

Sender for anvendt forskning. Rapport. 1985:8

Report no. 8/1985

Norges Handelshøyskole
BIBLIOTEKET

9 JAN. 1986

STABILITY, COLLECTIVE CHOICE
AND SEPARABLE WELFARE

BY

TERJE LENSBERG

CENTER FOR APPLIED RESEARCH
THE NORWEGIAN SCHOOL OF ECONOMICS
AND BUSINESS ADMINISTRATION

BERGEN, 1985

PREFACE

The work reported in this study began when I was attending William Thomson's course in social choice and game theory at Harvard in the spring of 1981. Inspired by Thomsons' results on the fair division of a fixed supply among a growing population, I came across some ideas related to his, which led to a joint project on "The axiomatic theory of bargaining with a variable population". The present study is part of my contribution to that project.

I would like to thank William Thomson for teaching me the gentle art of the axiomatic method and for his generous offer of collaboration, which has resulted in numerous stimulating discussions and extensive comments to my drafts. I am also grateful to my adviser, Karl Borch, who has contributed to my understanding of the subject matter of this study through several helpful discussions and critical comments.

Thanks are also due to Kåre P. Hagen, Jostein Lillestøl, Bjørn Sandvik, Eivind Stensholt, Lars Thorlund-Petersen and H. Peyton Young, with whom I have had the opportunity to discuss parts of the material presented in this study, and to Anne Kristin Wilhelmsen and Vibeke Farestvedt for professional and cooperative typing assistance.

Last, but not least, I am indebted to the Center for Applied Research for allowing me the time to complete my work with this study.

Bergen, august 1985

CONTENTS

1. BARGAINING AND FAIR ALLOCATION	1
1.1 Introduction	1
1.2 The model	5
1.3 The results	11
1.4 Conclusion	24
2. STABILITY AND THE NASH SOLUTION	26
2.1 Introduction	26
2.2 The model	28
2.3 The main result	38
2.4 Variants of the main result	47
2.5 Concluding remarks	53
3. STABILITY AND THE LEXIMIN SOLUTION	54
3.1 Introduction	54
3.2 The main result	63
3.3 Concluding remarks	72
4. STABILITY AND COLLECTIVE RATIONALITY	75
4.1 Introduction	75
4.2 The axioms	86
4.3 The main result	91
4.4 Variants of the main result	114
4.5 Concluding remarks	130
REFERENCES	131

CHAPTER 1

BARGAINING AND FAIR ALLOCATION

1.1 Introduction

This study is concerned with problems of collective choice, seen from the point of view of bargaining or fair division, a tradition initiated by John Nash's (1950) pioneering essay on cooperative bargaining. The present chapter contains an informal overview of the main results of the study, as an introduction to the more detailed exposition contained in chapters 2-4. In addition to that, this chapter outlines the implications of some of the results in subsequent chapters for a specific problem, namely the problem of attaining consistency of plans for the allocation of costs and benefits in a system of decentralized public decision making. We begin by discussing the relevance of Nash's bargaining paradigm to such allocation problems. Section 1.2 outlines the model that will be used throughout, section 1.3 contains the summary of results and section 1.4 some concluding remarks.

A typical allocation problem that would fit within Nash's framework is one in which a number of firms plan to undertake a joint venture if they can agree on how to share the profits from the project. The firms may settle for any agreement they could think of, as long as it is unanimous. If there is no agreement, there will be no project and no profits to share. Thus, we consider a situation where no proper subset of agents can accomplish anything on their own, i.e., there is no room for coalitions. The question then is: what will be the outcome to such a problem?

In Nash's framework, one attempts to gain some insight into the problem by means of the following type of model: First abstract from all physical characteristics of the particular problem at hand by representing it as a pair (S,d) , where S is a set of feasible von Neumann-Morgenstern utility vectors, one vector for each physical alternative or probability mixture of alternatives, and where d is the utility vector that will be the outcome in the case of no agreement between the agents (the status quo). Now one may consider whole families of such abstract bargaining problems and look for a set of general principles (axioms) that would describe the behavior of the agents in any given bargaining situation. Then by requiring that the solution outcome to any bargaining problem (denoted $F(S,d)$) should obey these axioms, one can hope to narrow down the set of possible candidates for a solution (the set of possible functions F) to a class with sufficient structure to yield some predictive power.

Clearly, nothing prevents us from looking at these principles from a normative instead of a descriptive viewpoint, e.g. as principles of fair division (Harsanyi (1955)), or the values of an arbiter (Raiffa (1953)). Here we will allow for both interpretations of the model and refer to the problems (S,d) under consideration as collective choice problems, or simply choice problems.

To take an example in the normative spirit, consider the problem of dividing the costs and benefits of a public utility among its clients. Is there a division scheme which is in some sense fair? Again, one way to attack the problem would be to identify a set of axioms that might reflect popular ideas of what constitutes a fair division, and use them to limit the opportunities for discretionary action by the management of the public utility. In some cases, as the one studied by Nash, the axioms will eliminate discretion altogether by singling out a unique division scheme.

As a third example, we may look at a society where public decisions are made in the following decentralized fashion: There are many publicly owned or regulated firms, each servicing a subset of the members of society; some municipal electricity companies, a number of public universities, a few airlines and some broadcasting companies. Clearly, the production and pricing decisions of each unit will affect the welfare of those individuals who consume and pay for its services. Suppose each unit is instructed to achieve a fair allocation of costs and benefits among its clients. An interesting question is then whether such decentralized public decision making will lead to allocations that are fair for society as a whole. Put differently, if such decentralized decision making is going to be consistent with some overall notion of fairness, then what are the implications as to the nature of the decision rules that would have to be followed by the decentralized units, and what restrictions, if any, would such a requirement impose on the notion of fairness itself?

This question will be a main topic of the present chapter. Our analysis is based on two ingredients, the first one is an axiom which was first used by Harsanyi (1959) in connection with the Nash bargaining solution, and which expresses the kind of consistency requirement mentioned in the previous paragraph.

Harsanyi's axiom differs from those that are usually studied within the tradition of bargaining and fair division in being a condition on the relationship between choice problems involving different sets of agents, while in the traditional model, the set of agents is fixed. Problems of collective choice with a variable number of agents was first studied in a systematic way by Thomson (1983a), and it is his model which is the second main ingredient in our analysis.

The results of this analysis are several new characterizations of familiar solutions which shed new light on the nature of those solutions. We discuss the relationship of Harsanyi's axiom to the problem of decentralized public decision making mentioned above. This axiom can be seen as a necessary condition for such decentralized decision making to be consistent with any overall notion of fair division, and we show that it has very precise implications as regards the nature of the decision rules that would have to be followed by the decentralized units.

1.2 The model

The classical axiomatic model of bargaining involves a fixed set P of agents, each equipped with a von Neumann-Morgenstern utility function. P may be taken to be a non-empty, finite subset of the natural integers. Let $|P|$ be the number of elements in P , and let \mathcal{R}^P be the $|P|$ -dimensional euclidean space, indexed by the members of P . \mathcal{R}_+^P denotes the non-negative orthant of \mathcal{R}^P . A $|P|$ -person bargaining problem is a pair (S, d) , where S is a subset of \mathcal{R}^P , d an element of S , and where S is compact and convex with at least one vector that strictly dominates d .

S is the set of utility allocations that can be achieved by the members of P through unanimous agreement, and d is the outcome that will result if they fail to agree. Thus, all subcoalitions of P can veto any outcome different from d , while cooperation by all agents is required in order to achieve another outcome. The existence of a point in S which strictly dominates d guarantees that all agents are non-trivially involved in the bargaining problem. The compactness of S is a technical assumption, convexity follows if the agents may jointly randomize between outcomes.

For simplicity, we will assume that the utility functions are normalized such that the vector d is always the origin of \mathcal{R}^P . We may then identify any bargaining problem (S, d) by the set S only. We will also restrict the family of bargaining problems under consideration to sets S such that S is a subset of \mathcal{R}_+^P and such that S is comprehensive, meaning that if x is a utility vector in S , then so is any non-negative vector that is weakly dominated by x . Comprehensiveness amounts to assuming free disposal of utility.

Let Σ^P be the class of bargaining problems S for the set P of agents, as defined above. Such problems will be referred to as choice problems, and will be denoted S, S', T etc. A typical choice problem S is illustrated in figure 1.1 for $P = \{1,2\}$.

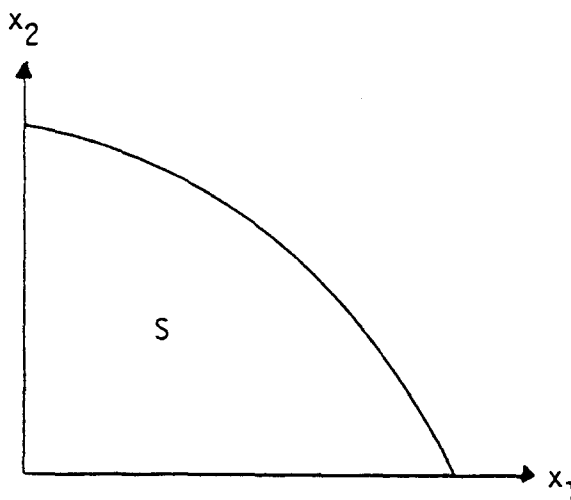


Figure 1.1

A typical choice problem

A solution is defined to be a function $F: \Sigma^P \rightarrow \mathbb{R}_+^P$, such that $F(S) \in S$ for all $S \in \Sigma^P$. Given $S \in \Sigma^P$, the vector $F(S)$ is called the solution outcome to S , interpreted as that compromise which is in some sense a best resolution of the conflict among the agents in P .

We have now given a description of the basic model of the collective choice problem. One can now proceed, as Nash (1950) did, to look for a set of axioms that would guide the agents in their search of a fair compromise. Nash suggested four such axioms, namely Pareto-optimality (PO), Symmetry (SY), Scale Invariance (S.INV) and Independence of Irrelevant Alternatives (IIA).

PO requires that the solution outcome to any choice problem S should be a Pareto-optimal point in S , and SY states that symmetric choice problems should have symmetric solution outcomes, i.e. if the geometry of a choice problem does not distinguish between the agents, then the solution should not do so either. A slightly stronger version of the symmetry axiom is Anonymity (AN) which states that the solution outcome should only depend on the geometry of the given choice problem, and not on the names of the agents.

S.INV requires that a rescaling of the utility representations of one or more agents by a positive linear transformation should rescale the solution outcome in the same way. It reflects the fact that von Neumann-Morgenstern utility functions are only unique up to positive affine transformations, the constant terms of these transformations having already been used to translate the disagreement point to the origin. Finally, IIA states that if one choice problem is obtained from another by narrowing down the set of feasible alternatives while keeping the solution outcome of the original problem a feasible alternative in the new problem, then the solution outcomes to the two problems should be the same. The idea is that if an alternative is "best" among a given set of alternatives, then it must also be "best" among any subset of those alternatives.

Nash showed that there exists one and only one solution that satisfies PO, SY, S.INV and IIA. It is the solution that for all S picks the unique outcome that maximizes the product of the agents' utility levels on S . Strictly speaking, Nash stated his result only for the two-person bargaining problem, presumably because in a situation with more than two agents, there might be room for coalitions, a feature which is not captured by the model.

Nash's model has been elaborated by Harsanyi (1959), (1963) and (1977), who showed how the problem of solving n-person bargaining problems could be reduced to the more familiar one of solving two-person problems. He argued that in any n-person bargaining problem, a particular payoff-vector "... will represent the equilibrium outcome of bargaining among the n players only if no pair of players i and j has any incentive to redistribute their payoffs between them, as long as the other players' payoffs are kept constant". (Harsanyi (1977) p. 196). This condition, which we will refer to as Bilateral Stability, was shown to imply that if, among a group of n agents, all two-person bargaining problems were solved by the two-person Nash solution, then all n-person bargaining problems had to be solved by the n-person Nash solution.

Before we give an illustration of Harsanyi's condition, which differs from the ones already introduced by involving a varying number of agents, it will be convenient to modify the basic model by following Thomson (1983a), who deals with this case in a more explicit fashion.

Let there be a fixed set I of agents that may potentially become involved in some collective choice problem, and let \mathcal{P} be the set of finite subsets of I. I may be taken to be the set of natural integers. Elements of \mathcal{P} are denoted P, P', Q etc. For all $P \in \mathcal{P}$, let Σ^P be the set of all choice problems for the set P of agents. We re-define a solution to be a function $F: \bigcup_{P \in \mathcal{P}} \Sigma^P \rightarrow \bigcup_{P \in \mathcal{P}} \mathbb{R}_+^P$, such that for all $P \in \mathcal{P}$ and all $S \in \Sigma^P$, $F(S)$ is an element of S. For all $P \in \mathcal{P}$, the restriction of F to Σ^P is called the P-component of F.

An illustration of Harsanyi's condition is given in figure 2.1, where T is a choice problem for the group $Q \equiv \{1,2,3\}$ of agents and where the solution outcome for T is x . By keeping the utility of agent 3 constant at x_3 , one obtains a choice problem S involving only agents 1 and 2, and the requirement made by B.STAB is that the solution outcome to this two-person problem should be (x_1, x_2) .

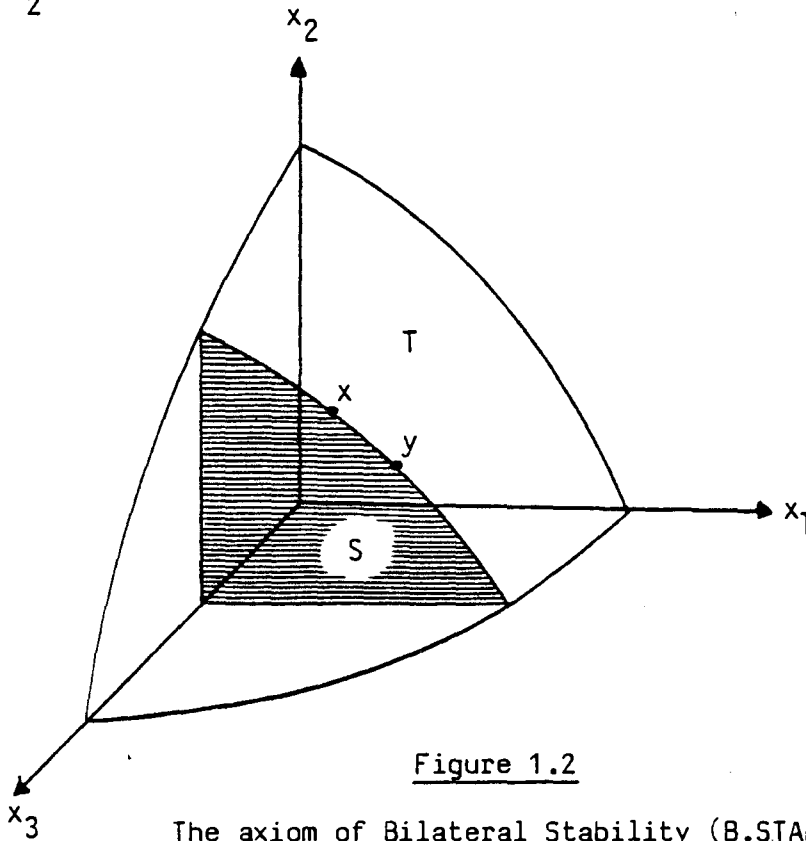


Figure 1.2

The axiom of Bilateral Stability (B.STAB)

Harsanyi motivates this condition by pointing out that a rational agent i will not accept a tentative agreement x for the bargaining problem T if he has reason to believe that he could successfully force some other agent j to make a concession in his favor: Suppose that the agents are all familiar with Nash's solution to the two-person bargaining problem, and that it is common knowledge among the agents that two-person problems are solved by that solution. Consider then agent 1, who is looking at the bargaining problem S

involving only agent 2 and himself, obtained by keeping the utility level of agent 3 fixed at x_3 .

If the Nash solution outcome to this two-person bargaining problem is (y_1, y_2) where $y_1 > x_1$, then agent 1 will not accept x_1 but will demand y_1 , arguing that agent 2 should lower his claim accordingly, by referring to their common knowledge of two-person bargaining theory. Whether agent 2 accepts or not does not really matter: If agent 1 rejects x , then x cannot be the solution outcome to the three-person bargaining problem.

Seen positively, this means that a utility vector x can be the solution outcome to the $|Q|$ -person choice problem T only if it agrees with the solution outcomes to all two-person subproblems S obtained from T by keeping the utility levels to all but two of the agents constant at the original outcome.

Because there seems to be no a priori reason why a dissatisfied agent should limit himself to challenging only one other agent at the time for concessions, it seems natural to consider the following generalization of B.STAB, that we call Multilateral Stability (M.STAB) and which states that the solution should be stable, not only with respect to two-person subproblems, but also with respect to subproblems involving any subset of the original group. As a principle of fair division, the axiom can be interpreted as a consistency requirement on the notion of fairness, saying that an allocation should not be declared a fair compromise for a given set of agents if it is unfair for some subset of those agents.

1.3 The results

In this section, we give an outline of our results involving the Stability axiom. In section 1.3.1 we present a new characterization of the Nash solution, section 1.3.2 is concerned with the Leximin solution, and in section 1.3.3 we discuss a family of collectively rational and "decentralizable" solutions. Proofs will not be given here, but may be found in chapters 2, 3 and 4, respectively.

1.3.1 Stability and the Nash solution

We begin by stating the following two theorems, due to Nash (1950) and Harsanyi (1959):

Theorem 1 (Nash): A solution F satisfies PO, SY, S.INV and IIA if and only if for all $P \in \mathcal{P}$ and all $S \in \Sigma^P$, $F(S) = N(S) \equiv \operatorname{argmax}_{x \in S} \{ \prod_{i \in P} x_i \}$.

Theorem 2 (Harsanyi): If a solution F satisfies CONT¹⁾ and B.STAB, and if F coincides with the Nash solution N for two-person problems, then $F = N$.

Theorem 2 demonstrates how B.STAB can be applied to reduce the problem of solving n -person problems to one of solving two-person problems. In particular, if it is known ex ante that two-person problems are solved by the Nash solution, one is left with no degrees of freedom as regards the choice of a suitable n -person solution.

1) CONT is a continuity assumption, stating that similar choice problems should have similar solution outcomes.

What then if nothing is known at the outset about the nature of the two-person solution? What kind of analytical power does the stability axiom have in this more general situation? It turns out that it can be used to give the following alternative characterization of the Nash solution:

Theorem 3: A solution F satisfies PO , AN , $S.INV$ and $M.STAB$ if and only if it is the Nash solution.

If we compare this result to Nash's own characterization, we see that except for a strengthening of SY to AN , the only difference is that IIA has been replaced by $M.STAB$. This is interesting, since the axiom of IIA has been somewhat controversial within the bargaining tradition (cf. Luce and Raiffa (1957)), and because of that, some authors (e.g. Kalai and Smorodinsky (1975), Roth (1977), Thomson (1981a)) have replaced it with other axioms, and have arrived at different solutions. Here, however, we replace IIA by a version of Harsanyi's stability condition and still arrive at the Nash solution. Thus, it seems that the Nash solution does not rest so heavily on the axiom of IIA as is often thought.

As regards the connection to Harsanyi's theorem, we observe that, except for dropping the hypothesis that two-person problems are solved by the Nash solution, Theorem 3 uses the stronger version of the Stability axiom, while $CONT$ is not needed. Alternatively, we could weaken $M.STAB$ to $B.STAB$, impose $CONT$ and obtain the following variant of Theorem 3:

Theorem 4: A solution F satisfies PO , AN , $S.INV$, $CONT$ and $B.STAB$ if and only if it is the Nash solution.

Going back to Theorem 3, it turns out that this is not the strongest result that one can prove. Specifically, the axiom of Pareto-optimality may be considerably weakened and still permit a characterization of the Nash solution. Suppose we weaken PO to require only that all one-person choice problems should be solved optimally:

Individual optimality (IO): For all $P \in \mathcal{P}$ with $|P| = 1$, for all $S \in \Sigma^P$,
 $F(S) = \max\{x \mid x \in S\}$.

Because IO and M.STAB together imply PO, we obtain:

Theorem 5: A solution F satisfies IO, AN, S.INV and M.STAB if and only if it is the Nash solution.

Let us also compare this result to Theorem 1. Nash's axioms seem to fall into two categories that are qualitatively quite different. The first category consists of SY and S.INV, which state that the solution outcome should not depend on information which is not contained in the model (Nash (1953), Roth (1979b)). In particular, S.INV is a reflection of the fact that von Neumann-Morgenstern utility functions are only unique up to positive affine transformation. In the second category are PO and IIA, which both demand some form of collective rationality of the agents. Theorem 5 employs slightly modified versions of Nash's axioms in the first category, and replaces those in the second category by IO and M.STAB, both of which express a kind of individual (rather than collective) rationality, when interpreted in a bargaining context.

The kind of collective rationality expressed by IIA can be seen more clearly by rephrasing it to say that if some alternative was declared to be "best" among a set of feasible alternatives, then it must also be "best" among any subset of those alternatives. This means in particular (Roth (1979b)) that if the set of feasible alternatives is expanded, then the solution either selects one of the new alternatives available, or it selects the solution outcome to the original problem. On the other hand, IIA does not say anything as to how the solution outcome should change if it changes as a result of an expansion in the set of feasible alternatives. For example, one might feel that if the set of alternatives is expanded in a direction which is particularly favorable to some agent, then that agent should gain, or at least should not be worse off, as a result of such a change in the problem.

Several authors have proposed and used axioms that express such a condition of individual monotonicity (Kalai and Smorodinsky (1975), Kalai (1977a), Roth (1979a) and Thomson and Myerson (1980)). In the next section we study the consequences of imposing such an axiom in conjunction with M.STAB.

1.3.2 Stability and the Leximin solution

An illustration of the axiom of Individual Monotonicity (I.MON), is given in figure 1.3, where the choice problem S' for the set $P = \{1,2\}$ of agents is obtained from S by expanding the set of feasible alternatives in agent 1's direction, while leaving the set of feasible utility levels for agent 2 unchanged. The requirement made by I.MON is that agent 1 should not lose as a result of this change, i.e. that the solution outcome to S' should lie in the shaded area of figure 1.3.

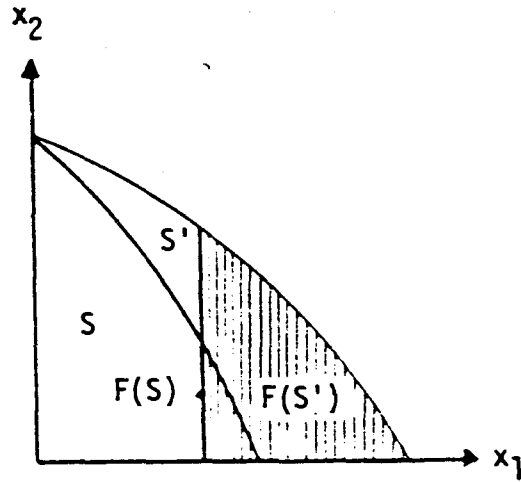


Figure 1.3

The axiom of Individual Monotonicity (I.MON)

What do we get if we add I.MON to the list of axioms in Theorem 3? The answer is nothing, because the Nash solution does not satisfy I.MON, as shown in figure 1.4.

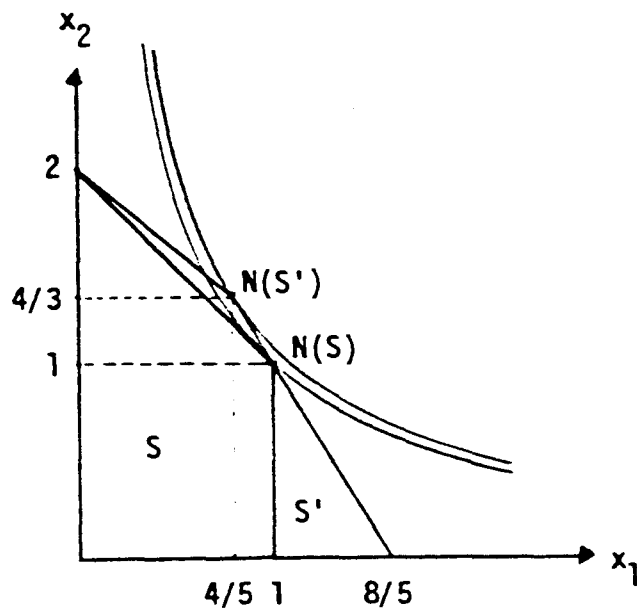


Figure 1.4

The Nash solution does not satisfy I.MON

Thus, if we want the solution to satisfy I.MON, then some other axiom in Theorem 3 must go. The question is which one. If we look at figure 1.4 again, we see that the reason why the Nash solution does not satisfy I.MON is that the level curves of the Nash product $\prod_{i \in P} x_i$ permit too much trade-off between the utility levels of agents 1 and 2. As it turns out, (see section 1.3.3) the axiom of Nash which is responsible for the particular shape of those level curves is S.INV, so it is this axiom that will have to go.

The question then is whether there are any solutions that satisfy PO, AN, I.MON and M.STAB. One possible candidate is the Egalitarian solution E , which to each choice problem S picks the unique point of equal coordinates in the upper boundary of S . (See Raiffa (1953), Myerson (1977), (1981) and Thomson (1983b)). This solution does not admit any trade-off between the utilities of different agents, and so it would not violate I.MON in the example given in figure 1.4. However, it satisfies neither PO nor M.STAB as is clear from figure 1.5, where $Q \equiv \{1,2,3\}$, $P \equiv \{2,3\}$ and $T \equiv \{x \in \mathbb{R}_+^Q \mid x \leq (1,2,3)\}$, and where the Egalitarian solution outcome for T is $E(T) = (1,1,1)$. By keeping the utility level for agent 1 constant at $E_1(T) = 1$, one obtains the two-person problem $S \equiv \{x \in \mathbb{R}_+^P \mid x \leq (2,2)\}$, whose Egalitarian solution outcome $E(S)$ is $(2,2)$. Because $(1,1,1)$ not a Pareto-optimal point in T , then E does not satisfy PO, and because $(2,2) \neq (1,1)$, then E does not satisfy M.STAB.

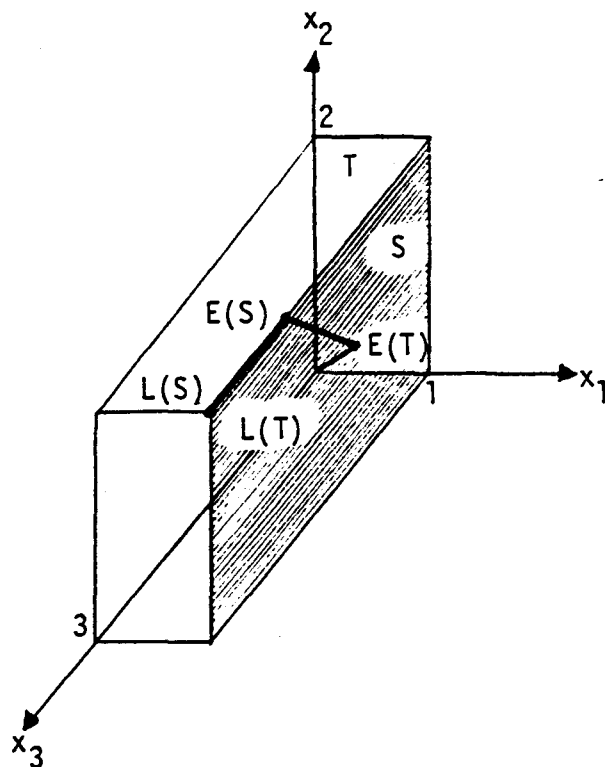


Figure 1.5

The Egalitarian solution satisfies neither PO nor M.STAB,
but the Leximin solution does

The Egalitarian solution is closely related to the Rawlsian maximin criterion (Rawls (1971)) by always selecting a feasible alternative which maximizes the utility of the worst off individual. In general, there may be more than one such alternative, as shown in figure 1.5, but Sen (1970) has suggested the following lexicographic extension of the Rawlsian maximin criterion which eliminates this indeterminacy. First maximize the utility of the worst-off individual, then do the same for the next to worst-off individual, and so on, until all possibilities for increasing the utility of any individual has been exhausted. The solution obtained in this way is called the Leximin solution and is denoted L. It is illustrated in figure 1.5, which also shows that L

satisfies both PO and M.STAB in the given example. In fact, we have the following theorem:

Theorem 6: A solution satisfies PO, AN, I.MON and M.STAB if and only if $F = L$, the Leximin solution.

Observe that the list of axioms used in Theorem 6 differs from the one used to characterize the Nash solution in Theorem 3 only in that S.INV has been replaced by I.MON.¹⁾ Now, S.INV can be interpreted as a condition which rules out interpersonal comparisons between agents whose preferences are represented by (cardinal) von Neumann-Morgenstern utility functions. Theorems 3 and 6 show that S.INV and I.MON are in a sense polar opposites when used in conjunction with the other three axioms: The Leximin solution exploits to a maximum degree the possibilities for interpersonal comparability of relative utility levels that become available when S.INV is dropped, by admitting no trade-off between the utility levels of different agents.

One problem with the Leximin solution is that it is not continuous, as can be seen by considering any sequence $\{T^v\}$ of choice problems converging to the problem T depicted in figure 1.5, such that each T^v is strictly convex (in \mathcal{R}_+^P). Then $L(T^v) = E(T^v)$ for all T^v in the sequence, which means that $\{L(T^v)\}$ converges to $E(T)$. Because $E(T) = (1,1,1)$ while $L(T) = (1,2,3)$, this is a violation of CONT.

1) Imai (1983) has given a characterization of the Leximin solution which parallels Theorem 1 in a similar way.

Thus, under the Leximin solution, it is not always the case that similar choice problems have similar solution outcomes. This might cause someone who was supposed to use it, say in a cost-benefit analysis, to worry about the quality of his data. Although there may be other problems to worry about in connection with implementing a collective decision rule, it will nevertheless be of interest to investigate the consequences of imposing continuity as an axiom in the model. This will be done in the next section.

1.3.3 Stability and Collective Rationality

Although the Nash solution and the Leximin solution are different in many respects, they have one thing in common: Both are consistent with the maximization of some ordering¹⁾ on the space of alternatives. In the terminology of Richter (1971), such solutions are said to be [collectively] rational.

Clearly, the Nash solution is collectively rational: For all P in \mathcal{P} and all S in Σ^P , the Nash solution outcome for S is obtained by maximizing the ordering \succsim_P^N over S , where \succsim_P^N is defined on \mathbb{R}_+^P by $x \succsim_P^N y$ if and only if $\prod_{i \in P} x_i \geq \prod_{i \in P} y_i$. As regards the Leximin solution, Imai (1983) has shown that it is consistent with the maximization of the ordering \succsim_P^L , where for each P , \succsim_P^L is the (symmetric) lexicographic extension of the ordering \succsim_P^E of \mathbb{R}_+^P defined by $x \succsim_P^E y$ if and only if $\min_{i \in P} x_i \geq \min_{i \in P} y_i$.

1) An ordering is a binary relation which is transitive, reflexive and complete.

The property of being collectively rational is enjoyed by many other solutions as well. To fix ideas, one may think of the orderings \succsim_p^N and \succsim_p^L as Bergson-Samuelson social welfare functions (swf) (Bergson (1938), Samuelson (1947)). It is then clear that from any Bergson-Samuelson swf one obtains a social choice function (a solution) provided the maxima for the swf always exist and are unique on the relevant domain of choice problems.

A condition that is often imposed on the social ordering is Separability or Independence of Unconcerned Individual as it is also sometimes called. This condition (due to Fleming (1952)) says that if the utility levels for a subset of the agents of society is the same for some pair of alternatives, then the social ordering of those alternatives should not depend on the utility levels of those agents. This means that if \succsim_Q is a social ordering of the utility space \mathcal{R}_+^Q for a group Q of agents, then for all subsets P of Q , the ordering \succsim_P obtained from \succsim_Q by restricting \succsim_Q to any hyperplane parallel to \mathcal{R}^P must be the same for all such hyperplanes. In particular, if the ordering \succsim_Q is continuous, then it has an additively separable numerical representation. In other words, there is a real-valued function f^Q on \mathcal{R}_+^Q such that $f^Q(x) \geq f^Q(y)$ if and only if $x \succsim_Q y$, where f^Q is of the form $f^Q(x) = \sum_{i \in Q} f_i(x_i)$ (Debreu (1960)).

The condition of Separability is indeed satisfied by many of the commonly used Bergson-Samuelson swf's, such as the Utilitarian swf (classical utilitarianism), the Leximin swf (Sen (1970), which is the lexicographic extension of the Rawlsian maximin criterion, as well as the Nash swf (Nash (1950), Kaneko and Nakamura (1979)).

Intuitively, the Stability axiom is a natural counterpart to Separability in the sense that it imposes on a solution much the same requirement that Separability imposes on a social ordering. What is more interesting, and less obvious, is that it imposes on the solution a fair amount of collective rationality as well, as the next theorem shows.

Let F be the family of all sequences $\{f_i\}_{i \in I}$ of strictly increasing, extended real-valued functions, where each f_i is defined on $\mathbb{R}_+^{\{i\}}$, such that for all $P \in \mathcal{P}$, the function $f^P \equiv \sum_{i \in P} f_i$ is strictly quasi-concave. We now have

Theorem 7: A solution F satisfies PO, CONT and B.STAB if and only if there exists a sequence of functions $\{f_i\}_{i \in I}$ from F such that for all $P \in \mathcal{P}$ and all $S \in \Sigma^P$, $F(S) = \operatorname{argmax}\left\{ \sum_{i \in P} f_i \mid x \in S \right\}$.

It is interesting to see this result in relation to the problem mentioned earlier of attaining consistency in a system of decentralized public decision making, where each decentralized unit is trying to achieve a fair allocation among its own clients. When interpreted in this context, the Stability axiom requires that if an allocation is to be considered globally fair, then each decentralized unit should also regard the allocation as fair when considering only its own clients. This is clearly a necessary condition for such decentralized public decision making to be consistent with some global notion of fairness: If it were not satisfied, then the globally fair allocation could never be obtained, because there would always be some local unit that would want to move away from it.

Theorem 7 shows that such a condition, when imposed in conjunction with PO and CONT, has very precise implications concerning the nature of the decision

rules that will have to be followed by the decentralized units: Firstly, the global (and local) notion of fairness must correspond to some Bergson-Samuelson swf, and the decision rules must collectively rational. Secondly, the global swf must be additively separable, which means that it can be "split up" and distributed among the decentralized units in such a way that each unit can make its decisions based on information about its own clients only. Clearly, there is in general no guarantee that this type of decentralized decision making will actually lead the society towards the globally fair allocation, but the point is that if the globally fair allocation exists at all (which it does, according to Theorem 7), then the decision rules will have to be of this form.

