C_f}}(\mathbf{0}, 1)$. Hence, $x_\theta \in D_f(\mathbf{0}, 1)$. That is, $f(x_\theta) \leq 1$. Also, $x = \theta \cdot x_\theta$. Since f is a pre–norm, we have $f(x) = \theta \cdot f(x_\theta) \leq \theta$. Thus, $x \in D_f(\mathbf{0}, \theta)$. Since $x \in D_{\|\cdot\|_{C_f}}(\mathbf{0}, \theta)$ is arbitrary, it follows that:

$$D_{\|\cdot\|_{C_f}}(\mathbf{0}, \theta) \subseteq D_f(\mathbf{0}, \theta).$$

The argument to establish the above set–inclusion relied only on the following two facts. First, the sets $D_{\|\cdot|_{C_f}}(\mathbf{0}, \theta)$ and $D_f(\mathbf{0}, \theta)$ are equal. Second, both $\|\cdot\|_{C_f}$ and f are pre–norms. Hence, a symmetric argument implies the reverse set–inclusion. Therefore, the two sets $D_f(\mathbf{0}, \theta)$ and $D_{\|\cdot\|_{C_f}}(\mathbf{0}, \theta)$ are equal for all $\theta > 0$.

Let $B_{\|\cdot\|_{C_f}}(\mathbf{0}, 1)$ be the set $\{x \in \mathbb{R}^n : \|x\|_{C_f} < 1\}$. Then, it is obvious that $B_{\|\cdot\|_{C_f}}(\mathbf{0}, 1) = \bigcup_{0<\theta<1} D_{\|\cdot\|_{C_f}}(\mathbf{0}, \theta)$. Moreover, recall that $B_f(\mathbf{0}, 1)$ is the set $\{x \in \mathbb{R}^n : f(x) < 1\}$. Hence, $B_f(\mathbf{0}, 1) = \bigcup_{0<\theta<1} D_f(\mathbf{0}, \theta)$. Since $D_{\|\cdot\|_{C_f}}(\mathbf{0}, \theta) = D_f(\mathbf{0}, \theta)$ for all $\theta > 0$, we have: $B_{\|\cdot\|_{C_f}}(\mathbf{0}, 1) = B_f(\mathbf{0}, 1)$. Then, $D_{\|\cdot\|_{C_f}}(\mathbf{0}, 1) = D_f(\mathbf{0}, 1)$ implies that the following holds:

$$D_{\|\cdot\|_{C_f}}(\mathbf{0}, 1) \setminus B_{\|\cdot\|_{C_f}}(\mathbf{0}, 1) = D_f(\mathbf{0}, 1) \setminus B_f(\mathbf{0}, 1).$$

That is, $\{x \in \mathbb{R}^n : \|x\|_{C_f} = 1\} = \{x \in \mathbb{R}^n : f(x) = 1\}$. Thus, $\|x\|_{C_f} = 1$ iff $f(x) = 1$. Then, for any $\theta > 0$, $\|x\|_{C_f} = \theta$ iff $f(x) = \theta$. To see why, let $x \in \mathbb{R}^n$ satisfies $\|x|_{C_f} = \theta$ for some $\theta > 0$. Then, $x_\theta \coloneqq (1/\theta) \cdot x$ satisfies $\|x_\theta\|_{C_f} = (1/\theta) \cdot \|x\|_{C_f} = 1$. Thus, $f(x_\theta) = 1$. As $x = \theta \cdot x_\theta$, $f(x) = \theta \cdot f(x_\theta) = \theta$. That is, $\|x\|_{C_f} = \theta$ implies $f(x) = \theta$. Similarly, the converse obtains. Also, $\|x\|_{C_f} = 0$ iff $x = \mathbf{0}$ iff $f(x) = 0$. As $\|\cdot\|_{C_f}$ and f are $\mathbb{R}_+$–valued, $\|x\|_{C_f} = f(x)$ for all $x \in \mathbb{R}^n$. ■

PROOF OF THEOREM 1: There are two parts. First, we establish the existence of $\mathcal{N}_*$–representations. So, let $\succ$ be weak order over $\mathbb{R}^n$ which satisfies Continuity, Homotheticity, Convexity and Scale Monotonicity. Since $\mathcal{N}_*$ is the class of all pre–norms on $\mathbb{R}^n$, we argue: there exists a pre–norm f on $\mathbb{R}^n$ such that f is a representation of $\succ$.

Throughout the rest of the proof, let $x_0 \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ be arbitrary but fixed. Based on this chosen x_0, we define the set $C \subseteq \mathbb{R}^n$ as:

$$C := \big\{x \in \mathbb{R}^n : x_0 \succsim x\big\}.$$

Moreover, we define the map $\|\cdot\|_C : \mathbb{R}^n \to \mathbb{R}_+ \cup \{\infty\}$ as follows:

$$\|x\|_C := \inf \big\{\kappa > 0 : x \in \kappa \cdot C\big\} \quad \text{for all } x \in \mathbb{R}^n.$$

The strategy of our proof of existence entails showing that $\|\cdot\|_C$ is a norm and that $\|\cdot\|_C$ is a representation of $\succ$. To show that $\|\cdot\|_C$ is a norm, we shall make use of Theorem 2 of section 3. For is, we shall have to argue that the set C is convex and compact with $\mathbf{0}$ in its interior. This is where the major force of all the axioms is required. The proof is organised via the following steps.

Step 1: We argue: C is convex. Let $x, y \in C$ and $\alpha \in (0,1)$. Since $\succsim$ is complete, at least one of $x \succsim y$ or $y \succsim x$ holds. Without loss of generality, assume $x \succsim y$. Let $z := \alpha \cdot x + (1-\alpha) \cdot y$. Since $\succ$ satisfies Convexity, we have $x \succsim z$. Also, $x_0 \succsim x$ as $x \in C$. Hence, $x_0 \succsim z$ by transitivity of $\succsim$. That is, $z \in C$. Thus, C is convex.

Step 2: We argue: if $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ then $x \succ \mathbf{0}$. Fix any $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. First, suppose $x \sim \mathbf{0}$. Pick an arbitrary $\alpha > 1$. Then, $\alpha \cdot x \succ x$ by Scale Monotonicity. By cross transitivity of $\succ$ and $\sim$, $\alpha \cdot x \succ x$ and $x \sim \mathbf{0}$ imply $\alpha \cdot x \succ \mathbf{0}$. Also, $\alpha \cdot x \sim \alpha \cdot \mathbf{0} = \mathbf{0}$. Thus, both $\alpha \cdot x \sim \mathbf{0}$ and $\alpha \cdot x \succ \mathbf{0}$ hold. However, this is a contradiction since $\succ$ is asymmetric and $\sim$ is symmetric. Hence, the supposition that $x \sim \mathbf{0}$ holds must be wrong. That is, $x \sim \mathbf{0}$ does *not* hold.

Now, suppose $\mathbf{0} \succ x$. Since $\succ$ satisfies Continuity, there exists $\varepsilon > 0$ such that (1) $\varepsilon < 2\|x\|_1$, and[42] (2) $y \in B_{\|\cdot\|_1}(\mathbf{0}, \varepsilon)$ implies $y \succ x$. Let $y_0 := (\varepsilon/2\|x\|_1) \cdot x$. Then, $\|y_0\|_1 = \varepsilon/2$. Thus, $y_0 \in B_{\|\cdot\|_1}(\mathbf{0}, \varepsilon)$. Hence, $y_0 \succ x$. Let $\alpha_0 := 2\|x\|_1/\varepsilon$. Then, $x = \alpha_0 \cdot y_0$. Since $\alpha_0 > 1$ and $y_0 \neq \mathbf{0}$, Scale Monotonicity implies $x \succ y_0$. Thus, both $x \succ y_0$ and $y_0 \succ x$ hold which is a contradiction as $\succ$ is asymmetric. Thus, $\mathbf{0} \succ x$ does *not* hold. Since $\succsim$ is complete, we obtain: $x \succ \mathbf{0}$.

[42]For any $x \in \mathbb{R}^n$ and $r > 0$, let $B_{\|\cdot\|_1}(x, r) := \{y \in \mathbb{R}^n : \|y - x\|_1 < r\}$.

Step 3: We argue: C is closed in $\mathbb{R}^n$, and $\mathbf{0}$ is in the interior of C. Recall, $x_0 \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ and $C = \{x \in \mathbb{R}^n : x_0 \succsim x\}$. Now, observe that $\{x \in \mathbb{R}^n : x_0 \succsim x\} = \mathbb{R}^n \setminus U_{\succ}(x_0)$ by completeness of $\succsim$, where $U_{\succ}(x_0)$ is the set $\{x \in \mathbb{R}^n : x \succ x_0\}$. By Continuity of $\succ$, $U_{\succ}(x_0)$ is an open subset of $\mathbb{R}^n$. Thus, C is a closed subset of $\mathbb{R}^n$.

Moreover, $x_0 \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ implies $x_0 \succ \mathbf{0}$ by step 2. Thus, $\mathbf{0} \in L_{\succ}(x_0)$, where $L_{\succ}(x_0)$ is the set $\{x \in \mathbb{R}^n : x_0 \succ x\}$. By Continuity of $\succ$, the set $L_{\succ}(x_0)$ is open in $\mathbb{R}^n$. Further, $L_{\succ}(x_0) \subseteq C$ since $x_0 \succ x \implies x_0 \succsim x$. Thus, $\mathbf{0}$ is in the interior of C.

Step 4: We argue: C is compact. Note, C is a closed subset of $\mathbb{R}^n$ in step 3. It is enough to show that C is bounded. Thus, suppose C is *not* bounded. Thus, there exists a C–valued sequence $\{x_k\}_{k \in \mathbb{N}}$ such that $\|x_k\|_1 > k$ for all $k \in \mathbb{N}$. Define $y_k := x_k/\|x_k\|_1$ for every $k \in \mathbb{N}$. Let K be the set $\{x \in \mathbb{N} : \|x\|_1 = 1\}$. Thus, $y_k \in K$ for any $k \in \mathbb{N}$. Note, K is a closed set as $\|\cdot\|_1$ is a continuous map being a norm. Clearly, K is bounded. Then, by the Heine–Borel Theorem, K is a compact set. Thus, there exists a subsequence $l \in \mathbb{N} \mapsto k_l \in \mathbb{N}$ (that is, $k_l < k_{l+1}$ for all $l \in \mathbb{N}$) and $y_* \in K$ such that $\lim_{l\to\infty}\|y_{k_l} - y_*\|_1 = 0$.

Consider $L := \{\lambda \cdot y_* : \lambda > 0\}$. Fix an arbitrary $x_* \in L$. That is, $x_* = [\lambda_*/(1+\varepsilon_*)] \cdot y_*$ for some $\lambda_* > 0$ and $\varepsilon_* > 0$. Note, $\|\lambda_* \cdot y_*\|_1 = \lambda_*$. Let $l_0 \in \mathbb{N}$ satisfy $k_{l_0} > \lambda_*$. Since $\|x_k\|_1 > k$ for all $k \in \mathbb{N}$ and $l \in \mathbb{N} \mapsto k_l \in \mathbb{N}$ is strictly increasing, it follows from $k_{l_0} > \lambda_*$ that $\|x_{k_l}\|_1 > \lambda_*$ for all $l \geq l_0$. Now, fix any $l \geq l_0$. Define $\lambda_l := \lambda_*/\|x_{k_l}\|_1$. Note that $\lambda_l < 1$ and recall $y_{k_l} = x_{k_l}/\|x_{k_l}\|_1$. By Scale Monotonicity, $x_{k_l} \succ \lambda_l \cdot x_{k_l}$. Thus, $x_{k_l} \succ \lambda_* \cdot y_{k_l}$. Now, $x_0 \succsim x_{k_l}$ as $x_{k_l} \in C$. If $x_0 \succ x_{k_l}$ then $x_0 \succ \lambda_* \cdot y_{k_l}$ by the transitivity of $\succ$. If $x_0 \sim x_{k_l}$ then $x_0 \succ \lambda_* \cdot y_{k_l}$ by the cross transitivity of $\sim$ and $\succ$. That is, $x_0 \succ \lambda_* \cdot y_{k_l}$ for all $l \geq l_0$. Recall, $\lim_{l\to\infty}\|y_{k_l} - y_*\|_1 = 0$. Thus, $\lim_{l\to\infty}\|\lambda_* \cdot y_{k_l} - \lambda_* \cdot y_*\|_1 = 0$. Since $\succ$ satisfies Continuity and $x_0 \succ \lambda_* \cdot y_{k_l}$ for all $l \geq l_0$, from $\lim_{l\to\infty}\|\lambda_* \cdot y_{k_l} - \lambda_* \cdot y_*\|_1 = 0$ we have $x_0 \succsim \lambda_* \cdot y_*$. As $\lambda_* \cdot y_* = (1+\varepsilon_*) \cdot x_*$ and $\varepsilon_* > 0$, $\lambda_* \cdot y_* \succ x_*$ by Scale Monotonicity. Then, $x_0 \succsim \lambda_* \cdot y_*$ and $\lambda_* \cdot y_* \succ x_*$ imply $x_0 \succ x_*$. As $x_* \in L$ is arbitrary, we have:

$$x \in L \implies x_0 \succ x.$$

Fix an arbitrary $x_* \in L$. Thus, $x_* \neq \mathbf{0}$ implying $x_* \succ \mathbf{0}$ by step 2. Also, $x_0 \succ x_*$. By Continuity of $\succ$, there exists $\alpha_* \in (0, 1)$ such that $x_* \sim \alpha_* \cdot x_0 + (1 - \alpha_*) \cdot \mathbf{0} = \alpha_* \cdot x_0$. By Homotheticity of $\succ$, we have $x_0 \sim (1/\alpha_*) \cdot x_*$. Observe, since $x_* \in L$ and $\alpha_* > 0$, it follows that $(1/\alpha_*) \cdot x_* \in L$. Then, $x_0 \succ (1/\alpha_*) \cdot x_*$ which is a contradiction. Thus, C is bounded. Hence, C is compact.

