



The Kalai–Smorodinsky bargaining solution with loss aversion[☆]

Bram Driesen^{*,1}, Andrés Perea¹, Hans Peters¹

Department of Quantitative Economics, University of Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands

ARTICLE INFO

Article history:

Received 2 November 2009

Received in revised form

14 September 2010

Accepted 18 October 2010

Available online 28 October 2010

JEL classification:

C78

Keywords:

Bargaining

Kalai–Smorodinsky solution

Loss aversion

Reference dependence

ABSTRACT

We consider bargaining problems under the assumption that players are loss averse, i.e., experience disutility from obtaining an outcome lower than some reference point. We follow the approach of Shalev (2002) by imposing the self-supporting condition on an outcome: an outcome z in a bargaining problem is self-supporting under a given bargaining solution, whenever transforming the problem using outcome z as a reference point, yields a transformed problem in which the solution is z .

We show that n -player bargaining problems have a unique self-supporting outcome under the Kalai–Smorodinsky solution. For all possible loss aversion coefficients we determine the bargaining solutions that give exactly these outcomes, and characterize them by the standard axioms of Scale Invariance, Individual Monotonicity, and Strong Individual Rationality, and a new axiom called Proportional Concession Invariance (PCI). A bargaining solution satisfies PCI if moving the utopia point in the direction of the solution outcome does not change this outcome.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

In the bargaining problem as defined by Nash (1950) two players try to find agreement on a set of feasible outcomes. Failure to cooperate results in a disagreement outcome, unfavorable to both. Nash proposed and axiomatized the well known Nash bargaining solution. A wide range of other solutions have been formulated since. One of the most prominent alternatives to the Nash bargaining solution is the Kalai–Smorodinsky solution, defined by Raiffa (1953) and characterized by Kalai and Smorodinsky (1975).

In economics the risk attitude of an agent tends to play an important role in how that agent behaves. In the bargaining literature, much attention has been paid to the influence of risk attitudes, and in particular risk aversion, on the outcomes assigned by specific bargaining solutions. Several studies (Kannai, 1977; Kihlstrom et al., 1981) find that the Nash bargaining solution favors the less risk averse player. Kihlstrom et al. (1981) find similar results for the Kalai–Smorodinsky solution. Köbberling and Peters (2003) also study the effect of risk aversion on the Kalai–Smorodinsky solution, but distinguish between probabilistic

risk aversion and utility risk aversion. They find that it is an advantage to have a more utility risk averse opponent, or a less probabilistically risk averse opponent.

In the present paper we investigate the Kalai–Smorodinsky solution under a related behavioral phenomenon, called loss aversion. Loss aversion was first introduced by Kahneman and Tversky (1979). It is based on the premise that losses with respect to some reference point weigh heavier than gains, and thus that a decision maker's utility function exhibits a relatively sharp decrease below this reference point.

To introduce this concept in the bargaining problem, we follow the approach of Shalev (2002). Each player's preference is represented by a von Neumann–Morgenstern utility function, a nonnegative loss aversion coefficient, and a reference point; if a player's utility level is below his reference point, then he experiences an extra disutility equal to the size of his incurred loss, multiplied by the loss aversion coefficient. This way, incorporating the players' loss aversion is equivalent to applying a particular transformation to the bargaining problem. Reference points are made endogenous by imposing the self-supporting condition. An outcome z is said to be self-supporting under a given solution, whenever transforming the bargaining problem using outcome z as reference point, yields a problem of which the solution is z . We may interpret a player's reference point as the expectation or the aspiration of the utility payoff which that player may realize given a certain bargaining solution. The self-supporting condition then imposes that the bargaining solution assigns to each player exactly the (initially) aspired utility level, without the need to correct it for loss aversion.

[☆] We thank two referees and an associate editor for their valuable comments.

^{*} Corresponding address: Óbuda University, Faculty of Economics, Tavaszmező 15–17, H-1084 Budapest, Hungary. Tel.: +36 666 5208; fax: +36 666 5219.

E-mail addresses: driesen.bram@gmail.com (B. Driesen), A.Perea@maastrichtuniversity.nl (A. Perea), H.Peters@maastrichtuniversity.nl (H. Peters).

¹ Tel.: +31 43 3883835; fax: +31 43 3884874.

In this paper we show that n -player bargaining problems, $n \geq 2$, have exactly one outcome that is self-supporting under the Kalai–Smorodinsky solution². Kalai and Smorodinsky (1975) defined their solution on two-player bargaining problems. We consider the set of all n -player bargaining problems defined by Peters and Tijs (1984), and on this set we define a class of asymmetric n -person Kalai–Smorodinsky solutions, as follows. Consider a Pareto optimal outcome and the line segments connecting that outcome to the disagreement point and to the utopia point. For any pair of players we may then project these line segments into the plane. Our solution is defined by the unique Pareto optimal outcome such that for any two players, the slopes of these projected line segments satisfy a given proportion which depends on the loss aversion coefficients. We show that for each loss aversion profile, there is a single bargaining solution in our class that yields the associated self-supporting outcome. This implies uniqueness of a self-supporting outcome under the Kalai–Smorodinsky solution for n -player bargaining games. Moreover, the bargaining solutions we define make it easy to find this outcome.

We next provide a characterization of this class of bargaining solutions, by the standard axioms of Strong Individual Rationality, Scale Invariance, and Individual Monotonicity, and a new axiom called Proportional Concession Invariance: this axiom says that if players make concessions with respect to their utopia values in such a way that the new utopia point is on the line segment connecting the solution outcome and the original utopia point, then the solution outcome is left unchanged.

The paper proceeds as follows. After some preliminaries in Section 2 we show in Section 3 how loss aversion is incorporated into the bargaining problem. Section 4 describes the concept of monotonic curves and the associated bargaining solutions, and defines the Kalai–Smorodinsky solution as a special case. In Section 5 we describe for each loss aversion profile the bargaining solution that gives exactly the self-supporting outcome under the Kalai–Smorodinsky solution. Section 6 contains the axiomatic characterization of this class of bargaining solutions. Section 7 concludes. All proofs are relegated to the Appendix.