It can be shown that the axioms used in Theorem 7 are independent, in the sense that removing any one of them will permit solutions that are not collectively rational. Conversely, because the theorem characterizes a whole family of solutions, it is a useful framework for analyzing the implications of adding more axioms to the list in Theorem 7.

Adding Symmetry (SY) to the list of axioms implies that all the functions f_i must be identical. Next, we consider a weaker version of S.INV, namely Homogeneity (HOM), which says that if two choice problems are identical, except for a scale change, then their solution outcomes should also be identical, except for the same scale change.

Adding HOM to the list of axioms in Theorem 7 implies that the functions $\sum_{i \in P} f_i$ must be homothetic for all $P \in \mathcal{P}$. This means (Eichhorn (1978) Theorem 2.2.1) that (except for arbitrary constant terms) there exists $\rho > -1$ and a

sequence $\{\alpha_i\}_{i \in I}$ of positive real numbers such that $f_i(x_i) = -(\alpha_i/\rho)x_i^{-\rho}$ for all $i \in I$ if $\rho \neq 0$, and $f_i(x_i) = \alpha_i \log x_i$ for all $i \in I$ if $\rho = 0$. Thus $\sum_{i \in P} f_i$ is a CES-function for all $P \in \mathcal{P}$. If SY is also imposed, then f_i , and hence α_i , must be the same for all i . As $\rho \rightarrow 0$, we then obtain the Nash swf, as $\rho \rightarrow -1$ we obtain classical utilitarianism and as $\rho \rightarrow \infty$ we get the Rawlsian maximin criterion.¹⁾ Alternatively, dropping SY and strengthening HOM to S.INV would imply that $\rho = 0$, yielding a whole family of non-symmetric Nash solutions.²⁾

It should be noted that the Utilitarian and the Rawlsian maximin swf's do not yield well defined solutions, since their maximizers are not always unique on the domain considered here. One may then consider single-valued selections, at the cost of relaxing either PO or CONT. For example, keeping PO and dropping CONT will admit the Leximin solution studied in the previous section.

1) See Roberts (1980) for related results in the Arrow tradition of social choice theory.

2) This family of solutions has been studied by Harsanyi and Selten (1972), Kalai (1977b) and Roth (1979b).

1.4 Conclusion

We have attempted in this chapter to view the problem of allocating costs and benefits among a group of individuals as one of bargaining or fair division. A generalization, due to Thomson (1983a), of Nash's (1950) model of the bargaining problem has been used to explore the consequences of an axiom, due to Harsanyi (1959), which can be seen as a requirement that the solution to the allocation problem should be decentralizable in a certain sense.

Although this way of looking at the problem is fairly abstract, it does give some insight that may be useful when trying to solve a concrete problem. Our main result is that a certain amount of collective rationality in the decision making process is a necessary prerequisite for an allocation procedure to be decentralizable. Thus, when faced with a practical problem, the theory tells us to look for a social welfare function in order to rank the given physical alternatives. In order to take care of the decentralization aspect, the social welfare function should be additively separable in individual utility levels. For practical purposes, this means that the composite function $f_i(x_i(\cdot))$, where x_i is agent i 's unobservable utility function and f_i is the i 'th component of the social welfare function, can be looked upon as a standard of living index for agent i , depending on the physical benefits or costs allocated to i .

This shows that the problem of solving allocation problems in a decentralized setting is similar to one of establishing a procedure for project evaluation in a public sector. In both cases, the basic problem consists in specifying an appropriate set of standard of living indices for (groups of) individuals to be used as a criterion for selecting among the physical alternatives

available in any given choice situation. Moreover, in a given choice situation, a description of the problem consists in specifying the effect of each physical alternative on the standard of living index for each individual or group of individuals. Thus, our results suggest that if one is interested in normative aspects of collective choice in a decentralized setting, then the problem can be attacked by means of the familiar tools of cost-benefit analysis.

CHAPTER 2

STABILITY AND THE NASH SOLUTION

2.1 Introduction

In his classic essay on the bargaining problem, Nash (1950) showed that under four axioms describing the behavior of the agents, there exists a unique solution to such a problem. Originally developed in the context of two-person cooperative bargaining, Nash's model has been elaborated by Harsanyi (1959), (1963) and (1977), who showed how the problem of solving n-person bargaining problems could be reduced to the more familiar one of solving two-person problems. He argued that in any n-person bargaining problem, a particular payoff-vector "...will represent the equilibrium outcome of bargaining among the n players only if no pair of players i and j has any incentive to redistribute their payoffs between them, as long as the other players' payoffs are kept constant." (Harsanyi (1977) p. 196). This condition, which Harsanyi calls Bilateral Equilibrium, was shown to imply that if, among a group of n agents, all two-person bargaining problems were solved by the two-person Nash solution, then all n-person bargaining problems had to be solved by the n-person Nash solution. Put differently, Harsanyi's condition states that in their search for a solution outcome to an n-person bargaining problem, the participants should look to the principles that would guide them in solving two-person problems, and his result shows that if these principles happen to be those of Nash, then this will completely determine the solution outcome to the n-person problem. In this way, the question of

how to solve an n -person bargaining problem is reduced to the question of whether or not Nash's axioms are acceptable as principles for solving two-person problems.

The one of Nash's axioms which has been most controversial is his Independence of Irrelevant Alternatives (IIA). Motivated by the objections that have been raised, several authors have investigated the consequences of replacing IIA with other assumptions, and have arrived at other solutions. (See e.g. Kalai and Smorodinsky (1975), Roth (1977) and Thomson (1981a). In this chapter, we replace IIA with Harsanyi's condition and give a new characterization of the Nash solution based on this condition and those of Nash's axioms that are usually accepted.

In section 2.2, we present the model and the axioms. Section 2.3 contains the main result, and in section 2.4 we discuss some variants of it. Section 2.5 contains some concluding remarks.

2.2 The model

The classical axiomatic model of bargaining involves a fixed set P of agents, each equipped with a von Neumann-Morgenstern utility function. P may be taken to be a non-empty, finite subset of the natural integers. Let $|P|$ be the number of elements in P , and let \mathcal{R}^P be the $|P|$ -dimensional euclidean space, indexed by the members of P . \mathcal{R}_+^P and \mathcal{R}_{++}^P denotes the non-negative and the strictly positive orthant of \mathcal{R}^P , respectively. Given $x, y \in \mathcal{R}^P$, we write $x \underline{\geq} y$ if $x - y \in \mathcal{R}_+^P$, $x > y$ if $x - y \in \mathcal{R}_{++}^P$ and $x \underline{>} y$ if $x \underline{\geq} y$ and $x \neq y$.

A $|P|$ -person bargaining problem is a pair (S, d) where S is a subset of \mathcal{R}^P , d an element of S , and where S satisfies the following properties:

A1: S is compact and convex.

A2: There exists $y \in S$ such that $y > d$.

S is the set of utility allocations that can be achieved by the members of P through unanimous agreement, and d is the outcome that will result if they fail to agree. Thus, all subcoalitions of P can veto any outcome different from d , while cooperation by all agents is required in order to achieve another outcome. The existence of a point in S which strictly dominates d guarantees that all agents are non-trivially involved in the bargaining problem. The compactness of S is a technical assumption, convexity follows if the agents may jointly randomize between outcomes.

Let (S, d) be a bargaining problem for the set P of agents. Since von Neumann-Morgenstern utility functions are only unique up to positive affine transformations, we may follow Nash (1950) and simplify the notation by using

utility representations such that $d_i = 0$ for all $i \in P$. We may then identify any bargaining problem (S, d) by the set S only. We will also restrict the family of bargaining problems under consideration to sets S with the following additional properties:

A3: S is a subset of \mathbb{R}_+^P .

A4: If $x \in S$ and $y \in \mathbb{R}_+^P$ such that $y \leq x$ then $y \in S$.

A4 is usually referred to as comprehensiveness and amounts to assuming free disposal of utility. Let Σ^P be the class of bargaining problems S for the set P of agents, such that S satisfies A1-A4. In general, we will refer to such problems as (collective) choice problems, because we shall consider other interpretations of the mathematical model in addition to that of bargaining.

A solution to the $|P|$ -person choice problem is a function $F: \Sigma^P \rightarrow \mathbb{R}_+^P$ such that for all S in Σ^P , $F(S)$ is an element of S .

Harsanyi's (1959) approach to the bargaining problem differs from the classical one in that he considers bargaining problems involving a varying number of agents. In order to formalize his idea of relating the solution outcomes for $|P|$ -person bargaining problems to the solution outcomes for two-person subproblems, it is convenient to use the following extended solution concept, due to Thomson (1983a):

Let I be the set of natural integers, and let \mathcal{P} be the family of non-empty, finite subsets of I . I may be thought as the set of potential agents. The members of \mathcal{P} will be denoted $P, P', Q \dots$ etc. For each $P \in \mathcal{P}$, let Σ^P denote the family of all choice problems S satisfying A1-A4, for the set P of agents. A solution is then redefined to be a function $F: \bigcup_{P \in \mathcal{P}} \Sigma^P \rightarrow \bigcup_{P \in \mathcal{P}} \mathbb{R}_+^P$ such that $F(S) \in S$ for all $P \in \mathcal{P}$ and all $S \in \Sigma^P$.

We next state four of Nash's axioms, slightly modified to fit this solution concept. A fifth assumption is implicit in Nash's original treatment of the bargaining problem, namely Individual rationality, which says that the disagreement point d is always weakly dominated by the solution outcome. In our model, this assumption is automatically satisfied as a result of assumption A3.

Pareto-optimality (PO): For all $P \in \mathcal{P}$, for all $S \in \Sigma^P$, for all $x \in \mathcal{R}_+^P$, if there exists $y \in S$ with $y \geq x$ then $F(S) \neq x$.

For all $P, P' \in \mathcal{P}$ with $|P| = |P'|$, let $\Gamma^{P, P'}$ be the family of one-to-one functions from P to P' . It will cause no confusion if we sometimes treat $\gamma \in \Gamma^{P, P'}$ as a function from \mathcal{R}^P to $\mathcal{R}^{P'}$, defined by $y \equiv \gamma(x)$ if $y_{\gamma(i)} = x_i$ for all $i \in P$.

Symmetry (SY): For all $P \in \mathcal{P}$, for all $S \in \Sigma^P$, if $\gamma(x) \in S$ for all $x \in S$ and all $\gamma \in \Gamma^{P, P}$, then $F_i(S) = F_j(S)$ for all $i, j \in P$.

A stronger version of the symmetry axiom, that we state for later use, is

Anonymity (AN): For all $P, P' \in \mathcal{P}$ with $|P| = |P'|$, for all $\gamma \in \Gamma^{P, P'}$, for all $S \in \Sigma^P$, $F(\gamma(S)) = \gamma(F(S))$.

For all $P \in \mathcal{P}$, let Λ^P be the family of functions from \mathcal{R}^P to \mathcal{R}^P such that for all $\lambda \in \Lambda^P$, there exists $a \in \mathcal{R}_{++}^P$ such that for all $i \in P$ and all $x \in \mathcal{R}^P$, $\lambda_i(x) = a_i x_i$.

Scale Invariance (S.INV): For all $P \in \mathcal{P}$, for all $S \in \Sigma^P$, for all $\lambda \in \Lambda^P$,
 $F(\lambda(S)) = \lambda(F(S))$.

Independence of Irrelevant Alternatives (IIA): For all $P \in \mathcal{P}$, for all
 $S, S' \in \Sigma^P$, if $S' \subset S$ and $F(S) \in S'$, then $F(S') = F(S)$.

For later reference, we state the following result:

Theorem 2.1 (Nash (1950)): There is a unique solution satisfying PO, SY,
S.INV and IIA. It is the solution N defined by $N(S) \equiv \operatorname{argmax}\{\prod_{i \in P} x_i \mid x \in S\}$
for all $P \in \mathcal{P}$ and all $S \in \Sigma^P$.

The four axioms that characterize the Nash solution seem to fall into two categories that are qualitatively quite different. The first category consists of SY and S.INV, which state that the solution should not depend on information which is not contained in the model. (Nash (1953), Roth (1979b). In particular, S.INV is a reflection of the fact that von Neumann-Morgenstern utility functions are only unique up to positive affine transformation. (The reason why, in the statement of S.INV, the linear transformations λ contain no constant terms, is that we have already used up this degree of freedom by fixing the disagreement outcome at the origin). In the second category are PO and IIA, which both demand some form of collective rationality of the agents. IIA says that if some feasible alternative was declared to be "best" among a set S of feasible alternatives, then it must also be "best" among any subset S' of those alternatives. While this seems to be a reasonable assumption about behavior in choice situations involving only one decision maker, or as a normative condition on collective choice, it may not be a good description of the strategic considerations involved in a bargaining situation, because it explicitly rules out the possibility that

narrowing down the set of feasible alternatives may affect some agent's bargaining position. The example reproduced in figure 2.1 below is often used to illustrate this point. In the example, S' is obtained from S by deleting all alternatives in S that give player 2 a higher payoff than the payoff ascribed to him by the solution outcome to S . To require then that the solution should satisfy IIA is to insist that the change from S to S' has not weakened player 2's bargaining position and that the solution outcome for S' should give him the maximum payoff he could hope for.

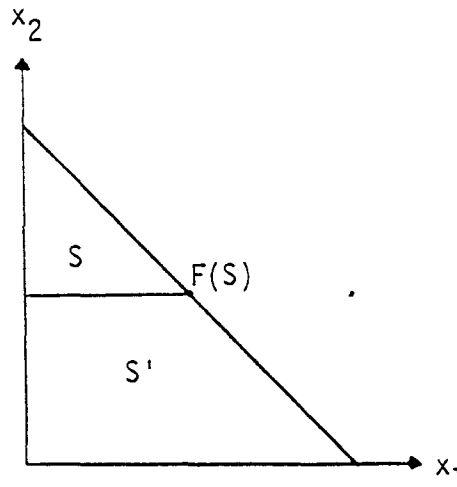


Figure 2.1

An objection to the independence axiom

In this chapter, we develop an alternative characterization of the Nash solution that makes no use of the independence axiom. Instead, we use a version of Harsanyi's condition of Bilateral Equilibrium, which is conceptually more appealing than IIA because it is a statement about individual and not collective rationality. We also show that by using a strengthened version of Harsanyi's condition, the axiom of Pareto-optimality may considerably weakened and still permit a characterization of the Nash solution.

In order to formalize Harsanyi's condition, that we will refer to as Bilateral Stability, the following notation is needed. Given $P, Q \in \mathcal{P}$ with $P \subset Q$ and $x \in \mathbb{R}^Q$ let x_P denote the projection of x on \mathbb{R}^P , and let H_P^x be the hyperplane in \mathbb{R}^Q defined by $H_P^x \equiv \{y \in \mathbb{R}^Q \mid y_{Q \setminus P} = x_{Q \setminus P}\}$. Given $A \subset \mathbb{R}^Q$ and $x \in A$, we denote by $t_P^x(A)$ the projection of $H_P^x \cap A$ on \mathbb{R}^P . We may now state the axiom of

Bilateral Stability (B.STAB): For all $P, Q \in \mathcal{P}$ with $P \subset Q$ and $|P| = 2$, for all $S \in \Sigma^P$ and all $T \in \Sigma^Q$, if $S = t_P^x(T)$ where $x = F(T)$, then $F(S) = x_P$.

An illustration of the axiom is given in figure 2.2 where $Q = \{1,2,3\}$ and $P = \{1,2\}$. It differs from Harsanyi's condition of Bilateral Equilibrium only in that we have explicitly taken account of the possibility that some subproblem $t_P^x(T)$ may not be well defined, by including the provision that $t_P^x(T) \in \Sigma^P$. The subproblems inherit properties A1, A3 and A4 from T , but not necessarily A2, the requirement that $t_P^x(T)$ should contain a strictly positive vector.

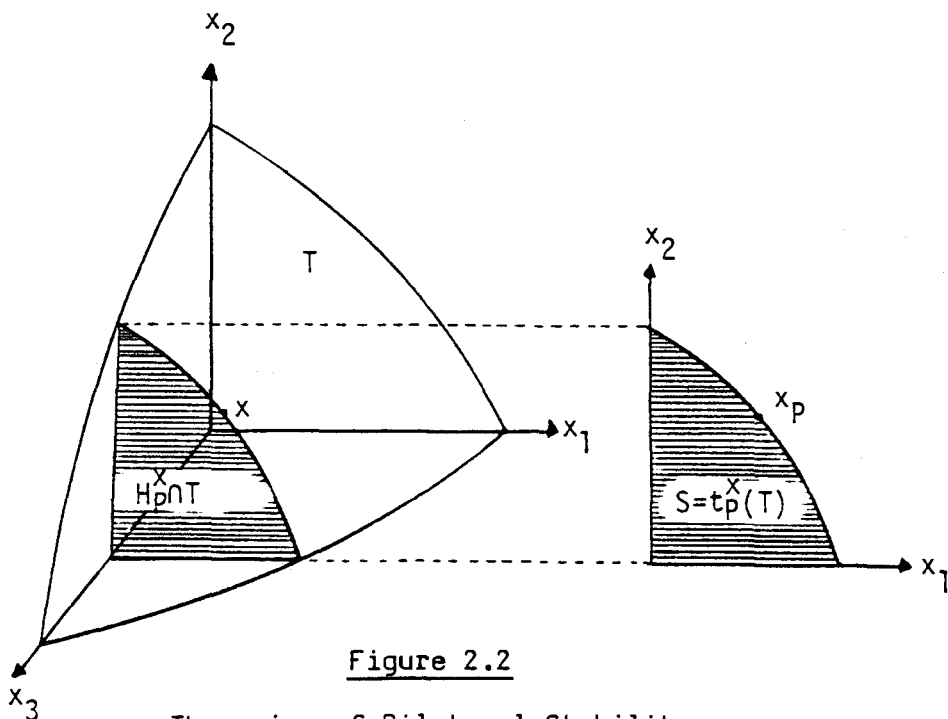


Figure 2.2

The axiom of Bilateral Stability

Harsanyi motivates this condition by pointing out that a rational player i will not accept a tentative agreement x for the $|Q|$ -person problem T if he has reason to believe that he may successfully threaten some other player j to make a concession in his favor. If i does not simultaneously challenge any of the other players for concessions, then i can base his beliefs concerning j 's willingness to concede on what i and j know about solving two-person bargaining problems. Since the situation is similar for all members of Q , then x can be the solution outcome to T only if x agrees with the solution outcomes to all two-person subproblems $t_p^x(T)$ obtained from T by keeping the payoffs to the other players constant at $x_{Q \setminus p}$.

This axiom was used by Harsanyi to show how the problem of solving n -person bargaining problems could be reduced to one of solving a set of two-person problems, provided these two-person problems were known to be solved by the Nash solution. Formally, we state

Theorem 2.2 (Harsanyi (1959)): If a [continuous] solution F satisfies B.STAB and if F coincides with the Nash solution N for all two-person problems, then $F = N$.

Continuity is an implicit assumption in Harsanyi's proof of the theorem. One version of continuity that would serve the purpose of Theorem 2.2 is

Continuity (CONT): For all $P \in \mathcal{P}$, for all $S \in \Sigma^P$, if $\{S^v\}$ is a sequence from Σ^P , converging in the Hausdorff-topology to S , then $\lim_{v \rightarrow \infty} F(S^v) = F(S)$.

It has been suggested by Raiffa (1953), that a bargaining solution can be thought of as a principle of fair division that an arbiter might use to solve

conflicts among agents with partly opposing interests. Harsanyi's theorem is a good illustration of the type of restriction that the stability axiom imposes on such a division principle, stating that the principle can not be allowed to vary with the number of agents involved in any given division problem.

As an example to illustrate the nature of such a consistency requirement for a principle of fair division, consider the following concrete situation involving a bricklayer, a carpenter and a painter, who have the option to build a house for a certain amount of money. In order to reach an agreement on how to share the money they decide to accept the judgement of an arbiter, who suggests a compromise agreeable to all of them. The contract is signed and the work proceeds in the obvious sequence; the bricklayer does his part of the work, collects his share of the money and leaves the scene. Now suppose that the carpenter refuses to carry out his part of the deal unless there is a redistribution of the remaining funds in his favor. It may then come as a surprise to the painter if the arbiter has changed his mind to support the new demands of the carpenter. Clearly, such a situation can only arise if the arbiter happens to violate the consistency requirement made by B.STAB, at least if only problems involving no more than three agents are considered. As a generalization of this requirement to choice problems involving any number of agents, we consider the following version of the stability axiom that we call

Multilateral Stability (M.STAB): For all $P, Q \in \mathcal{P}$ with $P \subset Q$, for all $S \in \Sigma^P$, for all $T \in \Sigma^Q$, if $S = t_P^x(T)$ where $x = F(T)$, then $F(S) = x_P$.

In other words, the solution should be stable, not only with respect to two-person subproblems, but also with respect to subproblems involving any subset of the original group of agents. Although formally, M.STAB is stronger than B.STAB, it may be a more natural condition to impose. For example, bearing in mind Harsanyi's motivation for B.STAB, there seems to be no a priori reason why a dissatisfied player should limit himself to challenging only one other agent at the time for concessions, if he believes that he may enter into multilateral renegotiations. Interpreted as a notion of fairness, the axiom of Multilateral Stability states that an allocation should not be declared a fair compromise for a given set of agents if it is unfair for some subset of those agents. As pointed out by Balinsky and Young (1982), this seems to be a very natural consistency requirement for any notion of fairness. Independently of Harsanyi's work, they have used an axiom in that spirit, called Uniformity, in their development of a theory of apportionment, e.g. for allocating seats in a parliament among political parties in agreement with the proportion of votes obtained by each party.

We close this section by introducing some additional concepts and notation.

Given a solution F and given $Q \in \mathcal{P}$, $T \in \Sigma^Q$ and $x \in T$, say that x is an F -multilaterally stable point in T if for all $P \subset Q$ with $P \neq Q$, either $t_P^x(T)$ is not a well defined choice problem or $F(t_P^x(T)) = x_P$. Letting $M_F(T)$ denote the set of F -multilaterally stable points in T , we observe that M.STAB is equivalent to requiring that $F(T)$ should belong to $M_F(T)$. The set $B_F(T)$ of F -bilaterally stable points in T is defined similarly, by adding the provision that $|P| = 2$. Clearly, B.STAB says that $F(T) \in B_F(T)$.

For all $P \in \mathcal{P}$ and all $i \in P$, we denote by e_i^P the i 'th unit vector in \mathcal{R}^P . Given $P, Q \in \mathcal{P}$ with $P \subset Q$, e_p^Q denotes the vector $\sum_{i \in P} e_i^Q$. Thus, e_p^Q has all of its coordinates in P equal to 1 and all of its coordinates in $Q \setminus P$ equal to 0. Whenever it is clear that $e_i^P \in \mathcal{R}^P$ and $e_p^Q \in \mathcal{R}^Q$, we drop the superscripts Q and P and write e_i and e_p , respectively.

Given $P \in \mathcal{P}$ and a subset A of \mathcal{R}_+^P , $\text{co}\{A\}$ denotes the convex hull of A , and $\text{cch}\{A\}$ denotes the convex and comprehensive hull of A , defined by $\text{cch}\{A\} \equiv \{x \in \mathcal{R}_+^P \mid \exists y \in \text{co}\{A\}, x \leq y\}$. Given $S \in \Sigma^P$, $\text{PO}(S)$ denotes the set of Pareto-optimal points in S , i.e. $\text{PO}(S) \equiv \{x \in S \mid \nexists y \in S, y \geq x\}$. Similarly, $\text{WPO}(S) \equiv \{x \in S \mid \nexists y \in S, y > x\}$ is the set of weakly Pareto-optimal points in S .

2.3 The main result

In this section, we show that the Nash solution satisfies PO, AN, S.INV and M.STAB, and that it is the only one to do so. We begin with

Proposition 2.1: The Nash solution satisfies PO, AN, S.INV. and M.STAB.

Proof: We need only show that the Nash solution satisfies M.STAB, since it is well known that it satisfies the other three axioms. Let $P, Q \in \mathcal{P}$ with $P \subset Q$ and $T \in \Sigma^Q$ be given, and let $z \equiv N(T)$ and $S \equiv t_P^Z(T)$. Then z maximizes the Nash product $\prod_{i \in Q} x_i$ on T , which implies that z maximizes $\prod_{i \in Q} x_i$ on $H_P^Z \cap T = S \times \{z_{Q \setminus P}\}$ (Cartesian product). Since T contains a strictly positive vector, then $z > 0$ and since $x_{Q \setminus P} = z_{Q \setminus P}$ for all $x \in S \times \{z_{Q \setminus P}\}$, it follows that z_P maximizes $\prod_{i \in P} x_i$ on S . Thus, $z_P = N(S)$, as required by M.STAB. QED.

For all $P \in \mathcal{P}$, let Σ_E^P be the family of choice problems S whose Nash solution outcome is Egalitarian, meaning that $N(S) = E(S)$, where $E(S)$ is the unique maximal point in S with equal coordinates. Observe that for all $P \in \mathcal{P}$ and all $S \in \Sigma_E^P$, there exists a positive linear transformation λ such that $\lambda(S) \in \Sigma_E^P$, thus if a solution F satisfies S.INV and if $F(S) = N(S)$ whenever $N(S)$ is Egalitarian, then $F = N$.

In order to prove the converse of Proposition 2.1, we first show in Lemma 2.1 that if a solution F satisfies PO, AN and B.STAB, then $F(S) = N(S)$ for most two-person problems S such that $N(S)$ is Egalitarian. In Lemma 2.2, we add S.INV to those axioms and extend this result to all $P \in \mathcal{P}$ and all $S \in \Sigma_E^P$. The meaning of the term "most" will be made precise after having presented the following outline of the main idea involved in the proof of Lemma 2.1.

In figure 2.3(a) is depicted a choice problem $S \in \Sigma_E^P$, where $P = \{1,2\}$, whose solution outcome is to be determined. Note that for S to be a member of Σ_E^P , then S must be supported at $E(S)$ by a hyperplane with normal e_p . We have chosen $S = cch\{\alpha e_p, 2\alpha e_1\}$ because it is a limit case with respect to that requirement. Figure 2.3(b) illustrates an attempt to show that $F(S) = N(S)$ by adding agent 3 and constructing a three-person choice problem T by taking the convex and comprehensive hull of the three slices S^1, S^2 and S^3 , where $S^1 = Sx\{\alpha e_3\}$ and where S^2 and S^3 are obtained from S^1 by counterclockwise permutations of coordinates.

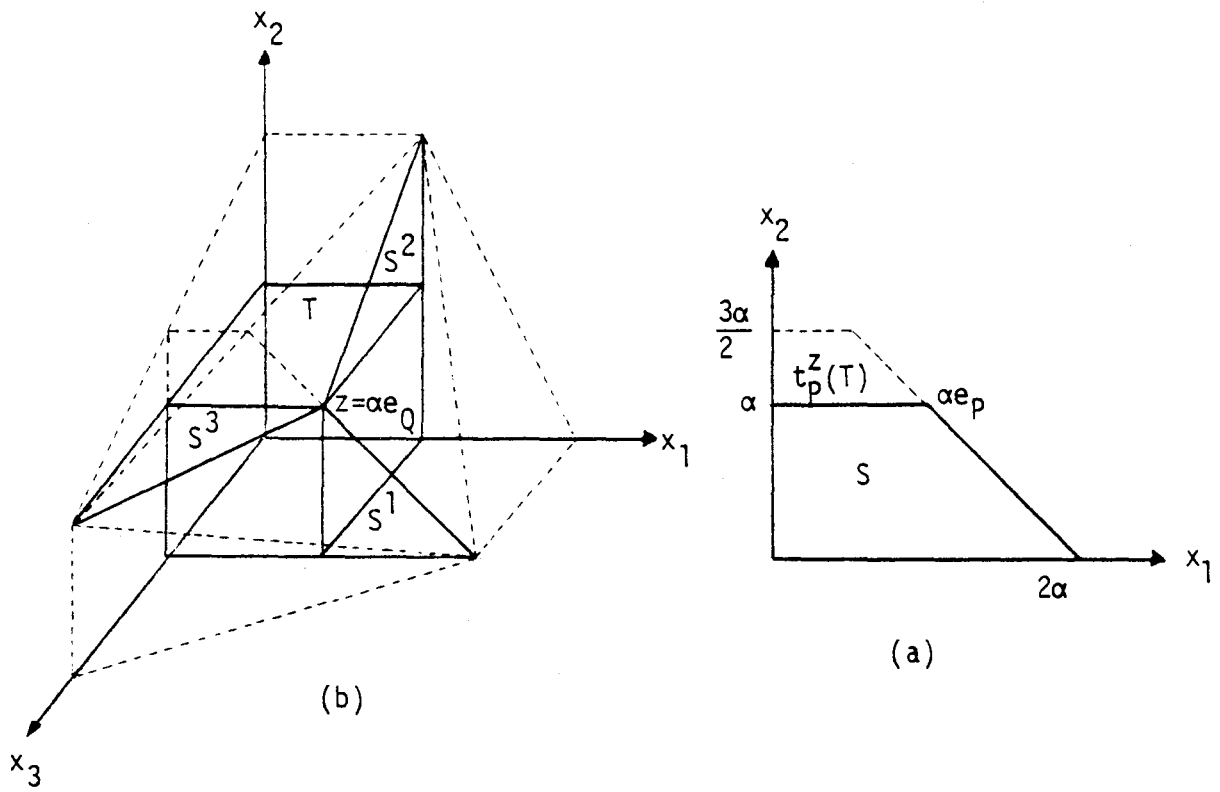


Figure 2.3

The proof of Lemma 2.1

Since all the members of $Q \equiv \{1,2,3\}$ play the same role in the construction of T , and since $z \equiv \alpha e_Q$ is the only Pareto-optimal point of T with equal coordinates, it follows by AN and PO that $F(T) = z$. Now, if it had been the case that $t_p^Z(T)$ were equal to S , then by B.STAB, we would obtain the desired conclusion that $F(S) = z_p = \alpha e_p = N(S)$. Unfortunately, however, by taking the convex and comprehensive hull of the three slices S^1, S^2 and S^3 , something has been added to S , as illustrated in figure 2.3, the reason being that S is a very asymmetric member of Σ_E^P . In fact, it is the worst possible case as regards the difference between $t_p^Z(T)$ and S .

Note that the difference between $t_p^Z(T)$ and S in terms of $\max\{x_2 \mid x \in t_p^Z(T)\} - \max\{x_2 \mid x \in S\}$ is $\alpha/2 = \alpha/(|Q|-1)$. We show in Lemma 2.1 that by adding more than one agent in the construction of T , this difference can be made arbitrarily small. Therefore, if $PO(S)$ happened to coincide in a neighbourhood of $E(S)$ with the symmetric hyperplane supporting S at $E(S)$, then for some finite number of additional agents one would get a choice problem T such that $t_p^Z(T) = S$.

Having outlined the idea of the proof of Lemma 2.1, we define for all $P \in \mathcal{P}$ the set $\Sigma_{EU}^P \subset \Sigma_E^P$ of choice problems such that for all $S \in \Sigma_{EU}^P$, there exists a neighbourhood U of $E(S)$ such that $PO(S) \cap U = H \cap U$, where H is the hyperplane with normal e_p through $E(S)$. (Observe that most choice problems in Σ_{EU}^P are contained in Σ_E^P in the sense that any $S \in \Sigma_E^P$ can be approximated by a sequence of problems from Σ_{EU}^P .) We may now state

Lemma 2.1: If a solution F satisfies PO, AN and B.STAB, then $F(S) = N(S)$ for all $P \in \mathcal{P}$ with $|P| = 2$ and all $S \in \Sigma_{EU}^P$.

Proof: Let $P \in \mathcal{P}$ with $|P| = 2$ and $S \in \Sigma_{EU}^P$ be given and let $\alpha e_P \equiv N(S)$. By AN, we can assume without loss of generality that $P = \{1,2\}$. By definition of Σ_{EU}^P , there exists $\delta > 0$ such that the line segment $[(\alpha+\delta, \alpha-\delta), (\alpha-\delta, \alpha+\delta)]$ is a subset of $PO(S)$. Let n be an integer such that $n \geq \alpha/\delta + 1$ and let $Q \equiv \{1,2,\dots,n\}$. For each $j \in Q$, let γ^j be the permutation of Q such that for all $i \in Q$, $\gamma^j(i) = i + \hat{j}$, where \hat{j} denotes sum modulo n . Let $S^1 \subset \mathcal{R}_+^Q$ be defined by $S^1 \equiv S \times \{\alpha e_{Q \setminus P}\}$. For all $j \in Q \setminus \{1\}$, let $S^j \equiv \gamma^j(S^1)$ and let $T \equiv \text{cch}\{\bigcup_{j=1}^n S^j\}$. Similarly, we construct a stylized version of T , denoted \bar{T} , as follows: Let $\bar{S} \equiv \{x \in \mathcal{R}_+^P \mid \sum_{i \in P} x_i = 2\alpha\}$ and $\bar{S}^1 \equiv \bar{S} \times \{\alpha e_{Q \setminus P}\}$. For all $j \in Q \setminus \{1\}$, let $\bar{S}^j \equiv \gamma^j(\bar{S}^1)$ and $\bar{T} \equiv \text{cch}\{S^1 \cup_{j=2}^n \bar{S}^j\}$. Note that $T \subset \bar{T}$ since $S \subset \bar{S}$. We claim that

(i) $F(T) = z \equiv \alpha e_Q$.

To see this, note first that since all members of Q play the same role in the construction of T , it follows by AN that $F(T)$ must have equal coordinates, which by PO implies that $F(T) = E(T)$. Next, observe that since S is supported at αe_P by the hyperplane $\{x \in \mathcal{R}^P \mid \sum_{i \in P} x_i = 2\alpha\}$, then for all $j \in Q$, S^j is supported at αe_Q by the hyperplane $\{x \in \mathcal{R}^Q \mid \sum_{i \in Q} x_i = n\alpha\}$, and therefore T , which is the convex and comprehensive hull of $\bigcup_{j \in Q} S^j$, is also supported at αe_Q by that hyperplane. This implies that $E(T) = \alpha e_Q$ which proves (i). Next we claim that

(ii) $t_p^Z(T) = S$.