Step 5: We argue: $\|x\|_C \cdot x_0 \sim x$ for any $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. So, fix any $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Since $x_0 \neq \mathbf{0}$, step 2 implies $x_0 \succ \mathbf{0}$. By Continuity of $\succ$, there exists $\varepsilon > 0$ such that:

$$y \in B_{\|\cdot\|_1}(\mathbf{0}, \varepsilon) \implies x_0 \succ y.$$

Let $y_x := (\varepsilon/2\|x\|_1) \cdot x$. Then, $\|y_x\|_1 = \varepsilon/2$. Thus, $y_x \in B_{\|\cdot\|_1}(\mathbf{0}, \varepsilon)$. Hence, $x_0 \succ y_x$. Moreover, $y_x \neq \mathbf{0}$ as $\|y_x\|_1 > 0$. Thus, $y_x \succ \mathbf{0}$ by step 2. Hence, we have $x_0 \succ y_x \succ \mathbf{0}$. By Continuity of $\succ$, there exists $\theta \in (0,1)$ such that $\theta \cdot x_0 = \theta \cdot x_0 + (1-\theta) \cdot \mathbf{0} \sim y_x$. Scale Monotonicity implies that this θ corresponding to y_x is unique. Now, observe that $x = (2\|x\|_1/\varepsilon) \cdot y_x$. Thus, $(2\theta\|x\|_1/\varepsilon) \cdot x_0 \sim x$ by Homotheticity of $\succ$. Define $\alpha_x := 2\theta\|x\|_1/\varepsilon$. Therefore, α_x is the unique element in $\mathbb{R}_{++}$ such that $\alpha_x \cdot x_0 \sim x$. We shall now argue: $\alpha_x = \|x\|_C$.

Recall, $C = \{y \in \mathbb{R}^n : x_0 \succsim y\}$. Thus, $\alpha \cdot C = \{y \in \mathbb{R}^n : \alpha \cdot x_0 \succsim y\}$ for any $\alpha > 0$ because $\succ$ satisfies Homotheticity. Since $\alpha_x \cdot x_0 \sim x$, Scale Monotonicity implies the following:

1. $x \in \alpha \cdot C$ for all $\alpha > \alpha_x$, and

2. $x \notin \alpha \cdot C$ for all $0 < \alpha < \alpha_x$.

Thus, $\inf\{\alpha > 0 : x \in \alpha \cdot C\} = \alpha_x$. Hence, $\alpha_x = \|x\|_C$. Thus, $\alpha_x \cdot x_0 \sim x$ implies $\|x\|_C \cdot x_0 \sim x$ as required.

Step 6: We argue: $\|\cdot\|_C$ is an $\mathcal{N}_*$–representation of $\succ$. Note, C is a compact convex set with $\mathbf{0}$ in its interior by steps 1, 3 and 4. Then, $\|\cdot\|_C$ is a pre–norm by Theorem 2. It remains to argue:

$$x \succ y \iff \|x\|_C > \|y\|_C.$$

Fix any $x, y \in \mathbb{R}^n$. If atleast one of x or y is $\mathbf{0}$ then the above equivalence is trivial. This is because, for any $z \in \mathbb{R}^n$, we have:

1. $\|z\|_C = 0$ iff $z = \mathbf{0}$,

2. $\|z\|_C > 0$ iff $z \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, and

3. $z \succ \mathbf{0}$ if $z \in \mathbb{R}^n \setminus \{\mathbf{0}\}$.

Henceforth, we assume $x, y \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. By step 5, $\|x\|_C \cdot x_0 \sim x$ and $\|y\|_C \cdot x_0 \sim y$. Thus, $x \succ y$ iff $\|x\|_C \cdot x_0 \succ \|y\|_C \cdot x_0$. Further, Scale Montonicity implies: $\|x\|_C \cdot x_0 \succ \|y\|_C \cdot x_0$ iff $\|x\|_C > \|y\|_C$. Hence, we obtain $\big(x \succ y \iff \|x\|_C > \|y\|_C\big)$ as required.

With the proof of existence complete, we now proceed to show the necessity of the axioms. So, assume $\succ$ is a binary relation over $\mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}_+$ is a pre–norm such that:

$$x \succ y \iff f(x) > f(y).$$

Clearly, $\succ$ satisfies asymmetry and negative transitivity. That is, $\succ$ is a weak order over $\mathbb{R}^n$. Moreover, f being a pre–norm is a continuous map. Thus, $\succ$ must satisfy Continuity. It remains to show that $\succ$ satisfies Homotheticity, Convexity and Scale Monotonicity.

That $\succ$ satisfies Homotheticity is an immediate consequence of the fact that f being a pre–norm is a homogenous function of degree one. Further, Scale Monotonicity of $\succ$ follows from (1) $f(x) > 0$ iff $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, and (2) $f(\alpha \cdot x) = \alpha \cdot f(x)$ for all $\alpha > 0$ and $x \in \mathbb{R}^n$. We now show that $\succ$ satisfies Convexity.

Assume $x \succsim y$ and $\alpha \in (0, 1)$. Let $z := \alpha \cdot x + (1 - \alpha) \cdot y$. We must show that $x \succsim z$. Equivalently, we shall argue: $f(x) \geq f(z)$. Note that $x \succsim y$ implies $f(x) \geq f(y)$. Let $u := \alpha \cdot x$ and $v := (1 - \alpha) \cdot y$. Now, f being a pre–norm is a homogenous function of degree one. Thus, $f(u) = \alpha \cdot f(x)$ and $f(v) = (1-\alpha) \cdot f(y)$. Since $\alpha < 1$ and $f(x) \geq f(y)$, we have $(1-\alpha) \cdot f(y) \leq (1-\alpha) \cdot f(x)$. Thus, $f(u) + f(v) \leq f(x)$. Note, $z = u + v$ holds. Since f is a pre–norm, we have: $f(z) \leq f(u) + f(v)$. Then, $f(z) \leq f(x)$ as required. Thus, $\succ$ satisfies Convexity. Hence, the necessity of the axioms has been demonstrated. ■

Proof of Proposition 3: Let $\succ$ be a binary relation over $\mathbb{R}^n$ and suppose $f : \mathbb{R}^n \to \mathbb{R}_+$ is a representation of $\succ$. That is,

$$x \succ y \iff f(x) > f(y).$$

First, assume f is a norm. Then, f is a pre–norm as every norm is a pre–norm by definition. Further, f satisfies $f(-x) = f(x)$ for all $x \in \mathbb{R}^n$. Thus, $x \sim -x$ holds for all $x \in \mathbb{R}^n$. That is, $\succ$ satisfies Reflection Symmetry. Hence, if f is an $\mathcal{N}$–representation of $\succ$ then f is an $\mathcal{N}_*$–representation of $\succ$. Moreover, $\succ$ must satisfy Reflection Symmetry if it admits an $\mathcal{N}$–representation.

Now, assume f is a pre–norm and $\succ$ satisfies Reflection Symmetry. Thus, $f(x) = f(-x)$ for all $x \in \mathbb{R}^n$. Fix any $x \in \mathbb{R}^n$ and $\alpha < 0$. Then, $f(\alpha \cdot x) = f([-\alpha] \cdot x)$. Further, $f([-\alpha] \cdot x) = [-\alpha] \cdot f(x)$ as $\alpha < 0$ and f is a pre–norm. Thus, $f(\alpha \cdot x) = [-\alpha] \cdot f(x)$ if $\alpha < 0$. Of course, $f(\alpha \cdot x) = \alpha \cdot f(x)$ if $\alpha > 0$. That is, $f(\alpha \cdot x) = |\alpha| \cdot f(x)$ for all $\alpha \in \mathbb{R}$. Hence, f is a norm. Thus, if f is an $\mathcal{N}_*$–representation and $\succ$ satisfies Reflection Symmetry, then f is an $\mathcal{N}$–representation. ■

The proofs of the results stated in subsection 3.2 are supplied in this subsection of the Appendix. The organization is as follows. First, we prove Theorem 5. Then, this result and Theorem 1 from subsection 3.1, which asserts the existence of pre–norms as representations, will be used to prove the remaining results of subsection 3.2. Throughout, we shall use Theorem 2 from subsection 3.1 which asserts the connections between pre–norms and compact convex sets that have the origin in their interior. However, we need four geometric results.

LEMMA *A.I.2(a)*: *Suppose K is a non–empty compact subset of $\mathbb{R}^n$. Then, the map $f_K : \mathbb{R}^n \to \mathbb{R}_+$ defined as*:

$$f_K(x) := \max_{y \in K} x \cdot y \quad \textit{for every } x \in \mathbb{R}^n$$

is a convex function which is homogenous of degree one. Additionally, if $\mathbf{0}$ is in the interior of K then f_K is a pre–norm.

PROOF: We note, at the outset, the map f_K is indeed $\mathbb{R}$–valued as K is compact and, for any $x \in \mathbb{R}^n$, the map $y \in \mathbb{R}^n \mapsto x \cdot y \in \mathbb{R}$ is continuous. Further, the map f_K is $\mathbb{R}_+$–valued as $\mathbf{0} \in K$ implying: $f_K(x) \geq x \cdot \mathbf{0} = 0$ for all $x \in \mathbb{R}^n$.

We now show: f_K is a convex function. Let $x_0, x_1 \in \mathbb{R}^n$ and $\alpha \in [0, 1]$ be arbitrary. Let $x_\alpha := \alpha \cdot x_1 + [1 - \alpha] \cdot x_0$. Pick any $y \in K$. Then, $x_\alpha \cdot y = \alpha(x_1 \cdot y) + [1 - \alpha](x_0 \cdot y)$. Now, $x_1 \cdot y \leq f_K(x_1)$ and $x_0 \cdot y \leq f_K(x_0)$ by definition of the map f_K. Thus, we have:

$$x_\alpha \cdot y \leq \alpha \cdot f_K(x_1) + [1 - \alpha] \cdot f_K(x_0) \quad \text{for all } y \in K.$$

Hence, $f_K(x_\alpha) \leq \alpha \cdot f_K(x_1) + [1 - \alpha] \cdot f_K(x_0)$. Since $x_0, x_1 \in \mathbb{R}^n$ and $\alpha \in [0, 1]$ are arbitrary, f_K is a convex function.

Now, fix any $\alpha > 0$ and $x \in \mathbb{R}^n$. Let $x_\alpha := \alpha \cdot x$. Let $y_*, y_{**} \in K$ be such that $f_K(x) = x \cdot y_*$ and $f_K(x_\alpha) = x_\alpha \cdot y_{**}$. Observe, $x \cdot y_* \geq x \cdot y_{**}$ and $x_\alpha \cdot y_{**} \geq x_\alpha \cdot y_*$. Also, note that $x_\alpha \cdot y_{**} \geq x_\alpha \cdot y_*$ is equivalent to $x \cdot y_{**} \geq x \cdot y_*$ because $x_\alpha = \alpha \cdot x$ and $\alpha > 0$. Thus, $x \cdot y_* = x \cdot y_{**}$ where $x_\alpha \cdot y_{**} = \alpha(x \cdot y_{**})$. Hence, $f_K(x_\alpha) = \alpha \cdot f_K(x)$. That is:

$$f_K(\alpha \cdot x) = \alpha \cdot f_K(x) \quad \text{for all } \alpha > 0 \text{ and } x \in \mathbb{R}^n.$$

Hence, f_K is a homogenous function of degree one. To show that f_K is a pre–norm, it remains to verify that f_K satsifies condition 2 and 4 of Definition 1 (section 2). For this, we shall make the additional assumption that $\mathbf{0}$ is in the interior of K.