2. Preliminaries

The set of players is denoted by $N = \{1, \dots, n\}$, with $n \geq 2$. For $x, y \in \mathbb{R}^N$ we write $x \geq y$ if $x_i \geq y_i$ for all $i \in N$, and $x > y$ if $x_i > y_i$ for all $i \in N$. The relations ' \leq ' and ' $<$ ' are defined similarly. We define $\mathbb{R}_+^N = \{z \in \mathbb{R}^N \mid z \geq 0\}$ and $\mathbb{R}_{++}^N = \{z \in \mathbb{R}^N \mid z > 0\}$. For $x, y \in \mathbb{R}^N$ we denote by xy the vector (x_1y_1, \dots, x_ny_n) , and for $S \subseteq \mathbb{R}^N$ we define $xS = \{xy \in \mathbb{R}^N \mid y \in S\}$. Similarly, we denote $(x_1 + y_1, \dots, x_n + y_n)$ as $x + y$, and the set $\{x + z \mid z \in S\}$ as $x + S$. The vector in \mathbb{R}^N that has the i -th coordinate equal to 1 and all other coordinates equal to 0 is denoted e^i . The vector e^M , $M \subseteq N$ and $M \neq \emptyset$, has all coordinates $i \in M$ equal to 1, and all coordinates $i \notin M$ equal to 0. For $x \in \mathbb{R}_{++}^N$, we denote $(\frac{1}{x_1}, \dots, \frac{1}{x_n})$ as x^{-1} . A set $S \subseteq \mathbb{R}^N$ is *comprehensive* if $x \in S$ and $x \geq y$ imply $y \in S$, for all $x, y \in \mathbb{R}^N$.

A bargaining problem for N is a pair (S, d) where

- $S \subseteq \mathbb{R}^N$ is non-empty, closed, convex, and comprehensive,
- $d \in S$,
- there exists a $z \in S$ such that $z > d$, and
- $S_d := \{z \in S \mid z \geq d\}$ is bounded.

The assumption of comprehensiveness can be interpreted as free disposal of utility in the sense that any player can choose a lower

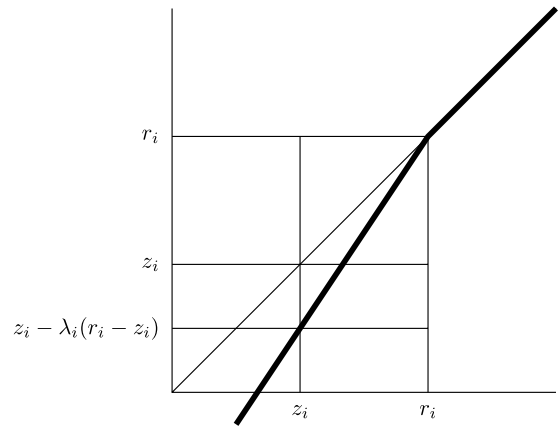


Fig. 1. The thick curve is (part of) the graph of the transformation w_i . In this diagram, $\lambda_i = 0.5$ and $z_i < r_i$.

utility payoff without this leading to an infeasible outcome. Players seek agreement on an outcome z in S , yielding utility z_i to player i . In case no agreement is reached the *disagreement outcome* d results.

The set of all bargaining problems is denoted by B^N . For $(S, d) \in B^N$ and each $i \in N$, we define

$$u_i(S, d) = \max\{z_i \mid z \in S_d\}.$$

This represents the highest possible utility payoff player i can attain in the bargaining problem (S, d) , given that no player $j \in N, j \neq i$, obtains a utility payoff lower than d_j . The vector $u(S, d) = (u_1(S, d), \dots, u_n(S, d))$ is called the *utopia point* of (S, d) . For all $(S, d) \in B^N$ we define the *Pareto set* of S as

$$P(S) = \{z \in S \mid \text{for all } x \in \mathbb{R}^N, \text{ if } x \geq z \text{ and } x \neq z, \text{ then } x \notin S\}.$$

A *bargaining solution* or in short, a *solution*, is a map $\varphi : B^N \rightarrow \mathbb{R}^N$ that assigns to each bargaining problem $(S, d) \in B^N$ a single point $\varphi(S, d) \in S$.

3. Bargaining with loss aversion

Shalev (2000, 2002) introduced a transformation that models loss aversion of players. Each player $i \in N$ has a non-negative *loss aversion coefficient* λ_i and a *reference point* r_i . We denote the vector $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}_+^N$ by λ . Similarly, $r = (r_1, \dots, r_n) \in \mathbb{R}^N$. Each player i evaluates a utility payoff $z_i \in \mathbb{R}$ by the transformation w_i , defined as

$$w_i(z_i, \lambda_i, r_i) = \begin{cases} z_i & \text{if } z_i \geq r_i \\ z_i - \lambda_i(r_i - z_i) & \text{if } z_i < r_i. \end{cases}$$

Thus, a player i who incurs a loss, that is, obtains a utility payoff z_i below his reference point r_i , experiences a disutility that is equal to his loss $r_i - z_i$, multiplied by the loss aversion coefficient λ_i . Payoffs above the reference point are left unchanged. See Fig. 1 for an illustration, and see Remark 3.3 at the end of this section for a short further discussion of this model of loss aversion.

For utility outcomes $z \in \mathbb{R}^N$ we write

$$w(z, \lambda, r) = (w_1(z_1, \lambda_1, r_1), \dots, w_n(z_n, \lambda_n, r_n)).$$

For sets $T \subseteq \mathbb{R}^N$ we write $w(T, \lambda, r) = \{w(z, \lambda, r) \mid z \in T\}$. Henceforth, the transformation $w : \mathbb{R}^N \times \mathbb{R}_+^N \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ is referred to as the *Shalev transformation*.

For bargaining problems $(S, d) \in B^N$ we write

$$w((S, d), \lambda, r) = (w(S, \lambda, r), w(d, \lambda, r)).$$

The following lemma shows that the set B^N is closed under the Shalev transformation.

Lemma 3.1. *Let $(S, d) \in B^N$, $\lambda \in \mathbb{R}_+^N$, and $r \in \mathbb{R}^N$. Then $w((S, d), \lambda, r) \in B^N$.*

² This was already noticed by Shalev (2002), without proof, for the case of two-player bargaining problems.