Since $S \subset t_p^Z(T)$ by construction of T , and since $T \subset \bar{T}$, it is sufficient to show that $t_p^Z(\bar{T}) \subset S$, i.e. that $x_p \in S$ for all $x \in H_p^Z \cap \bar{T}$. Since \bar{T} is comprehensive, we may without loss of generality take x to be a Pareto-optimal

point in $H_p^z \cap \bar{I}$. Then x is a convex combination of points (y^1, y^2, \dots, y^n) in $PO(S^1) \times PO(\bar{S}^2) \times \dots \times PO(\bar{S}^n)$, i.e. $x = \sum_{j \in Q} a_j y^j$ for some a in the unit simplex of \mathbb{R}^Q . Since y^1 and x both belong to H_p^z , then $y_{Q \setminus p}^1 = x_{Q \setminus p} = \alpha e_{Q \setminus p}$. For all $j \in Q \setminus \{1\}$, since $y^j \in PO(\bar{S}^j)$, then $y^j = \alpha e_Q + b_j(e_{j+1} - e_j)$ for some b_j in the interval $[-\alpha, \alpha]$. The system $x = \sum_{j \in Q} a_j y^j$ may then be written more explicitly as follows:

$$(1) \quad x_1 = \alpha + (y_1^1 - \alpha)a_1 + a_n b_n$$

$$(2) \quad x_2 = \alpha + (y_2^1 - \alpha)a_1 - a_2 b_2$$

$$(3) \quad 0 = a_2 b_2 - a_3 b_3$$

⋮

$$(n) \quad 0 = a_{n-1} b_{n-1} - a_n b_n$$

Note first that since $a_j \geq 0$ and $\sum_{j=1}^n a_j = 1$ and $-\alpha \leq b_j \leq \alpha$ for all $j \in Q \setminus \{1\}$,

then $-\alpha(1-a_1) \leq \sum_{j=2}^n a_j b_j \leq \alpha(1-a_1)$. Moreover, since $a_2 b_2 = a_3 b_3 = \dots = a_n b_n \equiv \beta$

by equations (3)-(n), it follows that $\sum_{j=2}^n a_j b_j = (n-1)\beta$. Consequently,

$-\alpha(1-a_1)/(n-1) \leq \beta \leq \alpha(1-a_1)/(n-1)$, which since $n \geq \alpha/\delta + 1$ implies that

$-\delta(1-a_1) \leq \beta \leq \delta(1-a_1)$. Using the fact that $\beta = a_2 b_2 = a_n b_n$, we may write

equations (1) and (2) as $(x_1, x_2) = x_p = a_1 y_p^1 + (1-a_1)\alpha e_p + (\beta, -\beta)$, where β

belongs to the interval $[-\delta(1-a_1), \delta(1-a_1)]$, or equivalently, as $x_p =$

$a_1 y_p^1 + (1-a_1)v$, where v is a vector in the segment $[(\alpha+\delta, \alpha-\delta), (\alpha-\delta, \alpha+\delta)]$.

By hypothesis, this segment is a subset of S . But then x_p is a convex combination of y_p^1 and v , which both belong to S , and by convexity of S , then

$x_p \in S$ as well, which proves (ii). That $F(S) = z_p = \alpha e_p$ now follows by (i),

(ii) and B.STAB.

QED.

Lemma 2.2: If a solution F satisfies PO, S.INV and M.STAB, and if $F(S) = N(S)$ for all $P \in \mathcal{P}$ such that $|P| = 2$ and all $S \in \Sigma_{EU}^P$, then $F = N$.

Proof: We must show that F coincides with N on Σ^P for all $P \in \mathcal{P}$. If $|P| = 1$, the proof is trivial. If $|P| \geq 2$, we distinguish between two cases, according to whether $|P| = 2$ or $|P| > 2$. First we show that

(i) For all $P \in \mathcal{P}$ with $|P| = 2$, for all $S \in \Sigma^P$, $F(S) = N(S)$.

Let $P \in \mathcal{P}$ with $|P| = 2$ and $S \in \Sigma^P$ be given, and suppose by way of contradiction that $F(S) \neq N(S)$. By S.INV, we may without loss of generality assume that $N(S) = e_p \neq F(S)$. Let k be an agent who is not a member of P , let $Q \equiv PU\{k\}$ and construct a choice problem $T \in \Sigma^Q$ as follows: Let $S^1 \equiv Sx\{e_k\}$, and let H be the hyperplane $\{x \in \mathbb{R}^Q \mid \sum_{i \in Q} x_i = n\}$. Let $\epsilon > 0$ be given, and let C^ϵ be the cone with vertex $(1+\epsilon)e_k$, spanned by S^1 . By choosing ϵ sufficiently close to 0, the point ne_i belongs to the interior of C^ϵ for all $i \in P$. Define $T \equiv C^\epsilon \cap \text{cch}\{H \cap \{x \in \mathbb{R}_+^Q \mid x_k \leq 1\}\}$. T is illustrated in figure 2.4 for $Q = \{i, j, k\}$. Note that H supports S^1 and T at e_Q and that for all $i \in P$, the open line segment $\sigma_i \equiv (e_Q, ne_i)$ lies in the interior of $H \cap \text{PO}(T)$, relative to H .

Let $y \equiv e_Q$. Since $F(S) \neq y_p$ and $t_p^y(T) = S$, it follows by B.STAB that $z \equiv F(T) \neq e_Q$. Therefore, since $\max\{x_k \mid x \in T\} = 1$, there exists $j \in P$ such that $z_j > 1$. If $z_j = n$, then $t_p^z(T) = \bar{S} \equiv \{x \in \mathbb{R}_+^P \mid \sum_{i \in P} x_i \leq n\} \in \Sigma_{EU}^P$, which by Lemma 2.1 implies that $F_j(\bar{S}) = N_j(\bar{S}) = n/2 \neq n$, in contradiction with B.STAB. Consequently, $z_j < n$. Let $\{i\} \equiv P \setminus \{j\}$ and let $P' \equiv \{i, k\}$. Since $1 < z_j < n$, there exists a point v in the open segment $\sigma_j = (e_Q, ne_j)$ such

that $v \in H_{p,NT}^z$. This is the situation depicted in figure 2.4. Since this segment is contained in the interior of $H \cap PO(T)$, relative to H , and since v_p , has equal coordinates, then $S' \equiv t_{p'}^z(T) \in \Sigma_{EU}^{P'}$ and therefore, by Lemma 2.1, $F(S') = N(S') = v_p$, which by B.STAB implies that $z_{p'} = v_{p'}$, i.e. that $z_i = z_k$. Letting $S'' \equiv t_p^z(T)$, it follows by a similar argument that $S'' \in \Sigma_{EU}^P$ and that $z_i = z_j$, thus z must have equal coordinates which by PO implies that $z = e_Q$, a contradiction. Consequently, $F(S) = e_p$, which proves (i).

To complete the proof of the lemma, we first repeat the argument in the first paragraph of the proof of (i), except for dropping the assumption that $|P| = 2$. Then we proceed as follows:

Observe first that by construction, T has the property that for all $i \in P$ and all $z \in T$, if $x \in \mathbb{R}_+^Q$ such that $x_k = 0$, $x_i = z_i + z_k$ and $x_j = z_j$ for all $j \in Q \setminus \{i,k\}$, then $x \in T$ as well.

Let $z \equiv F(T)$. We claim that $z_i \geq z_k$ for all $i \in P$. This is clearly the case if $z_k = 0$. Suppose $z_k > 0$, let $i \in P$ be given, let $P' \equiv \{i,k\}$ and $S' \equiv t_{p'}^z(T)$. By convexity of S' and the previous paragraph, it follows that all points in the segment $\sigma \equiv [(z_i, z_k), (z_i + z_k)e_i]$ belong to S' . This implies that $S' \in \Sigma^{P'}$ and therefore, by B.STAB and (i), then $F(S') = z_{p'} = N(S')$, which implies that $z_i z_k \geq x_i x_k$ for all $x \in \sigma$. Thus, $z_i z_k \geq [\alpha z_i + (1-\alpha)(z_i + z_k)] \cdot [\alpha z_k + (1-\alpha) \cdot 0]$ for all $\alpha \in [0,1]$, which implies that $z_i \geq z_k$, as claimed.

Since $z_i \geq z_k$ for all $i \in P$, it follows by PO that $z \in A \equiv \text{co}\{e_Q, (ne_i)_{i \in P}\}$, which is a subset of $H \cap PO(T)$. Let $y \equiv e_Q$. Since $F(S) \neq e_p$ and $t_p^y(T) = S$, it follows by B.STAB that $z \neq y$. Therefore, since $z \in A$, then $z_k < 1$, which together imply that z belongs to the interior of $H \cap PO(T)$, relative to H .

Since $z \in B_N(T)$ by B.STAB and (i), this implies that $z \in B_N(\bar{T})$, where $\bar{T} \equiv \text{cch}\{HNR_+^Q\}$. However, this is impossible, since y is the only member of $B_N(\bar{T})$ and since $z \neq y$. Thus, the assumption that $F(S) \neq N(S)$ has led to a contradiction. QED.

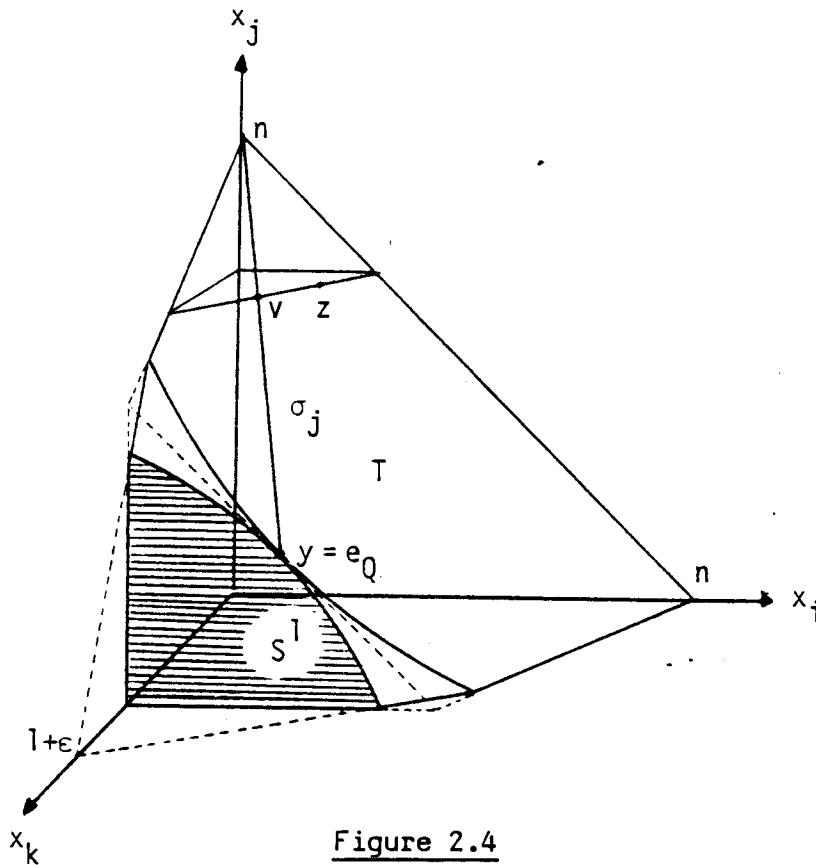


Figure 2.4

The proof of Lemma 2.2

From Lemma 2.1 and Lemma 2.2, we obtain the following converse of Proposition 2.1:

Proposition 2.2: If F satisfies PO, AN, S.INV and M.STAB, then $F = N$, the Nash solution.

Thus we have

Theorem 2.3: A solution satisfies PO, AN, S.INV and M.STAB if and only if it is the Nash solution.

If we compare this result to Nash's own characterization, (Theorem 2.1), we see that except for a strengthening of SY to AN, the only difference is that IIA has been replaced by M.STAB. As regards the connection to Harsanyi's Theorem 2.2, we observe that, except for dropping the hypothesis that two-person problems are solved by the Nash solution, Theorem 2.3 uses the stronger version of the stability axiom, while no continuity assumption is needed. The one result which allows us to dispense with the assumption that all two-person problems are solved by the Nash solution is Lemma 2.1, which together with S.INV and CONT would imply that assumption. Note however, that the proof of Lemma 2.1 relies heavily on the availability of an infinite number of agents, while only a finite number of potential agents is needed for Harsanyi's result. In the next section, we establish some variants of the main result, some of which involve only a finite number of potential agents.

2.4 Variants of the main result

In the proof of Lemma 2.1, as well as that of Lemma 2.2, essential use was made of the assumption that the set I of potential agents be infinite. Could Theorem 2.3 still be true if I were a finite set? The answer turns out to be negative, as will be shown by constructing a solution which satisfies the four axioms of Theorem 2.3, but which differs from the Nash solution.

Proposition 2.3: If the number of potential agents is finite, then the Nash solution is not the only one to satisfy PO, AN, S.INV and M.STAB.

Proof: Note first that if $|I| \leq 2$, then M.STAB loses most of its power, so that any solution satisfying PO, AN and S.INV will also satisfy M.STAB. An example of such a solution is the Kalai-Smorodinsky solution (Kalai and Smorodinsky (1975)).

Next, suppose that $I = \{1, 2, \dots, n\}$, where $n \geq 3$. Let the solution F be defined by $F(S) = N(S)$ for all $P \in \mathcal{P}$ and all $S \in \Sigma^P$, except for $S \in \Sigma_T^I$, where Σ_T^I is a family of choice problems to be defined next. Let $T \in \Sigma_T^I$ be defined by $T \equiv \text{cch}\{e_I, z\}$, where $z \equiv (3/2)e_{I \setminus \{n\}}$. T is illustrated in figure 2.5 for $n = 3$. It is straightforward to verify that all points in the segment $[e_I, z]$ are Nash-multilaterally stable points of T . Consequently, any point in that segment could be the solution outcome to T without violating M.STAB. For example, we may define $F(T) \equiv (e_I + z)/2$.

Let now Σ_T^I be the subset of Σ^I , such that for all $T' \in \Sigma_T^I$, there exists a linear transformation $\lambda \in \Lambda^I$ and a permutation $\gamma \in \Gamma^{I, I}$ such that $T' = \gamma(\lambda(T))$ and define $F(T') \equiv \gamma(\lambda(F(T)))$. Since $N(T) = e_I \neq F(T)$, we now

have a solution which satisfies PO, AN, S.INV and M.STAB, but which differs from the Nash solution. QED.

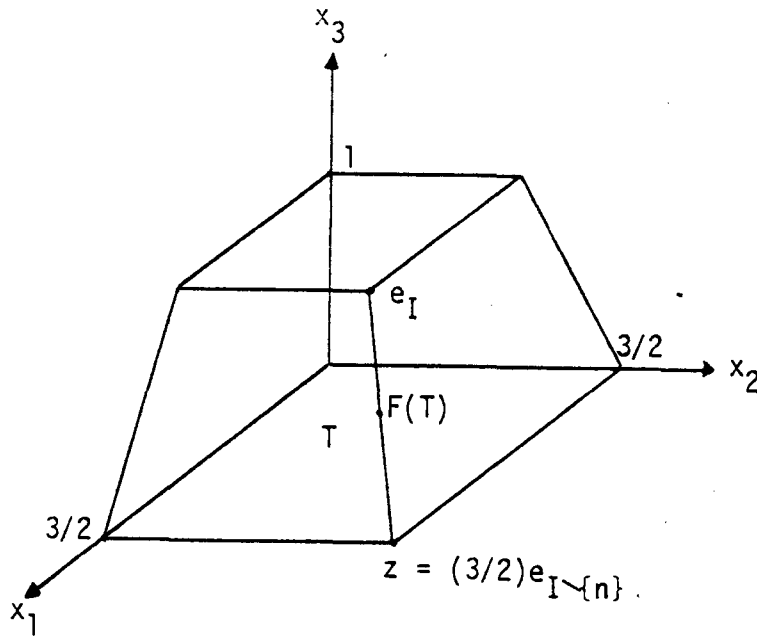


Figure 2.5

The proof of Proposition 2.3

A condition that would eliminate solutions such as F above, is Continuity. (Clearly, F is not continuous at T). What we do next is to show that by adding CONT to the list of axioms in Proposition 2.2, the conclusion of Proposition 2.2 still holds if I is finite, as long as $|I| \geq 3$, and in fact, that it continues to hold if M.STAB is weakened to B.STAB. We begin with Lemma 2.3, which says that if CONT is added to the list of axioms in Lemma 2.1, then the solution must coincide with the Nash solution for all two-person problems S such that $N(S)$ is Egalitarian, and not only for most such problems, as in the conclusion of Lemma 2.1.

Lemma 2.3: Suppose the number of potential agents is at least 3. If a solution F satisfies PO, AN, CONT and B.STAB, then for all $P \in \mathcal{P}$ with $|P| = 2$ and all $S \in \Sigma_E^P$, $F(S) = N(S)$.

Proof: See figure 2.6 for an illustration. Let $P \in \mathcal{P}$ with $|P| = 2$ and $S \in \Sigma_E^P$ be given, and let H be the hyperplane with normal e_p supporting S at $\alpha e_p \equiv E(S)$. To show that $F(S) = \alpha e_p$ also, we first assume that $S \in \Sigma_{EU}^P$, i.e. that

(i) $PO(S) \cap U = H \cap U$ for some neighbourhood U of αe_p .

Let now $S' \equiv \{x \in \mathbb{R}_+^P \mid \sum_{i \in P} x_i \leq 2\alpha\}$, and observe that $F(S') = \alpha e_p$ by PO and AN. Let k be an agent who is not a member of P and let $Q \equiv PU\{k\}$. Define $S^1 \equiv Sx\{\alpha e_k\}$ and for all $\varepsilon \geq 0$, let C^ε be the cone with vertex $(\alpha + \varepsilon)e_k$, spanned by S^1 . Define $T^\varepsilon \equiv C^\varepsilon \cap \text{cch}\{S^1\}$ and $U^1 \equiv Ux\{\alpha e_k\}$ and note that for all $\varepsilon \geq 0$, $U^1 \cap S^1 \subset T^\varepsilon$ and $T^\varepsilon \in \Sigma^Q$.

Let $z \equiv F(T^0)$. We claim that $z = \alpha e_Q$. To see this, note that whatever z is, it follows by construction of T^0 that $t_p^z(T^0) = S'$, which by B.STAB implies that $z_p = \alpha e_p$. Since $y \equiv \alpha e_Q$ is the only Pareto-optimal point in T^0 with the property that $y_p = \alpha e_p$, it follows by PO that $z = \alpha e_Q$.

Consider now $z^\varepsilon \equiv F(T^\varepsilon)$ as $\varepsilon \rightarrow 0$. Since $z^0 = z$ and $T^\varepsilon \rightarrow T^0$ as $\varepsilon \rightarrow 0$, it follows by CONT that $z^\varepsilon \rightarrow z$. Therefore, by PO, there exists $\bar{\varepsilon} > 0$ such that $z^\varepsilon \in U^1$ for all ε in $[0, \bar{\varepsilon}]$, which by B.STAB implies that z_p^ε is constant and equal to $F(S)$ for all such ε . But then $z^\varepsilon = z$ for all ε in $(0, \bar{\varepsilon})$ by the fact that $z^\varepsilon \rightarrow z$ in U^1 as $\varepsilon \rightarrow 0$, which implies that $F(S) = z_p = \alpha e_p$, the desired conclusion.

To complete the proof, it suffices to observe that if S does not satisfy condition (i) above, then we may approximate it by a sequence of elements from Σ_{EU}^P that does, and apply CONT once more to conclude that $F(S) = \alpha e_p$ in this case also. QED.

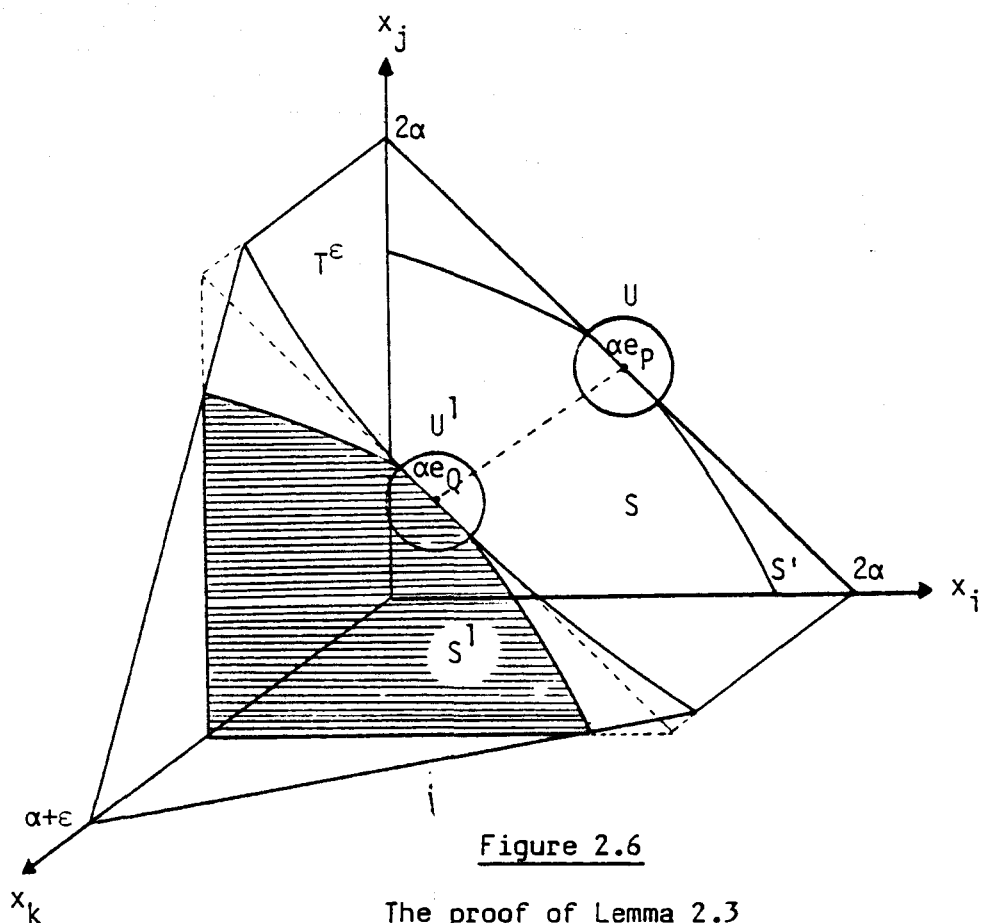


Figure 2.6

The proof of Lemma 2.3

Lemma 2.4: Suppose the number of potential agents is at least 3. If a solution F satisfies $P0$, $S.INV$, $CONT$ and $B.STAB$, and if $F(S) = N(S)$ for all $P \in \mathcal{P}$ such that $|P| = 2$ and all $S \in \Sigma_E^P$, then $F = N$.

Proof: We must show that F coincides with N on Σ^P for all $P \in \mathcal{P}$. If $|P| = 1$, the proof is trivial, if $|P| = 2$, it follows by $S.INV$ and the hypothesis that F coincides with N on Σ_E^P . Let $Q \in \mathcal{P}$ with $|Q| \geq 3$ and $T \in \Sigma^Q$ be given. In order to show that $F(T) = N(T)$, we first assume that

(i) $PO(T) = WPO(T)$ and

(ii) For all $x \in WPO(T) \cap \mathbb{R}_{++}^Q$, there exists a unique hyperplane supporting T at x .

Let $z \equiv F(T)$. We claim that $z > 0$ and prove it by contradiction. If $z_i = 0$ for some $i \in Q$, then by PO , there exists $j \in Q$ such that $y_j > 0$. Let $P \equiv \{i, j\}$ and $S \equiv t_P^Z(T)$. Since $z_P \neq 0$, it follows by (i) that S contains a strictly positive vector, which implies that S is a well defined choice problem. Since $|P| = 2$, it follows by the first part of the proof that $F(S) = N(S) > 0$, and by $B.STAB$ that $z_P > 0$, which is the announced contradiction.

Since $z > 0$, then by (ii), there exists a unique hyperplane supporting T at z . Let $G \equiv \{x \in \mathbb{R}_+^Q \mid \prod_{i \in Q} x_i \geq \prod_{i \in Q} z_i\}$, and let H_N be the unique hyperplane supporting G at z . We claim that $H_F = H_N$, and prove it by contradiction.

If $H_F \neq H_N$, then for some $P \subset Q$ with $|P| = 2$, $t_P^Z(H_F) \neq t_P^Z(H_N)$. Since $z > 0$ then (ii) implies that $t_P^Z(H_F)$ is the unique tangent to $S' \equiv t_P^Z(T)$ at z_P . Since $|P| = 2$, then $F(S') = N(S')$, which by $B.STAB$ implies that $z_P = N(S')$. Thus, z_P maximizes $\prod_{i \in P} x_i$ on S' , and therefore $t_P^Z(H_N)$ is tangent to S' at z_P . Since $t_P^Z(H_N) \neq t_P^Z(H_F)$, this is a contradiction which proves that $H_F = H_N$ and hence that $z = N(T)$.

The proof of the lemma is completed by observing that any $T \in \Sigma^Q$ can be approximated by a sequence of elements of Σ^Q that satisfy (i) and (ii), and by applying $CONT$. QED.

By Proposition 2.1, Lemma 2.3, Lemma 2.4 and the fact that the Nash solution satisfies CONT, we now obtain

Theorem 2.4: Suppose the number of potential agents is at least 3. A solution satisfies PO, AN, S.INV, CONT and B.STAB if and only if it is the Nash solution.

It should be mentioned that the only if-part of Theorem 2.3 can be strengthened by weakening AN to SY. This is possible since AN was only used once in the proof of Theorem 2.4, namely in Lemma 2.3 in order to determine the solution outcome for a symmetric choice problem.

Going back to Theorem 2.3, it turns out that the only if-part of that result can be strengthened as well, by replacing PO with the following condition, that we call

Individual Optimality (IO): For all $i \in I$, for all $S \in \Sigma^{\{i\}}$, $F(S) = \max\{x \mid x \in S\}$.

IO states that all choice problems involving only one agent are solved by maximizing the utility of that agent. Because IO and M.STAB together imply PO, we obtain

Theorem 2.5: A solution satisfies IO, AN, S.INV and M.STAB if and only if it is the Nash solution.

2.5 Concluding remarks

The Nash solution has often been criticized for its reliance on the axiom of Independence of Irrelevant Alternatives. This axiom, as well as Pareto-optimality, is an assumption of collective rationality in conflict situations. We have shown that the independence axiom can be dispensed with, by providing a characterization of the Nash solution which makes no use of it. Instead, we use a stability condition, due to Harsanyi (1959), which is conceptually more appealing, since it is motivated by considerations of individual, and not collective, rationality. Moreover, a version of this axiom also allows us to weaken Pareto-optimality to Individual Optimality and still get a characterization of the Nash solution.

CHAPTER 3

STABILITY AND THE LEXIMIN SOLUTION

3.1 Introduction

A characteristic feature of the Nash solution is that it does not assume utility to be interpersonally comparable. This is so because the Nash solution satisfies S.INV, which states that the solution outcome should be invariant with respect to changes in the unit of measurement for the utility of any agent. Throughout this study, the agents' utility functions are assumed to have cardinal significance, thus if two utility functions both represent the preferences of a given agent, then they can at most differ by a positive affine transformation. Since the constant terms of these positive affine transformation have been used up by selecting an origin of utility space, one is left with one degree of freedom for each utility function, namely the choice of a unit of measurement.

The restrictions that S.INV imposes on the response of the Nash solution to changes in the agents' utility functions automatically carry over to its response to changes in the underlying set of physical alternatives, because a solution depends on these physical alternatives only through their image in utility space. There is then a potential conflict between the requirement that the solution should not involve interpersonal comparisons, and a requirement that the response of the solution to changes in the set of physical alternatives should agree with observed behavior in bargaining situations (if the model is used descriptively) or with common notions of

fairness (if used normatively). This conflict has been discussed extensively by Roth (1979b), and also by Nydegger and Owen (1974) and Roth and Malouf (1979), (1982), who perform experiments showing that the axiom of Scale Invariance is systematically violated in bargaining situations whenever the players have complete information about the underlying sets of physical alternatives.

As an example in the normative spirit, consider a situation in which the set of physical alternatives is expanded, leading to a corresponding expansion in the set of feasible utility allocations. Because the Nash solution satisfies IIA, either the solution outcome does not change or one of the new alternatives is selected as the solution outcome to the new problem. However, IIA does not say anything about how the solution outcome should change, if it changes as a result of an expansion is the feasible set of alternatives. For example, if the expansion consists in adding new alternatives that are particularly favorable to some agent, it would be natural to require that this agent should gain, or at least should not be worse off as the result of such a change in the problem. Various axioms in that spirit have been proposed and used in the literature, e.g. by Kalai and Smorodinsky (1975), Kalai (1977a), Roth (1979a) and Thomson and Myerson (1980). Before we state Kalai's (1977a) version of this axiom, we define for all $P, Q \in \mathcal{P}$ with $P \subset Q$ and all $A \subset \mathcal{R}^Q$, the set A_P to be the projection of A on \mathcal{R}^P .

Individual Monotonicity (I.MON): For all $P \in \mathcal{P}$, for all $S, S' \in \Sigma^P$, for all $i \in P$, if $S \subset S'$ and $S_{P \setminus \{i\}} = S'_{P \setminus \{i\}}$, then $F_i(S') \geq F_i(S)$.

An illustration is given in figure 3.1, where $P = \{i, j\}$. The feasible set of utility vectors in S for the agents in $P \setminus \{i\}$ is $S_{P \setminus \{i\}}$ and this set is not affected by the expansion from S to S' . Moreover, for each such vector, the maximum feasible utility for agent 1 increases, or at least does not decrease. In fact, as is the case in figure 3.1, some of the new alternatives in S' may be unambiguously more favorable to agent i as compared to any alternative in S , while this will not be the case for the agents in $P \setminus \{i\}$, since $S_{P \setminus \{i\}} = S'_{P \setminus \{i\}}$. It seems reasonable therefore to require that agent 1 should not lose.

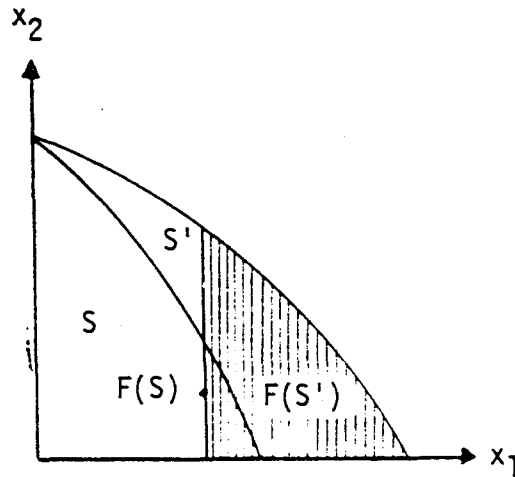


Figure 3.1

The axiom of Individual Monotonicity

In figure 3.2, an example is given which shows that the Nash solution does not satisfy I.MON. As is clear from the figure, the reason why I.MON is violated here is that the level curves of the Nash product $\prod_{i \in P} x_i$ are too responsive to changes in the trade-off between the utility levels of different agents along the Pareto-optimal boundary of the feasible set. The

level curves of the Nash product are characterized by symmetry and a constant elasticity of substitution of -1 , and it is the axiom of Scale Invariance which is responsible for these substitution possibilities.

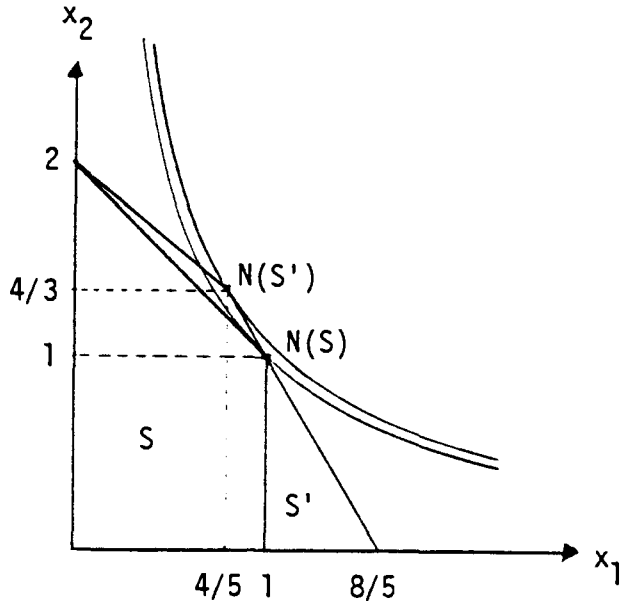


Figure 3.2

The Nash solution does not satisfy I.MON

Thus, while S.INV has the desirable property of eliminating the need for interpersonally comparable utility, it conflicts with certain reasonable criteria of fairness as expressed by I.MON when it is used in conjunction with PO, AN and M.STAB. Consequently, if one is willing to make interpersonal comparisons for the sake of fairness, while preserving as many as possible of the remaining properties of the Nash solution, it is natural to investigate the consequences of replacing S.INV by I.MON in the list of axioms used to characterize the Nash solution in Theorem 2.1.

The first question that will have to be answered, is whether there are any solutions that satisfy PO, AN, I.MON and M.STAB. One possible candidate would be the Egalitarian solution¹⁾ E, which to each choice problem associ-

1) Cf. Raiffa (1953), Myerson (1977), (1981) and Thomson (1983b).

ates the unique maximal point with equal coordinates. This solution does not allow any substitution between the utility levels of different agents, and therefore it would not violate I.MON in the example given in figure 3.2. However, it satisfies neither PO nor M.STAB, as is clear from the example in figure 3.3, where $Q \equiv \{1,2,3\}$, $P \equiv \{2,3\}$ and $T \equiv \text{cch}\{(1,2,2)\}$.

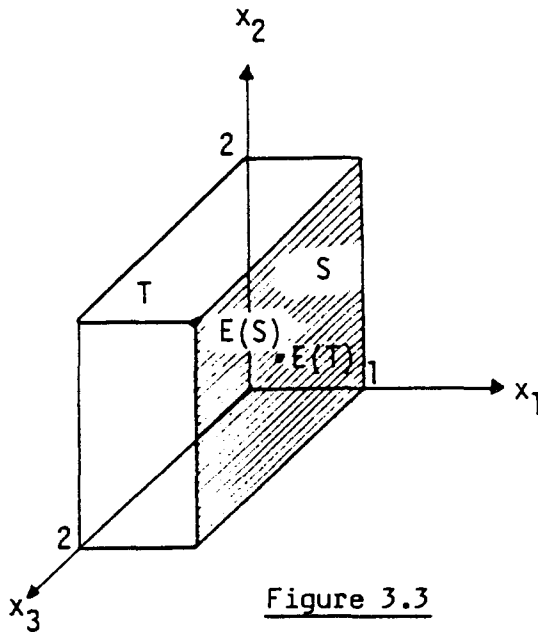


Figure 3.3

The Egalitarian solutions satisfies neither PO nor M.STAB

We have $E(T) = x \equiv e_Q$ and $S \equiv t_p^x(T) = \text{cch}\{(2,2)\}$, so that $E(S) = 2e_p$. Since e_Q is not a Pareto-optimal point in T , then E does not satisfy PO, and since $2e_p \neq z_p = e_p$, it does not satisfy M.STAB either.