Of course, $f_K(\mathbf{0}) = 0$ by definition of f_K. Let $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Since $\mathbf{0}$ is in the interior of K, there exists $\varepsilon > 0$ such that $B_{\|\cdot\|_2}(\mathbf{0}, \varepsilon) \subseteq K$. Note, $\|x\|_2 > 0$ as $\|\cdot\|_2$ is a norm and $x \neq \mathbf{0}$. Define $x_\varepsilon := (\varepsilon/2\|x\|_2) \cdot x$. Clearly, $\|x_\varepsilon\|_2 = \varepsilon/2$. Thus, $x_\varepsilon \in B_{\|\cdot\|_2}(\mathbf{0}, \varepsilon)$. Since $B_{\|\cdot\|_2}(\mathbf{0}, \varepsilon) \subseteq K$, it follows $x_\varepsilon \in K$. Hence, $f_K(x) \geq x \cdot x_\varepsilon$. Now, $x \cdot x_\varepsilon = \varepsilon\|x\|_2/2 > 0$. Thus, $f_K(x) > 0$ if $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Thus we have established:

$$f_K(x) = 0 \quad \text{iff} \quad x = \mathbf{0}.$$

That is, condition 2 of Definition 1 has been verified.

Let $x, y \in \mathbb{R}^n$, $\alpha := 1/2$, $\mu := 1/\alpha$ and $x_* := \mu \cdot x$ and $y_* := \mu \cdot y$. By condition 3, $f_K(x_*) = \mu \cdot f_K(x)$ and $f_K(y_*) = \mu \cdot f_K(y)$. Note, $\alpha \cdot \mu = (1 - \alpha) \cdot \mu = 1$. Thus, $\alpha \cdot f_K(x_*) \leq f_K(x)$ and $[1 - \alpha] \cdot f_K(y_*) = f_K(y)$. Also, $\alpha \cdot x_* + [1 - \alpha] \cdot y_* = x + y$. As f_K is convex,

$$f_K(\alpha \cdot x_* + [1 - \alpha] \cdot y_*) \leq \alpha \cdot f_K(x_*) + [1 - \alpha] \cdot f_K(y_*).$$

Thus, $f(x + y) \leq f(x) + f(y)$ for all $x, y \in \mathbb{R}^n$. That is, f_K satisfies condition 4 of Definition 1 as well. Hence, f_K is a pre–norm. ■

LEMMA A.I.2(b): *Let $K \subseteq \mathbb{R}^n$ be compact with $\mathbf{0}$ in the interior of K. Then, for any $\lambda > 0$, the set $L_{K,\lambda}$ defined as:*

$$L_{K,\lambda} := \left\{ x \in \mathbb{R}^n : \max_{y \in K} \; x \cdot y \leq \lambda \right\}$$

is compact and convex with $\mathbf{0}$ in the interior of $L_{K,\lambda}$.

PROOF: Fix any $\lambda > 0$. Define the map $f_K : \mathbb{R}^n \to \mathbb{R}_+$ as follows:

$$f_K(x) := \max_{y \in K} \; x \cdot y \quad \text{for every } x \in \mathbb{R}^n.$$

By Lemma A.I.2(a), f_K is a pre–norm and is a convex function. Let $D_\lambda \subseteq \mathbb{R}$ and $B_\lambda \subseteq \mathbb{R}$ be the intervals $[0, \lambda]$ and $[0, \lambda)$, respectively. Note, $L_{K,\lambda} = f_K^{-1}(D_\lambda)$. Since f_K is a pre–norm, Proposition 6 (subsection 5.1) implies that f_K is continuous. Thus, $f_K^{-1}(D_\lambda)$ is a closed subset of $\mathbb{R}^n$. Further, by the euqivalence of pre–norms according to Proposition 6, it follows that $f_K^{-1}(D_\lambda)$ is bounded. Then, by the Heine–Borel Theorem, $L_{K,\lambda}$ is a compact subset of $\mathbb{R}^n$. Further, $f_K^{-1}(B_\lambda)$ is an open subset of $\mathbb{R}^n$. Clearly, $\mathbf{0} \in f_K^{-1}(B_\lambda) \subseteq f_K^{-1}(D_\lambda)$. Thus, there exists an $\varepsilon > 0$ such that $B_{\|\cdot\|_1}(\mathbf{0}, \varepsilon) \subseteq f_K^{-1}(D_\lambda)$. Hence, $\mathbf{0}$ is in the interior of $L_{K,\lambda}$. The convexity of $L_{K,\lambda}$ is an immediate consequence of the fact that f_K being a convex function is quasi–convex. ■

Lemma *A.I.2(c)*: *Let $K \subseteq \mathbb{R}^n$ be compact and K_* be the closure of the convex hull of K. Then, for any $x \in \mathbb{R}^n$, the following holds*:

$$\max_{y \in K} \; x \cdot y \;=\; \max_{y \in K_*} \; x \cdot y.$$

Proof: Fix an $x \in \mathbb{R}^n$. Let $\theta \coloneqq \max_{y \in K} x \cdot y$ and $\theta_* \coloneqq \max_{y \in K_*} x \cdot y$. Since K_* is the closure of the convex hull of K, it follows that $K \subseteq K_*$. Hence, $\theta \leq \theta_*$ holds. It remains to argue: $\theta \geq \theta_*$.

Since K is compact, it is bounded. Thus, the convex hull of K is bounded. Since the closure of a bounded set must be bounded, K_* is bounded. Moreover, K_* is a closed set. Then, K_* is compact by the Heine–Borel Theorem. Also, the map $y \in \mathbb{R}^n \mapsto x \cdot y \in \mathbb{R}$ is continuous. Thus, there exists $y_* \in K_*$ such that $x \cdot y_* = \theta_*$. Let (y_m) be a sequence in the convex hull of K which converges to y_*. That is, the sequence (y_m) satisfies the following properties:

1. $\lim_{m \to \infty} \|y_* - y_m\|_1 = 0$, and
2. For each $m \in \mathbb{N}$, there exists:
   (a) $J_m \in \mathbb{N}$ (we define $[J_m] \coloneqq \{1, \ldots, J_m\}$),
   (b) $y_{jm} \in K$ for each $j \in [J_m]$, and
   (c) $\langle\, \alpha_{jm} \in \mathbb{R}_+ : j \in [J_m] \,\rangle$ such that $\sum_{j \in [J_m]} \alpha_{jm} = 1$

   such that: $y_m = \sum_{j \in [J_m]} \alpha_{jm} \cdot y_{jm}$ for all $m \in \mathbb{N}$.

Fix an arbitrary $m \in \mathbb{N}$. By 2(b) and the definition of θ, we have: $\theta \geq x \cdot y_{jm}$ for all $j \in [J_m]$. Then, $\theta \geq x \cdot \left(\sum_{j \in [J_m]} \alpha_{jm} \cdot y_{jm} \right)$ as 2(c) holds. Thus, $\theta \geq x \cdot y_m$. Since $m \in \mathbb{N}$ is arbitrary, we have:

$$x \cdot y_m - \theta \leq 0 \quad \text{for every } m \in \mathbb{N}.$$

Because $\lim_{m \to \infty} \|y_* - y_m\|_1 = 0$ (property 1 above) and the map $y \in \mathbb{R}^n \mapsto x \cdot y - \theta \in \mathbb{R}$ is continuous, we obtain: $x \cdot y_* \leq \theta$. That is, $\theta \geq x \cdot y_*$. Recall, $x \cdot y_* = \theta_*$. Thus, $\theta \geq \theta_*$. Since it has already been argued that $\theta \leq \theta_*$, we have: $\theta = \theta_*$. Now, recall that by definition $\theta = \max_{y \in K} x \cdot y$ and $\theta_* = \max_{y \in K_*} x \cdot y$. Therefore, $\theta = \theta_*$ implies that the following equality is true:

$$\max_{y \in K} \; x \cdot y \;=\; \max_{y \in K_*} \; x \cdot y.$$

This completes the proof of the lemma. ■

Lemma *A*.I.2(*d*): *Let $K \subseteq \mathbb{R}^n$ be compact, and K_* be the closure of the convex hull of K. Suppose $x_0 \in K$ satisfies:*

$$\theta \ := \max_{y \in K} \ x_0 \cdot y \geq \max_{y \in K} \ x \cdot y \quad \textit{for every } x \in K.$$

Then, K_* is compact convex with $x_0 \in K_*$ and the following holds:

$$\theta_* := \max_{y \in K_*} \ x_0 \cdot y \geq \max_{y \in K_*} \ x \cdot y \quad \textit{for every } x \in K_*.$$

Moreover, $\theta = \theta_*$ with the common value being $\|x_0\|_2^2$.

Proof: The compactness of K_* follows from the Heine–Borel Theorem. This is because (1) the convex hull of a bounded set is bounded, and (2) the closure of a bounded set is bounded. Thus, K_* is compact as K is compact. Also, $K \subseteq K_*$ and $x_0 \in K$ imply $x_0 \in K_*$. Since the closure of a convex set is convex, it follows that K_* is convex. Consider the map $f_{K_*} : \mathbb{R}^n \to \mathbb{R}$ defined as follows:

$$f_{K_*}(x) := \max_{y \in K_*} \ x \cdot y \quad \text{for every } x \in \mathbb{R}^n.$$

Fix any $x \in K_*$. Since K_* is the closure of the convex hull of K, there exists a K_*–valued sequence (x_m) satisfying:

1. $\lim_{m \to \infty} \|x - x_m\|_1 = 0$, and
2. For each $m \in \mathbb{N}$, there exists:
   (a) $J_m \in \mathbb{N}$ (we define $[J_m] := \{1, \ldots, J_m\}$),
   (b) $x_{jm} \in K$ for each $j \in [J_m]$, and
   (c) $\langle \alpha_{jm} \in \mathbb{R}_+ : j \in [J_m] \rangle$ such that $\sum_{j \in [J_m]} \alpha_{jm} = 1$

   such that: $x_m = \sum_{j \in [J_m]} \alpha_{jm} \cdot x_{jm}$ for all $m \in \mathbb{N}$.

Fix an arbitrary $m \in \mathbb{N}$. Then, $f_{K_*}(x_{jm}) = \max_{y \in K} x_{jm} \cdot y$ for all $j \in [J_m]$ by Lemma A.I.2(c). Thus, $f_{K_*}(x_{jm}) \leq \theta$ for all $j \in [J_m]$. By Lemma A.I.2(a), f_{K_*} is a convex function. Thus, we have:

$$f_{K_*}(x_m) - \theta \leq 0 \quad \text{for every } m \in \mathbb{N}.$$

Further, f_K being a convex function is continuous. Hence, the map $x \in \mathbb{R}^n \mapsto f_{K_*}(x) - \theta \in \mathbb{R}$ is continuous. Thus, $\lim_{m \to \infty} \|x - x_m\|_1 = 0$ implies $f_{K_*}(x) - \theta \leq 0$. Also, $\theta = \theta_*$ by Lemma A.I.2(c). Hence, $\max_{y \in K_*} \ x_0 \cdot y \geq \max_{y \in K_*} \ x \cdot y$. Therefore, to complete the proof, it remains to demonstrate that $\theta = \|x_0\|_2^2$.

Note, $x_0 \in K$ implies $\theta \geq x_0 \cdot x_0 = \|x_0\|_2^2$. Suppose $\theta > \|x_0\|_2^2$. Let $y_* \in K$ be such that $\theta = x_0 \cdot y_*$. Clearly, $\theta > 0$ implies $x_0 \cdot y_* = |x_0 \cdot y_*|$. Thus, $\theta = |x_0 \cdot y_*|$. Then, $\theta > \|x_0\|_2^2$ implies $|x_0 \cdot y_*| > \|x_0\|_2^2$. Also, by the Cauchy–Schwarz Inequality, we have:

$$|x_0 \cdot y_*| \leq \|x_0\|_2 \cdot \|y_*\|_2.$$

Hence, $\|x_0\|_2 \cdot \|y_*\|_2 > \|x_0\|_2^2$ which implies $\|y_*\|_2^2 > \|x_0\|_2 \cdot \|y_*\|_2$. Observe, $y_* \in K$ implies $\theta \geq y_* \cdot y_* = \|y_*\|_2^2$. Then, $\theta = |x_0 \cdot y_*|$ and $\|y_*\|_2^2 > \|x_0\|_2 \cdot \|y_*\|_2$ imply $|x_0 \cdot y_*| > \|x_0\|_2 \cdot \|y_*\|_2$. This contradicts the Cauchy–Schwarz Inequality. Hence, our supposition that $\theta > \|x_0\|_2^2$ must be wrong. Therefore, $\theta = \|x_0\|_2^2$ as required. ■

LEMMA A.I.2(e): *Suppose C is a compact subset of $\mathbb{R}^n$ with $\mathbf{0}$ in the interior of C. Let $x_0 \in \mathbb{R}^n$ be such that:*

$$\max_{y \in C} x_0 \cdot y \geq \max_{y \in C} x \cdot y \quad \text{for every } x \in C.$$

Then, the sets C_* and C_{**} defined as:

$$C_* := \left\{x \in \mathbb{R}^n : \max_{y \in C} x \cdot y \leq \|x_0\|_2^2\right\}, \text{ and}$$
$$C_{**} := \left\{x \in \mathbb{R}^n : \max_{y \in C_*} x \cdot y \leq \|x_0\|_2^2\right\}$$

are compact convex subsets of $\mathbb{R}^n$ with $\mathbf{0}$ in their interiors. Moreover, C_{**} is the closure of the convex hull of C.