Consider a bargaining problem $(S, d) \in B^N$ and a bargaining solution φ . If the players agree on φ as the solution to be used, but they are also loss averse in varying degrees λ_i ($i \in N$), then they should agree on a point $z \in \mathbb{R}^N$ which results from applying φ to the Shalev-transformed problem with z determining the reference points of the players. This gives rise to the following definition. For given λ and (S, d) , a point $z \in S$ is a *self-supporting outcome under φ* if

$$z = \varphi(w((S, d), \lambda, z)).$$

Hence, with a self-supporting outcome under φ determining the reference points of the players, they reach an agreement according to φ in the bargaining problem in which their loss aversion has been incorporated. Put somewhat differently, the players expect to obtain utility payoffs as described by the solution φ , and so these expectations determine their reference points: each player experiences anything lower as a loss. Thus, the final agreement is in general not given by $\varphi(S, d)$ – since $\varphi(S, d)$ is not necessarily equal to $\varphi(w((S, d), \lambda, \varphi(S, d)))$ – but by the outcome z in S such that $\varphi(w((S, d), \lambda, z)) = z$.

Remark 3.2. The notion of a self-supporting outcome was introduced in Shalev (2002). It is analogous to Shalev's (2000) notion of loss aversion equilibrium. A pair of (mixed) strategies in a bimatrix game is a loss aversion equilibrium if the following holds: there is a pair (z_1, z_2) of payoffs such that, in the Shalev-transformed bimatrix game with z_1 and z_2 as reference outcomes, this pair of strategies is a Nash equilibrium resulting in (z_1, z_2) as (expected) payoffs. The role of the bargaining solution φ is, thus, played by the solution concept of Nash equilibrium. The notion of loss aversion equilibrium is, in turn, analogous to the notion of personal equilibrium in Kőszegi and Rabin (2006); the latter notion is the individual decision making version of loss aversion equilibrium.

For a fixed loss aversion profile λ and bargaining problem (S, d) , we define the set of all self-supporting outcomes under the bargaining solution φ as

$$\text{Self}^\varphi((S, d), \lambda) = \{z \in S \mid z = \varphi(w((S, d), \lambda, z))\}.$$

Clearly, $\text{Self}^\varphi : B^N \times \mathbb{R}_+^N \rightarrow \mathbb{R}^N$ is a correspondence that assigns to each bargaining problem $(S, d) \in B^N$ and loss aversion profile $\lambda \geq 0$ a (possibly empty) subset of S .

For two-player bargaining games (S, d) , Shalev (2002) characterized the set of self-supporting outcomes under the Nash bargaining solution. Specifically, he showed that it is a closed, connected subset of the Pareto set $P(S)$.

In this paper, starting in the next section, we focus on the Kalai–Smorodinsky solution. We conclude the present section with a remark on a preference foundation for the Shalev loss aversion transformation.

Remark 3.3. A preference foundation for the loss aversion model proposed by Shalev (2000, 2002) is provided in Peters (2010). As a matter of fact, it is a quite common approach in theories of loss aversion to multiply utility losses by a constant loss aversion coefficient and subtract this from a basic utility function. See, for instance, Köbberling and Wakker (2005) and the references therein. Shalev extends this approach by assuming, conveniently, that the same loss aversion coefficient is used for different reference outcomes. Thus, the loss aversion coefficient is constant in two different respects: for losses compared to a fixed reference outcome, and across different reference outcomes. In the preference model of Peters (2010) an agent's preference is a triadic relation (as for instance in Sugden, 2003), involving triples consisting of two outcomes and a reference outcome. The imposed axioms make sure that the decision maker is an expected utility maximizer where the expected utility function is determined by a basic utility function and a reference point. For instance, the set of outcomes can be the real line and all finite lotteries on it, a case that may underly the bargaining model in this paper.

4. The Kalai–Smorodinsky solution

Raiffa (1953) and Kalai and Smorodinsky (1975) defined and characterized the Kalai–Smorodinsky solution (KS) for bargaining problems in $B^{(1,2)}$. Roth (1979) observed that the n -player extension of the KS solution is not Pareto optimal on all bargaining problems in B^N , i.e., does not assign an element of $P(S)$ to each $(S, d) \in B^N$. Therefore, Peters and Tijs (1984) introduced a subclass of bargaining problems in B^N for which this problem does not occur. Consider the following condition on a problem $(S, d) \in B^N$:

For all $x \in S$, $x \geq d$, $i \in N$:

$$x \notin P(S) \text{ and } x_i < u_i(S, d) \Rightarrow \exists \varepsilon > 0 \text{ with } x + \varepsilon e^i \in S. \quad (1)$$

This condition says that if a feasible outcome x is not Pareto optimal, then for any player i who receives less than his utopia payoff it is possible to increase his utility while all other players j still receive x_j . Let $I^N \subseteq B^N$ consist of all bargaining problems satisfying (1). The class of bargaining problems $(S, 0) \in I^N$ is denoted by I_0^N ; for bargaining problems in I_0^N we henceforth omit the disagreement point, i.e., we denote $(S, 0) \in I_0^N$ by S .

Peters and Tijs (1984) defined the n -player extension of the KS solution by making use of monotonic curves. A *monotonic curve* for N is a map

$$\vartheta : [1, n] \rightarrow \left\{ x \in \mathbb{R}_+^n \mid x_i \leq 1 \text{ for all } i \in N, \text{ and } 1 \leq \sum_{i \in N} x_i \right\}$$

such that for all $1 \leq s \leq t \leq n$ we have $\vartheta(s) \leq \vartheta(t)$ and $\sum_{i \in N} \vartheta_i(s) = s$. The set of all monotonic curves for N is denoted by Θ^N .

Lemma 4.1 (Peters and Tijs, 1984). For each $\vartheta \in \Theta^N$ and $S \in I_0^N$ with $u(S) = e^N$, the set

$$P(S) \cap \{\vartheta(t) \mid t \in [1, n]\}$$

contains exactly one point.

Let ϑ be some monotonic curve in Θ^N . In view of Lemma 4.1 we can define $\rho^\vartheta : I^N \rightarrow \mathbb{R}^N$, the solution associated with ϑ . Let $S \in I_0^N$; if $u(S) = e^N$, then

$$\{\rho^\vartheta(S)\} := P(S) \cap \{\vartheta(t) \mid t \in [1, n]\},$$

and if $u(S) = u$, then $\rho^\vartheta(S) := u\rho^\vartheta(u^{-1}S)$. For $(S, d) \in I^N$, we define $\rho^\vartheta(S, d) = d + \rho^\vartheta(S - d)$. The class of all solutions associated with a monotonic curve in Θ^N is referred to as the class of *individually monotonic bargaining solutions*. The KS solution is an element of this class, namely the solution $\rho^{\hat{\vartheta}}$, where $\hat{\vartheta}(t) = \frac{te^N}{n}$ for all $t \in [1, n]$. Observe that $\hat{\vartheta}$ defines a straight line in \mathbb{R}^N , which for bargaining games $S \in I_0^N$ with $u(S) = e^N$, coincides with the line connecting the disagreement point 0 and the utopia point e^N . For general bargaining problems $(S, d) \in I^N$, the KS solution is the intersection of the Pareto set $P(S)$ and the straight line that connects the disagreement point d and the utopia point $u(S, d)$. We also write KS instead of $\rho^{\hat{\vartheta}}$.