The Egalitarian solution is related to the Rawlsian maximin criterion (Rawls (1971)) by always selecting a feasible alternative which maximizes the utility of the worst-off individual. In general, there may be more than one such alternative, as is the case in the example of figure 3.3, where all points on the right side of the rectangle T maximize the utility of agent 1, who turns out to be the worst-off individual in T . However, Sen (1970) has

suggested the following lexicographic extension of the Rawlsian maximin criterion which overcomes this problem of non-uniqueness: First maximize the utility of the worst-off individual, then do the same for the next to worst-off individual and so on, until all possibilities for increasing the utility of any individual have been exhausted. The solution obtained in this way is called the Leximin solution and is denoted L . It is similar to the Egalitarian solution in that it does not permit any substitution between the utility of different agents, but differs from it by being both Pareto-optimal and multilaterally stable. In the example of figure 3.3, $L(T) \equiv z = (1,2,2)$, which is the only Pareto-optimal point in T , and $L(t_p^Z(T)) = (2,2) = z_p$ as required by M.STAB.

It will be shown in section 3.2 that the Leximin solution is the only one that satisfies PO, AN, I.MON and M.STAB. Observe that this list of axioms differs from the one used to characterize the Nash solution in Theorem 2.1 only in that S.INV has been replaced by I.MON. In that respect, our result in this chapter is similar to that of Imai (1983), who gives a characterization of the Leximin solution which parallels Nash's theorem (Theorem 2.1) in the same way. The Leximin solution is then a second example of a solution that can be characterized using either IIA or M.STAB, thus showing that the two axioms share much of the same analytical power.

In the remainder of this section a more formal definition of the Leximin solution will be given. In fact, we will give two equivalent definitions, and the first one is a description of an algorithm that generates the Leximin solution outcome.

Given $P \in \mathcal{P}$ and $S \in \Sigma^P$, the algorithm is initialized at iteration $\tau = 0$ by setting $P^0 = P$ and $x^0 = E(S)$. At any subsequent iteration τ , we set $P^\tau \equiv \{i \in P \mid \exists x \in S, x \geq x^{\tau-1}, x_i > x_i^{\tau-1}\}$ and let the iterate x^τ be given by $x_{P \setminus P^\tau}^\tau \equiv x_{P \setminus P^\tau}^{\tau-1}$ and $x_{P^\tau}^\tau \equiv E(t_{P^\tau}^{x^{\tau-1}}(S))$. Thus at each iteration, the algorithm first identifies the maximal subset P^τ of P such that for all $i \in P^\tau$, it is feasible to increase i 's utility beyond $x_i^{\tau-1}$ without reducing the utility of any other agent. Then the new iterate x^τ is obtained from $x^{\tau-1}$ by increasing the utility of each agent in P^τ by the same amount until no such further increase is feasible. This is equivalent to setting $x_{P^\tau}^\tau$ equal to the Egalitarian solution outcome for the subproblem $t_{P^\tau}^{x^{\tau-1}}(S)$, obtained from S by keeping $x_{P \setminus P^\tau}^{\tau-1}$ constant. The algorithm terminates at the first iteration $\bar{\tau}$ such that $x_{P \setminus P^\tau}^{\bar{\tau}} = x_{P \setminus P^\tau}^{\bar{\tau}-1}$, i.e. when $P^{\bar{\tau}} = \emptyset$. We then set $L(S) \equiv x^{\bar{\tau}}$. Since P is finite and P^τ is a proper subset of $P^{\tau-1}$ whenever $x^\tau \neq x^{\tau-1}$, the algorithm terminates in a finite number of iterations.

The algorithm is illustrated in figure 3.4 for $P = \{1,2,3\}$ and $S = \text{cch}\{(1,2,3)\}$. In this example, $P^0 = P$, $x^0 = (1,1,1)$, $P^1 = \{2,3\}$, $x^1 = (1,2,2)$, $P^2 = \{3\}$, $x^2 = (1,2,3)$, $P^3 = \emptyset$, $x^3 = (1,2,3)$ and $L(S) = (1,2,3)$.

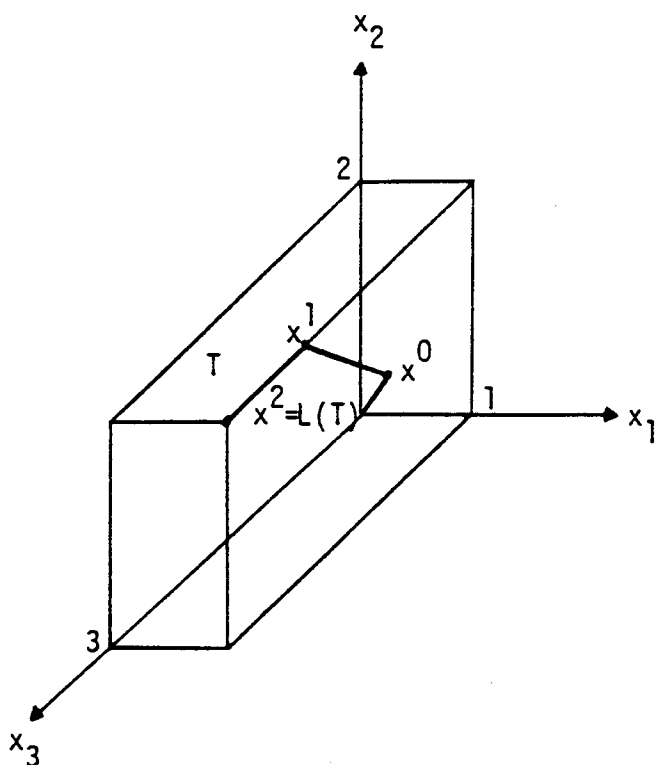


Figure 3.4

The algorithm generating the Leximin solution outcome

While this definition of the Leximin solution is useful as a visualization of its behavior, the next one is more convenient for analytical purposes:

Given $P \in \mathcal{P}$ and $x \in \mathcal{R}_+^P$, let $\bar{P} \equiv \{1, 2, \dots, |P|\}$ and let the vector $\gamma(x) \in \mathcal{R}_+^{\bar{P}}$ be a relabelling of the coordinates of x such that $\gamma_1(x) \leq \gamma_2(x) \leq \dots \leq \gamma_{|P|}(x)$. Let \succ_p be the ordering of \mathcal{R}_+^P such that for all $x, y \in \mathcal{R}_+^P$, $x \succ_p y$ if and only if there exists $k \in \bar{P}$ such that $\gamma_k(x) > \gamma_k(y)$ and $\gamma_i(x) = \gamma_i(y)$ for all $i \in \bar{P}$ such that $i < k$. If and only if $\gamma(x) = \gamma(y)$, then $x \sim_p y$. For all $S \in \Sigma^P$, let $\Lambda(S) \equiv \{y \in S \mid y \succ_p x \text{ for all } x \in S\}$. It follows by a simple adaptation of Lemma 3 and Lemma 4 in Imai (1983) that $\Lambda(S)$ consists of the single point $L(S)$ for all $S \in \Sigma^P$. Thus for all $P \in \mathcal{P}$, the Leximin solution outcome to any $S \in \Sigma^P$ can be found by either applying the algorithm previously described, or by maximizing the Leximin ordering \succ_p over S .

Let \succsim be the family of orderings $\{\succsim_P \mid P \in \mathcal{P}\}$, where each \succsim_P is defined on \mathbb{R}_+^P . It is a well known fact that the family \succsim satisfies the property that for all $P, Q \in \mathcal{P}$ with $P \subset Q$ and all $x, y \in \mathbb{R}_+^Q$ such that $x_{Q \setminus P} = y_{Q \setminus P}$, $x \succsim_Q y$ if and only if $x_P \succsim_P y_P$ (see e.g. d'Aspremont and Gevers (1977) and Deschamps and Gevers (1978)). This property, which is known in the social choice literature as Separability, is due to Fleming (1952), and is closely related to the axiom of Multilateral Stability. This relationship will be discussed in greater detail in chapter 4.

3.2 The main result

We show in this section that the Leximin solution satisfies Pareto-optimality, Anonymity, Individual Monotonicity and Multilateral Stability, and that it is the only one to do so. We begin with

Proposition 3.1: The Leximin solution satisfies PO, AN, I.MON and M.STAB.

Proof: It follows directly by the algorithm which generates the Leximin solution outcome that the Leximin solution satisfies PO and AN. A proof that L satisfies I.MON is given in Imai (1983). It remains to show that it satisfies M.STAB.

Let $Q, P \in \mathcal{P}$ with $P \subset Q$ and $T \in \Sigma^Q$ be given. Let $z \equiv L(T)$. Then z is a maximal element for the ordering \succsim_Q in T . Let $S \equiv t_P^z(T)$ and $S^z \equiv S \times \{z_{Q \setminus P}\}$. Since S^z is a subset of T containing z , it follows that z is a maximal element for \succsim_Q in S^z . Because $x_{Q \setminus P} = z_{Q \setminus P}$ for all $x \in S^z$, then by Separability of the ordering \succsim_Q , it follows that z_P is a maximal element for \succsim_P in S . Thus, $z_P \in \Lambda(S)$ and since $\Lambda(S)$ consists of only $L(S)$, we conclude that $L(S) = z_P$, as required by M.STAB. QED.

We next show that the Leximin solution is the only one to satisfy the four axioms. The proof is organized as follows: The point of departure is Lemma 2.1, restated here as Lemma 3.1, which says that if a solution F satisfies PO, AN and B.STAB, then for all $P \in \mathcal{P}$ with $|P| = 2$, F coincides with the Egalitarian solution E on Σ_{EU}^P , the set of choice problems S whose Pareto-optimal boundary coincides with the hyperplane through $E(S)$ with normal e_p in a neighbourhood of $E(S)$. We then use this result and I.MON in Lemma 3.2 to

show that F must coincide with L for two-person problems that are more and more asymmetric in the sense that the maximal utility available to one of the agents is proportionally greater and greater than the maximal utility available to the other agent. Finally, we extend this result to choice problems that involve more than two agents (Proposition 3.2).

Lemma 3.1 (Lemma 2.1 restated): If a solution F satisfies PO , AN and $B.STAB$, then $F(S) = E(S)$ for all $P \in \mathcal{P}$ with $|P| = 2$ and all $S \in \Sigma_{EU}^P$.

For all $P \in \mathcal{P}$ and all $S \in \Sigma^P$, let the vector $a(S)$ be defined by $a_i(S) \equiv \max\{x_i \mid x \in S\}$ for all $i \in P$. Note that if $P = \{i, j\}$, then for all $S, S' \in \Sigma^P$, we have $a_i(S) = a_i(S')$ if and only if $S_{\{i\}} = S'_{\{i\}}$, by comprehensiveness of S and S' . $I.MON$ then simplifies to saying that $F_j(S') \geq F_j(S)$ whenever $S \subset S'$ and $a_i(S) = a_i(S')$. As a measure of degree of asymmetry of any two-person problem S , we will use the smallest integer $v \equiv v(S)$ such that $a(S) < vE(S)$. Given $P \in \mathcal{P}$ and a subset A of \mathcal{R}_+^P , $\text{int}A$ denotes the interior of A .

Lemma 3.2: If a solution F satisfies PO , AN , $I.MON$ and $M.STAB$, then for all positive integers v , for all $P \in \mathcal{P}$ with $|P| = 2$ and all $S \in \Sigma^P$, if $v(S) = v$ then $F(S) = L(S)$.

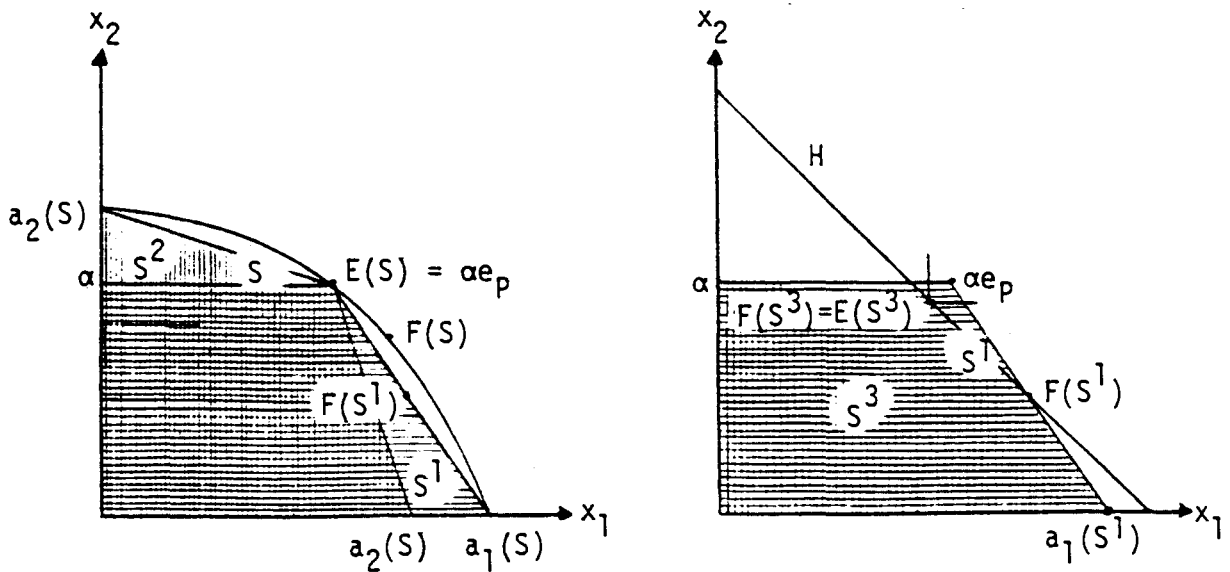
Proof: The proof is by induction on v . If $v = 1$ there is nothing to prove, so assume first that $v = 2$. Let $P \in \mathcal{P}$ with $|P| = 2$ and $S \in \Sigma^P$ with $v(S) = 2$ be given, and suppose, by way of contradiction, that

(1) $F(S) \neq L(S)$.

By AN, we can assume without loss of generality that $P = \{1,2\}$ and that $a_1(S) \geq a_2(S)$. Let $\alpha e_p \equiv E(S)$, define $S^1 \equiv \text{cch}\{a_1(S)e_1, \alpha e_p\}$, and note that $E(S^1) = E(S) = \alpha e_p$ and $v(S^1) = v(S) = 2$. We claim that

$$(2) \quad F_1(S^1) > \alpha \quad \text{and} \quad F_2(S^1) < \alpha.$$

The proof is illustrated in figure 3.5 (a). We first show that (2) holds for S instead of S^1 . The proof is by contradiction. Suppose first that $F_1(S) < \alpha$. Let $S^2 \equiv \text{cch}\{a_2(S)e_1, \alpha e_p, a_2(S)e_2\}$. I.MON implies that $F_1(S^2) \leq F_1(S) < \alpha$. However, S^2 is a symmetric problem, which by PO and AN implies that $F(S^2) = \alpha e_p$, a contradiction. Hence $F_1(S) \geq \alpha$. If $F_1(S) = \alpha$, then $F(S) \geq \alpha e_p$ by PO. Therefore, since $L(S)$ is the only Pareto-optimal point of S which dominates αe_p , it follows by PO that $F(S) = L(S)$, contradicting (1). Consequently, $F_1(S) > \alpha$ and $F_2(S) < \alpha$. Since $F_2(S^1) \leq F_2(S) < \alpha$ by I.MON, then $F_1(S^1) > \alpha$ by PO, which proves (2).



(a) Figure 3.5 (b)

The proof of Lemma 3.2 for $v = 2$

Next, we use Lemma 3.1 to obtain a contradiction to (2) and hence to (1). See figure 3.5 (b) for an illustration. Let H be the hyperplane through $F(S^1)$ with normal e_p and let $S^3 \equiv S^1 \cap \text{cch}\{H \cap \mathbb{R}_+^P\}$. Since $a(S^1) < 2E(S^1) = 2\alpha e_p$, it follows by (2) that αe_p lies strictly above H . This fact has two consequences: First, it implies that $E(S^3) \in H \cap \text{int} S^1$ and therefore $S^3 \in \Sigma_{EU}^P$, which by Lemma 3.1 implies that $F(S^3) = E(S^3)$. Secondly, it implies that $a(S^3) = a(S^1)$, which by I.MON applied twice implies that $F(S^1) \geq F(S^3)$. Since $F_1(S^3) < \alpha$ and $F_1(S^1) > \alpha$, this is in contradiction with (2).

Since (2) is not true, then (1) is not true either, which proves the lemma for $v = 2$. Let now $v \geq 2$ be given, and suppose the lemma holds for all $v' \leq v$. To prove it for $v(S) = v+1$, we first show that (1) implies (2) using the same argument as for $v(S) = 2$. Then we proceed as follows:

Let $Q \equiv \{1, 2, 3\}$ and for all $i \in Q$, let $P^i \equiv Q \setminus \{i\}$. Set $b \equiv (v+1)/v$ and $T \equiv \text{cch}\{S^1 \times \{\alpha e_3\}, \alpha b e_{P^2}\}$. T is depicted in figure 3.6. We are going to show that $F(T) = \alpha e_Q$. Let the segment σ_2 be defined by $\sigma_2 \equiv [\alpha e_Q, \alpha b e_{P^2}]$. We first show that

$$(3) \quad F(T) \in \sigma_2.$$

To see this, let x be any point in T and consider $S^x \equiv t_{P^2}^x(T)$. We claim that $a(S^x) < vE(S^x)$ for all $x \in T$. Note that S^x only depends on x_2 and that $x_2 \in [0, \alpha]$. If $x_2 = \alpha$, then $S^x = \text{cch}\{\alpha e_{P^2}\}$ and $a(S^x) = E(S^x)$, which since $v \geq 2$ implies that $a(S^x) < vE(S^x)$. If $x_2 = 0$, then $S^x = \text{cch}\{(a_1(S^1), \alpha), \alpha b e_{P^2}\}$, $a(S^x) = (a_1(S^1), \alpha b)$ and $E(S^x) = \alpha b e_{P^2}$, which since

$a_1(S^1) < (v+1)E_1(S^1) = (v+1)\alpha$ and $b = (v+1)/v$ and $v \geq 2$ implies that $a(S^X) < vE(S^X)$.

Since $a(S^X)$ and $E(S^X)$ are linear functions of x_2 , this proves the claim that $a(S^X) < vE(S^X)$ for all $x \in T$. It then follows by the induction hypothesis that $F(S^X) = L(S^X) \equiv z^X$ for all $x \in T$. Note that for all $x \in T$, $z^X = E(S^X)$ and therefore $(z_1^X, x_2, z_3^X) \in \sigma_2$. M.STAB then implies that $F(T) \in \sigma_2$, which proves (3).

Let now the segment σ_1 be defined by $\sigma_1 \equiv [\alpha e_0, \alpha(b, c, 1)]$, where $c \equiv (a_1(S^1) - \alpha b) / (a_1(S^1) - \alpha)$. Since $v(S^1) = v+1$ then $a_1(S^1) > \alpha b > \alpha$ and therefore $0 < c < 1$. We claim that

(4) If $F(T) \in \sigma_2$ then $F(T) \in \sigma_1$.

To see this, let x be any point in σ_2 and consider $S^X \equiv t_{p1}^X(T)$. We claim that $a(S^X) < vE(S^X)$ for all $x \in \sigma_2$. Note that S^X only depends on x_1 and that $x_1 \in [\alpha, \alpha b]$. If $x_1 = \alpha$, then $S^X = cch\{\alpha e_2, \alpha b e_3\}$, $a(S^X) = \alpha(1, b)$ and $E(S^X) = \alpha e_{p1}$, which since $b = (v+1)/v$ and $v \geq 2$ implies that $a(S^X) < vE(S^X)$.

If $x_1 = \alpha b$, then $S^X = cch\{\alpha(c, 1), \alpha b e_3\}$, $a(S^X) = \alpha(c, b)$ and $E(S^X) = \alpha(c, c)$ which since $v \geq 2$ implies that $a(S^X) < vE(S^X)$. Since $a(S^X)$ and $E(S^X)$ are linear functions of x_1 for $x_1 \in [\alpha, \alpha b]$, this proves the claim that $a(S^X) < vE(S^X)$ for all $x \in \sigma_2$. It then follows by the induction hypothesis that $F(S^X) = L(S^X) \equiv z^X$ for all $x \in \sigma_2$. Note that for all $x \in \sigma_2$, $z_3^X = \alpha$ and therefore $(x_1, z^X) \in \sigma_1$. Thus if $F(T) = x \in \sigma_2$, then M.STAB implies that $F(T) \in \sigma_1$, which proves (4).

Since $\sigma_1 \cap \sigma_2 = \{\alpha e_Q\}$, (3) and (4) imply that $F(T) = \alpha e_Q \equiv z$. Therefore, since $t_p^z(T) = S^1$, it follows by M.STAB that $F(S^1) = \alpha e_p$. This is a contradiction to (2) and hence to (1), and we conclude that $F(S) = L(S)$. QED.

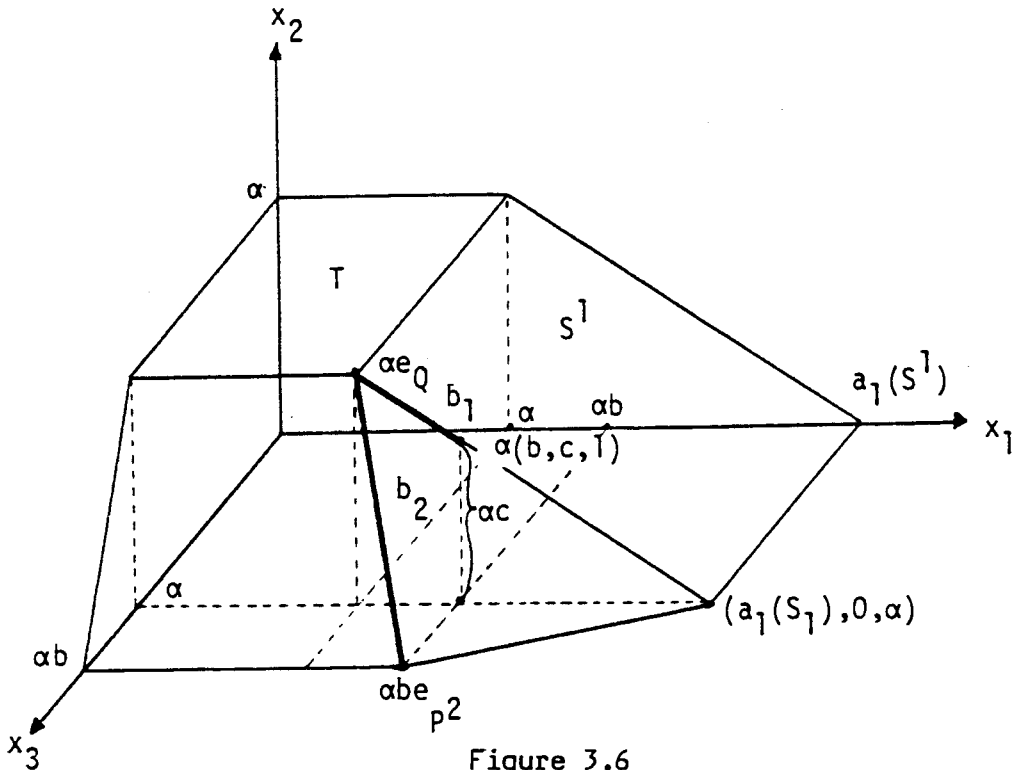


Figure 3.6

The proof of Lemma 3.2 for $v \geq 3$

So far, we have demonstrated that if a solution satisfies PO, AN, I.MON and M.STAB, then it coincides with the Leximin solution for all two-person problems. In order to extend this result to choice problems involving an arbitrary number of agents, we begin by establishing in Lemma 3.3 some useful properties of the Leximin solution. Given a solution F , a group Q of agents and a choice problem $T \in \Sigma^Q$, recall from chapter 2 the definition of $M_F(T)$ as the set of F -multilaterally stable points in T .

Lemma 3.3: For all $Q \in \mathcal{P}$, for all $T \in \Sigma^Q$, the following properties hold:

- (1) $L(T) \geq E(T)$.
- (2) If $z \in M_L(T)$ and $z \geq E(T)$, then $z = L(T)$.
- (3) If $z \in M_L(T)$ and $z \neq L(T)$, then there exists $k \in Q$ such that $z_k < E_k(T)$ and $z_i > E_i(T)$ for all $i \in Q \setminus \{k\}$.

Proof: It follows by the algorithmic definition of the Leximin solution that

(1) holds. To prove (2), let $Q \in \mathcal{P}$, $T \in \Sigma^Q$ and $z \in M_L(T)$ with $z \geq u \equiv E(T)$ be given. Let $Q^1 \equiv \{i \in Q \mid \exists x \in T, x \geq u, x_i > u_i\}$ and $P \equiv Q \setminus Q^1$. Note that $P \neq \emptyset$ and, since $z \geq u$, that $z_p = u_p$. Therefore, by definition of L , $L_P(T) = z_p$. Since $z \geq E(T)$, then $z > 0$, which implies that $S \equiv t_{Q^1}^z(T)$ is a well defined choice problem. Since $z \in M_L(T)$, then $L(S) = z_{Q^1}$, which by definition of L implies that $L(T) = (u_p, z_{Q^1})$. Since $u_p = z_p$, we conclude that $L(T) = z$, which proves (2).

Finally, to show that (3) holds, let $z \in M_L(T)$ with $z \neq L(T)$ be given. Then $z \geq u \equiv E(T)$ by (2). Consequently, there exists $k \in Q$ with $z_k < u_k$. Suppose, by way of contradiction, that $z_j \leq u_j$ for some $j \in Q \setminus \{k\}$. Let $P \equiv Q \setminus \{j\}$ and $S \equiv t_P^z(T)$. Since $z \in M_L(T)$, then $L(S) = z_p$. Since $z_j \leq u_j$ and T is comprehensive, then $u_p \in S$, which implies that $E(S) \geq u_p$. Therefore, by (1), $z_p \geq u_p$. Since $k \in P$ and $z_k < u_k$, this is the announced contradiction. Consequently, $z_i > u_i$ for all $i \in Q \setminus \{k\}$. QED.

Proposition 3.2: IF a solution F satisfies PO, AN, I.MON and M.STAB, then $F = L$, the Leximin solution.

Proof: We must show that F coincides with L on Σ^Q for all $Q \in \mathcal{P}$. The proof is by induction on $|Q|$, the number of elements in Q . If $|Q| = 1$, the proof is trivial, if $|Q| = 2$, it follows by Lemma 3.2. Suppose that F coincides with L for all $Q \in \mathcal{P}$ with $|Q| < n$. Let $Q \in \mathcal{P}$ with $|Q| = n \geq 3$ and $T \in \Sigma^Q$ be given, and let $z \equiv F(T)$. It follows by M.STAB and the induction hypothesis that $z \in M_L(T)$. Suppose, by way of contradiction, that $z \neq L(T)$. Letting $u \equiv E(T)$, it follows by (3) of Lemma 3.3 that there exists $k \in Q$ such that $z_k < u_k$ and $z_i > u_i$ for all $i \in P \equiv Q \setminus \{k\}$. Suppose without loss of generality that $Q = \{1, \dots, n\}$ and that $k = 1$.

Let $T^1 \equiv \text{cch}\{\alpha e_Q, T_P \times \{0e_1\}\}$ and let β be a real number such that $T^1 \subset \bar{T} \equiv \text{cch}\{\alpha e_Q, \beta e_2, \dots, \beta e_n\}$. For all $i \in P$, define T^i recursively by $T^i \equiv \text{cch}\{T^{i-1}, \beta e_i\}$. Note that $T^1 \subset T$ and $T_P^1 = T_P$; that $T^i \in \Sigma^Q$ and $E(T^i) = \alpha e_Q$ for all $i \in Q$, and that $T^{i-1} \subset T^i$ and $T_{Q \setminus \{i\}}^{i-1} = T_{Q \setminus \{i\}}^i$ for all $i \in P$. (See figure 3.7). For all $i \in Q$, let $z^i \equiv F(T^i)$. We claim that $z_1^i < \alpha$ and $z_p^i > \alpha$ for all $i \in Q$ and prove it by induction on i .

For $i = 1$, it follows by I.MON that $z_1^1 \leq z_1$. Since $E(T^1) = \alpha e_Q$ and $z_1^1 \leq z_1 < \alpha$, then (1) of Lemma 3.3 implies that $z^1 \neq L(T^1)$. Since $z^1 \in M_L(T^1)$ by M.STAB and the induction hypothesis for F , and since $z_1^1 < \alpha$, it follows by (3) of Lemma 3.3 that $z_p^1 > \alpha e_p$, which proves the claim for $i = 1$. Next, suppose the claim holds for $i = k \in \{1, \dots, n-1\}$. To prove it for $i = k+1$, observe first that $z_i^i \geq z_i^k$ and $z_i^k > \alpha$ by I.MON and the induction hypothesis for z^k , respectively. This implies that $z_1^i < \alpha$, for otherwise $z^i = \alpha e_Q$ by PO and construction of T^i . Since $E(T^i) = \alpha e_Q$ and $z_1^i < \alpha$, then (1) of Lemma (3.3) implies that $z^i \neq L(T^i)$. Since $z^i \in M_L(T^i)$ by M.STAB and the induction hypothesis for F , and since $z_1^i < \alpha$, it follows by (3) of Lemma 3.3 that $z_p^i > \alpha e_p$, which proves the claim.

Since $T^1 \subset \bar{T}$, it follows by construction of T^n that $T^n = T = \text{cch}\{\alpha e_Q, \beta e_2, \dots, \beta e_n\}$. It is easy to verify that $M_L(T^n)$ consists of the single point αe_Q , which by M.STAB and the induction hypothesis implies that $z^n = \alpha e_Q$. Thus, since $z_1^n < \alpha$, the assumption that $F(T) \neq L(T)$ has led to a contradiction. QED.

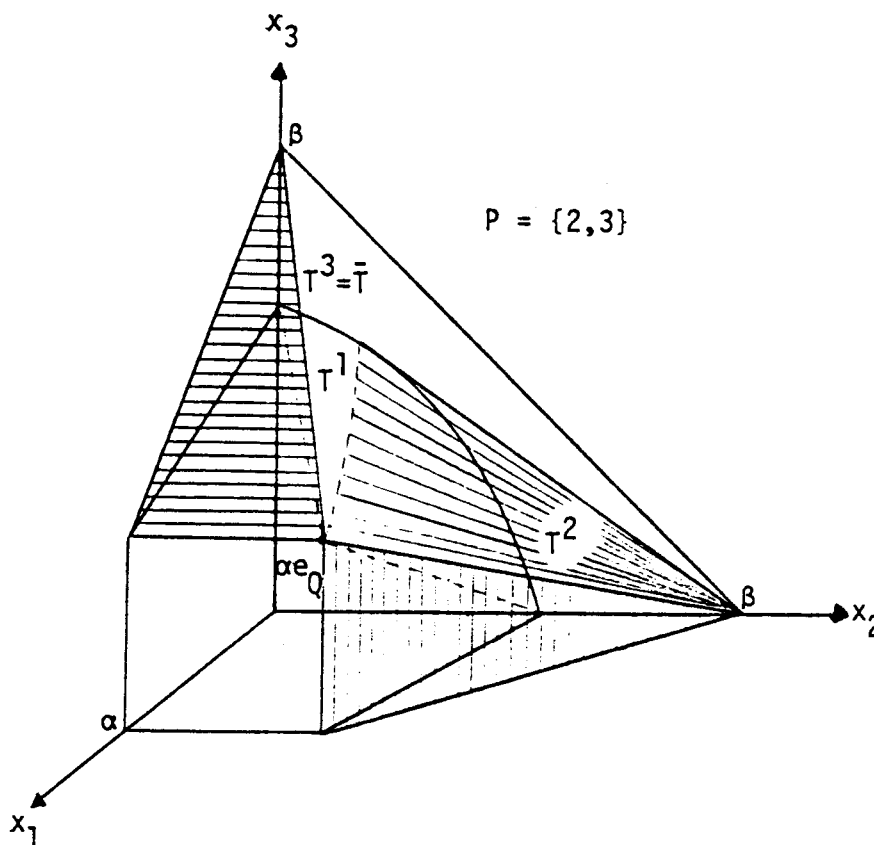


Figure 3.7

The proof of Proposition 3.2

Combining the results of Propositions 3.1 and 3.2, we obtain

Theorem 3.1: A solution satisfies PO, AN, I.MON and M.STAB if and only if it is the Leximin solution.

3.3 Concluding remarks

In this chapter, we have given a characterization of the Leximin solution using a list of four axioms which differ from those used to characterize the Nash solution in Theorem 2.1 only in that Scale Invariance has been replaced by Individual Monotonicity. Theorems 2.1 and 3.1 show that S.INV and I.MON are in a sense polar opposites when imposed in conjunction with the three other axioms: The Leximin solution exploits to a maximum degree the possibilities for interpersonal comparability of relative utility levels that become available when S.INV is dropped, by admitting no substitution between the utility levels of different agents.

The relationship between the Nash solution and the Leximin solution can be seen more clearly by observing that both of them are limit cases of a one-parameter family of solutions $\{F^\rho \mid \rho > -1\}$ defined as follows: Let $\rho > -1$ be given, and for all $P \in \mathcal{P}$, let the function $f^{\rho, P}$ on \mathcal{R}_+^P be defined by

$$f^{\rho, P}(x) \equiv \left(\sum_{i \in P} x_i^{-\rho} \right)^{-\frac{1}{\rho}} \text{ for all } x \in \mathcal{R}_+^P \text{ if } \rho \neq 0 \text{ and } f^{\rho, P}(x) \equiv \prod_{i \in P} x_i \text{ if } \rho = 0.$$

The function $f^{\rho, P}$ is a symmetric CES function and $-1/(1+\rho)$ is its constant elasticity of substitution. Since $\rho > -1$, the function $f^{\rho, P}$ is strictly quasi-concave for all $P \in \mathcal{P}$, and there exists a solution F^ρ defined by $F^\rho(S) \equiv \operatorname{argmax}\{f^{\rho, P}(x) \mid x \in S\}$ for all $P \in \mathcal{P}$ and all $S \in \Sigma^P$. As $\rho \rightarrow 0$, then $f^{\rho, P}$ converges¹⁾ to the Nash product $\prod_{i \in P} x_i$ and F^ρ converges to the Nash solution.

As $\rho \rightarrow \infty$, then $f^{\rho, P}(x)$ converges to $\min\{x_i \mid i \in P\}$, but the ordering of \mathcal{R}_+^P induced by this function is different from the Leximin ordering \succsim_P . However, the ordering \succsim_P^ρ of \mathcal{R}_{++}^P induced by the function $f^{\rho, P}$ does converge to the restriction of the Leximin ordering to \mathcal{R}_{++}^P in the sense that

1) Pointwise convergence.

for all $x, y \in \mathbb{R}_{++}^P$, if $x \succ_p y$, then there exists $\bar{\rho} \geq 0$ such that $x \succ_{\rho} y$ for all $\rho > \bar{\rho}$. Since $F^{\rho}(S) > 0$ for all choice problems S whenever $\rho \geq 0$, this can be used to show that F^{ρ} converges to the Leximin solution as $\rho \rightarrow \infty$.