PROOF: Let $C \subseteq \mathbb{R}^n$ be compact with $\mathbf{0}$ in the interior of C. Also, let $x_0 \in \mathbb{R}^n$ and $C_*, C_{**} \subseteq \mathbb{R}^n$ be as in the statement of the lemma. By Lemma A.I.2(b), C_* and C_{**} are compact convex with $\mathbf{0}$ in each of their interiors. Therefore, it remains to argue: if $C^\dagger$ is the closure of the convex hull of C then $C_{**} = C^\dagger$.

We begin with the following reduction. Observe, Lemma A.I.2(d) implies that $x_0 \in C^\dagger$ and satisfies the following:

$$\max_{y \in C^\dagger} x_0 \cdot y \geq \max_{y \in C^\dagger} x \cdot y \quad \text{for every } x \in C^\dagger.$$

Further, consider the sets $C_*^\dagger := \{x \in \mathbb{R}^n : \max_{y \in C^\dagger} x \cdot y \leq \|x_0\|_2^2\}$ and $C_{**}^\dagger := \{x \in \mathbb{R}^n : \max_{y \in C_*^\dagger} x \cdot y \leq \|x_0\|_2^2\}$. Then, $C_*^\dagger = C_*$ and $C_{**}^\dagger = C_{**}$ by Lemma A.I.2(c). Thus, the claim of the lemma under consideration will be established if it is proven that $C = C_{**}$ under the additional assumption that C is convex. Henceforth, we assume $C \subseteq \mathbb{R}^n$ to be compact *convex* with $\mathbf{0}$ in its interior. We argue: $C = C_{**}$.

First, we argue: $C \subseteq C_{**}$. Pick an arbitrary $x \in C$. Observe, if $y \in C_*$ then $x \cdot y \leq \|x_0\|_2^2$. This follows from the definition of C_* and that $x \in C$. Thus, $\max_{y \in C_*} x \cdot y \leq \|x_0\|_2^2$. Hence, $x \in C_{**}$ by definition of C_{**}. Since $x \in C$ is arbitrary, we have: $C \subseteq C_{**}$.

Now, we argue: $C_{**} = C$. Suppose $C_{**} \setminus C \neq \varnothing$. Pick any $x_1 \in C_{**}$ such that $x_1 \notin C$. As $\{x_1\}$ and C_* are disjoint and convex compact sets, the Separating Hyperplane Theorem asserts that there exists some $p \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ and $\alpha \in \mathbb{R}$ such that:

1. $p \cdot x_1 > \alpha$, and

2. $p \cdot x < \alpha$ for all $x \in C$.

Since $\mathbf{0}$ is in the interior of the set C, there exists $\varepsilon > 0$ such that $B_{\|\cdot\|_2}(\mathbf{0}, \varepsilon) \subseteq C$. Let $p_\varepsilon := (\varepsilon / 2\|p\|_2) \cdot p$. Clearly, $\|p_\varepsilon\|_2 = \varepsilon/2$. Thus, $p_\varepsilon \in B_{\|\cdot\|_2}(\mathbf{0}, \varepsilon)$ implying $p_\varepsilon \in C$. Note, $p \cdot p_\varepsilon = \varepsilon \|p\|_2 / 2$. As $p \neq \mathbf{0}$, it follows $p \cdot p_\varepsilon > 0$. Let $\theta := \max_{x \in C} p \cdot x$. Then, $\theta < \alpha$ by (2). Also, $p_\varepsilon \in C$ implies $\theta \geq p \cdot p_\varepsilon$. Thus, $\theta > 0$ and $\lambda_* := \|x_0\|_2^2 / \theta > 0$.

Consider $p_* := \lambda_* \cdot p$ and $\alpha_* := \lambda_* \cdot \alpha$. Clearly, $\theta < \alpha$ and $\lambda_* > 0$ imply $\lambda_* \cdot \theta < \alpha_*$. That is, $\|x_0\|_2^2 < \alpha_*$. Also, $\lambda_* > 0$ and (1) imply $p_* \cdot x_1 > \alpha_*$. Thus, $\|x_1\|_2^2 < p_* \cdot x_1$. Since $x_1 \in C_{**}$, it follows: $p_* \notin C_*$. Now, $\max_{x \in C} p_* \cdot x = \lambda_* \cdot \big(\max_{x \in C} p \cdot x\big) = \lambda \cdot \theta$ and $\lambda_* \cdot \theta = \|x_0\|_2^2$ imply $\max_{x \in C} p_* \cdot x = \|x_0\|_2^2$. By definition of C_*, we have: $p_* \in C_*$. However, $p_* \notin C_*$ and $p_* \in C_*$ is a contradiction. Thus, $C_{**} \setminus C = \varnothing$. Observe, $C \subseteq C_{**}$ and $C_{**} \setminus C = \varnothing$ imply $C = C_{**}$. ■

LEMMA *A*.I.2(*f*): *Let $C \subseteq \mathbb{R}^n$ be compact with $\mathbf{0}$ in the interior of C. Also, let $x_0 \in \mathbb{R}^n$ be such that:*

$$\max_{y \in C} \; x_0 \cdot y \;\geq\; \max_{y \in C} \; x \cdot y \qquad \text{for every } x \in C.$$

Let $D_{\|\cdot\|_2}(\mathbf{0}, \|x_0\|_2) := \big\{x \in \mathbb{R}^n : \|x\|_2 \leq \|x_0\|_2\big\}$ and C_* be defined as:

$$C_* \; := \Big\{x \in \mathbb{R}^n : \max_{y \in C} \; x \cdot y \leq \|x_0\|_2^2\Big\}$$

Then, $C = C_*$ if and only if $C = D_{\|\cdot\|_2}(\mathbf{0}, \|x_0\|_2)$.

PROOF: First, we argue: if $C = C_*$ then $C = D_{\|\cdot\|_2}(\mathbf{0}, \|x_0\|_2)$. So, assume $C = C_*$. By Lemma *A*.I.2(*b*), C_* is compact convex with $\mathbf{0}$ in its interior. Then, $C = C_*$ implies C is convex. Let $\theta := \max_{y \in C} x_0 \cdot y$. Then, $\theta = \|x_0\|_2^2$ by Lemma *A*.I.2(*d*).

Let $x \in C$ be arbitrary. Then, $\theta \geq x \cdot y$ for all $y \in C$. In particular, $\theta \geq x \cdot x = \|x\|_2^2$. Thus, $\|x_0\|_2 \geq \|x_0\|_2$. Hence, $x \in D_{\|\cdot\|_2}(\mathbf{0}, \|x_0\|_2)$. Since $x \in C$ is arbitrary, we have: $C \subseteq D_{\|\cdot\|_2}(\mathbf{0}, \|x_0\|_2)$.

We now argue: $D_{\|\cdot\|_2}(\mathbf{0}, \|x_0\|_2) \subseteq C$. Pick an arbitrary $u \in \mathbb{R}^n$ such that $\|u\|_2 = 1$. For any $\lambda \in \mathbb{R}$, let $H_\lambda := \{x \in \mathbb{R}^n : u \cdot x = \lambda\}$. Define $\Lambda := \{\lambda \in \mathbb{R} : H_\lambda \cap C \neq \varnothing\}$. Let $\lambda_* := \sup \Lambda$. Let us first show: $\lambda_* \in \mathbb{R}$. For this, it is enough to argue that Λ is non–empty and bounded above in $\mathbb{R}$. We shall do so by using the compactness of C.

Since $\mathbf{0} \in C$ and $u \cdot \mathbf{0} = 0$, we have $\mathbf{0} \in H_0 \cap C$. Then, $H_0 \cap C \neq \varnothing$ implies $0 \in \Lambda$. Thus, we have: $\Lambda \neq \varnothing$. Suppose Λ is *not* bounded above in $\mathbb{R}$. Thus, get a Λ–valued sequence (λ_k) such that $\lim_{k\to\infty} \lambda_k = +\infty$. Then, the definition of Λ implies, there exists a C–valued sequence (x_k) such that $u \cdot x_k = \lambda_k$ for all $k \in \mathbb{N}$. Since C is compact, there exists $x_* \in C$ and a subsequence $l \in \mathbb{N} \mapsto k_l \in \mathbb{N}$ such that (1) $k_l < k_{l+1}$ for all $l \in \mathbb{N}$, and (2) $\lim_{l\to\infty} \|x_{k_l} - x_*\|_2 = 0$. By continuity of the map $x \in \mathbb{R}^n \mapsto u \cdot x \in \mathbb{R}$, we have $\lim_{l\to\infty} u \cdot x_{k_l} = u \cdot x_*$. However, $u \cdot x_* \in \mathbb{R}$ and $\lim_{l\to\infty} u \cdot x_{k_l} = +\infty$ as (1) $u \cdot x_k = \lambda_k$ for all $k \in \mathbb{N}$, and (2) $\lim_{k\to\infty} \lambda_k = +\infty$. Thus, we have a contradiction. Hence, Λ is bounded above in $\mathbb{R}$. Therefore, we have: $\lambda_* \in \mathbb{R}$.

Now, consider any arbitrary $y \in C$ and let $\lambda_y := u \cdot y$. Clearly, $y \in H_{\lambda_y}$. Thus, $y \in H_{\lambda_y} \cap C$ implying $H_{\lambda_y} \cap C = \varnothing$. Hence, $\lambda_y \in \Lambda$. Then, $\lambda_* = \sup \Lambda$ implies $\lambda_* \geq \lambda_y$. That is, $\lambda_* \geq u \cdot y$. Since $y \in C$ is arbitrary, we have established the following:

$$y \in C \implies u \cdot y \leq \lambda_*.$$

We claim: $\lambda_* \geq \|x_0\|_2$. Suppose $\lambda_* < \|x_0\|_2$. Let $\varepsilon := \|x_0\| - \lambda_*$. Thus, $\varepsilon > 0$ by our supposition. Also, $\lambda_* + \varepsilon = \|x_0\|_2$ by the definition of ε. Let $x_\varepsilon := (\lambda_* + \varepsilon) \cdot u$. Then, $u \cdot x_\varepsilon = (\lambda_* + \varepsilon) \cdot \|u\|_2^2$. Since $\|u\|_2 = 1$, we have $u \cdot x_\varepsilon = \lambda_* + \varepsilon$. Then, $\varepsilon > 0$ implies $u \cdot x_\varepsilon > \lambda_*$. Thus, $x_\varepsilon \notin C$. Now, fix an arbitrary $y \in C$. Then, $x_\varepsilon \cdot y = (\lambda_* + \varepsilon)(u \cdot y)$. Also, $u \cdot y \leq \lambda_*$ as $y \in C$. Thus, $x_\varepsilon \cdot y \leq (\lambda_* + \varepsilon) \cdot \lambda_*$. Since $\lambda_* + \varepsilon = \|x_0\|_2$ and $\lambda_* < \|x_0\|_2$, we obtain $x_\varepsilon \cdot y \leq \|x_0\|_2^2$. Since $y \in C$ is arbitrary, it follows: $\max_{y \in C} x_\varepsilon \cdot y \leq \|x_0\|_2^2$. Thus, $x_\varepsilon \in C_*$ by the definition of C_*. Since $C_* = C$, we have: $x_\varepsilon \in C$. However, we have already concluded that $x_\varepsilon \notin C$. Thus, we have a contradiction implying our supposition that $\lambda_* < \|x_0\|_2$ is wrong. Hence, we have: $\lambda_* \geq \|x_0\|_2$.

We now claim: there exists $x_* \in C$ such that $u \cdot x_* = \lambda_*$. Since λ_* is $\sup \Lambda$, let (λ_k) be a Λ–valued sequence such that $\lim_{k\to\infty} \lambda_k = +\infty$. Thus, there exists a C–valued sequence (x_k) such that: $u \cdot x_k = \lambda_k$ for all $k \in \mathbb{N}$. Since C is compact, there exists $x_* \in C$ and a subsequence $l \in \mathbb{N} \mapsto k_l \in \mathbb{N}$ such that (1) $k_l < k_{l+1}$ for all $l \in \mathbb{N}$, and (2) $\lim_{l\to\infty} \|x_{k_l} - x_*\|_2 = 0$. By continuity of the map $x \in \mathbb{R}^n \mapsto u \cdot x \in \mathbb{R}$, we have $\lim_{l\to\infty} u \cdot x_k = u \cdot x_*$. Since $u \cdot x_k = \lambda_k$ for all $k \in \mathbb{N}$, from $\lim_{k\to\infty} \lambda_k = \lambda_*$ we have $u \cdot x_* = \lambda_*$. Since $x_* \in C$, we have shown: there exists $x_* \in C$ such that $u \cdot x_* = \lambda_*$.