5. The solution class \mathcal{D}^N

In this section we show that self-supporting outcomes under the KS solution are well-defined, and that each game in I^N has exactly one such outcome. Peters and Tijs (1984) show that $I^{(1,2)} = B^{(1,2)}$, which implies that our result generalizes Shalev's (2002) remark about the uniqueness of a self-supporting outcome under the KS solution for two-player bargaining problems. Furthermore, we introduce a class \mathcal{D}^N of bargaining solutions on I^N , such that

for any $\lambda \in \mathbb{R}_+^N$ there is a unique $\varphi \in \mathcal{D}^N$ such that $\varphi(S, d)$ is the unique self-supporting outcome of the problem (S, d) under the KS solution.³

From Lemma 3.1 and the fact that the Shalev transformation preserves the ordering of payoffs, we obtain that $(S, d) \in I^N$ implies $w((S, d), \lambda, r) \in I^N$ for all $\lambda \in \mathbb{R}_+^N$ and $r \in \mathbb{R}^N$. Therefore, $\text{Self}^{\text{KS}}((S, d), \lambda)$, the set of self-supporting outcomes under the KS solution, is well-defined.

We now introduce the class \mathcal{D}^N of bargaining solutions. Let $\bar{N} = N \setminus \{n\}$, and define the correspondence $D^k : I^N \rightarrow \mathbb{R}^N$ for all $k \in \mathbb{R}_{++}^{\bar{N}}$ and $(S, d) \in I^N$ by

$$D^k(S, d) = \{z \in P(S) \mid z \geq d \text{ and for all } i \in \bar{N}: (u_n(S, d) - z_n) \times (z_i - d_i) = k_i(u_i(S, d) - z_i)(z_n - d_n)\}. \quad (2)$$

It is not hard to verify that $D^k \neq D^{k'}$ whenever $k, k' \in \mathbb{R}_{++}^{\bar{N}}$ with $k \neq k'$. Then we define $\mathcal{D}^N = \{D^k \mid k \in \mathbb{R}_{++}^{\bar{N}}\}$. For $k \in \mathbb{R}_{++}^{\bar{N}}$, define

$$G^k = \{z \in \mathbb{R}_+^N \mid (1 - z_n)z_i = k_i(1 - z_i)z_n \text{ for all } i \in \bar{N}\},$$

and for $t \in [1, n]$, we define

$$\vartheta^k(t) = \left\{ z \in G^k \mid \sum_{i=1}^n z_i = t \right\}. \quad (3)$$

In Lemma A.1 in the Appendix, we show that each ϑ^k is a monotonic curve in Θ^N .

Theorem 5.1. $D^k(S, d) = \{\rho^{\vartheta^k}(S, d)\}$ for all $k \in \mathbb{R}_{++}^{\bar{N}}$ and $(S, d) \in I^N$.

It follows from Theorem 5.1 that the set $\mathcal{D}^N := \{D^k \mid k \in \mathbb{R}_{++}^{\bar{N}}\}$ is a subset of the class of individually monotonic bargaining solutions. The following theorem shows that the solutions in \mathcal{D}^N provide exactly the outcomes that are self-supporting under the Kalai–Smorodinsky solution.

Theorem 5.2. For all $(S, d) \in I^N$ we have

$$\text{Self}^{\text{KS}}((S, d), \lambda) = D^k(S, d)$$

$$\text{where } k = \left(\frac{1+\lambda_n}{1+\lambda_1}, \dots, \frac{1+\lambda_n}{1+\lambda_{n-1}} \right).$$

We henceforth write $D^k(S, d) = z$ if $D^k(S, d) = \{z\}$ and, thus, regard D^k as a bargaining solution rather than a correspondence. From Theorem 5.2 it follows that for each loss aversion profile $\lambda \in \mathbb{R}_+^N$, we may look at $\text{Self}^{\text{KS}}(\cdot, \lambda)$ as an asymmetric n -player Kalai–Smorodinsky solution where the asymmetry is fully determined by the players' degrees of loss aversion. In the following section we provide an axiomatic characterization of these solutions.

6. An axiomatic characterization of \mathcal{D}^N

From Theorem 5.1 it follows that \mathcal{D}^N is a subclass of the individual monotonic bargaining solutions, defined and characterized by Peters and Tijs (1984). Of their axioms we retain *Scale Invariance* and *Individual Monotonicity*.

(SI) $\varphi : B^N \rightarrow \mathbb{R}^N$ satisfies *Scale Invariance* if $t(\varphi(S, d)) = \varphi(t(S), t(d))$, where $t : \mathbb{R}^N \rightarrow \mathbb{R}^N$ is a linear transformation $t(x) := \alpha + \beta x$, with $\alpha \in \mathbb{R}^N$, $\beta \in \mathbb{R}_{++}^{\bar{N}}$, and $t(S) := \alpha + \beta S$ for $S \subseteq \mathbb{R}^N$.

(IM) $\varphi : B^N \rightarrow \mathbb{R}^N$ satisfies *Individual Monotonicity* if $\varphi_i(S, d) \leq \varphi_i(T, d)$ for all $(S, d), (T, d) \in B^N$ and $i \in N$ with $S \subseteq T$ and $u_j(S) = u_j(T)$ for all $j \in N \setminus \{i\}$.

The axiom SI is consistent with the premise that players' preferences are representable by von Neumann–Morgenstern utility functions: this assumption is still valid under loss aversion, cf. Remark 3.3. Kalai and Smorodinsky (1975) introduced IM as an alternative to Nash's (1950) Independence of Irrelevant Alternatives (IIA). A third well known axiom we impose is the following.