These convergence problems in the limit as $\rho \rightarrow \infty$ arise because the Leximin ordering is not continuous. As a result of that, the Leximin solution itself is not continuous, as is clear from the example depicted in figure 3.8, where $P = \{1, 2\}$ and $S^u \equiv \text{cch}\{(2, 1-1/u), e_2\}$ for each positive integer u . As $u \rightarrow \infty$, then S^u converges to $S \equiv \text{cch}\{(2, 1)\}$ and $L(S^u)$ converges to $E(S) = (1, 1)$, while $L(S) = (2, 1)$, thus L is not continuous at S .

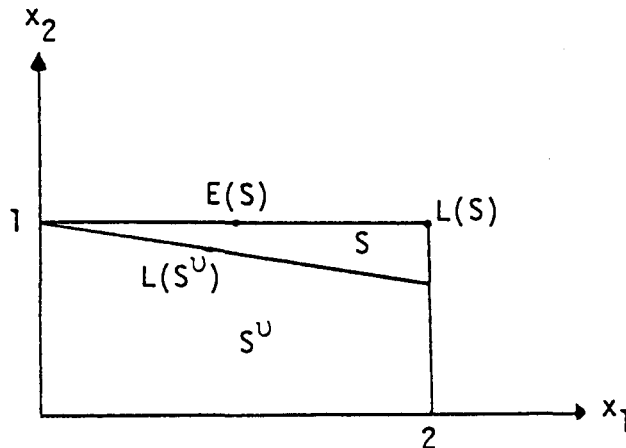


Figure 3.8

The Leximin solution is not continuous

Given that a solution F belongs to (the closure of) the family $\{F^{\rho} \mid \rho > -1\}$, it is easy to see the roles played by the axioms of S.INV and I.MON in singling out the Nash solution and the Leximin solution, respectively. If F satisfies S.INV, then the elasticity of substitution of the functions $f^{\rho, P}$

must be -1 , and therefore F must be the Nash solution. Next, suppose that F satisfies I.MON. As long as the (constant) elasticity of substitution for the functions $f^{\rho, P}$ differs from zero, then the corresponding solution F^{ρ} would violate I.MON in the example given in figure 3.2. Hence F must be the solution obtained by letting the elasticity of substitution approach zero, and therefore F must be the Leximin solution.

It should be noted that this is only a partial account of the roles played by S.INV and I.MON in the characterizations of the Nash solution and the Leximin solution respectively, because there remains the question of why the lists of axioms in Theorems 2.1 and 4.1 both lead to solutions that are consistent with the maximization of some ordering that is related to the family of CES functions. In fact, why should the solutions derived from these axioms be consistent with the maximization of any ordering or function, be it CES or not? Is this maximization property a joint result of all the axioms in each lists, or is it a consequence only of those axioms that are common to the Nash solution and the Leximin solution? An attempt at answering some of these questions will be made in the next chapter, which is concerned with the identification and characterization of a minimal family of solutions that satisfies the abovementioned maximization property and that includes the Nash solution and the Leximin solution as special or limit cases.

CHAPTER 4

STABILITY AND COLLECTIVE RATIONALITY

4.1 Introduction

A major part of the theory of social choice is based on the assumption of collective rationality - that social decisions be made consistently with the maximization of some ordering on the space of alternatives, where by an ordering is meant a binary relation which is reflexive, transitive and complete. For example, collective rationality is a defining property of an Arrow (1951) social welfare function (Arrow swf), which produces social orderings of a given set of physical alternatives from profiles of individual orderings of those alternatives. Another well-known example is the family of Bergson-Samuelson social welfare functions (Bergson (1938), Samuelson (1947)).

By a Bergson-Samuelson swf is usually meant a real-valued function defined on utility space, representing a social ordering of all possible utility allocations for the society under consideration. As pointed out by Sen (1970) pp. 34-35, it is in general the social ordering itself that is of primary interest for social decision making. Consequently, there is no need to distinguish between social orderings according to whether or not they have real-valued representations. Following Sen (1970), any ordering of utility space will here be referred to as a Bergson-Samuelson swf, regardless of whether it has a real-valued representation.

Observe that the domain of definition of a Bergson-Samuelson swf differs from that of an Arrow swf: A Bergson-Samuelson swf is defined on a space of utility allocations, while an Arrow swf is defined on a space of individual orderings, e.g. a space of utility functions.

A condition that is often imposed on the Bergson-Samuelson swf is Separability or Independence of Unconcerned Individuals as it is also sometimes called. This condition (due to Fleming (1952)) says that if the utility levels for a subset of the agents of society is the same for some pair of alternatives, then the social ordering of those alternatives should not depend on the utility levels of those agents. This means that if \succsim_Q is a social ordering of utility space \mathbb{R}_+^Q for a group Q of agents, then if $P \subset Q$, the ordering \succsim_P obtained from \succsim_Q by restricting \succsim_Q to any hyperplane parallel to \mathbb{R}^P must be the same for all such hyperplanes. If, in addition, the ordering \succsim_Q is continuous, then it has an additively separable numerical representation, i.e. there exists a real-valued function f on \mathbb{R}_+^Q such that $f(x) \geq f(y)$ if and only if $x \succsim_Q y$, where f is of the form $f(x) = \sum_{i \in Q} f_i(x_i)$ (Debreu (1960)).

The condition of Separability is indeed satisfied by most of the commonly used Bergson-Samuelson swf's, such as the Utilitarian swf (classical utilitarianism), the Nash swf¹⁾ (Nash (1950)) as well as the Leximin swf (Sen (1970)), which is the symmetric lexicographic extension of the Rawlsian (maximin) swf (Rawls (1971)).

1) To be precise, the Nash swf, \succsim^N , satisfies Separability only on the strictly positive orthant of utility space: For each $P \in \mathcal{P}$, \succsim_P^N is defined by $x \succsim_P^N y$ if and only if $\prod_{i \in P} x_i \geq \prod_{i \in P} y_i$, hence given $P, Q \in \mathcal{P}$ with $P \subset Q$ and $P \neq Q$, all points in $\mathbb{R}_+^Q \times \{\alpha \in \mathbb{R}_{Q \setminus P}\}^P$ are indifferent for \succsim_Q^N if $\alpha = 0$, while this is not the case if $\alpha > 0$.

Clearly, from any Bergson-Samuelson swf, one obtains a social choice function¹⁾, provided existence and uniqueness of the maxima for the swf on the relevant domain of problems. For example, the Utilitarian solution U is derived from the ordering \succeq^U defined by $x \succeq^U y$ if and only if $\sum_i x_i \geq \sum_i y_i$, and is well defined on the domain of strictly convex problems.

On the other hand, the existence of a social ordering is not a necessary prerequisite for social decision making; it is only necessary that there be a solution outcome to every choice problem. The concept of a solution that we use here is based on this minimal requirement and thus makes no presumption about collective rationality. Nevertheless, it turned out in previous chapters that many sets of axioms do lead to solutions that are collectively rational, thus one could look at such characterization results as being related to the integrability problem in demand theory, which concerns the identification of necessary and sufficient conditions for a demand function to come from utility maximization.

By analogy with the integrability problem, it would be of interest to establish conditions under which a solution is (1) collectively rational, i.e. consistent with the maximization of a Bergson-Samuelson swf, and (2) consistent with the maximization of a Separable Bergson-Samuelson swf. An answer to the first question has been given by Richter (1971), who showed that a choice function is rational if and only if it satisfies a generalized form of Houthakker's (1950) Strong Axiom of Revealed Preference (SARP). The second question is the topic of the present chapter.

1) A choice function is a rule which for every set in a collection of feasible sets selects a unique element from that set. A particular choice function is obtained by (i) specifying the collection of feasible sets (the domain, e.g. convex and comprehensive subsets of \mathbb{R}_+^n) and (ii) specifying the rule (e.g. select the point of equal coordinates in the upper boundary of the feasible set). Thus, a solution is a particular example of a choice function.

Earlier, the axiom of Multilateral Stability (M.STAB) was used to characterize the Nash solution (chapter 2) and the Leximin solution (chapter 3). As already noted, to each of these solutions there corresponds a Bergson-Samuelson swf which happens to be Separable.

This parallelism between M.STAB and Separability is in fact much more general: M.STAB implies that if two choice problems T and T' for a group Q of agents yield the same subproblem S for some subgroup P when intersected with hyperplanes parallel to \mathbb{R}^P through their solution outcomes x and x' , then $x_P = x'_P$. Thus M.STAB seems to be the natural counterpart to Separability in the sense that it imposes on a solution much the same requirement that Separability imposes on a Bergson-Samuelson swf. In fact, any solution obtained from a Separable Bergson-Samuelson swf can easily be shown to satisfy M.STAB.

What is more interesting, and less obvious, is that M.STAB imposes on a solution a fair amount of collective rationality as well. We show here that M.STAB is a necessary and sufficient condition for a Pareto-optimal and continuous solution to be consistent with the maximization of an additively separable, strictly quasi-concave Bergson-Samuelson social welfare function¹⁾. This is the main result of this chapter. We also show that the three axioms are independent in the sense that removing any one of them will permit solutions that are not collectively rational.

It may be worthwhile at this point to outline the basic structure of the proof of this characterization result to serve as a guide through some of the technical details involved in the argument. In particular, the proof

1) The weaker version of the stability axiom (B.STAB) will be sufficient.

is related to, and draws heavily upon concepts and ideas from demand theory, in particular from integrability theory. The connection with demand theory is established by considering the restriction \bar{F} of a solution F to the domain of linear budget problems known from demand theory. If the number of agents is held constant, the model then becomes formally equivalent to the standard model of demand theory which is made up of commodities, budgets and demand functions instead of agents, choice problems and solutions. These new terms will be used freely throughout in order to emphasize the relationship with demand theory, a relationship that will be exploited in the main part of the characterization proof which consists in establishing sufficient conditions for a demand function \bar{F} , defined for a variable number of commodities, to be consistent with the maximization of an additively separable utility function. Thus, as opposed to previous chapters where the full domain was used throughout the analysis, the main part of the analysis of the present chapter is conducted by considering only that subdomain which is relevant for demand theory, thereby yielding a result on the integrability of demand functions as a by-product of independent interest.

There are essentially two approaches to the integrability problem in demand theory (Chipman et. al. (1971), pp. 3-6): The first one is set theoretic and uses some version of SARP and some demand continuity assumption¹⁾. The second one is analytic and uses symmetry and negative semi-definiteness of the Slutsky-matrix (direct demand function) or of the Antonelli-matrix (indirect demand function)²⁾.

1) Exceptions are (1) Uzawa (1971), who uses Samuelson's (1938) Weak Axiom of Revealed Preference (WARP) and a regularity condition on the revealed preference relation and (2) Hurwicz and Richter (1971), who do not assume continuity of the demand function.

2) An exception is Hurwicz and Richter (1979), who use WARP and an axiom of Ville (1951).

The approach followed here is analytic, based on the indirect demand function, although we make no explicit use of the Antonelli-matrix, the reason being that our axioms do not imply that the indirect demand function is differentiable. In fact, the indirect demand function may even fail to be single-valued and continuous. The reason why the indirect demand function is used, despite its apparent irregularity, is that M.STAB imposes a lot of structure on its local behaviour, but has little or nothing to say about the local behaviour of the direct demand function. This is so because integrating an indirect demand function yields a direct utility function, while integrating a direct demand function yields an indirect utility function¹⁾. M.STAB is related to additive separability of the direct utility function, and it is in general not true that the indirect demand function is additively separable if the direct one has that property.

The point of departure for the characterization theorem is Lemma 4.1, which is a slight modification of Lemma 2.3, the main step in the characterization of the Nash solution that uses Continuity. The proof of Lemma 2.3 is essentially a proof that PO, CONT and B.STAB together imply IIA for all two-person components of F , a result that extends to any n -person component of F by strengthening B.STAB to M.STAB.

This result is useful for two reasons: First, it can be used to show that every strictly positive vector is the solution outcome to some budget problem. This fact allows one to restrict attention to the domain of budget problems, because most problems can then be solved by applying IIA to some budget problem that contains it. Second, it turns out that the demand function \bar{F} satisfies WARP whenever the solution F satisfies IIA. Of course,

1) See Hurwicz (1971) for an excellent exposition and more details.

WARP is not a sufficiently strong condition to obtain integrability results, as the well-known counter-example by Gale (1960) has shown¹⁾. However, WARP has two other useful consequences: First, it implies that the inverse demand correspondence (from quantities to prices) is convex-valued, which can be used to show that the normalized inverse demand correspondence (from quantities to prices in the unit simplex) must be single-valued except on a negligible set. Second, it implies that the real-valued representation for \bar{F} , if one exists, is quasi-concave.

To sum up, the main role played by the stability axiom so far has been to show that WARP must hold for the demand function \bar{F} , thereby establishing properties that correspond to negative semidefiniteness of the Antonelli-matrix under conditions of differentiability. What we do next is to use M.STAB to establish a property of the inverse demand correspondence that would imply symmetry of the Antonelli-matrix under differentiability.²⁾

An outline of the main idea can be given by assuming, for convenience, that the normalized inverse demand function is single-valued everywhere on the strictly positive orthant. The situation is illustrated in figure 4.1, which depicts a two-dimensional budget problem S solved at $x > 0$. Φ is the inverse demand correspondence and ϕ is the normalized inverse demand correspondence. Thus, $\Phi(x)$ is the set of normals to all budget problems solved at x and $\phi(x)$ is the intersection between $\Phi(x)$ and the unit simplex. Since

1) See also Kihlstrom, Mas-Colell and Sonnenschein (1976), whose results yield an infinity of demand functions that satisfy WARP but not SARP.

2) Observe that symmetry of the Antonelli-matrix in the mathematical integrability condition, i.e. it is used to prove that level curves exist, while negative semi-definiteness serves to give these level curves the right curvature, i.e. convexity. Again the reader is referred to Hurwicz (1971) for more details.

$x > 0$, then $\phi(x)$ is a singleton, $\Phi(x)$ is a ray and S is the only budget problem solved at x . The quantity $M(x) \equiv \phi_1(x)/\phi_2(x)$ is the price of a unit of the first commodity in terms of the second. Thus, it could be (the negative of) the slope of a level curve of some underlying utility function at x , i.e. a marginal rate of substitution (MRS) at x . Note that $M(x_1, x_2) = 1/M(x_2, x_1)$.

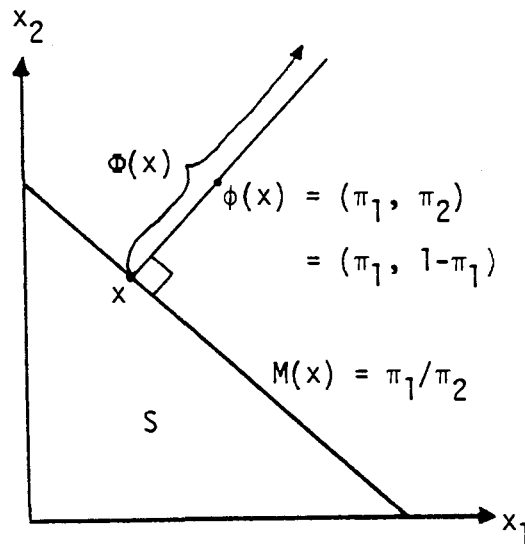


Figure 4.1

Inverse demand correspondences and the MRS

Next, let $Q \in \mathcal{P}$ with $|Q| \geq 3$ be given. M.STAB implies that for all $x \in \mathbb{R}_+^Q$, if $P \subset Q$ then $\Phi_P(x) \subset \Phi(x_P)$, where as usual, $\Phi_P(x)$ denotes the projection of $\Phi(x)$ on \mathbb{R}^P . In fact, this inclusion property is exactly what M.STAB says about the restricted solution \bar{F} , thus it could serve as an alternative definition of M.STAB for \bar{F} . Note that if $x > 0$, the reverse inclusion also

holds, since by assumption, Φ is ray valued on the strictly positive orthant. Hence $\Phi_P(x) = \Phi(x_P)$ for all $P \subset Q$ and all $x \in \mathbb{R}_{++}^Q$ or equivalently, $\phi_P(x)$ is proportional to $\phi(x_P)$. Because $\phi_i(x_P)/\phi_j(x_P) = M(x_i, x_j)$ for $P = \{i, j\}$, this implies that

$$(1) \quad \phi_i(x)/\phi_j(x) = M(x_i, x_j) \text{ for all } Q \in \mathcal{P}, \text{ all } i, j \in Q \text{ and all } x \in \mathbb{R}_{++}^Q.$$

(1) implies that the ratio of any two components i and j of the normalized inverse demand function is independent of all but the i 'th and j 'th arguments. Setting $Q \equiv \{i, j, k\}$ and noting that

$$(2) \quad \phi_i(x)/\phi_j(x) = (\phi_i(x)/\phi_k(x))/(\phi_j(x)/\phi_k(x))$$

identically, it follows (see figure 4.2) that

$$(3) \quad M(x_i, x_j) = M(x_i, x_k)/M(x_j, x_k) \text{ for all } Q \equiv \{i, j, k\} \text{ and all } x \in \mathbb{R}_{++}^Q.$$

Expression (3) reveals the structure imposed by M.STAB on the local behavior of the indirect demand function, stating that the MRS between x_i and x_j can be written as a quotient, where the numerator does not depend on x_j and the denominator does not depend on x_i . This allows us to fix $x_k = x_k^0$ arbitrarily and integrate the two terms on the RHS of (3) separately to obtain two functions f_i and f_j such that the LHS of (3) is (the negative of) the slope of a level curve of the function $f_i + f_j$ wherever the latter is differentiable. Continuity of the direct demand function \bar{F} and single-valuedness of the indirect demand function ϕ , except on a negligible set, will guarantee that the MRS-functions are sufficiently well-behaved to be integrable.

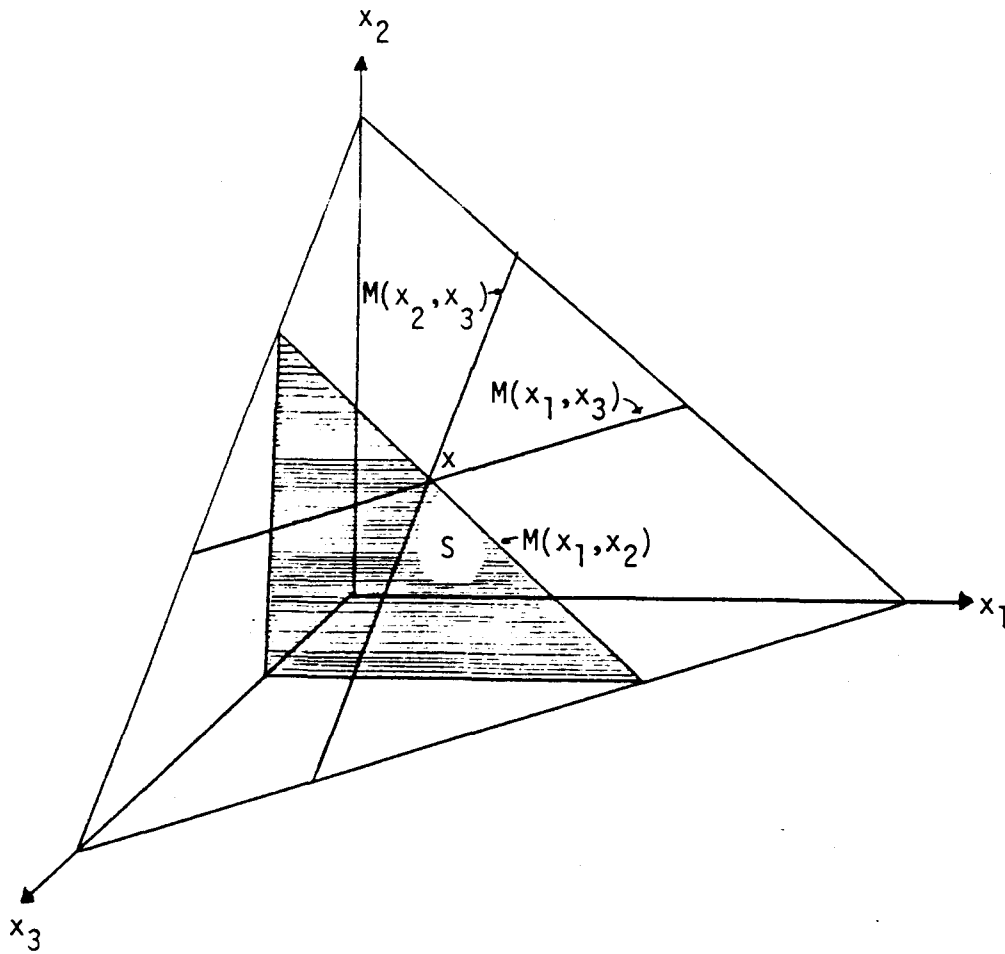


Figure 4.2

Relationship between MRS-functions

This integrability result generalizes easily to higher dimensions by defining f_i as indicated in the previous paragraph for all $i \neq k$. It then follows by repeated use of (3) and (1) that for all $Q \in \mathcal{P}$ that do not contain k and all $x \in \mathbb{R}_{++}^Q$, if the gradient of the function $\sum_{i \in Q} f_i$ exists at x , then it is normal to the budget problem solved at x . It is shown in Lemma 4.12 that this result extends to all $Q \in \mathcal{P}$ by a suitable choice of the remaining function f_k . Finally, as mentioned earlier, WARP will guarantee that the level curves of the function $\sum_{i \in Q} f_i$ are strictly convex for all Q .

Having outlined the main ideas in the characterization proof, it is indicated in figure 4.3 below how the result in this chapter, seen as a result on the integrability of demand functions, fits in with other results in this tradition. The left part of the figure concerns representable demand functions and the right part concerns demand functions that have additively separable representations.

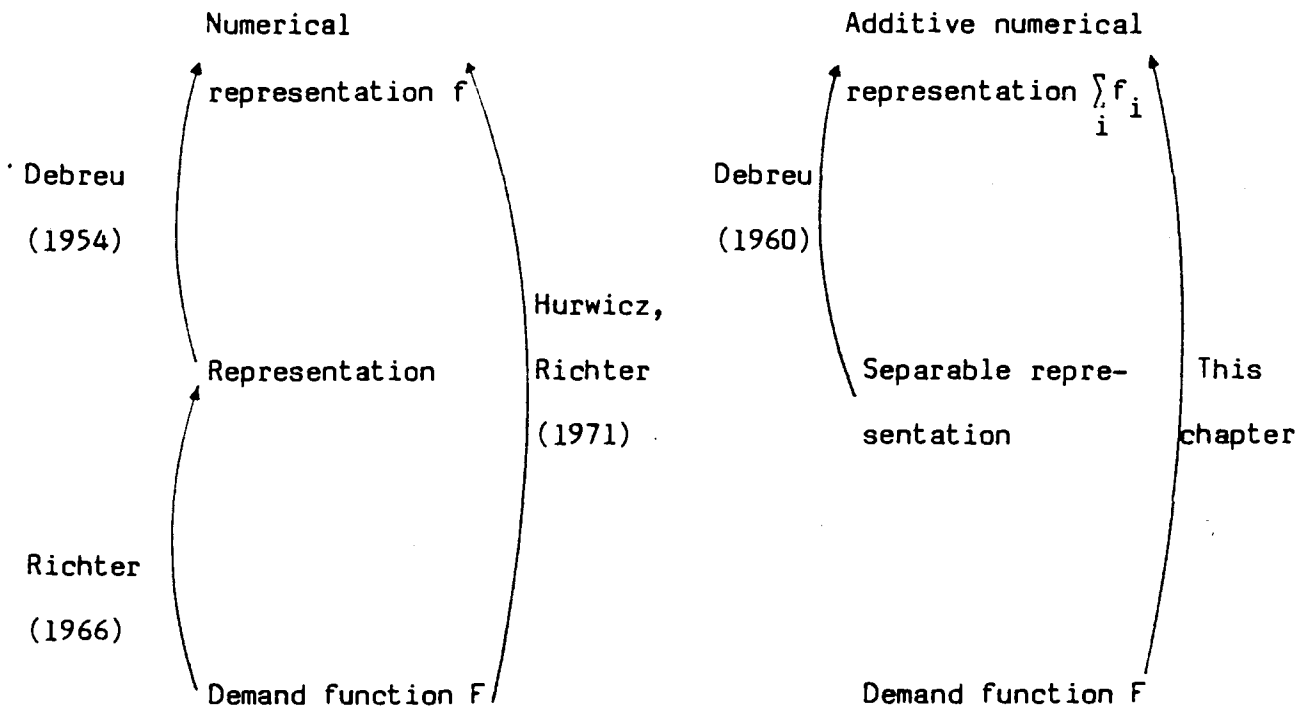


Figure 4.3

Results on the integrability of demand functions

The next section introduces some new notation and some new properties (axioms) that will be seen to follow from the main axioms.

4.2 The axioms

We first introduce the family of collectively rational solutions. A solution F belongs to this family if there exists a list of orderings $\succ \equiv \{\succ_P \mid P \in \mathcal{P}\}$, where each \succ_P is defined on \mathcal{R}_+^P , such that for all $P \in \mathcal{P}$ and all $S \in \Sigma^P$, $F(S)$ is the unique maximal element for \succ_P in S . The list \succ will be referred to as a representation for F . If for each component \succ_P of \succ , there exists an extended real-valued function f^P defined on \mathcal{R}_+^P , with the property that $f^P(x) \geq f^P(y)$ if and only if $x \succ_P y$, then the list $f \equiv \{f^P \mid P \in \mathcal{P}\}$ is said to be a numerical representation for F .

Our concern here will be with solutions that have additively separable numerical representations: Let F be the family of all sequences $\{f_i\}_{i \in I}$ of strictly increasing and continuous functions $f_i: \mathcal{R}_+^{\{i\}} \rightarrow \mathcal{R}_+^{\{i\}} \cup \{-\infty\}$, such that for all $P \in \mathcal{P}$, the function $\sum_{i \in P} f_i$ is strictly quasi-concave¹⁾. A solution F has an additively separable numerical representation if there exists a sequence from F of functions $\{f_i\}_{i \in I}$ such that for all $P \in \mathcal{P}$ and all $S \in \Sigma^P$, $F(S) = \operatorname{argmax}\{\sum_{i \in P} f_i(x) \mid x \in S\}$. A collectively rational solution that admits of an additively separable numerical representation will be referred to as a CRS solution.

As mentioned earlier, the main part of the characterization proof consists in showing that if a solution F satisfies certain axioms then its restriction to the family of budget problems is of the CRS type. For each $P \in \mathcal{P}$,

1) An extended real-valued function h on \mathcal{R}_+^n is strictly quasi-concave if for all $y \in \mathcal{R}_+^n$, the upper contour set $G(y) \equiv \{x \in \mathcal{R}_+^n \mid h(x) \geq h(y)\}$ is strictly convex in \mathcal{R}_+^n , meaning that for all $x, z \in G(y)$ with $x \neq z$ and all $\lambda \in (0, 1)$, the point $\lambda x + (1-\lambda)z$ is an interior point of $G(y)$, relative to \mathcal{R}_+^n . An extended real-valued function g on \mathcal{R}_+ is strictly increasing if $g(x) > g(y)$ for all x, y such that $x \geq y$.

the family of budget problems for the set P of agents is denoted Σ_{Δ}^P and consists of those $S \in \Sigma^P$ such that $S = \{x \in \mathcal{R}_+^P \mid \pi x \leq \pi\omega\}$ for some $(\pi, \omega) \in \mathcal{R}_{++}^P \times \mathcal{R}_+^P \setminus \{0\}$.

It will be convenient to consider the set of such price-endowment pairs (π, ω) as the domain of definition for the restriction \bar{F} of a solution F to the family of budget problems: For all $P \in \mathcal{P}$ and all $(\pi, \omega) \in \mathcal{R}_{++}^P \times \mathcal{R}_+^P \setminus \{0\}$, let $b(\pi, \omega) \equiv \{x \in \mathcal{R}_+^P \mid \pi x \leq \pi\omega\}$, define $\bar{F}(\pi, \omega) \equiv F(b(\pi, \omega))$ if $\omega \neq 0$ and $\bar{F}(\pi, 0) \equiv 0$. The convention that $\bar{F}(\pi, 0) = 0$ is adopted for notational convenience, although strictly speaking, the set $b(\pi, 0) = \{0\}$ is not a well defined problem as it does not contain a strictly positive vector. Thus, \bar{F} is a function from $\bigcup_{P \in \mathcal{P}} \mathcal{R}_{++}^P \times \mathcal{R}_+^P$ to $\bigcup_{P \in \mathcal{P}} \mathcal{R}_+^P$, such that for all $P \in \mathcal{P}$ and all $(\pi, \omega) \in \mathcal{R}_{++}^P \times \mathcal{R}_+^P$, $\bar{F}(\pi, \omega)$ satisfies the budget constraint $\pi \bar{F}(\pi, \omega) \leq \pi\omega$.

Recall that for all $P \in \mathcal{P}$, $\Delta^P \equiv \{\pi \in \mathcal{R}_+^P \mid \sum_{i \in P} \pi_i = 1\}$ is the unit simplex in \mathcal{R}_+^P . Because \bar{F} is homogeneous of degree 0 in π , Δ^P may be used to normalize the domain of definition for F with respect to prices π .

We next restate the axioms of PO, CONT, B.STAB and M.STAB for the restricted solution F as follows:

Pareto-optimality (PO): For all $P \in \mathcal{P}$, for all $(\pi, \omega) \in \mathcal{R}_{++}^P \times \mathcal{R}_+^P$, $\pi \bar{F}(\pi, \omega) = \pi\omega$.

Continuity (CONT): For all $P \in \mathcal{P}$, if $\{(\pi^u, \omega^u)\}$ is a sequence from $\mathcal{R}_{++}^P \times \mathcal{R}_+^P$ converging to $(\pi, \omega) \in \mathcal{R}_{++}^P \times \mathcal{R}_+^P$, then $\lim_{u \rightarrow \infty} \bar{F}(\pi^u, \omega^u) = \bar{F}(\pi, \omega)$.

Multilateral stability (M.STAB): For all $P, Q \in \mathcal{P}$ with $P \subset Q$, for all $(\pi, \omega) \in \mathcal{R}_{++}^Q \times \mathcal{R}_+^Q$ and all $\omega' \in \mathcal{R}_+^P$, if $\bar{F}(\pi, \omega) = x$ and $\pi_P \omega' = \pi \omega - \pi_{Q \setminus P} x_{Q \setminus P}$, then $\bar{F}(\pi_P, \omega') = x_P$.

Recall that in the statement of M.STAB for F in chapter 2, a provision was required in order to deal with the possibility that some subproblem $t_p^x(T)$ might not be well defined. What could go wrong was that $t_p^x(T)$ might not contain a strictly positive vector. Such a provision is not needed in the above statement of M.STAB for \bar{F} because the only way that some subproblem $b(\pi_P, \omega')$ can fail to contain a strictly positive vector is by having $\omega' = 0$, in which case $b(\pi_P, \omega') = \{0\}$. This case has already been dealt with by explicitly including all such trivial problems in the domain of \bar{F} .

Bilateral stability (B.STAB): Same as M.STAB, except for adding the provision that $|P| = 2$.

Pursuing the interpretation of \bar{F} as a demand function, P0 states that the whole budget is spent. CONT needs no comment. M.STAB says that if it is optimal to allocate $\pi_P \omega'$ of the total budget $\pi \omega$ to the commodities in a subset P of the set Q of all commodities, then the income $\pi_P \omega'$ can be spent optimally on the commodities in P without having to worry about how to spend the remaining income on the commodities in the set $Q \setminus P$. Thus, a consumer whose demand function satisfies M.STAB is one for whom two-stage budgeting is optimal for any partition of the set of commodities.

Observe that if \bar{F} satisfies P0, then $\bar{F}(\pi, \omega) = x$ implies $\bar{F}(\pi, x) = x$. Moreover, M.STAB simplifies to saying that if $\bar{F}(\pi, \omega) = x$ for some $(\pi, \omega) \in \mathcal{R}_{++}^Q \times \mathcal{R}_+^Q$, then $\bar{F}(\pi_P, x_P) = x_P$ for all $P \subset Q$.

We next introduce two additional properties that will be used in connection with the restricted solution \bar{F} . The first one is the Weak Axiom of Revealed Preference:

Let \bar{F} be a restricted solution and let $P \in \mathcal{P}$ be given. Given $x, y \in \mathcal{R}_+^P$ with $x \neq y$, say that x is revealed preferred to y (by \bar{F}), written $x \underset{\bar{F}}{R} y$, if $x = \bar{F}(\pi, \omega)$ and $\pi y \leq \pi \omega$ for some $(\pi, \omega) \in \mathcal{R}_{++}^P \times \mathcal{R}_+^P$.

Weak Axiom of Revealed Preference (WARP): For all $P \in \mathcal{P}$, for all $x, y \in \mathcal{R}_+^P$, if $x \underset{\bar{F}}{R} y$ then not $y \underset{\bar{F}}{R} x$.

The final property is the following boundary condition, which states that if, in a sequence of budget problems, some price approaches zero, then the corresponding sequence of demand vectors must be unbounded (with respect to the euclidean norm).

Boundary Condition (BOUND): For all $P \in \mathcal{P}$, for all $\omega \in \mathcal{R}_{++}^P$ and all sequences $\{\pi^v\}$ from Δ^P , if $\pi_i^v \rightarrow 0$ for some $i \in P$, then $\lim_{v \rightarrow \infty} \|\bar{F}(\pi^v, \omega)\| = \infty$.

This condition is identical to one that is often imposed on market excess demand functions in general equilibrium theory in order to guarantee the existence of a strictly positive equilibrium price vector¹⁾. Another version of that boundary condition states that if the price of some commodity approaches zero, then the market excess demand for that commodity approaches infinity. The weaker version only requires that the demand for some commodity must become unbounded. This is a strictly weaker requirement

1) See e.g. Varian (1981) for an exposition.

in the case where more than one price approach zero at the same time. Here we need only the weak version of this boundary condition, although the strong version does hold for \bar{F} if F satisfies PO, CONT and M.STAB.

4.3 The main result

In this section, we show that PO, CONT and M.STAB characterize the family of CRS solutions. First it is demonstrated that a CRS solution is well defined and satisfies the three axioms.

Proposition 4.1: For all $\{f_i\}_{i \in I} \in F$, there exists a CRS solution F for which $\{f_i\}$ is an additively separable numerical representation, such that F satisfies PO, CONT and M.STAB.

Proof: The following intermediate result will be used:

(i) Let $\{f_i\}_{i \in I}$ be a sequence of strictly increasing functions $f_i: \mathbb{R}_+^i \rightarrow \mathbb{R}^i \cup \{-\infty\}$. For all $Q \in \mathcal{P}$ and all $T \in \Sigma^Q$, if $z \in \operatorname{argmax}\{\sum_{i \in Q} f_i(x_i) \mid x \in T\}$, then for all $P \subset Q$, $z_P \in \operatorname{argmax}\{\sum_{i \in P} f_i(x_i) \mid x \in t_P^Z(T)\}$.