Henceforth, let $x_* \in C$ be such that $u \cdot x_* = \lambda_*$. Then, $\lambda_* \geq \|x_0\|_2$ implies $u \cdot x_* \geq \|x_0\|_2$. Since $x_* \in C$, note that Lemma $A.I.2(d)$ implies $\|x_0\|_2^2 \geq \max_{y \in C} x_* \cdot y$. In particular, $\|x_0\|_2^2 \geq x_* \cdot x_* = \|x_*\|_2^2$ which implies $\|x_0\|_2 \geq \|x_*\|_2$. Then, $u \cdot x_* \geq \|x_*\|_2$ as $\|x_0\|_2 \geq \|x_*\|_2$. Since $\|u\|_2 = 1$, it follows: $u \cdot x_* \geq \|u\|_2 \cdot \|x_*\|_2$. Note, $u \cdot x_* = |u \cdot x_*|$ as $\|\cdot\|_2$ is $\mathbb{R}_+$–valued. Thus, we have: $|u \cdot x_*| \geq \|u\|_2 \cdot \|x_*\|_2$. However, the Cauchy–Schwarz Inequality asserts:

$$|u \cdot x_*| \leq \|u\|_2 \cdot \|x_*\|_2,$$

with equality iff $x_* = \lambda \cdot u$ for some $\lambda \neq 0$. Thus, $|u \cdot x_*| \geq \|u\|_2 \cdot \|x_*\|_2$ implies, there exists $\lambda^\dagger \neq 0$ such that $x_* = \lambda_\dagger \cdot u$. Then, $u \cdot x_* = \lambda^\dagger (u \cdot u) = \lambda^\dagger \|u\|_2^2$. As $\|u\|_2 = 1$, we have $u \cdot x_* = \lambda^\dagger$. Since $u \cdot x_* \geq \|x_0\|_2$, we obtain: $\lambda^\dagger \geq \|x_0\|_2$. As $x_* \in C$ and $x_* = \lambda^\dagger \cdot u$, we have:

$$\big(\, \exists\, \lambda \in \mathbb{R} \,\big) \big[\; \lambda^\dagger \geq \|x_0\|_2 \;;\; \lambda^\dagger \cdot u \in C \;\big].$$

Henceforth, assume $\lambda^\dagger \in \mathbb{R}$ is such that $\lambda^\dagger \geq \|x_0\|_2$ and $\lambda^\dagger \cdot u \in C$. Let $\alpha \coloneqq \|x_0\|_2 / \lambda^\dagger$. Thus, $\alpha \in (0, 1)$. Define $x_\alpha \coloneqq \alpha(\lambda^\dagger \cdot u) + (1 - \alpha) \cdot \mathbf{0}$. Since $\lambda^\dagger \cdot u$ and $\mathbf{0}$ are in C, the convexity of C implies $x_\alpha \in C$. Also, $x_\alpha = \|x_0\|_2 \cdot u$ by definition of α and x_α. Thus, $\|x_0\|_2 \cdot u \in C$. Since $u \in \mathbb{R}^n$ is arbitrary such that $\|u\|_2 = 1$, we have established:

$$\big(\, \forall\, u \in \mathbb{R}^n \,\big) \big[\; \|u\|_2 = 1 \implies \|x_0\|_2 \cdot u \in C \;\big].$$

Now, pick an arbitrary $x \in D_{\|\cdot\|_2}(\mathbf{0}, \|x_0\|_2)$. That is, $\|x\|_2 \leq \|x_0\|_2$. If $\|x\|_2 = 0$ then $x = \mathbf{0}$. Then, $\mathbf{0} \in C$, we have: if $\|x\|_2 = 0$ then $x \in C$. Henceforth, assume $\|x\|_2 > 0$. Let $u \coloneqq x / \|x\|_2$. Clearly, $\|u\|_2 = 1$. Thus, $\|x_0\|_2 \cdot u \in C$. Let $\alpha \coloneqq \|x\|_2 / \|x_0\|_2$. Clearly, $\alpha \in (0, 1)$. Define $x_\alpha \coloneqq \alpha(\|x_0\|_2 \cdot u) + (1 - \alpha) \cdot \mathbf{0}$. Since $\|x_0\|_2 \cdot u$ and $\mathbf{0}$ are in C, the convexity of C implies $x_\alpha \in C$. Also, $x_\alpha = \|x\|_2 \cdot u$ by definition of α and x_α. Thus, $\|x\|_2 \cdot u \in C$. Since $\|x\|_2 \cdot u = x$, we obtain: $x \in C$. That is, $D_{\|\cdot\|_2}(\mathbf{0}, \|x_0\|_2) \subseteq C$. Hence, we have established:

$$\big[\; C = C_* \;\big] \implies \big[\; C = D_{\|\cdot\|_2}(\mathbf{0}, \|x_0\|_2) \;\big].$$

For the converse, assume $C = D_{\|\cdot\|_2}(\mathbf{0}, \|x_0\|_2)$. First, we shall argue: $C_* \subseteq C$. For this, suppose $x \in C_*$ and $x \notin C$. Then, $\|x\|_2 > \|x_0\|_2$. Let $y_x \coloneqq (\|x_0\|_2 / \|x\|_2) \cdot x$. Clearly, $\|y_x\|_2 = \|x_0\|_2$. Thus, $y_x \in C$. Then, $x \cdot y_x \leq \max_{y \in C} x \cdot y$. However, $\max_{y \in C} x \cdot y \leq \|x_0\|_2^2$ by the definition of C_* and that $x \in C_*$. Thus, $x \cdot y_x \leq \|x_0\|_2^2$. Now, note that $x \cdot y_x = (\|x_0\|_2 / \|x\|_2)(x \cdot x)$. That is, $x \cdot y_x = \|x_0\|_2 \cdot \|x\|_2$. Also, $\|x\|_2 > \|x_0\|_2$ implies $\|x_0\|_2 \cdot \|x\|_2 > \|x_0\|_2^2$. Thus, $x \cdot y_x > \|x_0\|_2^2$ which contradicts to $x \cdot y_x \leq \|x_0\|_2^2$. Thus, we have: $C_* \subseteq C$.

It remains to argue: $C \subseteq C_*$. Pick an arbitrary $x \in C$. Since $C = D_{\|\cdot\|_2}(\mathbf{0}, \|x_0\|_2)$, we have $\|x\|_2 \le \|x_0\|_2$. Consider an arbitrary $y \in C$. Again, $\|y\|_2 \le \|x_0\|_2$. Thus, $\|x\|_2 \cdot \|y\|_2 \le \|x_0\|_2^2$. Further, the Cauchy–Schwarz Inequality implies $|x \cdot y| \le \|x\|_2 \cdot \|y\|_2$. Thus, $|x \cdot y| \le \|\|_2^2$. Clearly, $x \cdot y \le |x \cdot y|$ which implies $x \cdot y \le \|x_0\|_2^2$. Since $y \in C$ is arbitrary, we have $\max_{y \in C} x \cdot y \le \|x_0\|_2^2$. By definition of C_*, we obtain: $x \in C_*$. Since $x \in C$ is arbitrary, we have: $C \subseteq C_*$. Thus, the converse has been established. ■

Proof of Theorem 5: A pre–norm f *on* $\mathbb{R}^n$ is said to be *regular* if,

$$\max_{f(x) \le 1} \|x\|_2 = 1.$$

Regularity of a pre–norm on $\mathbb{R}^n$ is a "normalization" requirement. Consider an arbitrary pre–norm f. Let $\alpha \coloneqq \max_{f(x) \le 1} \|x\|_2$ and the map $f_* : \mathbb{R}^n \to \mathbb{R}_+$ be defined as $f_* \coloneqq \alpha \cdot f$. Define $\alpha_* \coloneqq \max_{f_*(x) \le 1} \|x\|_2$. Observe, $\alpha_* = \max_{f(\alpha \cdot x) \le 1} \|x\|_2 = 1$ as f and $\|\cdot\|_2$ are homogenous. Thus, f_* is regular. The rest of the proof is as follows.

Step 1: We argue: if $f \in \mathcal{N}_*$ then $g_f \in \mathcal{N}_*$. So, let f be a pre–norm on $\mathbb{R}^n$. Consider the set $C_f \coloneqq \{x \in \mathbb{R}^n : f(x) \le 1\}$. Theorem 2 implies that C_f is compact with $\mathbf{0}$ in its interior. Then, Lemma A.I.2(a) implies that the map $f_{C_f} : \mathbb{R}^n \to \mathbb{R}_+ \cup \{\infty\}$ defined as:

$$f_{C_f}(x) \coloneqq \max_{y \in C_f} x \cdot y \quad \text{for every } x \in \mathbb{R}^n$$

is a pre–norm over $\mathbb{R}^n$. However, observe that $g_f = f_{C_f}$ by definition of the map g_f and the set C_f. Hence, g_f is a pre–norm over $\mathbb{R}^n$.

Step 2: Suppose f is a *regular* pre–norm on $\mathbb{R}^n$, and consider the set $C_f \coloneqq \{x \in \mathbb{R}^n : f(x) \le 1\}$. Let $x_0 \in C_f$ be such that:

$$\theta_f \coloneqq \max_{y \in C_f} x_0 \cdot y \ge \max_{y \in C_f} x \cdot y \quad \text{for all } x \in C_f.$$

We argue: $\|x_0\|_2 = 1 = \theta_f$. Fix an arbitrary $x \in C_f$. As $\|x\|_2^2 = x \cdot x$, it follows that $\|x\|_2^2 \le \max_{y \in C_f} x \cdot y \le \theta_f$. However, x is an arbitrary element in C_f. Thus, we obtain: $\max_{x \in C_f} \|x\|_2^2 \le \theta_f$. Now, Lemma A.I.2(d) implies that $\theta_f = \|x_0\|_2^2$. Hence, $\max_{x \in C_f} \|x\|_2^2 \le \|x_0\|_2^2$ holds. However, $x_0 \in C_f$ implies $\max_{x \in C_f} \|x\|_2^2 \ge \|x_0\|_2^2$. Therefore, we obtain: $\max_{x \in C_f} \|x\|_2^2 = \|x_0\|_2^2$. Since f is regular, we have: $\max_{x \in C_f} \|x\|_2^2 = 1$. Thus, $\|x_0\|_2 = 1$ which also implies $\theta_f = 1$ as $\theta_f = \|x_0\|_2^2$.

Step 3: We argue: if f is a regular pre–norm, then $[T \circ T](f) = f$. Recall, the map $T : \mathcal{N}_* \to \mathcal{N}_*$ is defined as follows:

$$T(f) \coloneqq g_f \quad \text{for every } f \in \mathcal{N}_*,$$

where $g_f \in \mathcal{N}_*$ satisfies, $g_f(x) \coloneqq \max_{f(y) \le 1} \; x \cdot y$ for all $x \in \mathbb{R}^n$.

Suppose f is a regular pre–norm, and let $C_f \coloneqq \{x \in \mathbb{R}^n : f(x) \le 1\}$. Further, let $x_0 \in C_f$ satisfy the following:

$$\theta_f \coloneqq \max_{y \in C_f} \; x_0 \cdot y \; \ge \; \max_{y \in C_f} \; x \cdot y \quad \text{for all } x \in C_f.$$

Let $f_* \coloneqq g_f$ and $f_{**} \coloneqq g_{f_*}$. Since f is a regular pre–norm, step 2 implies $\|x_0\|_2 = 1 = \theta_f$. Define $C_f^* \coloneqq \{x \in \mathbb{R}^n : \max_{y \in C_f} \; x \cdot y \le \|x_0\|_2^2\}$. Then, the definition of C_f and g_f implies $C_f^* = x \in \mathbb{R}^n : g_f(x) \le \|x_0\|_2^2$. As $g_f = f_*$ and $\|x_0\|_2 = 1$, we have: $C_f^* = \{x \in \mathbb{R}^n : f_*(x) \le 1\}$. Also, define $C_f^{**} \coloneqq \{x \in \mathbb{R}^n : \max_{y \in C_f^*} \; x \cdot y \le \|x_0\|_2^2\}$ and observe:

$$C_f^{**} = \{x \in \mathbb{R}^n : f_{**}(x) \le 1\}.$$

Since f is a pre–norm, the set C_f is compact and convex with $\mathbf{0}$ in its interior. In particular, the convexity of C_f ensures that the convex hull of C_f is the set C_f. Then, Lemma A.I.2(e) implies: $C_f = C_f^{**}$. That is, $\{x \in \mathbb{R}^n : f(x) \le 1\} = \{x \in \mathbb{R}^n : f_{**}(x) \le 1\}$. Define $A_{f,\xi} \coloneqq \{x \in \mathbb{R}^n : f(x) \le \xi\}$ and $A_{f_{**},\xi} \coloneqq \{x \in \mathbb{R}^n : f_{**}(x) \le \xi\}$ for every $\xi > 0$. As f and f_{**} are homogenous, we obtain:

$$A_{f,\xi} = A_{f_{**},\xi} \quad \text{for all } \xi > 0.$$

Hence, $\{x \in \mathbb{R}^n : f(x) = \xi\} = \{x \in \mathbb{R}^n : f_{**}(x) = \xi\}$ for every $\xi > 0$. Thus, $f(x) = f_{**}(x)$ for all $x \in \mathbb{R}^n$. That is, $f_{**} = f$. Observe, $f_{**} = [T \circ T](f)$ by definition. Thus, $[T \circ T](f) = f$.