(SIR) $\varphi : B^N \rightarrow \mathbb{R}^N$ satisfies *Strong Individual Rationality* if $\varphi(S, d) > d$ for all $(S, d) \in B^N$.

We next introduce a new axiom called *Proportional Concession Invariance* (PCI). One can regard a solution outcome $\varphi(S, d)$ as expressing the concessions that players make with respect to the utopia point $u(S, d)$. The PCI axiom says that if we replace $u(S, d)$ by a point \hat{u} on the line segment connecting $\varphi(S, d)$ and $u(S, d)$, and shrink the bargaining set accordingly, then the solution outcome should not change. Put differently, if the players' utopia values are reduced in such a way that their concessions with respect to the original solution outcome change proportionally, then this solution outcome is left unchanged. Formally,

(PCI) $\varphi : B^N \rightarrow \mathbb{R}^N$ satisfies *Proportional Concession Invariance* if for each bargaining problem $(S, d) \in B^N$ with solution $\varphi(S, d)$, and each bargaining problem (\hat{S}, d) with

$$\hat{S} := \{z \in S \mid z \leq \hat{u}\},$$

where $\hat{u} = \alpha\varphi(S, d) + (1 - \alpha)u(S, d)$ for some $\alpha \in [0, 1]$, we have $\varphi(\hat{S}, d) = \varphi(S, d)$.

See Fig. 2 for an illustration of PCI. This property can also be seen as a very weak form of IIA, and is therefore satisfied by the Nash bargaining solution and its asymmetric variants. Furthermore, PCI is the counterpart to Disagreement Point Convexity (DPC), one of the axioms used by Peters and van Damme (1991) to characterize the class of asymmetric Nash bargaining solutions. This property requires that the solution outcome $\varphi(S, d)$ remains unchanged if we replace d by a point \hat{d} on the line segment connecting the disagreement point d and the solution outcome $\varphi(S, d)$.

The characterization result is as follows.

Theorem 6.1. Let $\varphi : I^N \rightarrow \mathbb{R}^N$ be a bargaining solution. Then $\varphi \in \mathcal{D}^N$ if and only if it satisfies SIR, SI, IM, and PCI.

Note that Theorem 6.1 does not use Pareto optimality. In the Appendix, we use an argument similar as in Roth (1977) to show that Pareto optimality is implied by our axioms.⁴

7. Concluding remarks

We have established that bargaining problems with loss averse players have exactly one self-supporting outcome under the Kalai–Smorodinsky solution. Moreover, we have determined the asymmetric n -player versions of the KS solution which directly capture the asymmetry resulting from the players' degrees of loss aversion. We summarize this in the following corollary.

Corollary 7.1. Let $\lambda \in \mathbb{R}_+^N$ be the profile of loss aversion. For every $(S, d) \in I^N$, an outcome $z \in S$ is self-supporting under the KS solution if and only if $z = D^k(S, d)$, where $k_i := \frac{1+\lambda_n}{1+\lambda_i}$ for all $i \in N$.

We have characterized the class \mathcal{D}^N of all bargaining solutions D^k by Strong Individual Rationality, Scale Invariance, Individual Monotonicity, and Proportional Concession Invariance. While the

³ Unless of course $S = \{x \in \mathbb{R}^N \mid x \leq u(S, d)\}$, in which case all Pareto optimal solutions assign $u(S, d)$.

⁴ Although formally not required, the class I^N of bargaining problems is not only closed under taking loss aversion transformations but also does not shrink: for zero loss aversion coefficients the loss aversion transformation is just identity.

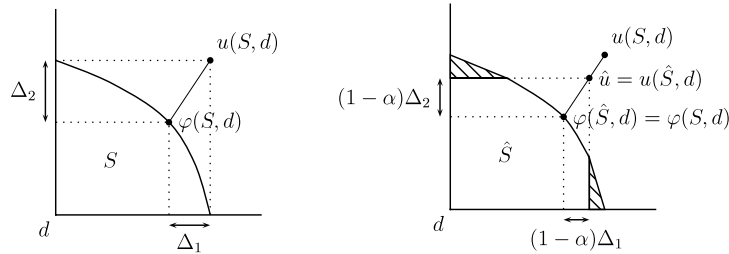


Fig. 2. A visual illustration of PCI in two-player bargaining problems.

first three properties are standard in the axiomatic bargaining literature, the last one is new.

A careful comparison between the axiomatic characterizations of KS and \mathcal{D}^N reveals that the correction needed to account for players' loss aversion is equivalent to relaxing the axiom of Symmetry, and imposing that the solution remains unchanged when players make concessions with respect to their utopia values in a proportion that is determined, uniquely, by the loss aversion coefficients.

As a final remark, it is not hard to see that it is advantageous for the opponents of a player i if i becomes more loss averse, i.e., if λ_i increases while the $\lambda_j, j \neq i$, stay the same. This can be deduced from Corollary 7.1 and inspection of the monotonic curves underlying the solutions D^k .

Appendix. Proofs

A.1. Proof of Lemma 3.1

Let $(S, d) \in B^N, \lambda \in \mathbb{R}_+^N$, and $r \in \mathbb{R}^N$. Since the Shalev transformation is continuous, one-to-one, and preserves the ordering of payoffs, it follows that

- $w(S, \lambda, r)$ is non-empty, closed, and comprehensive,
- $w(d, \lambda, r) \in w(S, \lambda, r)$,
- there are $z \in w(S, \lambda, r)$ with $z > w(d, \lambda, r)$, and
- $w(S_d, \lambda, r)$ is bounded.

It is left to show that $w(S, \lambda, r)$ is convex. Let $x, y \in S$ and $t \in [0, 1]$. By convexity of S , we have $tx + (1 - t)y \in S$. By the fact that w is concave in the first coordinate, it follows that

$$w(tx + (1 - t)y, \lambda, r) \geq tw(x, \lambda, r) + (1 - t)w(y, \lambda, r).$$

By comprehensiveness of $w(S, \lambda, r)$ this implies

$$tw(x, \lambda, r) + (1 - t)w(y, \lambda, r) \in w(S, \lambda, r).$$

This implies convexity of $w(S, \lambda, r)$. □

A.2. Proof of Theorem 5.1

We start with a lemma.

Lemma A.1. Let $k \in \mathbb{R}_{++}^{\bar{N}}$. Then the correspondence $\vartheta^k(t), t \in [1, n]$, is a monotonic curve.