To prove (i), let $\{f_i\}_{i \in I}$, Q , T and z satisfy its hypothesis. Let $P \subset Q$ be given and define $S \equiv t_P^Z(T)$. By assumption,

$$(1) \sum_{i \in Q} f_i(z_i) \geq \sum_{i \in Q} f_i(x_i) \text{ for all } x \text{ in any subset of } T,$$

and in particular for all $x \in Sx\{z_{Q \setminus P}\}$. Since each f_i is strictly increasing and T contains a strictly positive vector, it follows by (1) that

$$\sum_{i \in Q} f_i(z_i) > -\infty. \text{ Since } T \text{ is bounded, then } \sum_{i \in Q} f_i(z_i) < \infty. \text{ These two in-$$

equalities imply that $-\infty < f_i(x_i) < \infty$ for all $i \in Q$. Since $x_{Q \setminus P} = z_{Q \setminus P}$ for

all $x \in Sx\{z_{Q \setminus P}\}$, the finite quantity $\sum_{i \in Q \setminus P} f_i(z_i)$ may be subtracted from

$$\text{both sides of (1) to obtain that } \sum_{i \in P} f_i(z_i) \geq \sum_{i \in P} f_i(x_i) \text{ for all } x \in S. \text{ This}$$

completes the proof of (i).

Next we show that a CRS solution is well defined, i.e. that it associates one and only one solution outcome to every choice problem. To see this, let $\{f_i\}_{i \in I} \in F$ be given and let F be the CRS solution associated with $\{f_i\}_{i \in I}$. Thus, for all $P \in \mathcal{P}$ and all $S \in \Sigma^P$, $F(S) \equiv \operatorname{argmax}\{\sum_{i \in P} f_i(x_i) \mid x \in S\}$.

Let $Q \in \mathcal{P}$ and $T \in \Sigma^Q$ be given. Since T is compact and $\sum_{i \in Q} f_i$ is continuous, it follows that $F(T) \neq \emptyset$. To show that $F(T)$ is a singleton, suppose by way of contradiction that $F(T)$ contains two distinct points x^0 and x^2 . Let $x^1 \equiv \lambda x^0 + (1-\lambda)x^2$ for some $\lambda \in (0,1)$ and let P be the maximal subset of Q such that $x_p^1 > 0$. Since $x^0 \neq x^2$ and $\lambda \in (0,1)$, then $P \neq \emptyset$ and $x_{Q \setminus P}^v = 0$ for $v = 0,1,2$, which implies that $t_p^x(T) = t_p^x(T) = t_p^x(T) \equiv S$. It follows by (i) that $x_p^0, x_p^2 \in \operatorname{argmax}\{\sum_{i \in P} f_i(x_i) \mid x \in S\}$, which since $\sum_{i \in P} f_i$ is strictly quasi-concave implies that x_p^1 belongs to the interior of $G(x_p^1) = G(x_p^0) \equiv \{x \in \mathcal{R}_+^P \mid \sum_{i \in P} f_i(x_i) \geq \sum_{i \in P} f_i(x_i^0)\}$, relative to \mathcal{R}_+^P . Because $x_p^1 > 0$, this implies that $\sum_{i \in P} f_i(x_i^1) > \sum_{i \in P} f_i(x_i^0)$, in contradiction with the facts that $x_p^1 \in S$ and $x_p^0 \in \operatorname{argmax}\{\sum_{i \in P} f_i(x_i) \mid x \in S\}$. This completes the proof that a CRS solution is well defined.

Finally, we show that PO, CONT and M.STAB are satisfied by any CRS solution. Let F be a CRS solution. Then there exists a sequence from F of functions $\{f_i\}_{i \in I}$ which represents F . F satisfies PO since each f_i is strictly increasing; moreover, F satisfies CONT since $\sum_{i \in P} f_i$ is strictly quasi-concave on \mathcal{R}_+^P for all $P \in \mathcal{P}$. It remains to show that F satisfies M.STAB.

Let $P, Q \in \mathcal{P}$ with $P \subset Q$ and $T \in \Sigma^Q$ be given. Let $z \equiv F(T)$ and $S \equiv t_p^z(T)$. Assume without loss of generality that S is a well defined member of Σ^P (if not, then M.STAB holds trivially) and let $y \equiv F(S)$. Since $\{f_i\}_{i \in I}$ repre-

sents F , then $z = \operatorname{argmax}_{i \in Q} \{ \sum_i f_i(x_i) \mid x \in T \}$ and $y = \operatorname{argmax}_{i \in P} \{ \sum_i f_i(x_i) \mid x \in S \}$, which by (i) implies that $y = z_p$, the desired conclusion. QED.

The proof of the converse of Proposition 4.1 is in several steps. The first one consists of Lemma 4.1 through Lemma 4.3 where it is shown that if a solution F satisfies PO, CONT and M.STAB then its restriction \bar{F} also satisfies WARP and BOUND.

Lemma 4.1: If a solution satisfies PO, CONT and M.STAB, then it satisfies IIA.

Proof: The proof, which is illustrated in figure 4.4, is a simple adaptation of the proof of Lemma 2.3, the main step in the characterization of the Nash solution that uses Continuity. Let F be a solution satisfying PO, CONT and M.STAB, and let $P \in \mathcal{P}$ be given. Let S and S' be two members of Σ^P such that $S' \subset S$ and $y \equiv F(S) \in S'$. In the figure, $P = \{i, j\}$. We must show that $y = F(S')$ also. To do this, assume first that

(i) $S' \cap U = S \cap U$ for some neighbourhood U of y .

Let now k be an agent who is not a member of P and let $Q \equiv P \cup \{k\}$. Define $S^1 \equiv S' \times \{e_k\}$ and for all $\varepsilon \geq 0$, let C^ε be the cone with vertex $(1 + \varepsilon)e_k$, spanned by S^1 . Define $T^\varepsilon \equiv C^\varepsilon \cap \operatorname{cch}\{S^1\}$ and $U^1 \equiv U \times \{e_k\}$, and note that for all $\varepsilon \geq 0$, $U^1 \cap S^1 \subset T^\varepsilon$ and $T^\varepsilon \in \Sigma^Q$.

Let $z \equiv F(T^0)$. We claim that $z = w \equiv (y, 1)$. To see this, note that whatever z is, it follows by PO and construction of T^0 that $t_p^z(T^0) = S$, which by M.STAB implies that $z_p = y$. Since w is the only Pareto-optimal point of T^0 with property that $w_p = y$, we conclude by PO that $z = w$.

Consider now $z^\epsilon \equiv F(T^\epsilon)$ as $\epsilon \rightarrow 0$. Since $z^0 = z$ and $T^\epsilon \rightarrow T^0$ as $\epsilon \rightarrow 0$, it follows by CONT that $z^\epsilon \rightarrow z$. Therefore, by P0, there exists $\bar{\epsilon} > 0$ such that $z^\epsilon \in U^1$ for all ϵ in $[0, \bar{\epsilon}]$, which by M.STAB implies that z_p^ϵ is constant and equal to $F(S')$ for all such ϵ . But then $z^\epsilon = z$ for all ϵ in $(0, \bar{\epsilon})$ by the fact that $z^\epsilon \rightarrow z$ in U^1 as $\epsilon \rightarrow 0$, which implies that $F(S') = z_p = y$, the desired conclusion.

To complete the proof, it suffices to observe that if S' does not satisfy condition (i) above, then it can be approximated by a sequence of elements from Σ^P that does. CONT may then be applied once more to conclude that $F(S') = F(S)$ in this case also. QED.

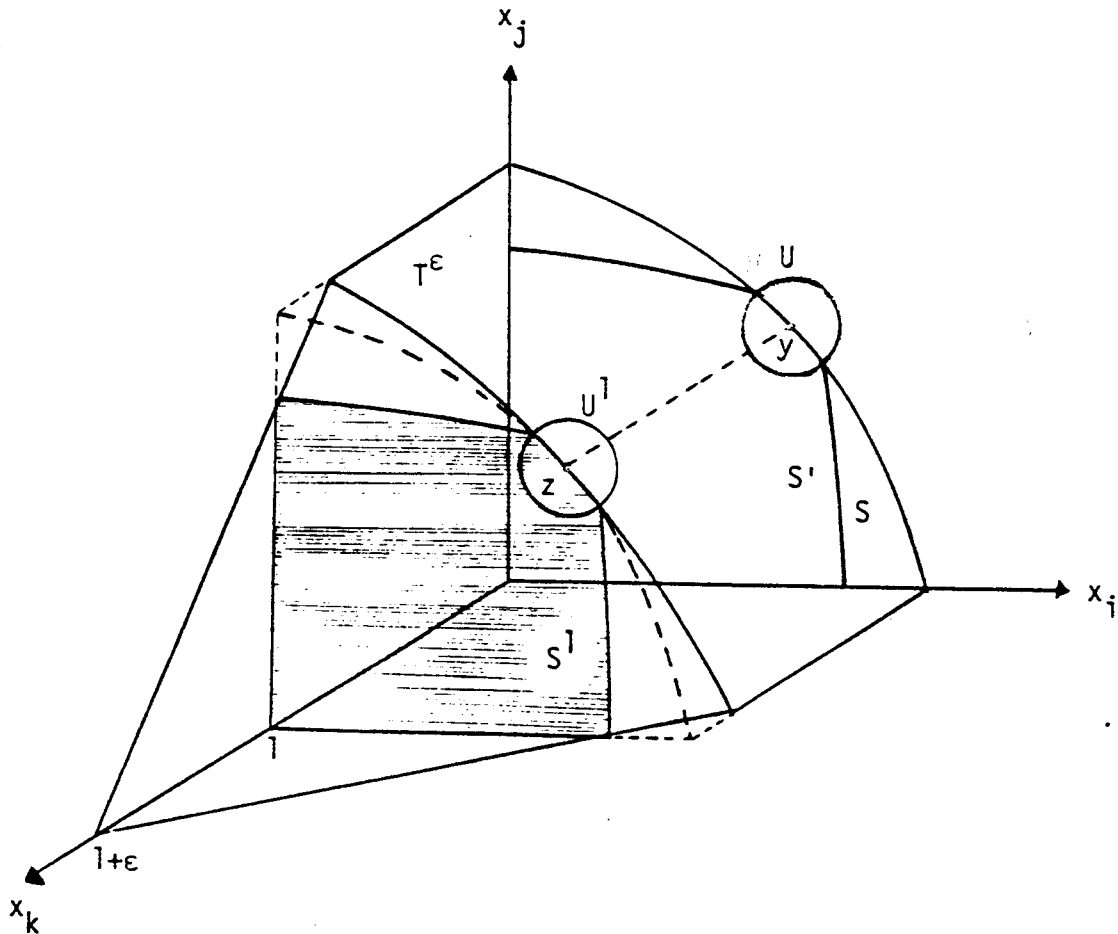


Figure 4.4

The proof of Lemma 4.1

Lemma 4.2: If F satisfies IIA then \bar{F} satisfies WARP.

Proof: Let F be a solution that satisfies IIA and suppose, by way of contradiction that \bar{F} does not satisfy WARP. Since F agrees with \bar{F} on the family of budget problems, then there exists $P \in \mathcal{P}$, and $S, S' \in \Sigma_{\Delta}^P$ such that $F(S) \in S'$, $F(S') \in S$ and $F(S) \neq F(S')$. Let $S'' \equiv S \cap S'$ and note that $S'' \in \Sigma^P$. By IIA applied twice, $F(S'') = F(S)$ and $F(S'') = F(S')$, a contradiction, since $F(S) \neq F(S')$. QED.

The relationship between IIA and WARP is closer than indicated by the one-sided implication of Lemma 4.2. In fact, it is straightforward to show that IIA is equivalent to WARP if the latter axiom is restated for F instead of \bar{F} . No use will be made of this more general result here, however.

Lemma 4.3: If F satisfies PO, CONT and IIA, then \bar{F} satisfies BOUND.

Proof: (See figure 4.5 for an illustration). Let $P \in \mathcal{P}$ and $\omega \in \mathcal{R}_{++}^P$ be given, and let $\{\pi^u\}$ be a sequence from Δ^P and i a member of P such that $\pi_i^u \rightarrow 0$. For each u , let $z^u \equiv \bar{F}(\pi^u, \omega)$, and suppose, by way of contradiction, that $\|z^u\|$ does not converge to infinity. Then there exists a subsequence of $\{z^u\}$ which converges to a point $z \in \mathcal{R}_+^P$ and a point $a \in \mathcal{R}_{++}^P$ such that $z < a$ and $z^u < a$ for all z^u in that subsequence. Assume without loss of generality that the sequence $\{z^u\}$ itself has that property.

Because $\{\pi^u\}$ is a subset of the compact set Δ^P , it has a subsequence that converges to some $\pi \in \Delta^P$. Assume without loss of generality that $\{\pi^u\}$ itself has that property. For each u , let $S^u \equiv b(\pi^u, \omega) \cap \text{cch}\{a\}$. Then $S^u \subset b(\pi^u, \omega)$ and since $z^u < a$, then $z^u \in S^u$. Therefore by IIA, $z^u = F(S^u)$ for all u . Since $\pi^u \rightarrow \pi$, the sequence $\{S^u\}$ converges to $S \equiv b(\pi, \omega) \cap \text{cch}\{a\}$

and since $z^v \rightarrow z$, it follows by CONT that $F(S) = z$. On the other hand, PO requires that $F_i(S) = a_i$ because $\pi_i = 0$. Since $a_i > z_i$, this is the announced contradiction. QED.

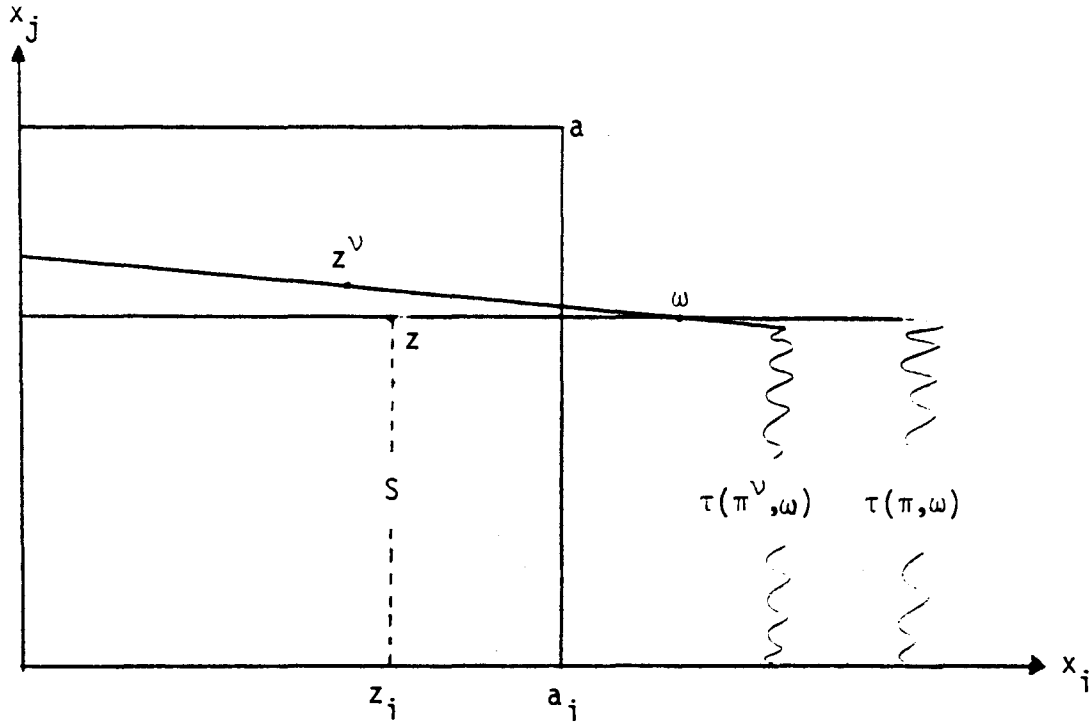


Figure 4.5

The proof of Lemma 4.3

So far we have demonstrated that if a solution F satisfies PO, CONT and M.STAB, then its restriction \bar{F} satisfies the two additional properties of WARP and BOUND. The remainder of the proof of the main theorem consists in showing that \bar{F} has an additively separable representation (Lemma 4.4 through Lemma 4.14 and Proposition 4.2); the extension of that result to F (Proposition 4.3) will then follow easily from the fact that F satisfies IIA.

In the statements of Lemma 4.4 through Lemma 4.14, it is assumed that the restricted solution \bar{F} satisfies PO, CONT, M.STAB, WARP and BOUND.

In what follows, extensive use will be made of the inverse choice correspondence $\Phi: \bigcup_{P \in \mathcal{P}} \mathcal{R}_+^P \rightarrow \bigcup_{P \in \mathcal{P}} \mathcal{R}_{++}^P$, associated with the restricted solution \bar{F} . Given $P \in \mathcal{P}$ and $x \in \mathcal{R}_+^P$, $\Phi(x)$ is defined by $\Phi(x) \equiv \{\pi \in \mathcal{R}_{++}^P \mid \bar{F}(\pi, \omega) = x \text{ for some } \omega \in \mathcal{R}_+^P\}$. Thus, Φ is the inverse of \bar{F} projected on the price space. Note that for all x , either $\Phi(x)$ is empty or it is a cone with the vertex 0 removed, because \bar{F} is homogeneous of degree 0 in π . Also note that for all $P, Q \in \mathcal{P}$, with $P \subset Q$, if $x \in \mathcal{R}_+^Q$, then $\Phi_P(x) \subset \Phi(x_P)$ by M.STAB. As already mentioned, this inclusion property is exactly what M.STAB says about the restricted solution \bar{F} . In Lemma 4.5, it will be shown that the reverse inclusion also holds if PO, CONT and BOUND are imposed as well. First, we show in Lemma 4.4 that Φ is convex-valued. This is a consequence of WARP alone.

Lemma 4.4: The correspondence Φ is convex-valued.

Proof: Let $P \in \mathcal{P}$ and $x \in \mathcal{R}_{++}^P$ be given. Let π^1 and π^2 be two members of $\Phi(x)$ and let $\pi^3 \equiv \alpha\pi^1 + (1-\alpha)\pi^2$ for some $\alpha \in (0,1)$. We must show that $\pi^3 \in \Phi(x)$ and the proof is by contradiction. If $\pi^3 \notin \Phi(x)$, then in particular, $x \neq y \equiv \bar{F}(\pi^3, x)$. Then $y \bar{R}_x$, which by WARP implies that $\pi^1 y > \pi^1 x$ and $\pi^2 y > \pi^2 x$. Since π^3 is a convex combination of π^1 and π^2 , it follows that $\pi^3 y > \pi^3 x$, which is impossible in view of the budget constraint $\pi^3 y \leq \pi^3 x$.
QED.

Lemma 4.5: For all $P, Q \in \mathcal{P}$ with $P \subset Q$, for all $x \in \mathcal{R}_{++}^Q$, $\Phi_P(x) = \Phi(x_P)$, i.e. $\pi^0 \in \Phi(x_P)$ if and only if there exists $\pi \in \Phi(x)$ such that $\pi_P = \pi^0$.

Proof: (See Figure 4.6). Let P, Q and x satisfy the hypothesis of the lemma. If $P = Q$, then $\Phi_P(x) = \Phi(x_P)$ identically. If $P \subset Q$, it follows by M.STAB that $\Phi_P(x) \subset \Phi(x_P)$. In order to establish the reverse inclusion, we consider first the case where $|P| = |Q| - 1$.

Let π^0 be a member of $\Phi(x_P)$. We must show that there exists $\pi \in \Phi(x)$ with $\pi_P = \pi^0$. Let $\{k\} \equiv Q \setminus P$, and let the function $\pi: (0,1) \rightarrow \Delta^Q$ be defined by $\pi_k(\alpha) \equiv \alpha$ and $\pi_P(\alpha) \equiv ((1-\alpha) / \sum_{i \in P} \pi_i^0) \pi^0$. Note that $\pi_P(\alpha) \in \Phi(x_P)$ for all $\alpha \in (0,1)$ since $\pi_P(\alpha)$ is proportional to π^0 and $\Phi(x_P)$ is a cone. Therefore, since $\Phi(x)$ is also a cone, it is sufficient to show that $\pi(\bar{\alpha}) \in \Phi(x)$ for some $\bar{\alpha} \in (0,1)$.

To see this, let $z(\alpha) \equiv \bar{F}(\pi(\alpha), x)$ for $\alpha \in (0,1)$. As $\alpha \rightarrow 1$, then $\pi_P(\alpha) \rightarrow 0$, and BOUND implies that $\|z(\alpha)\| \rightarrow \infty$. Because $\pi_k(\alpha) = \alpha$ and $\pi(\alpha)z(\alpha) \leq \pi(\alpha)x$, this implies that $z_k(\alpha) < x_k$ for α sufficiently close to 1. By a similar argument, $z_k(\alpha) > x_k$ for α sufficiently close to 0. Then, by the intermediate value theorem, there exists $\bar{\alpha} \in (0,1)$ such that $z_k(\bar{\alpha}) = x_k$, and it follows that

$$\begin{aligned} z_P(\bar{\alpha}) &= \bar{F}(\pi_P(\bar{\alpha}), z_P(\bar{\alpha})) \text{ (by PO and M.STAB)} \\ &= \bar{F}(\pi_P(\bar{\alpha}), x_P) \text{ (because } z_k(\bar{\alpha}) = x_k \Rightarrow \pi_P(\bar{\alpha})z_P(\bar{\alpha}) = \pi_P(\bar{\alpha})x_P) \\ &= x_P \text{ (since } \pi_P(\bar{\alpha}) \in \Phi(x_P)). \end{aligned}$$

Thus, $\bar{F}(\pi(\bar{\alpha}), x) = x$, which proves that $\pi(\bar{\alpha}) \in \Phi(x)$. Hence $\Phi(x_P) \subset \Phi_P(x)$ if $|P| = |Q| - 1$.

In order to show that $\Phi(x_P) \subset \Phi_P(x)$ for all $P \subset Q$, we pick $\pi^0 \in \Phi(x_P)$, set $P^1 \equiv P \cup \{i_1\}$, where $i_1 \in Q \setminus P$, and conclude by the first part of the proof

that $\pi_p^1 = \pi^0$ for some $\pi^1 \in \Phi(x_{p_1})$. Repeating this argument by adding $i_2 \in Q \setminus P^1$ to P^1 etc, we conclude that $\pi_p = \pi^0$ for some $\pi \in \Phi_p(x)$. QED.

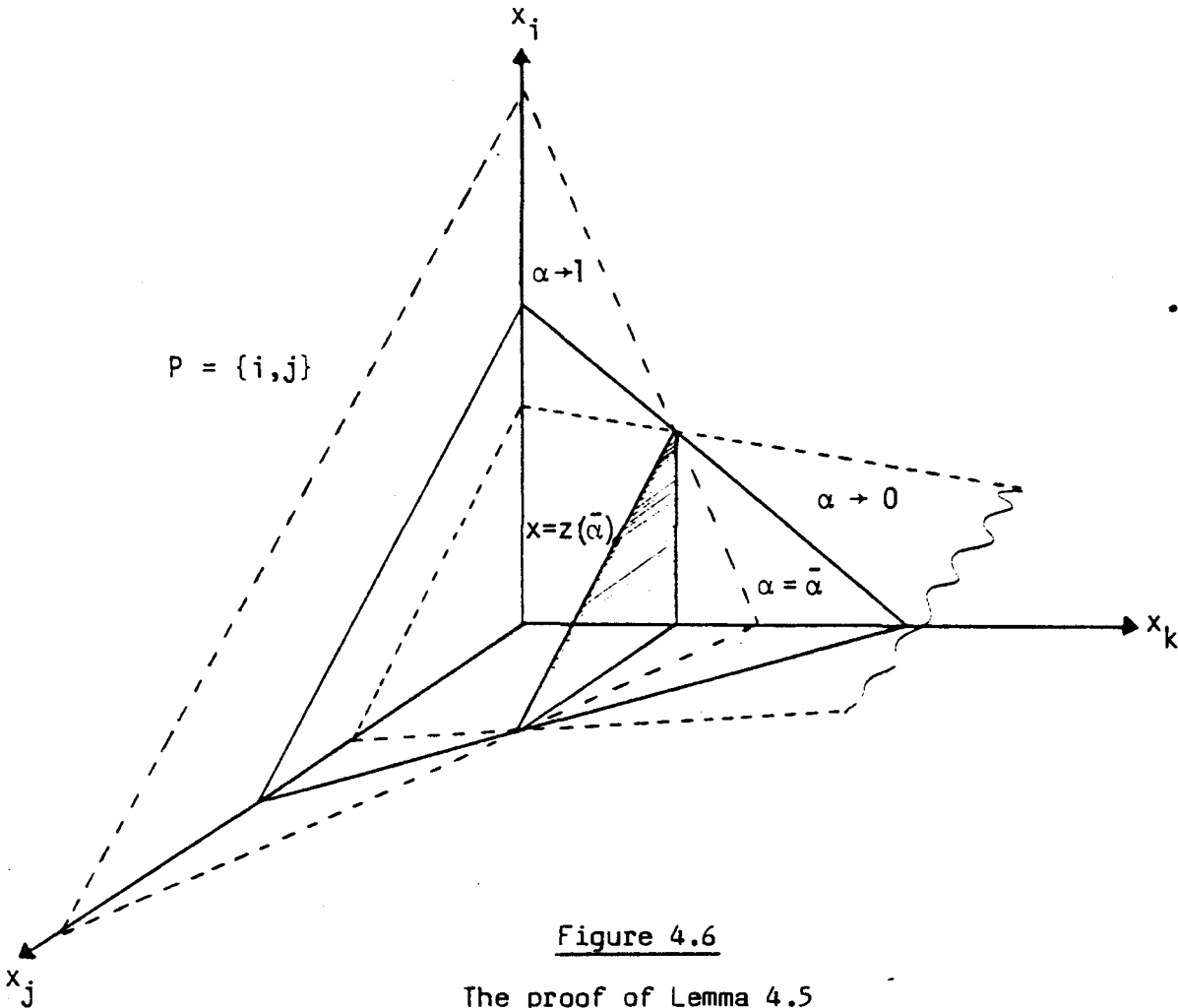


Figure 4.6

The proof of Lemma 4.5

Lemma 4.6: For all $P \in \mathcal{P}$, $\mathcal{R}_{++}^P \subset \bar{F}(\mathcal{R}_{++}^P, \mathcal{R}_+^P)$, i.e., every positive vector is the solution outcome to some budget problem.

Proof: Let $P \in \mathcal{P}$, $x \in \mathcal{R}_{++}^P$ and $i \in P$ be given. Because $\bar{F}(\pi_i, x_i) = x_i$ for $\pi_i > 0$ by P0, $\Phi(x_i) \neq \emptyset$, which by Lemma 4.5 implies that $\Phi(x) \neq \emptyset$, meaning that $\Phi(x)$ contains π such that $\bar{F}(\pi, \omega) = x$ for some $\omega \in \mathcal{R}_{++}^P$.

QED.

For all $P \in \mathcal{P}$, let \mathcal{D}^P be the set of points $x \in \mathcal{R}_{++}^P$ such that $\Phi(x)$ is a ray, and let $\bar{\mathcal{D}}^P \equiv \mathcal{R}_{++}^P \setminus \mathcal{D}^P$. Also, define $\mathcal{D} \equiv \bigcup_{P \in \mathcal{P}} \mathcal{D}^P$ and $\bar{\mathcal{D}} \equiv \bigcup_{P \in \mathcal{P}} \bar{\mathcal{D}}^P$. The set \mathcal{D} is of particular interest, because if \bar{F} has a numerical representation f , then $\Phi(x)$ is the set of tangent normals to the level curve of f at x . Thus, if $x \in \mathcal{D}^P$, then the level curve through x has a unique tangent at x , and therefore f is differentiable at x if f is additively separable.

In Lemma 4.7 through Lemma 4.9, we show that the set $\bar{\mathcal{D}}^P$ is negligible in the measure theoretic sense and that it has the following simple structure: Letting μ_n denote the n -dimensional Lebesgue-measure, there exists a sequence $\{\bar{\mathcal{D}}_i\}_{i \in I}$ of subsets of the non-negative real numbers, such that $\mu_1(\bar{\mathcal{D}}_i) = 0$ for all $i \in I$, and such that $\bar{\mathcal{D}}^P \subset \bigcup_{i \in P} (\bar{\mathcal{D}}_i \times \mathcal{R}_{++}^{P \setminus \{i\}})$ for all $P \in \mathcal{P}$.

Observe that if \bar{F} has an additively separable numerical representation, then $\bar{\mathcal{D}}^P$ must necessarily be of this form.

Lemma 4.7: For all $P \in \mathcal{P}$ with $|P| = 2$, $\mu_2(\bar{\mathcal{D}}^P) = 0$.

Proof: An illustration is given in Figure 4.7. Let $P \in \mathcal{P}$ with $|P| = 2$ be given. Since \mathcal{R}_{++}^P can be covered by a countable collection of rectangles, it is sufficient to show that $\mu_2(\bar{\mathcal{D}}^P \cap A) = 0$ for all rectangles $A \subset \mathcal{R}_{++}^P$. Let such an A be given, let $B \equiv \{\pi \in \mathcal{R}_{++}^P \mid \sum_{i \in P} \pi_i \leq 1\}$ and let

$$(1) \quad G \equiv \{(x, \pi) \in A \times B \mid x = F(\pi, \omega) \text{ for some } \omega \in \mathcal{R}_{++}^P\}.$$

For all $x \in A$ and all $\pi \in B$, let $G_x \equiv \{\pi' \mid (x, \pi') \in G\}$ and $G_\pi \equiv \{x' \mid (x', \pi) \in G\}$. Since $A \times B$ is bounded, then $\mu_4(G) < \infty$ and therefore

$$(2) \quad \int_A \mu_2(G_x) d\mu_2 = \mu_4(G) = \int_B \mu_2(G_\pi) d\mu_2.$$

The sets G_x and G_π are illustrated in figure 4.7, where G_x is a truncated ray of points proportional to π , i.e. a set of μ_2 -measure zero, and where $G_{x''}$ is a truncated convex cone with a non-empty interior, i.e. a set of positive measure. It will be shown next that the first case is the generic one.

The set G_π is an income-consumption path for the continuous demand function \bar{F} , restricted to A . Therefore $\mu_2(G_\pi) = 0$ for all $\pi \in B$, which by the second equality in (2) implies that $\mu_4(G) = 0$. It then follows by the first equality in (2) that $\mu_2(G_x) = 0$ for μ_2 -almost all $x \in A$.

Now, $G_x = \Phi(x) \cap B$ for all $x \in A$. Since $A \subset \mathbb{R}_{++}^P$, then $\Phi(x) \neq \emptyset$ for all $x \in A$ by Lemma 4.6. Therefore, since Φ is convex-valued by Lemma 4.4, then $\mu_2(\Phi(x) \cap B) > 0$ whenever $\Phi(x)$ is not a ray. Since $\mu_2(G_x) = 0$ for μ_2 -almost all $x \in A$, it follows that $\mu_2(\bar{D} \cap A) = 0$.

QED.

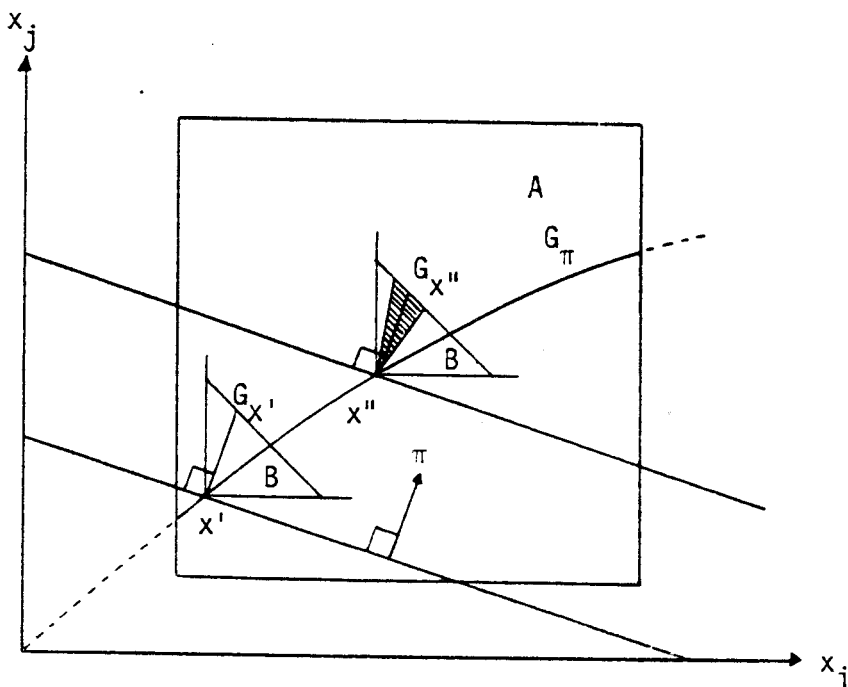


Figure 4.7

The proof of Lemma 4.7

Given $Q \in \mathcal{P}$ and a collection $\{P^1, \dots, P^n\}$ of subsets of Q , say that $\{P^1, \dots, P^n\}$ is a chain if $P^v \cap P^{v+1} \neq \emptyset$ for $v = 1, \dots, n-1$. If, in addition, $\{P^1, \dots, P^n\}$ covers Q (i.e. if $\bigcup_{v=1}^n P^v = Q$), say that $\{P^1, \dots, P^n\}$ is a Q-chain. For example, if $Q = \{1, 2, 3\}$, $P^1 = \{1, 2\}$ and $P^2 = \{2, 3\}$, then $\{P^1, P^2\}$ is a Q-chain.

Lemma 4.8: For all $Q \in \mathcal{P}$, for all $x \in \mathcal{R}_{++}^Q$, for all Q-chains $\{P^1, \dots, P^n\}$, if $x_{P^v} \in \mathcal{D}$ for all P^v in the chain, then $x_P \in \mathcal{D}$ for all $P \subset Q$.

Proof: Let Q, x and $\{P^1, \dots, P^n\}$ satisfy the hypothesis of the lemma. It is sufficient to show that $x \in \mathcal{D}$, for then $x_P \in \mathcal{D}$ for all $P \subset Q$ by Lemma 4.5. To this end, we define for each $v = 1, \dots, n$, the set $Q^v \equiv \bigcup_{i=1}^v P^i$ and show by induction on v that $x_{Q^v} \in \mathcal{D}$ for $v = 1, \dots, n$.