Step 4: We argue: $T(\alpha \cdot f) = (1/\alpha) \cdot T(f)$ for all $\alpha > 0$ and $f \in \mathcal{N}_*$. Let f be a pre–norm and $\alpha > 0$. Fix an arbitrary $x \in \mathbb{R}^n$. Then, $[T(\alpha \cdot f)](x) = \max_{\alpha \cdot f(y) \le 1} x \cdot y$. Also, f is homogenous of degree one. Further, $y \in \mathbb{R}^n \mapsto x \cdot y \in \mathbb{R}$ is a linear map. Thus, we have:

$$\max_{\alpha \cdot f(y) \le 1} \; x \cdot y = (1/\alpha) \max_{f(y) \le 1} x \cdot y.$$

Thus, $[T(\alpha \cdot f)](x) = (1/\alpha) \max_{f(y) \le 1} x \cdot y$. Since $\max_{f(y) \le 1} x \cdot y = [T(f)](x)$, we have: $[T(\alpha \cdot f)](x) = (1/\alpha) \cdot [T(f)](x) = [(1/\alpha) \cdot T(f)](x)$. Since $x \in \mathbb{R}^n$ is arbitrary, we have: $T(\alpha \cdot f) = (1/\alpha) \cdot T(f)$.

Step 5: We argue: $[T \circ T](f) = f$ for every $f \in \mathcal{N}_*$. Let f be any pre–norm and $\alpha := \max_{f(x) \leq 1} \|x\|_2$. Define $f_* := \alpha \cdot f$. Then, f_* is a pre–norm which is regular. By step 4, $T(f_*) = (1/\alpha) \cdot T(f)$. Define $f_{**} := (1/\alpha) \cdot T(f)$. Thus, $T(f_*) = f_{**}$. Also, let $\alpha_* := 1/\alpha$. Thus, $f_{**} = \alpha_* \cdot T(f)$. By step 4, $T(f_{**}) = (1/\alpha_*) \cdot T(f)$. Since $\alpha_* = 1/\alpha$, we have: $T(f_{**}) = \alpha \cdot [T \circ T](f)$. Moreover, $T(f_{**}) = [T \circ T](f_*)$ because $f_{**} = T(f_*)$. Hence, $\alpha \cdot [T \circ T](f) = [T \circ T](f_*)$. Since f_* is regular, step 3 implies $[T \circ T](f_*) = f_*$. Thus, $\alpha \cdot [T \circ T](f) = f_*$. Recall, $f_* = \alpha \cdot f$. Since $\alpha > 0$, we obtain: $[T \circ T](f) = f$.

Step 6: We argue: if $f \in \mathcal{N}_*$ then, $T(f) = f$ implies f is regular. So, assume f is a pre–norm on $\mathbb{R}^n$ that satisfies $T(f) = f$. Also, let $C_f := \{x \in \mathbb{R}^n : f(x) \leq 1\}$. Further, let $x_0 \in C_f$ satisfy:

$$\theta_f := \max_{y \in C_f} \, x_0 \cdot y \geq \max_{y \in C_f} \, x \cdot y \quad \text{for every } x \in C_f.$$

First, we show: $\max_{x \in C_f} \|x\|_2^2 = \|x_0\|_2^2 = \theta_f$. Consider an arbitrary $x \in C_f$. Since $\|x\|_2^2 = x \cdot x$, we have: $\|x\|_2^2 \leq \max_{y \in C_f} x \cdot y$. Thus, $\|x\|_2^2 \leq \theta_f$. As $x \in C_f$ is arbitrary, it follows: $\max_{x \in C_f} \|x\|_2^2 \leq \theta_f$. Now, C_f is a compact set with $\mathbf{0}$ in its interior because f is a pre–norm. This is due to Theorem 2 (subsection 3.1). Hence, Lemma A.I.2(d) implies: $\theta_f = \|x_0\|_2^2$. However, $\|x_0\|_2^2 \leq \max_{x \in C_f} \|x\|_2^2$ because $x_0 \in C_f$. Thus, $\theta_f \leq \max_{x \in C_f} \|x\|_2^2$. Hence, we obtain: $\max_{x \in C_f} \|x\|_2^2 = \theta_f$.

It remains to argue: $\|x_0\|_2 = 1$. Note, $[T(f)](x_0) = \max_{y \in C_f} x_0 \cdot y$ by definition of the map T. That is, $[T(f)](x_0) = \theta_f$. Since $T(f) = f$ and $\theta_f = \|x_0\|_2^2$, it follows: $f(x_0) = \|x_0\|_2^2$. Now, $x_0 \in C_f$ implies $f(x_0) \leq 1$ by definition of C_f. Thus, $\|x_0\|_2^2 \leq 1$.

Suppose $\|x_0\|_2^2 < 1$. That is, $f(x_0) < 1$. Since $\mathbf{0}$ is in the *interior* of C_f, $\max_{x \in C_f} \|x\|_2^2 = \|x_0\|_2^2$ implies $x_0 \neq \mathbf{0}$. Define $u_0 := x_0/\|x_0\|_2$ and $x_1 := x_0/\|x_0\|_2^2$. Then, $f(x_1) = (1/\|x_0\|_2^2) \cdot f(x_0) = 1$ because f is homogenous map and $f(x_0) = \|x_0\|_2^2$. Thus, $x_1 \in C_f$. Further, $\|x_1\|_2 = 1/\|x_0\|_2$. Since $\|x_0\|_2 < 1$, we have $\|x_1\|_2 > 1 \geq \|x_0\|_2$. Thus, $\|x_1\|_2^2 > \max_{x \in C_f} \|x\|_2^2$. However, x_1 is in C_f resulting in a contradiction. Hence, our supposition. Thus, $\|x_0\|_2 \geq 1$. Then, $\|x_0\|_2 \leq 1$ implies: $\|x_0\|_2 = 1$. Hence, $\max_{x \in C_f} \|x\|_2$. That is, f is regular.

Step 7: We shall argue: if $f \in \mathcal{N}_*$ then, $T(f) = f$ implies $f = \|\cdot\|_2^2$. Let f be a pre–norm on $\mathbb{R}^n$ such that $T(f) = f$. Consider the set $C_f := \{x \in \mathbb{R}^n : f(x) \leq 1\}$, and let $x_0 \in C_f$ satisfy:

$$\theta_f := \max_{y \in C_f} \, x_0 \cdot y \geq \max_{y \in C_f} \, x \cdot y \quad \text{for every } x \in C_f.$$

Now, define $C_f^* := \{x \in \mathbb{R}^n : \max_{y \in C_f} x \cdot y \leq \|x_0\|_2^2\}$ and note that $C_f^* = \{x \in \mathbb{R}^n : [T(f)](x) \leq \|x_0\|_2^2\}$ by definition of the map T. Then, $T(f) = f$ implies $C_f^* = \{x \in \mathbb{R}^n : f(x) \leq \|x_0\|_2^2\}$. Further, $T(f) = f$ and step 6 implies f is regular. As was shown in step 6, this is equivalent to asserting $\|x_0\|_2 = 1$. Thus, $C_f^* = \{x \in \mathbb{R}^n : f(x) \leq 1\}$. That is, $C_f = C_f^*$. By Lemma A.I.2(f), we have: $C_f = D_{\|\cdot\|_2}(\mathbf{0}, \|x_0\|_2)$. Since $\|x_0\|_2 = 1$, we obtain the following:

$$\big\{x \in \mathbb{R}^n : f(x) \leq 1\big\} = \big\{x \in \mathbb{R}^n : \|x\|_2 \leq 1\big\}.$$

Let $A_{f,\xi} := \{x \in \mathbb{R}^n : f(x) \leq \xi\}$ and $A_{\|\cdot\|_2,\xi} := \{x \in \mathbb{R}^n : \|\cdot\|_2(x) \leq \xi\}$ for all $\xi > 0$. As f and $\|\cdot\|_2$ are homogenous, we obtain:

$$A_{f,\xi} = A_{\|\cdot\|_2,\xi} \quad \text{for every } \xi > 0.$$

Thus, $\{x \in \mathbb{R}^n : f(x) = \xi\} = \{x \in \mathbb{R}^n : \|x\|_2 = \xi\}$ for all $\xi > 0$. That is, $f(x) = \|x\|_2$ for all $x \in \mathbb{R}^n$. Hence, $f = \|\cdot\|_2$ as required.

Step 8: We argue: if $f \in \mathcal{N}_*$ then the following inequality holds:

$$x \cdot y \leq f(x) \cdot [T \circ f](y) \quad \text{for all } x, y \in \mathbb{R}^n.$$

Let $x, y \in \mathbb{R}^n$ be arbitrary such that $x \neq \mathbf{0}$. Note, $f(x) > 0$ and let $u := x/f(x)$. Then, $f(u) = 1$ as f is homogenous of degree one. Consider the set $C_f := \{z \in \mathbb{R}^n : f(z) \leq 1\}$. The definition of the map T implies: $[T \circ f](y) = \max_{z \in C_f} y \cdot z$. Note, $u \in C_f$ as $f(u) = 1$. Thus, $\max_{z \in C_f} y \cdot z \geq y \cdot u$. Since $y \cdot u = u \cdot y$, we have: $u \cdot y \leq [T \circ f](y)$. Now, $x = f(x) \cdot u$ and $f(x) > 0$. Thus, we obtain:

$$x \cdot y \leq f(x) \cdot [T \circ f](y) \quad \text{for all } x \in \mathbb{R}^n \setminus \{\mathbf{0}\} \text{ and } y \in \mathbb{R}^n.$$

Thus, the inequality holds if $x \neq \mathbf{0}$. However, when $x = \mathbf{0}$, it holds trivially as then both $x \cdot y$ and $f(x)$ are 0.

This completes the proof of the theorem. ■

Proof of Corollary 1: Suppose f is a norm and $x, y \in \mathbb{R}^n$. Two cases arise. First, suppose $x \cdot y \geq 0$. Then, $|x \cdot y| = x \cdot y$. Since $x \cdot y \leq f(x) \cdot [T \circ f](y)$ by Theorem 5, we have: $|x \cdot y| \leq f(x) \cdot [T \circ f](y)$. Now, suppose $x \cdot y < 0$. Then, $|x \cdot y| = -(x \cdot y) = (-x) \cdot y$. By Theorem 5, $(-x) \cdot y \leq f(-x) \cdot [T \circ f](y)$. This implies $|x \cdot y| \leq f(-x) \cdot [T \circ f](y)$. As f is a norm $f(-x) = f(x)$ which then implies: $|x \cdot y| \leq f(x) \cdot [T \circ f](y)$. Since the cases are exhaustive, the proof is complete. ■

Proof of Proposition 4: Suppose $\succ$ is in $\mathcal{P}$. Thus, there exists a pre–norm g on $\mathbb{R}^n$ such that g represents $\succ$. Recall, $f_{\succ} : \mathbb{R}^n \to \mathbb{R}$ is defined, by the choice of an $x_0 \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, as follows:

$$f_{\succ}(y) \coloneqq \max_{x_0 \succsim x} \; x \cdot y \quad \text{for every } y \in \mathbb{R}^n.$$

Let $C \coloneqq \{x \in \mathbb{R}^n : x_0 \succ x\}$. Since g is a representation of $\succ$, we have $C = \{x \in \mathbb{R}^n : g(x) \leq g(x_0)\}$. Also, $x_0 \neq \mathbf{0}$ implies $g(x_0) > 0$ as g is a pre–norm. Then, the map $g_* : \mathbb{R}^n \to \mathbb{R}$ defined by $g_* \coloneqq g/g(x_0)$ is also pre–norm as g is homogenous. Clearly, g_* represents $\succ$. Observe, $C = \{x \in \mathbb{R}^n : g_*(x) \leq 1\}$. As g_* is a pre–norm, Theorem 2 implies C is compact convex set with $\mathbf{0}$ in its interior. Now, observe:

$$f_{\succ}(x) = \max_{y \in C} \; x \cdot y \quad \text{for every } x \in \mathbb{R}^n.$$

Then, Lemma A.I.2(a) implies that $f_{\succ}$ is a pre–norm. This completes the proof of the proposition. ■

Proof of Theorem 6: Suppose $\succ$ is in $\mathcal{P}$ and $\succ^*$ is its dual. Recall, $f_{\succ} : \mathbb{R}^n \to \mathbb{R}$ represents $\succ^*$, where $f_{\succ}$ is defined as:

$$f_{\succ}(y) \coloneqq \max_{x_0 \succsim x} \; x \cdot y \quad \text{for every } y \in \mathbb{R}^n.$$

Note, $x_0 \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ in the above definition. Now, let f be a pre–norm that represents $\succ$. First, we argue: $T(f)$ represents $\succ^*$.