Proof. Let $\bar{s} \in [1, n]$. We show that there is a unique point $z^* \in \mathbb{R}_+^N$ such that $\vartheta^k(\bar{s}) = z^*$. That is, we show that the system of equations

$$(1 - z_n)z_i = k_i(1 - z_i)z_n \quad \text{for all } i \in \bar{N} \tag{4}$$

$$\sum_{j=1}^n z_j = \bar{s}, \tag{5}$$

has exactly one solution z^* in \mathbb{R}_+^N . Suppose the system has a solution $z \in \mathbb{R}_+^N$, and suppose $z_n > 1$. Then for each $i \in \bar{N}$ we either have $z_i > 1$ or $z_i < 0$, which by $z \in \mathbb{R}_+^N$ implies $z_i > 1$ for all $i \in N$. Since this is a violation of (5), we must have $z_n \leq 1$. Since

$(1 - k_i)^{-1} \notin [0, 1]$ for all $i \in \bar{N}$, we can write

$$z_i = \frac{k_i z_n}{1 - z_n(1 - k_i)} \quad \text{for all } i \in \bar{N}.$$

Then for all $i \in \bar{N}$, z_i is strictly increasing in z_n on the domain $[0, 1]$. The observations that

$$\sum_{j=1}^n z_j = 0 \leq \bar{s} \quad \text{for } z_n = 0, \quad \text{and}$$

$$\sum_{j=1}^n z_j = n \geq \bar{s} \quad \text{for } z_n = 1,$$

together with the continuity of $\sum_{j=1}^n z_j$ in z_n then imply that there is exactly one $z^* \in \mathbb{R}_+^N$ that solves the system of Eqs. (4) and (5). It follows that the correspondence ϑ^k is single-valued. Moreover, for all $1 \leq s \leq t \leq n$ we have $0 < \vartheta^k(s) \leq \vartheta^k(t) \leq e^N$ and $\sum_{i \in \bar{N}} \vartheta_i^k(s) = s$. It follows that ϑ^k is a monotonic curve. □

To prove Theorem 5.1, we show that each map $D^k \in \mathcal{D}^N$ is the bargaining solution associated with the monotonic curve ϑ^k as defined in (3).

Proof of Theorem 5.1. Consider a normalized bargaining problem $T \in I_0^N$, i.e., $u(T) = e^N$. Let $k \in \mathbb{R}_{++}^{\bar{N}}$ and observe that by (2) we have $D^k(T) = P(T) \cap G^k$.

By convexity of T we have $P(T) \subseteq \{z \in \mathbb{R}_+^N \mid \sum_{j \in N} z_j \geq 1\}$, implying

$$D^k(T) = P(T) \cap \{\vartheta^k(t) \mid t \in [1, n]\},$$

where ϑ^k is defined by (3). It follows that $D^k(T) = \{\rho^{\vartheta^k}(T)\}$. From this it is easily established that

$$D^k(S, d) = \{\rho^{\vartheta^k}(S, d)\}$$

for all $(S, d) \in I^N$. □

A.3. Proof of Theorem 5.2

We start with the following lemma.

Lemma A.2. Let $(S, d) \in I^N$. Then $D^k(S, d) = KS(S, d)$ if and only if $k = e^{\bar{N}}$.

Proof. It is easy to show that for any $t \in [1, n]$, we have $\vartheta^k(t) = \hat{\vartheta}(t)$ if and only if $k = e^{\bar{N}}$. The result then follows from Theorem 5.1 and the definition of KS. □

Proof of Theorem 5.2. Let $(S, d) \in I^N$ and $\lambda \in \mathbb{R}_+^N$, and write $u = u(S, d)$. By Lemma A.2 and the fact that $KS(S, d) \geq d$, we have

$$\begin{aligned} KS(w((S, d), \lambda, z)) &= \{x \in P(w(S, \lambda, z)) \mid \text{for all } i \in \bar{N}: \\ & (u_n - x_n)(x_i - (1 + \lambda_i)d_i + \lambda_i z_i) \\ &= (u_i - x_i)(x_n - (1 + \lambda_n)d_n + \lambda_n z_n)\}. \end{aligned}$$

Observe that $z \in \text{Self}^{\text{KS}}((S, d), \lambda)$ if and only if $z = \text{KS}(w((S, d), \lambda, z))$. That is, $z \in \text{Self}^{\text{KS}}((S, d), \lambda)$ iff

- (i) $z \in P(w(S, \lambda, z))$, and
- (ii) $(u_n - z_n)(z_i - (1 + \lambda_i)d_i + \lambda_i z_i) = (u_i - z_i)(z_n - (1 + \lambda_n)d_n + \lambda_n z_n)$ for all $i \in \bar{N}$.

From $z = w(z, \lambda, z)$, $P(w(S, \lambda, z)) = w(P(S), \lambda, z)$, and the fact that the Shalev transformation is one-to-one, it follows that $z \in P(w(S, \lambda, z))$ is equivalent to $z \in P(S)$.

Define $k = (k_1, \dots, k_{n-1})$ where $k_i := \frac{1 + \lambda_n}{1 + \lambda_i}$ for all $i \in \bar{N}$. Then the statement in (ii) is equivalent to

$$(u_n - z_n)(z_i - d_i) = k_i(u_i - z_i)(z_n - d_n) \text{ for all } i \in \bar{N}.$$

It follows that $z \in \text{Self}^{\text{KS}}((S, d), \lambda)$ is equivalent to

$$z \in \{x \in P(S) \mid (u_n - x_n)(x_i - d_i) = k_i(u_i - x_i)(x_n - d_n) \text{ for all } i \in \bar{N}\}.$$

Hence, $\text{Self}^{\text{KS}}((S, d), \lambda) = D^k(S, d)$. \square

A.4. Proof of Theorem 6.1

The axiom of Pareto Optimality is useful for the proof.

(PO) $\varphi : B^N \rightarrow \mathbb{R}^N$ satisfies Pareto Optimality if $\varphi(S, d) \in P(S)$ for all $(S, d) \in B^N$.

From Peters and Tijs (1984) we obtain the following lemma.

Lemma A.3 (Peters and Tijs, 1984). Let $\varphi : I^N \rightarrow \mathbb{R}^N$ be a bargaining solution. Then φ satisfies PO, SI, and IM, if and only if $\varphi = \rho^\vartheta$ for some $\vartheta \in \Theta^N$.