For $v = 1$, then $Q^1 = P^1$ and $x_{Q^1} \in \mathcal{D}$ by hypothesis. Suppose now that $x_{Q^v} \in \mathcal{D}$ for $v < n$. We must show that $x_{Q^{v+1}} \in \mathcal{D}$ also. Because $x_{Q^v} \in \mathcal{D}$ and $x_{P^{v+1}} \in \mathcal{D}$, then $\Phi(x_{Q^v})$ and $\Phi(x_{P^{v+1}})$ are both rays, which by Lemma 4.5 implies that $\Phi_{Q^v}(x_{Q^{v+1}})$ and $\Phi_{P^{v+1}}(x_{Q^{v+1}})$ are rays. Consequently, for all $\pi^1, \pi^2 \in \Phi(x_{Q^{v+1}})$, there exist positive real numbers α and β such that $\pi_{Q^v}^1 = \alpha \pi_{Q^v}^2$ and $\pi_{P^{v+1}}^1 = \beta \pi_{P^{v+1}}^2$. Since $P^v \subset Q^v$ and $P^v \cap P^{v+1} \neq \emptyset$, then $Q^v \cap P^{v+1} \neq \emptyset$, which implies that $\alpha = \beta$. Therefore, since $Q^v \cup P^{v+1} = Q^{v+1}$, it follows that $\pi^1 = \alpha \pi^2$, which proves that $\Phi(x_{Q^{v+1}})$ is a ray and hence that $x_{Q^{v+1}} \in \mathcal{D}$.

Thus, $x_{Q^v} \in \mathcal{D}$ for all $v = 1, \dots, n$, and since $Q^n = Q$ then $x_Q = x \in \mathcal{D}$.

QED.

Lemma 4.9: For all $i \in I$, there exists a set $\bar{D}_i \subset \mathbb{R}_+^{\{i\}}$ such that $\mu_1(D_i) = 0$, and such that for all $P \in \mathcal{P}$, $\bar{D}^P \subset \bigcup_{i \in P} (\bar{D}_i \times \mathbb{R}_+^{P \setminus \{i\}})$.

Proof: Let $\{z_i\}_{i \in I}$ be a sequence of positive real numbers such that for all distinct $i, j \in I$, $(z_i, z_j) \in \mathcal{D}$ and $(z_i, x_j) \in \mathcal{D}$ for μ_1 -almost all $x_j \geq 0$. Such a sequence exists by Lemma 4.7 because I is a countable set. For each $i \in I$, define the sets \bar{D}_i and D_i by

$$(1) \quad \bar{D}_i \equiv \{\bar{x}_i \geq 0 \mid (\bar{x}_i, z_j) \in \bar{\mathcal{D}} \text{ for some } j \in I \setminus \{i\}\},$$

and

$$(2) \quad D_i \equiv \mathbb{R}_+^{\{i\}} \setminus \bar{D}_i.$$

We first show that

$$(3) \quad \bar{D}_i = \{\bar{x}_i \geq 0 \mid (\bar{x}_i, z_j) \in \bar{\mathcal{D}} \text{ for all } j \in I \setminus \{i\}\}.$$

To see this, let $i \in I$ and $\bar{x}_i \in \bar{D}_i$ be given and let j be a member of $I \setminus \{i\}$ such that $(\bar{x}_i, z_j) \in \bar{\mathcal{D}}$. If k is a member of $I \setminus \{i, j\}$ such that $(\bar{x}_i, z_k) \in \mathcal{D}$, then since $(z_k, z_j) \in \mathcal{D}$, it follows by Lemma 4.8 that $(\bar{x}_i, z_j) \in \mathcal{D}$, a contradiction. Hence, (3) holds.

Next, we claim that if $\bar{x}_i \in \bar{D}_i$ for some $i \in I$, then $(\bar{x}_i, x_j) \in \bar{\mathcal{D}}$ for all $j \in I \setminus \{i\}$ and μ_1 -almost all $x_j \geq 0$. To prove the claim, observe that if $(\bar{x}_i, x_j) \in \mathcal{D}$ and $(x_j, z_k) \in \mathcal{D}$ for some $k \in I \setminus \{i, j\}$, then $(\bar{x}_i, z_k) \in \mathcal{D}$ by Lemma 4.8. Because $(\bar{x}_i, z_k) \in \bar{\mathcal{D}}$ by (3) and $(x_j, z_k) \in \mathcal{D}$ for μ_1 -almost all $x_j \geq 0$, this proves the claim.

We can now show that $\mu_1(\bar{D}_i) = 0$ for all $i \in I$. Because $\mu_2(\mathcal{D}^{\{i,j\}}) = 0$ for all $j \in I \setminus \{i\}$ by Lemma 4.7 and $(\bar{x}_i, x_j) \in \mathcal{D}^{\{i,j\}}$ for μ_1 -almost all $x_j \geq 0$, whenever $\bar{x}_i \in \bar{D}_i$, it follows that $\mu_1(\bar{D}_i) = 0$.

To complete the proof, let $P \in \mathcal{P}$ and $x \in \mathcal{R}_{++}^P$ be given. It is sufficient to show that if $x_i \in D_i$ for all $i \in P$, then $x \in \mathcal{D}^P$. If $x_i \in D_i$ for all $i \in P$, then by (3), for all distinct $i, j \in P$, there exists $k \in I \setminus \{i, j\}$ such that $(x_i, z_k) \in \mathcal{D}$ and $(z_k, x_j) \in \mathcal{D}$, which by Lemma 4.8 implies that $(x_i, x_j) \in \mathcal{D}$. Since this holds for all $i, j \in P$, it follows by Lemma 4.8 that $x \in \mathcal{D}$.

QED.

For all $i \in I$, let \bar{D}_i and D_i be defined as in (1) and (2) in the proof of Lemma 4.9. For all $P \in \mathcal{P}$, let $\bar{D}_P \equiv \bigcup_{i \in P} (\bar{D}_i \times \mathcal{R}_+^{P \setminus \{i\}})$ and $D_P \equiv \mathcal{R}_{++}^P \setminus \bar{D}_P$. Next, let ϕ be a single-valued selection from Φ , such that $\phi(x) \in \Phi(x) \cap \Delta^P$ for all $P \in \mathcal{P}$ and all $x \in \mathcal{R}_{++}^P$, where Δ^P is the P -dimensional unit simplex. Note that by Lemma 4.9, all such single-valued selections coincide on the sets D_P .

For all $P \equiv \{i, j\} \in \mathcal{P}$ and all $(x_i, x_j) \in D_i \times D_j$, let $M(x_i, x_j) \equiv \phi_i(x_i, x_j) / \phi_j(x_i, x_j)$ and similarly, $M(x_j, x_i) \equiv 1 / M(x_i, x_j) = \phi_j(x_i, x_j) / \phi_i(x_i, x_j)$. Observe that if the P -component of \bar{F} has a real-valued representation f which is differentiable at (x_i, x_j) , then $M(x_i, x_j)$ is equal to $f'_i(x_i, x_j) / f'_j(x_i, x_j)$ - the marginal rate of substitution (MRS) at (x_i, x_j) with respect to f .

In Lemma 4.10 through Lemma 4.12, we show that the MRS-functions, defined in this way, can be integrated to obtain a sequence from F of functions $\{f_i\}_{i \in I}$, such that each f_i is differentiable on D_i and such that for all $P \in \mathcal{P}$ and all $x \in D_P$, the budget plane of the unique budget problem solved at x is tangent to a level curve of the function $\sum_{i \in P} f_i$ at x . The first step is Lemma 4.10,

where it is shown that the function ϕ is sufficiently well-behaved for the MRS-function to be integrable.

Lemma 4.10: For all $P \in \mathcal{P}$, the P -component of the function ϕ is continuous on D_P , and for all $i \in P$, ϕ_i is bounded away from zero on compact subsets of \mathcal{R}_{++}^P .

Proof: Let $P \in \mathcal{P}$ be given. To show that ϕ is continuous on D_P , let $\{x^u\}$ be a sequence from \mathcal{R}_{++}^P converging to $x \in D_P$, and suppose, by way of contradiction, that $\phi(x^u)$ does not converge to $\phi(x)$. Then, because ϕ takes its values in Δ^P , which is a compact set, $\{x^u\}$ has a subsequence, which without loss of generality we can assume to be $\{x^u\}$ itself, such that $\phi(x^u) \rightarrow \pi \neq \phi(x)$.

Because $\bar{F}(\phi(x^u), x^u) = x^u$ for all u by P0 and construction of ϕ , it follows by CONT that $\bar{F}(\pi, x) = x$, unless $\pi_i = 0$ for some $i \in P$. However, the latter possibility is ruled out by BOUND, because $\|x^u\| \rightarrow \|x\| < \infty$. Since $\pi \in \Delta^P$ and $x \in D_P$ and since $\bar{F}(\pi, x) = x$ implies that $\pi \in \Phi(x)$, it follows that $\pi = \phi(x)$. This contradiction completes the proof that ϕ is continuous on D_P .

To prove the second part of the lemma, let A be a compact subset of \mathcal{R}_{++}^P and suppose, by way of contradiction, that ϕ_i is not bounded away from zero on A for some $i \in P$. Then there exists a sequence $\{x^u\}$ from A such that $\phi_i(x^u) \rightarrow 0$. Because $\bar{F}(\phi(x^u), x^u) = x^u$ for all u , it follows by BOUND that $\|x^u\| \rightarrow \infty$, a contradiction, since $\{x^u\} \subset A$ which is a compact subset of \mathcal{R}_{++}^P .

QED.

Lemma 4.11: For all $Q \in \mathcal{P}$ with $|Q| \geq 3$, for all $x \in D_Q$, for all $\{i, j, k\} \subset Q$, $M(x_i, x_j) = M(x_i, x_k) \cdot M(x_k, x_j) = \phi_i(x) / \phi_j(x)$.

Proof: Let Q and x satisfy the hypothesis of the lemma. Since $x \in D_Q$, then $x \in D^Q$ by Lemma 4.9, and thus $\Phi(x)$ is a ray. By Lemma 4.5, $\Phi(x_P) = \Phi_P(x)$ for all $P \subset Q$, thus $\Phi(x_P)$ is a ray for all $P \subset Q$. Therefore, by definition of the function ϕ , $\phi(x_P)$ is proportional to $\phi_P(x)$ for all $P \subset Q$, and since $M(x_i, x_j) \equiv \phi_i(x_i, x_j)/\phi_j(x_i, x_j)$ for all $P = \{i, j\}$, it follows that $M(x_i, x_j) = \phi_i(x)/\phi_j(x)$ for all such P . Since $\phi_i(x)/\phi_j(x) = (\phi_i(x)/\phi_k(x)) \cdot (\phi_k(x)/\phi_j(x))$, the conclusion follows. QED.

Lemma 4.12: There exists a sequence $\{f_i\}_{i \in I}$ of functions $f_i: \mathcal{R}_+^{\{i\}} \rightarrow \mathcal{R}^{\{i\}} \cup \{-\infty\}$, where each f_i is continuous and strictly increasing on $\mathcal{R}_+^{\{i\}}$ and absolutely continuous on compact intervals in $\mathcal{R}_{++}^{\{i\}}$, such that for all $P \in \mathcal{P}$, all $x \in D_P$ and all $\{i, j\} \subset P$, $f'_i(x_i)/f'_j(x_j) = \phi_i(x)/\phi_j(x)$ (where f'_k denotes the derivative of f_k).

Proof: We begin by constructing for each $i \in I$ a function $g_i: \mathcal{R}_{++}^{\{i\}} \rightarrow \mathcal{R}_{++}^{\{i\}}$, whose indefinite integral will be f_i . For each $i \in I$, let x_i^0 be a member of D_i . Define

$$(1) \quad g_1(x_1) \equiv M(x_1, x_2^0) \cdot M(x_2^0, x_1^0), \text{ and}$$

$$(2) \quad g_i(x_i) \equiv M(x_i, x_1^0) \text{ for } i \in I \setminus \{1\}.$$

For all $i \in I$, let the function $f_i: \mathcal{R}_+^{\{i\}} \rightarrow \mathcal{R}^{\{i\}} \cup \{-\infty\}$ be defined by

$$f_i(x_i) \equiv \int_{x_i^0}^{x_i} g_i(\alpha) d\alpha \text{ for all } x_i > 0, \text{ and } f_i(0) \equiv \lim_{x_i \rightarrow 0} f_i(x_i).$$

To see that the functions f_i are well defined, observe first that g_i is

continuous on D_i and bounded on any compact interval of $\mathbb{R}_{++}^{\{i\}}$. This follows by (1), (2) and Lemma 4.10 since $M(x_i, x_j^0) \equiv \phi_i(x_i, x_j^0)/\phi_j(x_i, x_j^0)$ for all $i, j \in I$ and since $x_i^0 \in D_i \subset \mathbb{R}_{++}^{\{i\}}$ for all $i \in I$. Therefore g_i is Riemann-integrable and f_i is absolutely continuous. Because $\phi(x) > 0$ whenever $x > 0$, it follows that f_i is strictly increasing, which implies that $\lim_{x_i \rightarrow 0} f_i(x_i)$ exists as a real number or $-\infty$.

To complete the proof, we must show that $f'_i(x_i)/f'_j(x_j) = \phi_i(x)/\phi_j(x)$ for all $P \in \mathcal{P}$, all $x \in D_P$ and all $\{i, j\} \subset P$. Observe first that the derivative $f'_i(x_i)$ exists and is equal to $g_i(x_i)$ for all $x_i \in D_i$ because g_i is continuous at such points. Thus, because $\phi_i(x)/\phi_j(x) = M(x_i, x_j)$ whenever $x \in D_P$ by Lemma 4.11, it is sufficient to show that $g_i(x_i)/g_j(x_j) = M(x_i, x_j)$ for all distinct $i, j \in I$ and all $(x_i, x_j) \in D_i \times D_j$.

Let distinct $i, j \in I$ and $(x_i, x_j) \in D_i \times D_j$ be given. Recall that $x_k^0 \in D_k$ for all $k \in I$. Suppose first that $i \neq 1 \neq j$. Then

$$\begin{aligned} (3.1) \quad M(x_i, x_j) &= M(x_i, x_1^0) \cdot M(x_1^0, x_j) && \text{(by Lemma 4.11)} \\ (3.2) \quad &\equiv M(x_i, x_1^0)/M(x_j, x_1^0) \\ (3.3) \quad &= g_i(x_i)/g_j(x_j) && \text{(by (2))} \end{aligned}$$

Next suppose that $i = 1$ and $j \neq 2$. Then

$$\begin{aligned} (4.1) \quad M(x_1, x_j) &= M(x_1, x_2^0) \cdot M(x_2^0, x_j) && \text{(by Lemma 4.11)} \\ (4.2) \quad &= M(x_1, x_2^0) \cdot M(x_2^0, x_1^0) \cdot M(x_1^0, x_j) && \text{(by Lemma 4.11)} \\ (4.3) \quad &= g_1(x_1) \cdot M(x_1^0, x_j) && \text{(by (1))} \\ (4.4) \quad &= g_1(x_1)/M(x_j, x_1^0) \\ (4.5) \quad &= g_1(x_1)/g_j(x_j) && \text{(by (2))} \end{aligned}$$

Finally, if $i = 1$ and $j = 2$, we let k be a member of $I \setminus \{1, 2\}$. Then

$$(5.1) \quad M(x_1, x_2) = M(x_1, x_k^0) \cdot M(x_k^0, x_2) \quad (\text{by Lemma 4.11})$$

$$(5.2) \quad = (g_1(x_1)/g_k(x_k^0)) \cdot (g_k(x_k^0)/g_2(x_2)) \quad (\text{by (4.5) and (3.3)})$$

$$(5.3) \quad = g_1(x_1)/g_2(x_2)$$

Thus, $g_i(x_i)/g_j(x_j) = M(x_i, x_j)$ for all distinct $i, j \in I$ and all $(x_i, x_j) \in D_i \times D_j$. QED.

Having established that the single-valued selection ϕ from the inverse choice correspondence Φ can be integrated to yield the family of functions $\{f_i\}_{i \in I}$, the next step is to show that the function $\sum_{i \in P} f_i$ has the right curvature, i.e. strict quasi-concavity, for all $P \in \mathcal{P}$. To this end, we first demonstrate in Lemma 4.13 that the ordering of \mathcal{R}_+^P induced by the function $\sum_{i \in P} f_i$ agrees with the revealed preference relation R_F^P induced by \bar{F} in the sense that $\sum_{i \in P} f_i(x_i^1) \geq \sum_{i \in P} f_i(x_i^0)$ whenever $x^1 R_F^P x^0$, with strict inequality if $\sum_{i \in P} f_i(x_i^0)$ is finite. Quasi-concavity of the function $\sum_{i \in P} f_i$ is then established in Lemma 4.14.

Lemma 4.13: For all $Q \in \mathcal{P}$, for all $x^0, x^1 \in \mathcal{R}_+^Q$, if $x^1 R_F^Q x^0$ and $\sum_Q f_i(x_i^0) > -\infty$, then $\sum_Q f_i(x_i^1) > \sum_Q f_i(x_i^0)$.

Proof: Let Q , x^0 and x^1 satisfy the hypothesis of the lemma. We distinguish between two cases, according to whether $x_i^0 \neq x_i^1$ for all $i \in Q$ or not. See figure 4.8 for an illustration.

(i) $x_i^0 \neq x_i^1$ for all $i \in Q$. For all $t \in (0,1)$, let $x(t) \equiv tx^1 + (1-t)x^0$.

 Since $x_i^0 \neq x_i^1$ for all $i \in Q$, then $x(t) > 0$ for all such t which by Lemma 4.6 implies that $\phi(x(t)) \neq \emptyset$. Since $\bar{F}(\phi(x(t)), x(t)) = x(t)$ by P0, WARP implies that $\phi(x(t)) \cdot (x^1 - x(t)) > 0$, and since $x^1 - x(t) = (1-t)(x^1 - x^0)$, it follows that

$$(1) \phi(x(t)) \cdot (x^1 - x^0) > 0 \text{ for all } t \in (0,1).$$

Since $x_i^0 \neq x_i^1$ for all $i \in Q$, each function $x_i(\cdot)$ is linear and not constant. Lemma 4.9 then implies that $x(t) \in D_Q$ for almost all $t \in (0,1)$. By Lemma 4.12, the vector $(f'_i(x_i))_{i \in Q}$ is well defined and proportional to $\phi(x)$ whenever $x \in D_Q$, thus by (1),

$$(2) \sum_{i \in Q} f'_i(x_i(t)) \cdot (x_i^1 - x_i^0) > 0 \text{ for almost all } t \in (0,1).$$

By Lemma 4.12, each f_i is absolutely continuous on compact intervals of $\mathcal{R}_{++}^{\{i\}}$. Therefore, since each $x_i(\cdot)$ is linear and strictly positive on $(0,1)$, the LHS of (2) integrable on $[\varepsilon, 1-\varepsilon]$ for small $\varepsilon > 0$. Thus, by (2), for all such ε , it follows that

$$(3.1) \quad 0 < \int_{\varepsilon}^{1-\varepsilon} \sum_{i \in Q} f'_i(x_i(t)) \cdot (x_i^1 - x_i^0) dt$$

$$(3.2) \quad = \sum_{i \in Q} \int_{\varepsilon}^{1-\varepsilon} f'_i(x_i(t)) \cdot (x_i^1 - x_i^0) dt$$

$$(3.3) \quad = \sum_{i \in Q} f_i(x_i(1-\varepsilon)) - \sum_{i \in Q} f_i(x_i(\varepsilon))$$

As $\varepsilon \rightarrow 0$, $x_i(1-\varepsilon) \rightarrow x_i^1$ and $x_i(\varepsilon) \rightarrow x_i^0$. Since each f_i is continuous on $\mathcal{R}_{+}^{\{i\}}$ by Lemma 4.12 and since the RHS of (3.3) is a strictly increasing function

of by (2), it follows by (3) that $\sum_{i \in Q} f_i(x_i^1) - \sum_{i \in Q} f_i(x_i^0) > 0$. This completes the proof for case (i).

(ii) $x_i^0 = x_i^1$ for some $i \in Q$. Let P be the subset of Q such that $x_i^0 \neq x_i^1$ for all $i \in P$ and $x_i^0 = x_i^1$ for all $i \in Q \setminus P$. Since $x^1 \underset{F}{R} x^0$, then $P \neq \emptyset$, moreover since $x_{Q \setminus P}^1 = x_{Q \setminus P}^0$, then M.STAB implies that $x_P^1 \underset{F}{R} x_P^0$. Since each f_i is strictly increasing, then $f_i(x_i^0) < \infty$ for all $i \in Q$ and since $\sum_{i \in Q} f_i(x_i^0) > \infty$ by hypothesis, then $f_i(x_i^0)$ is finite for all $i \in Q$. Since $x_{Q \setminus P}^1 = x_{Q \setminus P}^0$, it follows that

$$(4) \sum_{i \in P} f_i(x_i^1) - \sum_{i \in P} f_i(x_i^0) = \sum_{i \in Q} f_i(x_i^1) - \sum_{i \in Q} f_i(x_i^0).$$

Since $\sum_{i \in P} f_i(x_i^0) > \infty$ and $x_P^1 \underset{F}{R} x_P^0$, it follows by (4) and the proof of case (i) that $\sum_{i \in Q} f_i(x_i^1) > \sum_{i \in Q} f_i(x_i^0)$. QED.

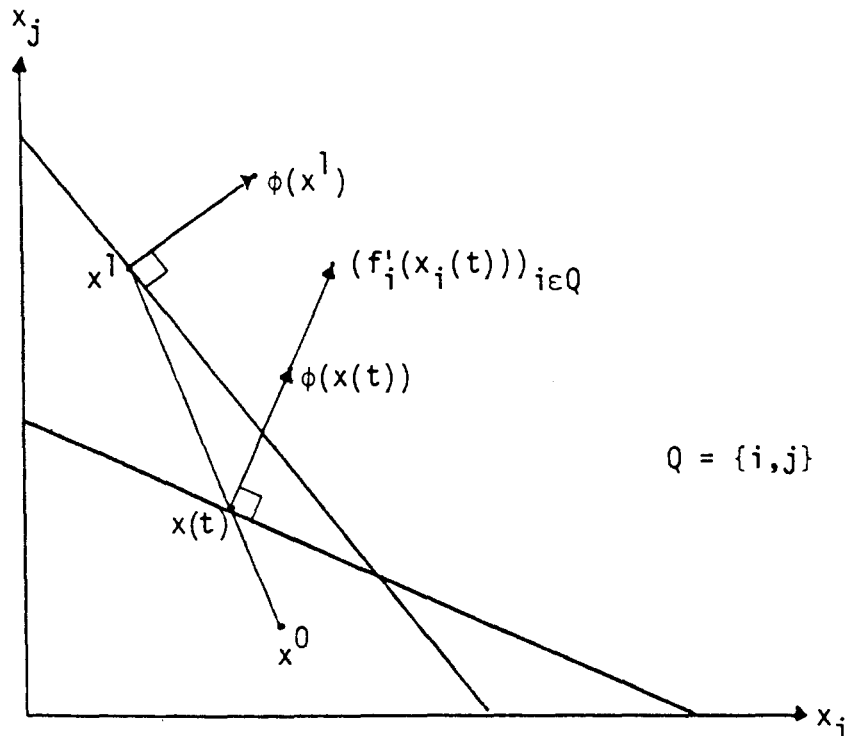


Figure 4.8

The proof of Lemma 4.13

Lemma 4.14: For all $Q \in \mathcal{P}$, the function $\sum_{i \in Q} f_i$ is strictly quasi-concave.

Proof: Let $Q \in \mathcal{P}$ and $y \in \mathbb{R}_+^Q$ be given, and let $c \equiv \sum_{i \in Q} f_i(y_i)$. We must show that the set $G(y) \equiv \{x \in \mathbb{R}_+^Q \mid \sum_{i \in Q} f_i(x_i) \geq c\}$ is strictly convex in \mathbb{R}_+^Q . This is clearly the case if $c = -\infty$, for then $G(y) = \mathbb{R}_+^Q$. Suppose $c > -\infty$, let x^0 and x^2 be two distinct points in $G(y)$ and let $x^1 \equiv \alpha x^0 + (1-\alpha)x^2$ for some $\alpha \in (0,1)$. To show that $\sum_{i \in Q} f_i(x_i^1) > c$, we distinguish between two cases, according to whether $x_i^0 \neq x_i^2$ for all $i \in Q$ or not.

(i) If $x_i^0 \neq x_i^2$ for all $i \in Q$, then $x^1 > 0$, which by PO and Lemma 4.6 implies that $\bar{F}(\pi, x^1) = x^1$ for some $\pi > 0$. Because x^1 is a convex combination of x^0 and x^2 , either $\pi x^1 \geq \pi x^0$ or $\pi x^1 \geq \pi x^2$. Assume without loss of generality that $\pi x^1 \geq \pi x^0$. Then $x^1 \underset{F}{R} x^0$, and since $\sum_{i \in Q} f_i(x_i^0) \geq c > -\infty$ because $x^0 \in G(y)$, it follows by Lemma 4.13 that $\sum_{i \in Q} f_i(x_i^1) > \sum_{i \in Q} f_i(x_i^0) \geq c$.

(ii) If $x_i^0 = x_i^2$ for some $i \in Q$, one may use the fact that $c > -\infty$ in an argument similar to the one used to prove (ii) of Lemma 4.13 to obtain the same conclusion as in (i). QED.

The main result concerning the restricted solution \bar{F} can now be established.

Proposition 4.2: If \bar{F} satisfies PO, CONT, M.STAB, WARP and BOUND, then it has an additively separable numerical representation.

Proof: Let $\{f_i\}_{i \in I}$ be the sequence of functions introduced in Lemma 4.12. By Lemma 4.12, each f_i is continuous and strictly increasing, and by Lemma

4.13, $\sum_{i \in P} f_i$ is strictly quasi-concave for all $P \in \mathcal{P}$. It remains to show that $\sum_{i \in P} f_i$ represents the P -component of F for all $P \in \mathcal{P}$.

To see this, let $P \in \mathcal{P}$ and $(\pi, \omega) \in \mathbb{R}_{++}^P \times \mathbb{R}_+^P$ be given, and let $S \equiv b(\pi, \omega)$. We must show that $z \equiv \operatorname{argmax}\{\sum_{i \in P} f_i(x_i) \mid x \in S\} = \bar{F}(\pi, \omega)$. If $\omega = 0$, the proof is trivial, so assume that $\omega > 0$. Then S contains a strictly positive vector y because $\pi > 0$. Therefore, since S is compact and convex then, by the existence part of the proof of Proposition 4.1, z exists and is unique. Moreover, $\sum_{i \in P} f_i(z_i) > \sum_{i \in P} f_i(y_i) > -\infty$ because each f_i is strictly increasing and $y > 0$. It then follows that $\bar{F}(\pi, \omega) = z$, for otherwise $\bar{F}(\pi, \omega) \underset{F}{R} z$ and

$\sum_{i \in P} f_i(z_i) > -\infty$, which is impossible in view of Lemma 4.13 since z maximizes $\sum_{i \in P} f_i$ on S . QED.

Proposition 4.2 is an interesting by-product of the analysis, as it gives sufficient conditions for a demand function to be consistent with the maximization of an additively separable utility function. It may well be that Proposition 4.2 is not the strongest result that can be proved in that respect. A question that is left open for further investigation, is whether WARP is implied by the other four axioms in Proposition 4.2. This is indeed the case if the correspondence between budget problems and solution outcomes is one-to-one, i.e. if the normalized inverse choice correspondence is single-valued everywhere. Note however, that PO, CONT and M.STAB alone are not sufficient conditions for utility maximization, as these conditions are also consistent with the minimization of a strictly quasi-convex and additively separable function on the budget plane.

Our next task is to show that the result obtained in Proposition 4.2 extends from budget problems to choice problems in general:

Proposition 4.3: If a solution satisfies PO, CONT and M.STAB, then it has an additively separable numerical representation.

Proof: Let F be a solution satisfying PO, CONT and M.STAB. By Lemma 4.1, F satisfies IIA and by Lemma 4.2 and Lemma 4.3, the restricted solution \bar{F} satisfies the additional properties of WARP and BOUND. By Proposition 4.2, \bar{F} has an additively separable numerical representation $\{f_i\}_{i \in I}$. We must show that $\{f_i\}_{i \in I}$ represents F also.

Let $P \in \mathcal{P}$ and $S \in \Sigma^P$ be given, and let $z \equiv \operatorname{argmax}\{\sum_{i \in P} f_i(x_i) \mid x \in S\}$. Proposition 4.1 implies that z exists and is unique. By convexity of S and quasi-concavity of $\sum_{i \in P} f_i$, there exists a hyperplane H in \mathcal{R}^P with normal π that separates S and $\{x \in \mathcal{R}_+^P \mid \sum_{i \in P} f_i(x_i) > \sum_{i \in P} f_i(z_i)\}$. π is strictly positive since $\sum_{i \in P} f_i$ is strictly increasing and since S contains a strictly positive vector. This implies that $S' \equiv \operatorname{cch}\{H \cap \mathcal{R}_+^P\}$ is a member of Σ_Δ^P . By Proposition 4.2, $F(S') \equiv \bar{F}(\pi, z) = z$, and since $z \in S \subset S'$, it follows by IIA that $F(S) = z$.

QED.

The announced characterization of the family of CRS solutions is now obtained by combining Propositions 4.1 and 4.3:

Theorem 4.1: A solution satisfies PO, CONT and M.STAB if and only if it is a CRS solution.

4.4 Variants of the main result.

In this section, we first show that the three axioms PO, CONT and M.STAB are independent, in the sense that removing any one of them will yield solutions that are not collectively rational. Then we investigate the consequences of adding more axioms to this list. Finally, we remove the requirement that the number of agents be infinite, relax M.STAB to B.STAB, and show that the conclusion of Proposition 4.3 still holds.

In order to show that the three axioms in Theorem 4.1 are independent, we begin with the following example of a solution which satisfies CONT and M.STAB, but not PO, and which is not collectively rational: Let $k \in I$ and $\lambda \in (0,1)$ be given. For all $P \in \mathcal{P}$ and all $S \in \Sigma^P$, set $F_i(S) \equiv 0$ if $i \neq k$ and $F_i(S) \equiv \lambda \cdot \max\{x_i \mid x \in S\}$ if $i = k$. Clearly, F satisfies CONT and M.STAB but not IIA, and therefore F is not collectively rational.

Next, we give an example of a solution which satisfies PO and M.STAB but not CONT, and which is not collectively rational. Let $P \equiv \{1,2\}$ and let \succsim be the usual lexicographic ordering of $R^{I \setminus P}$. Extend \succsim to R^I by defining $x \sim y$ iff $x = y$ and $x \succ y$ iff either of the following conditions holds:

- (i) $x_{I \setminus P} \succ y_{I \setminus P}$
- (ii) $x_{I \setminus P} = y_{I \setminus P}$ and $\min(x_1, x_2) > \min(y_1, y_2)$
- (iii) $x_{I \setminus P} = y_{I \setminus P}$ and $\min(x_1, x_2) = \min(y_1, y_2)$ and $\max(x_1, x_2) > \max(y_1, y_2)$
- (iv) $x_{I \setminus P} = y_{I \setminus P}$ and $\min(x_1, x_2) = \min(y_1, y_2)$ and $\max(x_1, x_2) = \max(y_1, y_2)$ and $x_1 > y_1$.

For all $Q \in \mathcal{P}$, let \succsim_Q denote the restriction of \succsim to \mathbb{R}^Q and define the solution F by $F(S) \equiv \{x \in PO(S) \mid \nexists y \in PO(S), x \succ_Q y\}$ for all $Q \in \mathcal{P}$ and all $S \in \Sigma^Q$. Thus, $F(S)$ is obtained by minimizing the ordering \succsim_Q on $PO(S)$. This is illustrated in Figure 4.9, which depicts a problem S such that $PO(S)$ is a proper subset of $WPO(S)$, and where the arrows indicate the direction of increased preference of the ordering \succsim . It is easy to check that the minimizer exists, and since $x \sim y$ iff $x = y$, it is unique. Moreover, F satisfies PO by construction and $M.STAB$ by an argument similar to that of Proposition 4.1. However, F does not satisfy IIA , and therefore it is not collectively rational.

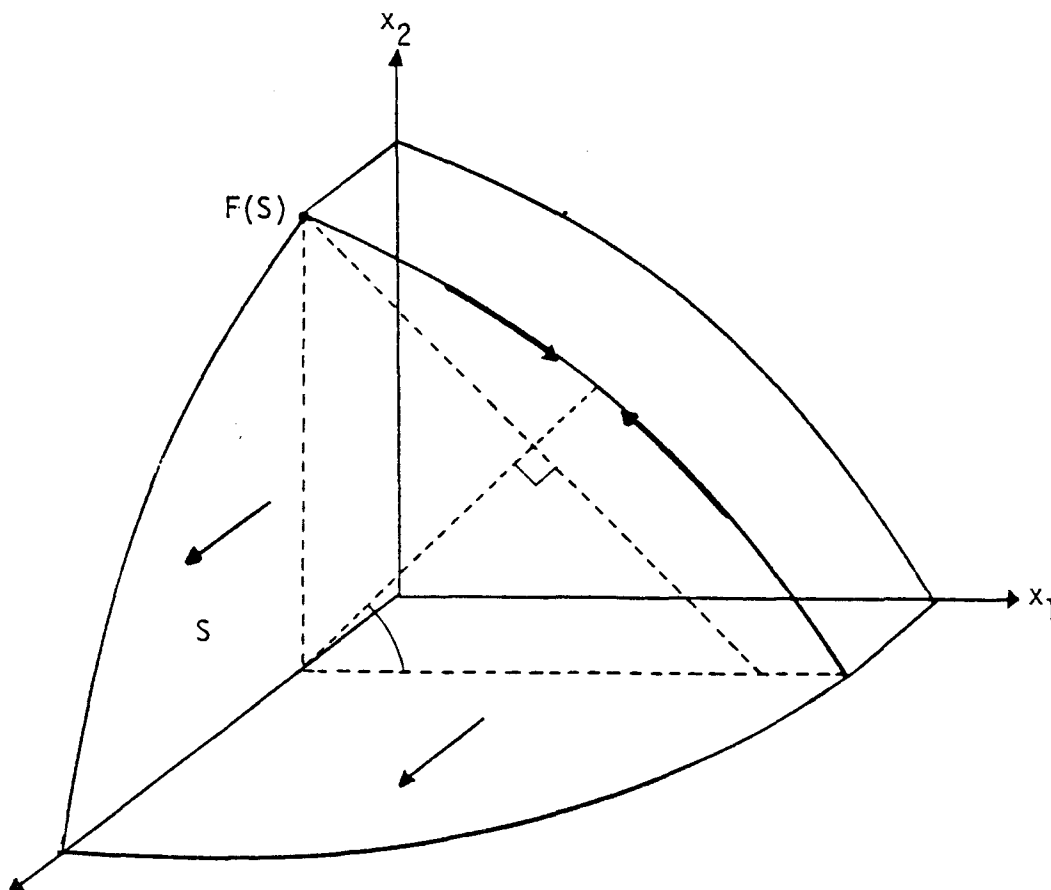


Figure 4.9

A solution satisfying PO and $M.STAB$ but not IIA

These two examples show that PO, CONT and M.STAB constitute a minimal set of conditions for a solution to be a member of the CRS-family. (Removing M.STAB would admit a large set of solutions that are not collectively rational).