Let $C_1 \coloneqq \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$, and fix an arbitrary $y \in \mathbb{R}^n$. By definition of the map $f_{\succ}$ and that f is a representation of $\succ^*$, we have: $f_{\succ}(y) = \max_{x \in C_1} x \cdot y$. Moreover, $[T(f)](y) = \max_{x \in C_0} x \cdot y$, where $C_0 \coloneqq \{x \in \mathbb{R}^n : f(x) \leq 1\}$. Since f is a pre–norm, $x_0 \neq \mathbf{0}$ implies $f(x_0) > 0$. Let $\kappa \coloneqq f(x_0)$. Now, being a pre–norm, the map f is homogenous of degree one. Hence, we have: $C_1 = \kappa \cdot C_0$. Further, the map $x \in \mathbb{R}^n \mapsto x \cdot y \in \mathbb{R}$ is linear. Hence, $f_{\succ}(y) = \kappa \cdot [T(f)](y)$. Since $y \in \mathbb{R}^n$ is arbitrary, we obtain: $f_{\succ} = \kappa \cdot T(f)$. Since $f_{\succ}$ represents $\succ^*$, it follows from $\kappa > 0$ that: $T(f)$ represents $\succ^*$.

Now, we shall argue: a pre–norm g represents $\succ^*$, if and only if, $g = \alpha \cdot T(f)$ for a unique $\alpha > 0$. Let g be an arbitrary pre–norm. First, suppose g represents $\succ^*$. Then, Proposition 2 (subsection 3.1) implies that $g = \alpha \cdot T(f)$ for some unique $\alpha > 0$. Thus, we have: if g is an $\mathcal{N}_*$–representation of $\succ^*$ then $g = \alpha \cdot T(f)$ for some unique $\alpha > 0$. Moreover, if $g = \alpha \cdot T(f)$ to being with, then g is a clearly a representation of $\succ^*$. This proves the converse. ■

With Theorems 5 and 6 proven, the proofs of Theorems 3 and 4 follow.

Proof of Theorem 3: Suppose $\succ$ is in $\mathcal{P}$. Let $\succ^*$ and $\succ^{**}$ be the dual and the second dual of $\succ$. Since $\succ$ is in $\mathcal{P}$, there exists a pre–norm f which represents $\succ$. Then, Theorem 6 implies that the pre–norm $T(f)$ represents $\succ^*$. Further, $\succ^{**}$ is the dual of $\succ^*$. Thus, Theorem 6 implies that $T(T(f))$ is a representation of $\succ^{**}$. That is, $[T \circ T](f)$ represents $\succ^{**}$. However, $[T \circ T](f) = f$ by Theorem 5. Hence, f represents both $\succ$ and $\succ^{**}$. Thus, $\succ^{**}$ is equal to $\succ$. ■

Proof of Theorem 4: Let $\succ$ be in $\mathcal{P}$ and $\succ^*$ be its dual. Suppose $\succ^*$ is equal to $\succ$. We must argue: $\|\cdot\|_2$ represents $\succ$. However, we first show: $T(\beta \cdot g) = (1/\beta) \cdot T(g)$ for any pre–norm g and $\beta > 0$.

Let g be a pre–norm and $\beta > 0$. Fix an arbitrary $x \in \mathbb{R}^n$. Then, $[T(\beta \cdot g)](x) = \max_{\beta \cdot g(y) \leq 1} x \cdot y$. Also, g is homogenous of degree one. Further, $y \in \mathbb{R}^n \mapsto x \cdot y \in \mathbb{R}$ is a linear map. Thus, we have:

$$\max_{\beta \cdot g(y) \leq 1} x \cdot y = (1/\beta) \max_{g(y) \leq 1} x \cdot y.$$

Now, g is homogenous of degree one. Then, as $y \in \mathbb{R}^n \mapsto \beta \cdot y \in \mathbb{R}^n$ is a bijection, we have: $[T(\beta \cdot g)](x) = (1/\beta) \max_{g(y) \leq 1} x \cdot y$. Since $\max_{g(y) \leq 1} x \cdot y = [T(g)](x)$, if follows:

$$[T(\beta \cdot g)](x) = (1/\beta) \cdot [T(g)](x) = [(1/\beta) \cdot T(g)](x).$$

As $x \in \mathbb{R}^n$ is arbitrary, we have: $T(\beta \cdot f) = (1/\beta) \cdot T(g)$. Now, we are ready to establish: $\|\cdot\|_2$ represents $\succ$.

Let f be a pre–norm which represents $\succ$. Then, Theorem 6 implies that $T(f)$ represents $\succ^*$. Since $\succ^*$ is equal to $\succ$, it follows that $T(f)$ represents $\succ$. Note, $T(f)$ is a pre–norm. Since both f and $T(f)$ are pre–norms which represent $\succ$, Proposition 2 implies: $T(f) = \alpha \cdot f$ for some $\alpha > 0$. Let $\beta := \alpha^{1/2}$ and $f_\dagger := \beta \cdot f$. Thus, $T(f_\dagger) = (1/\beta) \cdot T(f)$. Then, $T(f) = \alpha \cdot f$ implies $T(f_\dagger) = (\alpha/\beta) \cdot f$. That is, $T(f_\dagger) = \beta \cdot f = f_\dagger$ as $\beta = \alpha^{1/2}$ by definition. Note, $f_\dagger$ is a pre–norm which represents $\succ$ as $f_\dagger = \beta \cdot f$ where $\beta > 0$. Since $f_\dagger$ is a pre–norm such that $T(f_\dagger) = f_\dagger$, Theorem 5 implies: $f_\dagger = \|\cdot\|_2$. Since $f_\dagger$ is a representation of $\succ$, we obtain: $\|\cdot\|_2$ is a representation of $\succ$.

For the converse, assume $\succ$ admits $\|\cdot\|_2$ as a representation. Let $f := \|\cdot\|_2$. Then, Theorem 5 implies $T(f) = f$. That is, $T(f) = \|\cdot\|_2$. Further, Theorem 6 implies that $T(f)$ represents $\succ^*$. Thus, $\|\cdot\|_2$ is a representation of $\succ^*$. Since $\|\cdot\|_2$ represents both $\succ$ and $\succ^*$, it follows that $\succ^*$ equals $\succ$. This completes the proof. ■.

We prove the "existence" claim in Theorem 7 which is as follows: if $\succ$ admits a norm as a representation and satisfies separability, then there exists $\theta \in \mathbb{R}^n_{++}$ and $p \geq 1$ such that $\|\cdot\|_{(\theta,p)}$ represents $\succ$. So, assume $\succ$ admits a norm as representation and satisfies separability. Since $\succ$ admits a norm as a representation, Theorem 1 and Proposition 3 (subsection 3.1) imply that $\succ$ also satisfies:

1. Weak order.
2. Continuity.
3. Homotheticity.
4. Convexity.
5. Scale Monotonicity.
6. Reflection Symmetry.

Also, recall that $N := \{1, \ldots, n\}$. We now proceed to the proof.

Proof of Theorem 7: Since $n \geq 3$ and $\succ$ satisfies separability, Debreu's theorem (see Theorem 5.3 of Fishburn [1970]) asserts the existence of an n–tuple of *continuous* functions $h_1, \ldots, h_n : \mathbb{R} \to \mathbb{R}$ such that the map $u : \mathbb{R}^n \to \mathbb{R}$ defined as:

$$u(x) := \sum_{i=1}^{n} h_i(x_i) \quad \text{for all } x \equiv (x_1, \ldots, x_n) \in \mathbb{R}^n, \tag{3}$$

is a representation of $\succ$. Moreover, Scale Monotonicity implies that $\succ$ is non–trivial. Thus, all such "additive" representations of $\succ$ are unique up to *similar* positive affine transformations. Formally, if there exists maps $h'_1, \ldots, h'_n : \mathbb{R} \to \mathbb{R}$ such that $v : \mathbb{R}^n \to \mathbb{R}$ defined as:

$$v(x) := \sum_{i=1}^{n} h'_i(x_i) \quad \text{for all } x \equiv (x_1, \ldots, x_n) \in \mathbb{R}^n,$$

also represents $\succ$, then there exists $\alpha > 0$ and $\beta_1, \ldots, \beta_n \in \mathbb{R}$ such that:

$$h'_i(x) = \alpha h_i(x) + \beta_i \quad \text{for all } x \in \mathbb{R}^n \text{ and all } i \in N. \tag{4}$$

Note, the choice of $\alpha := 1$ and $\beta_i := h_i(0)$ for all $i \in N$ implies: $h'_i := \alpha h_i + \beta_i$ satisfies $h'_i(0) = 0$ for every $i \in N$. Therefore, we shall henceforth assume: $h_i(0) = 0$ for all $i \in N$.

For any $i \in N$, we argue: $h_i(-\xi) = h_i(\xi)$ for all $\xi \in \mathbb{R}$. Fix an arbitrary $\xi \in \mathbb{R}$. Let $x \equiv (x_1, \ldots, x_n) \in \mathbb{R}^n$ satisfy (a) $x_i \coloneqq \xi$, and (b) $x_j \coloneqq 0$ for all $j \in N \setminus \{i\}$. Then, $u(x) = h_i(\xi) + \sum_{j \in N \setminus \{i\}} h_j(0)$. Similarly, $u(-x) = h_i(-\xi) + \sum_{j \in N \setminus \{i\}} h_j(0)$. Now, $-x \sim x$ as $\succ$ satisfies reflection symmetry. Since u represents $\succ$, $-x \sim x$ implies $u(-x) = u(x)$. Then, $u(-x) = u(x)$ implies $h_i(-\xi) = h_i(\xi)$. Since $\xi \in \mathbb{R}$ is arbitrary, we obtain: $h_i(-\xi) = h_i(\xi)$ for all $\xi \in \mathbb{R}$.

For each $i \in N$, let $f_i : \mathbb{R}_+ \to \mathbb{R}$ be defined as: $f_i(\xi) \coloneqq h_i(\xi)$ for all $\xi \in \mathbb{R}_+$. Observe, $h_i(\xi) = f_i(|\xi|)$ for all $\xi \in \mathbb{R}$. For any $i \in N$, we argue: f_i is increasing. Pick arbitrary $\xi, \eta \in \mathbb{R}_{++}$ such that $\xi > \eta$. Consider $x \equiv (x_1, \ldots, x_n)$ such that (a) $x_i \coloneqq \eta$, and (b) $x_j \coloneqq 0$ for all $j \in N \setminus \{i\}$. Clearly, $x \neq \mathbf{0}$. Define $\alpha \coloneqq \xi/\eta$. Then, $\xi > \eta$ implies $\alpha > 1$. Since $\succ$ satisfies Scale Monotonicity, $x \neq \mathbf{0}$ and $\alpha > 1$ imply $\alpha \cdot x \succ x$. Since u represents $\succ$, we have: $u(\alpha \cdot x) > u(x)$. Observe, $u(\alpha \cdot x) = h_i(\alpha\eta) + \sum_{j \in N \setminus \{i\}} h_j(0)$ and $u(x) = h_i(\eta) + \sum_{j \in N \setminus \{i\}} h_j(0)$. Hence, $u(\alpha \cdot x) > u(x)$ implies: $h_i(\alpha\eta) > h_i(\eta)$. Then, $\alpha = \xi/\eta$ implies: $h_i(\xi) > h_i(\eta)$. Thus, we obtain the following:

$$\xi > \eta > 0 \implies h_i(\xi) > h_i(\eta). \tag{5}$$

We now argue: $h_i(\xi) > h_i(0)$ if $\xi > 0$. Consider $x \equiv (x_1, \ldots, x_n)$ that satisfies (a) $x_i \coloneqq \xi$, and (b) $x_j \coloneqq 0$ for all $j \in N \setminus \{i\}$. Since $\succ$ admits a norm as representation and $x \neq \mathbf{0}$, we have: $x \succ \mathbf{0}$. Then, $u(x) > u(\mathbf{0})$ as u represents $\succ$. Again, $u(x) = h_i(\xi) + \sum_{j \in N \setminus \{i\}} h_j(0)$ and $u(\mathbf{0}) = h_i(0) + \sum_{j \in N \setminus \{i\}} h_j(0)$. Thus, $u(x) > u(\mathbf{0})$ implies $h_i(\xi) > h_i(0)$. With (5), we obtain: $h_i(\xi) > h_i(\eta)$ for all $\xi > \eta \geq 0$. Now, recall that the domain of f_i is $\mathbb{R}_+$ and, for any $\xi \in \mathbb{R}_+$, $f_i(\xi) = h_i(\xi)$ by definition. Thus, for each $i \in N$, f_i is a *continuous* and *increasing* function such that $f_i(0) = 0$. Further, the representation u satisfies:

$$u(x) = \sum_{i=1}^{n} f_i(|x_i|) \quad \text{for all } x \equiv (x_1, \ldots, x_n) \in \mathbb{R}^n. \tag{6}$$

Now, for any $i \in N$, define $g_i : \mathbb{R}_{++} \to \mathbb{R}_{++}$ as: $g_i(\xi) \coloneqq f_i(\xi)/f_i(1)$ for all $\xi \in \mathbb{R}_{++}$. Then, $g_i(\xi\eta) = g_i(\xi)g_i(\eta)$ for all $\xi, \eta \in \mathbb{R}_{++}$. To see why, note $\kappa > 0$ implies $\big(x \succ y \iff \kappa \cdot x \succ \kappa \cdot y\big)$ as $\succ$ satisfies Homotheticity. Thus, the map $v : \mathbb{R}^n \to \mathbb{R}$ defined as:

$$v(x) \coloneqq \sum_{i=1}^{n} f_i(\kappa|x_i|) \quad \text{for all } x \equiv (x_1, \ldots, x_n) \in \mathbb{R}^n,$$

is also a representation of $\succ$. Note, u and v are additive.