This we use to establish the following result.

Proposition A.4. Let $\varphi : I^N \rightarrow \mathbb{R}^N$ be a bargaining solution in \mathcal{D}^N . Then φ satisfies SIR, SI, IM, and PCI.

Proof. Since $\varphi \in \mathcal{D}^N$ we have $\varphi = D^k$ for some $k \in \mathbb{R}_{++}^{\bar{N}}$. By Theorem 5.1, we have $\varphi = \rho^\vartheta$ where $\vartheta \in \Theta^N$, which by Lemma A.3 implies that φ satisfies SI and IM.

Consider a bargaining problem $(S, d) \in I^N$, and write $z = \varphi(S, d)$. By definition, $z \geq d$. To see that φ satisfies SIR, suppose there is an $i \in N$ such that $z_i = d_i$. Observe that $z \in P(S)$ and

$$(u_n(S, d) - z_n)(z_j - d_j) = k_j(u_j(S, d) - z_j)(z_n - d_n) \text{ for all } j \in \bar{N}.$$

If $i = n$, then $z_j = d_j$ for all $j \in \bar{N}$, implying $z = d$. Let $i \in \bar{N}$, and observe that $z_i = d_i$ implies $z_n = d_n$, and thus $z = d$. Since $d \notin P(S)$, we arrive at a contradiction. It follows that $z > d$.

To see that φ satisfies PCI, consider the problem $(\hat{S}, d) \in I^N$ where

$$\hat{S} := \{x \in S \mid x \leq \hat{u}\},$$

with $\hat{u} = \alpha z + (1 - \alpha)u(S, d)$ for some $\alpha \in (0, 1)$. Then $\hat{u} - z = (1 - \alpha)(u(S, d) - z)$, implying that

$$(\hat{u}_n - z_n)(z_i - d_i) = k_i(\hat{u}_i - z_i)(z_n - d_n) \text{ for all } i \in \bar{N}. \tag{6}$$

Since $z \in \hat{S}$ and $z \in P(S)$, we have $z \in P(\hat{S})$. This and (6) together imply $D^k(\hat{S}, d) = z$. Hence, φ satisfies PCI. \square

For the converse implication we need two additional lemmas.

Lemma A.5. Let $(S, d) \in I^N$, and $z \in S \setminus P(S)$. Then for the function $f : [0, 1] \rightarrow \mathbb{R}^N$ defined as

$$f(\alpha) := (1 - \alpha)z + \alpha u(S, d),$$

there is exactly one $\alpha^* \in [0, 1]$ such that $f(\alpha^*) \in P(S)$.

Proof. By compactness of S we have that

$$\alpha^* := \max\{\alpha \mid f(\alpha) \in S\}$$

is well defined. We now show that $f(\alpha^*) \in P(S)$. Suppose $f(\alpha^*) \notin P(S)$. By condition (1), it follows that for each $i \in N$ with $z_i < u_i(S, d)$, there is an $\varepsilon_i > 0$ such that $f(\alpha^*) + \varepsilon_i e^i \in S$. Then by convexity of S there is an $\varepsilon > 0$, such that

$$f(\alpha^*) + \varepsilon(u(S, d) - z) \in S.$$

But then there is a $\beta > \alpha^*$ with $f(\beta) \in S$. This is a contradiction.

To show uniqueness, let $\alpha_1, \alpha_2 \in [0, 1]$ with $\alpha_1 \neq \alpha_2$, and suppose $f(\alpha_1), f(\alpha_2) \in P(S)$. Without loss of generality, assume $\alpha_2 > \alpha_1$. Then since $u(S, d) \geq z$ and $u(S, d) \neq z$, we have $f(\alpha_2) \geq f(\alpha_1)$ and $f(\alpha_2) \neq f(\alpha_1)$. Since $f(\alpha_1) \in P(S)$, this implies $f(\alpha_2) \notin P(S)$, a contradiction. \square

Lemma A.6. Let $\varphi : I^N \rightarrow \mathbb{R}^N$ be a solution satisfying SIR, SI, and PCI. Then φ satisfies PO.

Proof. Let $\varphi : I^N \rightarrow \mathbb{R}^N$ be a bargaining solution satisfying SIR, SI, and PCI. By SI it is sufficient to restrict attention to bargaining problems in I_0^N . Let $S \in I_0^N$. By SIR we have $\varphi(S) > 0$. Now assume $\varphi(S) \notin P(S)$. By Lemma A.5, there is a single $z^* \in P(S)$, such that

$$z^* = (1 - \alpha)\varphi(S) + \alpha u(S) \text{ for some } \alpha \in (0, 1].$$

Define $\hat{S} := \{x \in S \mid x \leq z^*\}$, and observe that by PCI we have

$$\varphi(\hat{S}) = \varphi(S).$$

Similarly, for the set $T := \{x \in S \mid x \leq \varphi(S)\}$ we have $\varphi(T) = \varphi(S)$. Now observe that $\hat{S} = [z^*(\varphi(S))^{-1}T]$. Then by SI we have

$$\varphi(\hat{S}) = z^*(\varphi(S))^{-1}\varphi(T) = z^*(\varphi(S))^{-1}\varphi(S) = z^*.$$

This contradicts $\varphi(\hat{S}) = \varphi(S)$. \square

Proposition A.7. Let $\varphi : I^N \rightarrow \mathbb{R}^N$ be a bargaining solution satisfying SIR, SI, IM, and PCI. Then $\varphi \in \mathcal{D}^N$.

Proof. Since φ satisfies SIR, SI, and PCI, it follows from Lemma A.6 that φ satisfies PO. Then by Lemma A.3 it follows that $\varphi = \rho^{\vartheta^*}$ for some monotonic curve $\vartheta^* \in \Theta^N$.

Consider the problem

$$H := \text{conv}(\{e^i \mid i \in N\} \cup \{0\}),$$

(where ‘conv’ denotes the convex hull) and observe that $H \in I_0^N$. If there is an $i \in N$ with $\varphi_i(H) = 1$, then $\varphi_j(H) = 0$ for all $j \neq i$, which is a violation SIR. Hence, $0 < \varphi(H) < e^N$. It follows that $k = (k_1, \dots, k_{n-1})$, where

$$k_i = \frac{1 - \varphi_n(H)}{1 - \varphi_i(H)} \cdot \frac{\varphi_i(H)}{\varphi_n(H)}, \tag{7}$$

is well defined and $k \in \mathbb{R}_{++}^{\bar{N}}$.

In what follows, we show that

$$\vartheta^*(t) = \vartheta^k(t) \tag{8}$$

for all $1 \leq t \leq n$. We do this in three steps. Let $\bar{t} \in [1, n]$. Then

1. we construct a specific problem $S \in I_0^N$,
2. we show that $\varphi(S) = \vartheta^k(\bar{t})$, and
3. we show that $\varphi(S) = \vartheta^*(\bar{t})$.

By (8) we have $\rho^{\vartheta^k} = \rho^{\vartheta^*}$, which together with $\varphi = \rho^{\vartheta^*}$ and $D^k = \rho^{\vartheta^k}$ establishes $\varphi = D^k$.

Step 1: Define the function $g : [0, 1] \rightarrow [1, n]$ by

$$g(\alpha) = \sum_{i \in N} \frac{\varphi_i(H)}{\alpha \varphi_i(H) + (1 - \alpha)}. \tag{9}$$

From the fact that g is strictly increasing⁵ and continuous, and the fact that $g(0) = 1$ and $g(1) = n$, it follows that for each $t \in [1, n]$ there is a unique $\alpha \in [0, 1]$ such that $g(\alpha) = t$.

Let $\beta := \bar{\alpha}\varphi(H) + (1 - \bar{\alpha})e^N$, where $\bar{\alpha}$ is such that $g(\bar{\alpha}) = \bar{t}$. Then define the problem $S \in I_0^N$ by

$$S = \{\beta^{-1}z \mid z \in H \text{ and } z \leq \beta\}.$$

Since φ satisfies SI and PCI, we have

$$\varphi(S) = \beta^{-1}\varphi(H). \quad (10)$$

It follows from (10) and (9), and the fact that $g(\bar{\alpha}) = \bar{t}$ that $\sum_{i \in N} \varphi_i(S) = \bar{t}$.

Step 2: Rewriting (7) yields

$$(1 - \varphi_n(H))\varphi_i(H) = k_i(1 - \varphi_i(H))\varphi_n(H) \quad \text{for all } i \in \bar{N}.$$

From the definition of β we have $\beta - \varphi(H) = (1 - \bar{\alpha})(e^N - \varphi(H))$. Thus,

$$(\beta_n - \varphi_n(H))\varphi_i(H) = k_i(\beta_i - \varphi_i(H))\varphi_n(H) \quad \text{for all } i \in \bar{N}.$$

For each $i \in \bar{N}$, we can multiply both sides of the equation by $\frac{1}{\beta_n \beta_i}$. By (10) this yields

$$(1 - \varphi_n(S))\varphi_i(S) = k_i(1 - \varphi_i(S))\varphi_n(S) \quad \text{for all } i \in \bar{N}.$$

It follows that $\varphi(S) \in G^k$. Since $\sum_{i \in N} \varphi_i(S) = \bar{t}$, we have

$$\varphi(S) = \vartheta^k(\bar{t}). \quad (11)$$

Step 3: Since $\varphi = \rho^{\vartheta^*}$, and $S \in I_0^N$ with $u(S) = e^N$, we have

$$\{\varphi(S)\} = P(S) \cap \{\vartheta^*(t) \mid t \in [1, n]\},$$

which implies $\varphi(S) = \vartheta^*(t^*)$ for some $t^* \in [1, n]$. From the definition of monotonic curves we obtain

$$t^* = \sum_{i \in N} \vartheta_i^*(t^*) = \sum_{i \in N} \varphi_i(S) = \bar{t}.$$

Hence,

$$\varphi(S) = \vartheta^*(\bar{t}). \quad (12)$$

Combining (11) and (12) yields the desired result. \square

Theorem 6.1 now follows from Propositions A.4 and A.7.

References

- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of choice under risk. *Econometrica* 47 (2), 263–291.
- Kalai, E., Smorodinsky, M., 1975. Other solutions to Nash's bargaining problem. *Econometrica* 43 (3), 513–518.
- Kannai, Y., 1977. Concavifiability and constructions of concave utility functions. *Journal of Mathematical Economics* 4 (1), 1–56.
- Kihlstrom, R.E., Roth, A.E., Schmeidler, D., 1981. Risk aversion and solutions to Nash's bargaining problem. In: O. Moeschlin and D. Pallaschke (Eds.), *Game Theory and Mathematical Economics*. North-Holland, Amsterdam.
- Köbberling, V., Peters, H., 2003. The effect of decision weights in bargaining problems. *Journal of Economic Theory* 110 (1), 154–175.
- Köbberling, V., Wakker, P., 2005. An index of loss aversion. *Journal of Economic Theory* 122, 119–131.
- Köszegi, B., Rabin, M., 2006. A model of reference-dependent preferences. *Quarterly Journal of Economics* 121, 1133–1165.
- Nash, J., 1950. The bargaining problem. *Econometrica* 18 (2), 155–162.
- Peters, H., 2010. A preference foundation for constant loss aversion. Working Paper. Maastricht University.
- Peters, H., Tijs, S., 1984. Individually monotonic bargaining solutions for n -person bargaining games. *Methods of Operations Research* 51, 377–384.
- Peters, H., van Damme, E., 1991. Characterizing the Nash and Raiffa bargaining solutions by disagreement point axioms. *Mathematics of Operations Research* 16 (3), 447–461.
- Raiffa, H., 1953. Arbitration schemes for generalized two-person games. In: Kuhn, H.W., Tucker, A.W. (Eds.), *Contributions to the Theory of Games II*. In: *Annals of Mathematics Studies*, vol. 28. pp. 361–387.
- Roth, A.E., 1977. Individual rationality and Nash's solution to the bargaining problem. *Mathematics of Operations Research* 2 (1), 64–65.
- Roth, A.E., 1979. An impossibility result concerning n -person bargaining games. *International Journal of Game Theory* 8 (3), 129–132.
- Shalev, J., 2000. Loss aversion equilibrium. *International Journal of Game Theory* 29 (2), 269–287.
- Shalev, J., 2002. Loss aversion and bargaining. *Theory and Decision* 52 (3), 201–232.
- Sugden, R., 2003. Reference-dependent subjective expected utility. *Journal of Economic Theory* 111, 172–191.

⁵ This follows from $\varphi(H) \neq 0$ and $\varphi(H) \neq e^N$.