Thus, there exists $\alpha > 0$ and $\beta_1, \ldots, \beta_n \in \mathbb{R}$ such that, for every $i \in N$, $f_i(\kappa|\xi|) = \alpha f_i(|\xi|) + \beta_i$ for all $\xi \in \mathbb{R}$. Fix any $i \in N$. Then, $f_i(0) = 0$ implies $\beta_i = 0$. Thus, $f_i(\kappa\xi) = \alpha f_i(\xi)$ for all $\xi \geq 0$. In particular, evaluation at $\xi = 1$ implies $\alpha f_i(1) = f_i(\kappa)$. Thus, we obtain $f_i(\kappa\xi) = f_i(\kappa) f_i(\xi)/f_i(1)$. Then, κ equal to η implies:

$$g_i(\xi\eta) = g_i(\xi) g_i(\eta) \quad \text{for all } \xi, \eta \in \mathbb{R}_{++}. \tag{7}$$

The continuity of f_i implies the continuity of g_i. Moreover, g_i is increasing because f_i is increasing. That is, the map $g_i : \mathbb{R}_{++} \to \mathbb{R}_{++}$ is a continuous, increasing and satisfies (7). Consider the map $\Gamma_i : \mathbb{R} \to \mathbb{R}$ which is defined as follows:

$$\Gamma_i(\mu) := \log\big(g_i[\exp(\mu)]\big) \quad \text{for all } \mu \in \mathbb{R}. \tag{8}$$

Being a composition of continuous maps, Γ_i is continuous. Further, being the composition of increasing maps, Γ_i is increasing. Note that $\xi \in \mathbb{R}_{++} \mapsto \log \xi \in \mathbb{R}$ is a homeomorphism. Thus, observe:

$$\Gamma_i\big(\log \xi\big) = \log\big(g_i(\xi)\big) \quad \text{for all } \xi \in \mathbb{R}_{++}. \tag{9}$$

Now, (7) and (9) imply: $\Gamma_i(\mu + \nu) = \Gamma_i(\mu) + \Gamma_i(\nu)$ for all $\mu, \nu \in \mathbb{R}$. That is, Γ_i is a continuous and increasing map which satisfies the Cauchy functional equation. Then, by Corollary 2 of chapter 4 in Aczél & Dhombres (1989), there exists $\pi_i \in \mathbb{R}$ such that:

$$\Gamma_i(\mu) = \pi_i \mu \quad \text{for every } \mu \in \mathbb{R}. \tag{10}$$

Since Γ_i is increasing, it must be that $\pi_i > 0$. Further, (9) and (10) imply: $g_i(\xi) = \xi^{\pi_i}$ for all $\xi \in \mathbb{R}_{++}$. Define $\theta_i := f_i(1)$. Since $f_i(0) = 0$ and f_i is increasing, we have $f_i(1) > 0$. That is, $\theta_i > 0$. Thus, $f_i(\xi) = \theta_i \xi^{\pi_i}$ for all $\xi \in \mathbb{R}_{++}$. Since $\pi_i > 0$ and $f_i(0) = 0$, it follows that: $f_i(\xi) = \theta_i \xi^{\pi_i}$ for all $\xi \in \mathbb{R}_+$. We now argue: $\pi_i = \pi_j$ for all $i, j \in N$. Observe, the argument to establish (7) involved showing: for any $\kappa > 0$, there exists $\alpha > 0$ such that $\alpha f_i(1) = f_i(\kappa)$ for all $i \in N$. Since $f_i(\xi) = \theta_i \xi^{\pi_i}$ for all $\xi \in \mathbb{R}_{++}$, we obtain: $\pi_i = (1/\kappa) \log \alpha$ for all $i \in N$. Thus, $\pi_i = \pi_j$ for every $i, j \in N$. Now, pick any $i_* \in N$ and let $p := \pi_{i_*}$. Since $\pi_i = \pi_j$ for all $i, j \in N$, (6) implies:

$$u(x) = \sum_{i=1}^{n} \theta_i |x_i|^p \quad \text{for all } x \equiv (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Recall, from definition 2, the map $\|\cdot\|_{(\theta,p)} : \mathbb{R}^n \to \mathbb{R}_+$ is defined as: $\|x\|_{(\theta,p)} := \big(\sum_{i=1}^{n} \theta_i |x_i|^p\big)^{1/p}$ for all $x \in \mathbb{R}^n$.

Now, $p > 0$ implies $\xi \in \mathbb{R}_+ \mapsto \xi^{1/p}$ is increasing. Then, the map $x \in \mathbb{R}^n \mapsto [u(x)]^{1/p}$ represents $\succ$ because u represents $\succ$. Also, note that $[u(x)] = \|x\|_{(\theta,p)}$ for all $x \in \mathbb{R}^n$. Hence, $\|\cdot\|_{(\theta,p)}$ represents $\succ$. Recall, we have already obtained that $\theta_i > 0$ for all $i \in N$. This is because $\theta_i = f_i(1)$ by definition, where $f_i(0) = 0$ and f_i is increasing. Hence, it only remains to argue: $p \geq 1$.

Suppose $p < 1$. Let e_i be the ith standard basis vector of $\mathbb{R}^n$, and $C := \{x \in \mathbb{R}^n : u(x) \leq 1\}$. Define $x^{(i)} := (1/\theta_i^{1/p}) \cdot e_i$ for all $i \in N$. Also, let $x^* := (1/n) \cdot \sum_{i \in N} x^{(i)}$. Note, $u(x^{(i)}) = 1$ for every $i \in N$. Thus, $x^{(i)} \in C$ for each $i \in N$. Further, $u(x^*) = (1/n^p) \sum_{i \in N} 1 = n^{1-p}$. Then, $p < 1$ implies $u(x^*) > 1$. Thus, $x^* \notin C$. Hence, C is *not* convex. However, u represents $\succ$ which satisfies Convexity. Thus, we have a contradiction. Hence, $p \geq 1$ as required. ■

REFERENCES

Aczél, J. & J. Dhombres (1989): *Functional Equations in Several Variables*, New York: Cambridge University Press.

Barberà, S., F. Gul & E. Stacchetti (1993): "Generalized Median Voter Schemes and Committees", *Journal of Economic Theory*, vol. 61, no. 2, pp. 262–289.

Bogomolnaia, A. & J.–F. Laslier (2007): "Euclidean Preferences", *Journal of Mathematical Economics*, vol. 43, pp. 87–98.

Border, K. C. & J. S. Jordan (1983): "Straightforward Elections, Unanimity and Phantom Voters", *The Review of Economic Studies*, vol. 50, no. 1, pp. 153–170.

Chambers, C. P. & F. Echenique (2020): "Spherical Preferences", *Journal of Economic Theory*, vol. 189, 105086.

Chambers, C. P., F. Echenique & E. Shmaya (2014): "The Axiomatic Structure of Empirical Content", *American Economic Review*, vol. 104, No. 8, pp. 2303–2319.

D'Agostino, M. & V. Dardanoni (2009): "What's so special about Euclidean distance?", *Social Choice and Welfare*, vol. 33, pp. 211–233.

Davis, O. A, M. H. DeGroot & M. J. Hinich (1972): "Social Preference Orderings and Majority Rule", *Econometrica*, vol. 40, pp. 147–157.

Debreu, G. (1959): "Topological Methods in Cardinal Utility", in K. J. Arrow, S. Karlin and P. Suppes (editors) *Mathematical Methods in the Social Sciences*, Stanford: Stanford University Press.

Dekel, E. & B. L. Lipman (2010): "How (Not) to Do Decision Theory", *Annual Review of Economics*, vol. 2, pp. 257–282.

Echenique, F. & M. B. Yenmez (2015): "How to Control Controlled School Choice", *American Economic Review*, vol. 105, no. 8, pp. 2679–2694.

Eguia, J. X. (2011): "Foundations of Spatial Preferences", *Journal of Mathematical Economics*, vol. 47, pp. 200–205

Enelow, J. M. & M. Hinich (1982): "Nonspatial Candidate Characteristics and Electoral Competition", *The Journal of Politics*, vol. 44, pp. 115–130.

Enelow, J. M. & M. Hinich (1984): *The Spatial Theory of Voting: An Introduction*, New York: Cambridge University Press.

Enelow, J. M., M. Hinich & N. Mendell (1986): "An Empirical Evaluation of Alternative Spatial Models of Election", *The Journal of Politics*, vol. 48, pp. 675–693.

Enelow, J. M., N. R. Mendell & S. Ramesh (1988): "A Comparison of Two Distance Metrics through Regression", *The Journal of Politics*, vol. 50, no. 4, pp. 1057–1071.

Epstein, L. G. & U. Segal (1992): "Quadratic Social Welfare Functions", *Journal of Political Economy*, vol. 100, pp. 691–712.

Fields, G. S. & E. A. Ok (1996): "The Meaning and Measurement of Income Mobility", *Journal of Economic Theory*, vol. 71, pp. 349–377.

Fishburn, P. C. (1970): *Utility Theory for Decision Making*, New York: John Wiley and Sons, Inc.

Gershkov, A., B. Moldovanu & X. Shi (2019): "Voting on multiple issues: What to put on the ballot?", *Theoretical Economics*, vol. 14, pp. 555–596.

Gershkov, A., B. Moldovanu & X. Shi (2020): "Monotonic Norms and Orthogonal Issues in Multidimensional Voting", *Journal of Economic Theory*, vol. 189, 105103.

Kannai, Y. (1977): "Concavifiability and Constructions of Concave Utility Functions", *Journal of Mathematical Economics*, vol. 4, no. 1, pp. 1–56.

Krantz, D. H., R. D. Luce, P. Suppes & A. Tversky (1971): *Foundations of Measurement*, San Diego: Academic Press.

Machina, M. J. & S. M. Müller (1987): "Moment Preferences and Polynomial Utility", *Economic Letters*, vol. 23, pp. 349–353.

McKelvey, R. D. & R. E. Wendell (1976): "Voting Equilibria in Multidimensional Choice Spaces", *Mathematics of Operations Research*, vol. 1, no. 2, pp. 144–158.

Mitra, T. & E. A. Ok (1996): "The Measurement of Income Mobility: A Partial Ordering Approach", *Economic Theory*, vol. 12, pp. 77–102.

Peters, H., H. van der Stel & T. Storcken (1993): "Pareto Optimality, Anonymity and Strategy–Proofness in Location Problems", *International Journal of Game Theory*, vol. 21, pp. 221–235.

Plott, C. R. (1967): "A Notion of Equilibrium and its Possibility under Majority Rule", *American Economic Review*, vol. 57, pp. 787–806.

Tversky, A., & D. H. Krantz (1970): "The Dimensional Representation and Metric Structure of Similarity Data", *Journal of Mathematical Psychology*, vol. 7, pp. 572–596.

Wendell, R. E. & S. J. Thorson (1974): "Some Generalizations of Social Decisions under Majority Rule", *Econometrica*, vol. 42, no. 5, pp. 893–912.

Zhou, L. (1991): "Impossibility of Strategy–Proof Mechanisms in Economies with Pure Public Goods", *The Review of Economic Studies*, vol. 58, no. 1, pp. 107–119.