The examples given above of solutions violating either PO or CONT may indicate that the main role of these two axioms is to rule out some peculiarities that M.STAB permits. However, a certain peculiarity turns out to survive all three axioms, as shown next.

For all $P \in \mathcal{P}$, all $S \in \Sigma^P$ and all real numbers $\alpha > 0$, let αS denote the choice problem $\{x \in \mathcal{R}_+^P \mid x/\alpha \in S\}$. Given $S \in \Sigma^P$ and $\alpha > 1$, the symmetric expansion of the set of feasible alternatives from S to αS could come about as the result of what may be called a welfare neutral growth in the underlying set of physical alternatives. It turns out that the response of a solution to welfare neutral growth may be quite pathological. Given a solution F and an agent k , say that k is an eventual dictator for F if for all $P \in \mathcal{P}$ with $k \in P$ and all $S \in \Sigma^P$, $F_k(\alpha S) \rightarrow \max\{x_k \mid x \in \alpha S\}$ as $\alpha \rightarrow \infty$. In other words, agent k is an eventual dictator for F if persistent welfare neutral growth always causes the income distribution to approach his preferred alternative.

As an example of a CRS solution that creates an eventual dictator, let $f_1 \equiv e^{x_1}$ and $f_j \equiv (1/j)\log x_j$ for all $j \neq 1$. Then each f_i is continuous and strictly increasing, moreover, $\sum_{i \in P} f_i$ is strictly quasi-concave for all $P \in \mathcal{P}$ by Theorem 11 in Debreu and Koopmans (1982), even though f_1 is strictly convex. Thus, there exists a CRS solution F which is represented by the sequence of functions $\{f_i\}_{i \in I}$. The eventual dictator is agent 1, and it is

the strict convexity of the function f_1 which is responsible for this fact. The response of the solution F to welfare neutral growth is illustrated in figure 4.10, where the solid curve shows the income expansion path for $P = \{1,2\}$ and $\alpha S = \{x \in \mathbb{R}_+^P \mid (x_1/\alpha)^2 + (x_2/\alpha)^2 \leq 1\}$. For low levels of aggregate income, measured by the parameter α , both agents gain from an increase in income. However, at some point, agent 1's dictatorial tendencies start to dominate, and in the limit he will end up with his preferred alternative, leaving nothing to agent 2. In order to eliminate this type of phenomenon, one might want to impose additional restrictions on the solution.

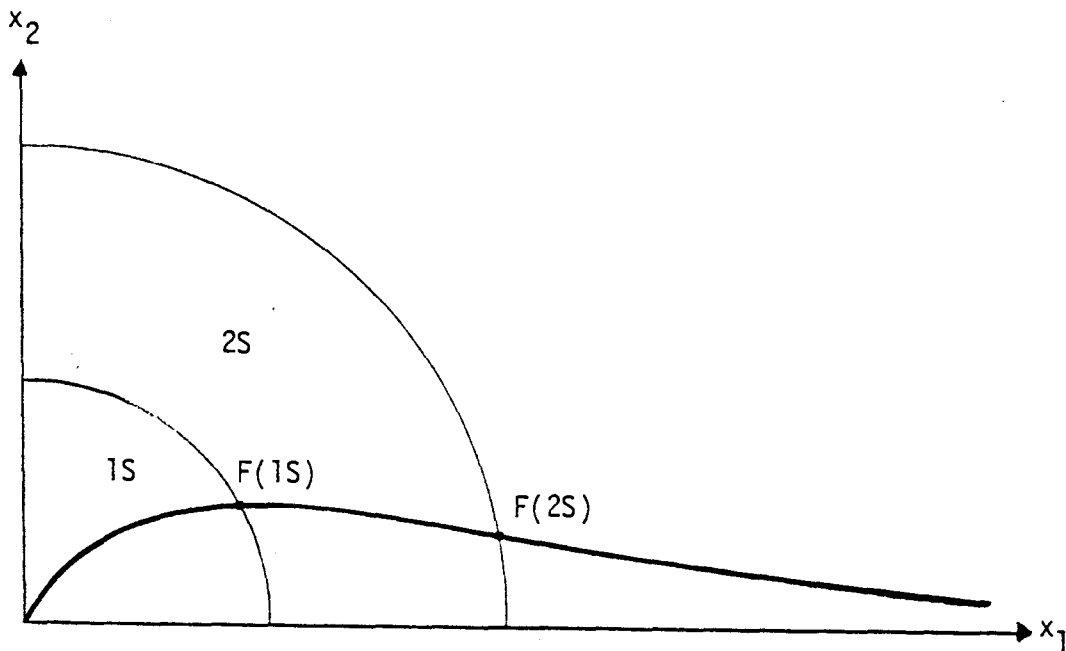


Figure 4.10

Agent 1 is an eventual dictator

Two conditions that are satisfied by the solutions associated with the four Bergson-Samuelson swf's mentioned in the introduction of this chapter (the

Utilitarian, Rawlsian, Leximin and Nash swf's) are Homogeneity (HOM) and Symmetry (SY). HOM says that if two problems S and S' have the property that $S' = \alpha S$ for some $\alpha > 0$, then $F(S') = \alpha F(S)$. Any of these two axioms will prevent a solution from creating an eventual dictator.

Adding HOM to the list of axioms in Theorem 4.1 implies that the functions $\sum_{i \in P} f_i$ must be homothetic for all $P \in \mathcal{P}$. This means (Eichhorn (1978) Theorem 2.2.1) that (except for arbitrary constant terms) there exists $\rho > -1$ and a sequence $\{\alpha_i\}_{i \in I}$ of positive real numbers such that $f_i(x_i) = -(\alpha_i/\rho)x_i^{-\rho}$ for all $i \in I$ if $\rho \neq 0$, and $f_i(x_i) = \alpha_i \log x_i$ for all $i \in I$ if $\rho = 0$. Thus $\sum_{i \in P} f_i$ is a CES function for all $P \in \mathcal{P}$. If SY is also imposed, then f_i , and hence α_i , must be the same for all i . The Nash swf is obtained when $\rho \rightarrow 0$, the Utilitarian swf when $\rho \rightarrow -1$ and the Rawlsian swf when $\rho \rightarrow \infty$. See Roberts (1980) for related results in the Arrow tradition of social choice theory.

Alternatively, dropping SY and strengthening HOM to S.INV would imply that $\rho = 0$, yielding a whole family of non-symmetric Nash solutions, cf. Harsanyi and Selten (1972), Kalai (1977b) and Roth (1979b).

It should be noted that the Utilitarian and Rawlsian swf's do not themselves yield well defined solutions, since their maximizers are not always unique on the domain considered here. One may then consider single-valued selections, but this will be at the necessary cost of relaxing either PO or CONT. For example, keeping PO and dropping CONT will admit the lexicographic extension of the Rawlsian swf as shown in Chapter 3.

In the proof of Theorem 4.1, the requirement that the set I of potential agents be infinite was only used in the proof of Lemma 4.1, which says that

PO, CONT and M.STAB imply IIA. If I were finite, the rest of the proof of the theorem would still go through, as long as $|I| \geq 3$.

In the remainder of this section, it is assumed that the number of potential agents is at least 3.

We show in Lemma 4.19, which differs from Lemma 4.1 by using B.STAB instead of M.STAB, that the conclusion of Lemma 4.1 still holds. As part of the proof, we show that PO, CONT and B.STAB imply that the restricted solution F must satisfy M.STAB as well. Since the proof of Proposition 4.3 makes no use of M.STAB, this result strengthens the only if-part of Theorem 4.1 in two ways - first, by reducing the minimum number of potential agents from infinite to finite, and second, by weakening M.STAB to B.STAB.

We begin by introducing some new notation and a dual version of the axiom of IIA which will be useful for the proof of Lemma 4.19.

Given a solution F , $Q \in \mathcal{P}$ and given $T \in \Sigma^Q$, recall from chapter 2 the definition of $B_F(T)$ as the set of F -bilaterally stable points of T . $B_F(T)$ is the set of points $x \in T$ such that for all $P, Q \in \mathcal{P}$ with $P \subset Q$ and $|P| = 2$ and for all $S \in \Sigma^P$, if $S = t_P^x(T)$ then $F(S) = x_P$. Given $P \in \mathcal{P}$ and a subset A of \mathcal{R}^P , $cl(A)$ denotes the closure of A .

Given $P \in \mathcal{P}$, and $S, S' \in \Sigma^P$, say that S and S' coincide in a neighbourhood U if $S \cap U = S' \cap U$. Given $x \in \mathcal{R}_+^P$, if there exists a sequence $\{S^u\}$ from Σ^P converging to S , such that for all u , S^u and S' coincide in a neighbourhood U^u of x , say that S coincides with S' at x .

The following axiom, which is illustrated in Figure 4.11 below, is a generalization of an axiom due to Thomson (1981):

Independence of Irrelevant Expansions (IIE): For all $P \in \mathcal{P}$, for all $S, S' \in \Sigma^P$, if $S' \subset S$ and S coincides with S' at $F(S')$ then $F(S) = F(S')$.

In Thomson's original version of this axiom, the sets S and S' were assumed to have a smooth weakly Pareto-optimal boundary, and the notion of S coinciding with S' at $F(S')$ was formulated in terms of supporting hyperplanes to S and S' at $F(S')$. The formulation adopted here allows one to dispense with the smoothness assumption.

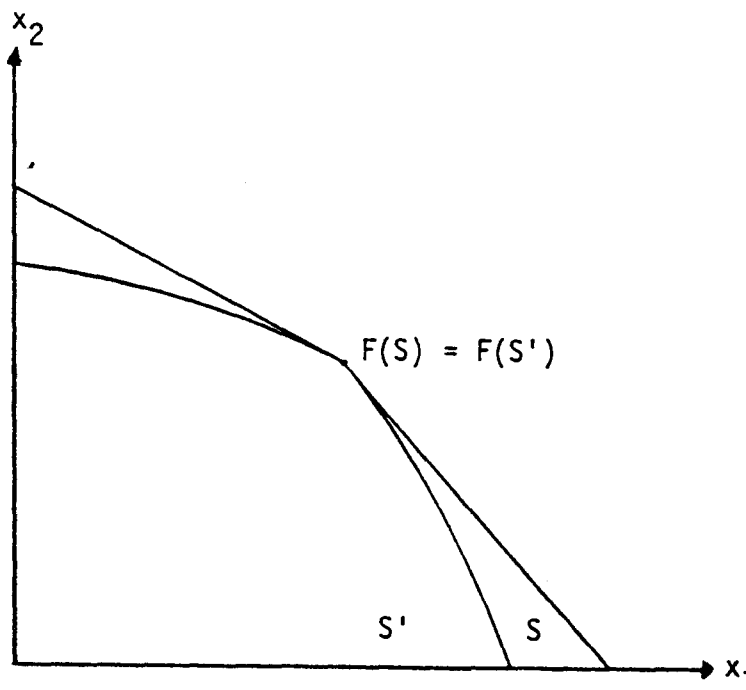


Figure 4.11

The axiom of IIE

As pointed out by Thomson (1981b), IIE is essentially the dual of IIA: IIE says that if, from a given choice problem S' , the set of feasible alternatives is expanded in such a way that the new problem S is locally identical

to S' at $F(S')$, then the solution outcome should not change. It has been showed by Inomson (1981) that IIE can replace IIA in Nash's characterization of his solution, i.e. that the two axioms are equivalent when imposed in conjunction with PO , SY and $S.INV$. Here we show that they are equivalent when imposed in conjunction with $CONT$.

Lemma 4.15: A continuous solution F satisfies IIA if and only if it satisfies IIE.

Proof: We first show that if F satisfies IIE and $CONT$, then it satisfies IIA. Let $P \in \mathcal{P}$ and $S, S' \in \Sigma^P$ be given such that $S' \subset S$ and $z \equiv F(S) \in S'$. For all $\nu \geq 1$, let U^ν be an open ball with center z and radius $1/\nu$, let $V^\nu \equiv \text{Sncl}(U^\nu)$, $S^\nu \equiv \text{cch}\{S' \cup V^\nu\}$, and note that $S^\nu \cap U^\nu = \text{Sn}U^\nu$ and $S' \subset S^\nu \subset S$.

We claim that $F(S^\nu) = z$ for all ν . To see this, let ν be given, and for all $\alpha \in [0, 1]$ let $S(\alpha) \equiv \alpha S^\nu + (1-\alpha)S$ and $z(\alpha) \equiv F(S(\alpha))$. Since $S^\nu \subset S$ and $S^\nu \cap U^\nu = \text{Sn}U^\nu$, then $S(\alpha) \subset S$ and $S(\alpha) \cap U^\nu = \text{Sn}U^\nu$ for all $\alpha \in [0, 1]$. The proof that $z(1) = z$ is by contradiction. If $z(1) \neq z$, then $z(\alpha) \in U^\nu \setminus \{z\}$ for some $\alpha \in (0, 1)$ since $z(0) = z \in U^\nu$ and since the function $z(\cdot)$ is continuous by continuity of F . Because $z(\alpha) \in U^\nu$ and $S(\alpha) \cap U^\nu = \text{Sn}U^\nu$, it follows by IIE that $F(S) = z(\alpha)$, a contradiction, since $z(\alpha) \neq z = F(S)$. Hence $F(S^\nu) = z$.

Because $F(S^\nu) = z$ for all ν and $S^\nu \rightarrow S'$ as $\nu \rightarrow \infty$, it follows by $CONT$ that $F(S') = z = F(S)$. Thus, F satisfies IIA if it satisfies IIE and $CONT$.

Next, we show that if F satisfies $CONT$ and IIA then it satisfies IIE. Let $P \in \mathcal{P}$ and $S, S' \in \Sigma^P$ be given such that $S' \subset S$ and such that S coincides with S' at $z \equiv F(S')$. Then there exists a sequence $\{S^\nu\}$ from Σ^P converging to S and a sequence $\{U^\nu\}$ of neighbourhoods of z such that $S^\nu \cap U^\nu = S' \cap U^\nu$ for all ν .

We claim that $F(S^v) = z$ for all v . To see this, let v be given, and for all $\alpha \in [0,1]$, let $S(\alpha) \equiv \alpha S^v + (1-\alpha)S'$ and $z(\alpha) \equiv F(S(\alpha))$. Since $S^v \cap U^v = S' \cap U^v$, then $S(\alpha) \cap U^v = S' \cap U^v$ for all $\alpha \in [0,1]$. The proof that $z(1) = z$ is by contradiction. If $z(1) \neq z$, then $z(\alpha) \in U^v \setminus \{z\}$ for some $\alpha \in (0,1)$, since $z(0) = z \in U^v$ and since the function $z(\cdot)$ is continuous. Let $S'' \equiv S(\alpha) \cap S'$. Because $z \in S' \cap U^v$, $z(\alpha) \in S(\alpha) \cap U^v$ and $S(\alpha) \cap U^v = S' \cap U^v$, it follows that z and $z(\alpha)$ both belong to S'' . Then by IIA applied twice, $F(S'') = z(\alpha)$ and $F(S'') = z$, a contradiction, since $z(\alpha) \neq z$. Hence $F(S^v) = z$.

Because $F(S^v) = z$ for all v and $S^v \rightarrow S$ as $v \rightarrow \infty$, it follows by CONT that $F(S) = z = F(S')$. Thus, F satisfies IIE if it satisfies IIA and CONT.

QED.

Lemma 4.16: If F satisfies PO, CONT and B.STAB, then

(1) for all $Q \in \mathcal{P}$, F satisfies M.STAB on Σ_{Δ}^Q , and

(2) for all $Q \in \mathcal{P}$ with $|Q| \geq 2$, for all $T \in \Sigma_{\Delta}^Q$, $B_F(T)$ is a singleton.

Proof: For the purpose of the proof, the convention is adopted that $\{0\} \in \Sigma_{\Delta}^P$ and $F(\{0\}) = 0$ for all $P \in \mathcal{P}$. If the lemma holds for this extension of F , then by definition of M.STAB and $B_F(\cdot)$, it holds for F as well.

We must show that (1) and (2) hold for each Q -component of F . The proof is by induction on the number of elements in Q . Clearly, the desired conclusion holds for $|Q| \leq 2$. Suppose it holds for $|Q| = n \geq 3$ and let $Q \in \mathcal{P}$ with $|Q| = n + 1$ be given. We claim that

(i) For all $T \in \Sigma_{\Delta}^Q$, for all $x \in B_F(T)$, for all $P \subset Q$ with $|P| < |Q|$, $F(t_P^x(T)) = x_P$.

To see this, let T , x and P satisfy the hypothesis of (i). If $|P| = 1$, then $F(t_P^x(T)) = x_P$ by $P0$, so assume that $|P| \geq 2$. Since $x \in B_F(T)$, then $x_P \in B_F(t_P^x(T))$, which is a singleton by induction hypothesis (2). B.STAB then implies that $F(t_P^x(T)) = x_P$, thus (i) holds. This proves (1) as well, for if $F(T) = x$ then $x \in B_F(T)$ by B.STAB.

Next, we show that (2) holds for this Q . The proof, which is by contradiction, is illustrated in Figure 4.12. Let $T \in \Sigma_{\Delta}^Q$ be given, let $z \equiv F(T)$ and suppose that $B_F(T)$ contains a point $y \neq z$. Then $y_{Q^i} \in B_F(t_{Q^i}^y(T))$ for all $i \in Q$, where $Q^i \equiv Q \setminus \{i\}$. Moreover, $z, y \in PO(T)$ by $P0$. Let π be the normal to $PO(T)$. For all $\alpha > 0$, let $z(\alpha) \equiv F(\alpha T)$, and for all $i \in Q$, let $\underline{K}^i(\alpha) \equiv \{x \in PO(\alpha T) \mid \pi_{Q^i} x_{Q^i} \geq \pi_{Q^i} y_{Q^i}\}$ and $\bar{K}^i(\alpha) \equiv \{x \in PO(\alpha T) \mid \pi_{Q^i} x_{Q^i} \leq \pi_{Q^i} y_{Q^i}\}$. Note that $z(1) = z$ and that $\alpha T = \{x \in \mathcal{R}_+^Q \mid \pi x \leq \alpha \pi y\}$ for all $\alpha > 0$ since $y \in PO(T)$. If $z(\alpha) \in \underline{K}^i(\alpha) \cap \bar{K}^i(\alpha)$, then $t_{Q^i}^{z(\alpha)}(\alpha T) = t_{Q^i}^y(T)$. Therefore, since $y_{Q^i} \in B(t_{Q^i}^y(T))$, it follows by (1) that $F(t_{Q^i}^{z(\alpha)}(\alpha T)) = y_{Q^i}$. Thus, by M.STAB, it follows that

(ii) For all $\alpha > 0$, for all $i \in Q$, if $z(\alpha) \in \underline{K}^i(\alpha) \cap \bar{K}^i(\alpha)$ then $z_{Q^i}(\alpha) = y_{Q^i}$.

Because $|Q| \geq 3$, there exists $P \subset Q$ with $|P| = 2$ such that $z_P \geq y_P$ or $z_P \leq y_P$. Suppose first that $z_P \geq y_P$. We prove by contradiction that $z(\alpha) \in \bigcap_{i \in P} \bar{K}^i(\alpha)$ for all $\alpha \geq 1$. Consider the path $\sigma \equiv z([1, \infty))$ traced out by $z(\cdot)$ as α increases from 1. $P0$ implies that $z(\alpha) \in PO(\alpha T)$ for all α and CONT implies that σ is a continuous curve. $z(\alpha)$ starts in $\bigcap_{i \in P} \bar{K}^i(1)$ for $\alpha = 1$

since $z_p \geq y_p$, and if it does not stay in $\bigcap_{i \in P} \bar{K}^i(\alpha)$ for all $\alpha \geq 1$, then there exists $\bar{\alpha} \geq 1$ and $j \in P$ such that $z(\bar{\alpha}) \in (\bigcap_{i \in P} \bar{K}^i(\bar{\alpha})) \cap \underline{K}^j(\bar{\alpha})$. This is the situation represented in figure 4.12. Let i be the other member of P . It follows by (ii) that $z_{Q^j}(\bar{\alpha}) = y_{Q^j}$. Because $z(1)$ and y belong to $PO(T)$, this implies that $\bar{\alpha} > 1$, otherwise $z(1) = y$, contrary to assumption. Since $z_{Q^j}(\bar{\alpha}) = y_{Q^j}$ and $\bar{\alpha} > 1$, it follows that $z_j(\bar{\alpha}) > y_j$. But then $\pi_{Q^i} z_{Q^i}(\bar{\alpha}) > \pi_{Q^i} y_{Q^i}$ since $z(\bar{\alpha}) \in PO(\bar{\alpha}T)$, and therefore $z(\bar{\alpha}) \notin \bar{K}^j(\bar{\alpha})$, a contradiction. Thus, $z(\alpha) \in \bigcap_{i \in P} \bar{K}^i(\alpha)$ for all $\alpha \geq 1$, as claimed.

This completes the proof for the case where $z_p \geq y_p$, because $\bigcap_{i \in P} \bar{K}^i(\alpha)$ is empty for sufficiently large α . If $z_p \leq y_p$, then we use the argument in the previous paragraph to show that $z(\alpha) \in \bigcap_{i \in P^-} \bar{K}^i(\alpha)$ for all $\alpha \in (0, 1]$, and obtain a similar contradiction by the fact that $\bigcap_{i \in P^-} \bar{K}^i(\alpha)$ is empty for α sufficiently close to 0. QED.

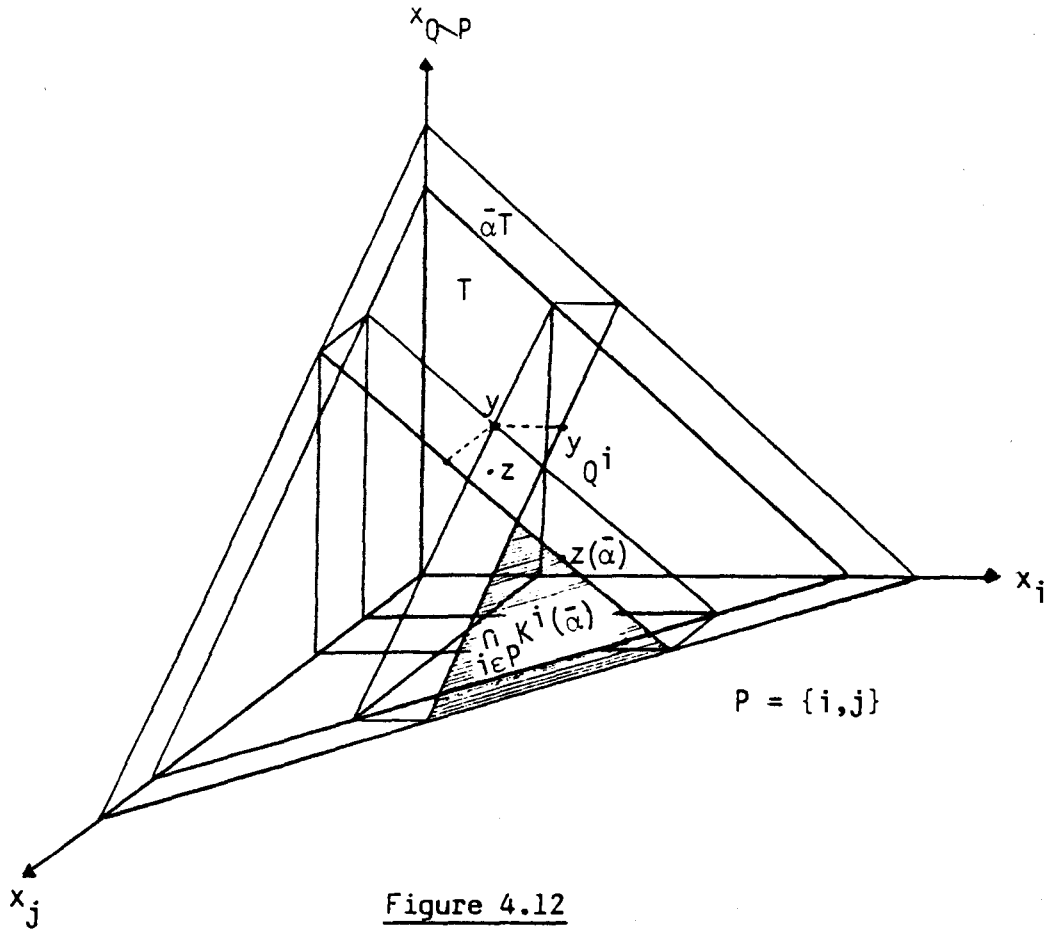


Figure 4.12

The proof of Lemma 4.16

Lemma 4.17: If F satisfies PO, CONT and B.STAB, then for all $Q \in \mathcal{P}$ and all $T^1, T^2 \in \Sigma^Q$, if there exists $T \in \Sigma_\Delta^Q$ such that T, T^1 and T^2 coincide in a neighbourhood U of $F(T^1)$, then $F(T) = F(T^2) = F(T^1)$.

Proof: Note first that since F satisfies PO, CONT and B.STAB, it follows by the same argument used to prove Lemma 4.1 that each two-person component of F must satisfy IIA. Next, let Q, T, T^1, T^2 and U satisfy the hypothesis of the present lemma, and let $z \equiv F(T^1)$. We first show that $F(T) = z$ also.

Since T is a budget problem that coincides with T^1 in U , it follows that $T^1 \subset T$ and $z \in T$. Moreover, $z \in B_F(T^1)$ by B.STAB. Because each two-person

component of F satisfies IIA, it satisfies IIE by Lemma 4.15. Consequently, $z \in B(T)$. Since T is a budget problem, it follows by Lemma 4.16 that $B_F(T)$ is a singleton, hence by B.STAB, $F(T) = z$.

Next, we show that $F(T^2) = z$. To this end, define for each $\alpha \in [0,1]$ the problem $T(\alpha) \equiv \alpha T^2 + (1-\alpha)T$ and set $z(\alpha) \equiv F(T(\alpha))$. We prove by contradiction that $z(1) = z$. If $z(1) \neq z$, then $z(\alpha) \in U \setminus \{z\}$ for some $\alpha \in (0,1)$ since $z(0) = z \in U$ and since the function $z(\cdot)$ is continuous. Because T and T^2 coincide in U , so do T and $T(\alpha)$, and therefore $F(T) = F(T(\alpha)) = z(\alpha)$ by the same argument used to prove that $F(T) = F(T^1)$. Since $z(\alpha) \neq z = F(T)$, this is the announced contradiction. Hence $z = F(T^2)$.

QED.

Lemma 4.18: If F satisfies PO, CONT and B.STAB, then for all $P \in \mathcal{P}$ and all $S \in \Sigma^P$, there exists $S' \in \Sigma_{\Delta}^P$ such that $S \subset S'$ and $F(S') \in S$.

Proof: We use a fixed-point argument, and the mapping is illustrated in Figure 4.13. Let $P \in \mathcal{P}$ and $S \in \Sigma^P$ be given. Let a be a point in \mathcal{R}_{++}^P such that $x < a$ for all $x \in S$ and let $A \equiv \text{cch}\{a\}$. We construct a correspondence G from the unit simplex Δ^P to itself as follows:

For all $\pi \in \Delta^P$, let $H(\pi)$ be the unique hyperplane with normal π supporting S at some point in $\text{WPO}(S)$, and let $S(\pi) \equiv \text{cch}\{H(\pi) \cap A\}$. Next, let $g(\pi)$ be the point of intersection between $\text{WPO}(S)$ and the line segment $[0, F(S(\pi))]$ and define $G(\pi)$ to be the set of points $\pi' \in \Delta^P$ such that π' is the normal to a hyperplane H' supporting S at $g(\pi)$. Clearly, the function $S(\cdot)$ is continuous. Since $S \subset S(\pi)$ for all $\pi \in \Delta^P$, it follows by PO and comprehensiveness of S that $g(\pi)$ exists and is unique. Moreover, g is continuous

since $S(\cdot)$ and F are continuous. Therefore, since S is convex then G is upper hemi-continuous and convex-valued. By Kakutani's fixed-point theorem, there exists $\bar{\pi} \in \Delta^P$ such that $\bar{\pi} \in G(\bar{\pi})$.

Since $\bar{\pi} \in G(\bar{\pi})$, the hyperplane $H(\bar{\pi})$ supports S at $g(\bar{\pi})$. This in turn implies that $F(S(\bar{\pi})) = z \equiv g(\bar{\pi})$. It follows by P0 that $\bar{\pi} > 0$, otherwise z would belong to the boundary of A , which is impossible since $A = \text{cch}\{a\}$ and $a > x$ for all $x \in S$. Let $S' \equiv \text{cch}\{H(\bar{\pi}) \cap \mathcal{R}_+^P\}$. S' is a well defined budget problem since $\bar{\pi} > 0$; moreover, S' and $S(\bar{\pi})$ coincide in a neighbourhood of z since $z < a$. Lemma 4.17 then implies that $F(S') = z \in S$.

QED.

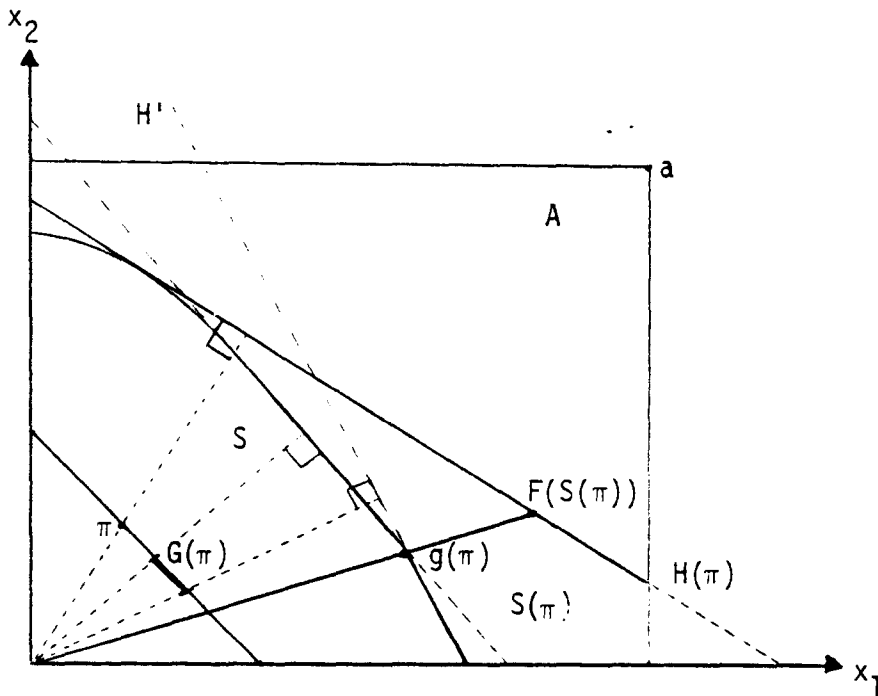


Figure 4.13

The proof of Lemma 4.18

Lemma 4.19: If a solution satisfies PO, CONT and B.STAB, then it satisfies IIA.

Proof: Let F be a solution satisfying PO, CONT and B.STAB. Let $P \in \mathcal{P}$ and $S', S \in \Sigma^P$ be given such that $S' \subset S$ and $F(S) \in S'$. We must show that $F(S') = F(S)$.

By Lemma 4.18, there exists $S'' \in \Sigma_{\Delta}^Q$ with $S \subset S''$ and $z \equiv F(S'') \in S$. We claim that $F(S) = z$ also.

For all $v \geq 1$, let U^v be the closed ball in \mathcal{R}^P with center z and radius $1/v$. Let $V^v \equiv S'' \cap U^v$ and $S^v \equiv \text{cch}\{S \cup V^v\}$. Because S^v and S'' coincide in U^v for all v , it follows by Lemma 4.17 that $F(S^v) = z$ for all v , and by CONT that $F(S) = z$, since $S^v \rightarrow S$ as $v \rightarrow \infty$. This proves the claim.

Since $S' \subset S \subset S''$ and $F(S'') = F(S) \in S'$, the argument in the previous paragraph may be applied to S' as well, and therefore $F(S') = z = F(S)$.

QED.

The main result of this section can now be stated:

Theorem 4.2: Suppose the number of potential agents is at least 3. A solution satisfies PO, CONT and B.STAB if and only if it is a CRS solution.

Proof: Because M.STAB implies B.STAB, it follows by Proposition 4.1 that a CRS solution satisfies PO, CONT and B.STAB.

To prove the converse, let F satisfy PO, CONT and B.STAB. Then F satisfies IIA by Lemma 4.19. Moreover, the restricted solution \bar{F} satisfies M.STAB by Lemma 4.16, WARP by Lemma 4.2 and BOUND by Lemma 4.3. By Proposition 4.2, \bar{F} has an additively separable numerical representation which, by the second paragraph in the proof of Proposition 4.3, represents F as well.

QED.

4.5 Concluding Remarks

Although conceptually, the axiom of Bilateral Stability does not seem to have much to do with collective rationality, it turned out to have quite strong implications in that respect. What then does a bilaterally stable solution have in common with a collectively rational one, that can explain this fact?

Intuitively, both types of solutions are based on some version of what may be called "the principle of pairwise comparisons." Consider first a collectively rational solution. The social ordering corresponding to such a solution provides a way of determining the solution outcome to all problems involving any pair of alternatives, and a utility vector z is the solution outcome to a particular choice problem S only if z agrees with the solution outcome to any subproblem of S involving a pair of alternatives $\{z,x\}$. This is a requirement of a similar type as the one expressed by B.STAB: The generalized solution concept that is used here provides a way of solving all choice problems involving any pair of agents, and B.STAB says that a utility allocation z is the solution outcome to a particular problem S only if z agrees with the solution outcome to any subproblem of S involving a pair of agents $\{i,j\}$.

Our characterization result, in particular theorem 4.2, shows that the second version of this principle is closely related to the first one, although, as shown in section 4.4, there is no direct implication.

References

- Arrow, K.J. (1951) Social Choice and Individual Values (New Haven: Yale University Press).
- Balinsky, M. and Young, H.P. (1982) Fair Representation: Meeting the Ideal of one Man one Vote. (New Haven: Yale University Press).
- Bergson, A. (1938) A reformulation of certain aspects of welfare economics. Quarterly Journal of Economics 52: 310-334.
- Chipman, J.S., Hurwicz, L., Richter, M.K., and Sonnenschein, H.F. (Eds) (1971) Preferences, Utility and Demand (New York: Harcourt Brace Jovanovich).
- d'Aspremont, C. and Gevers, L. (1977) Equity and the informational basis of collective choice. Review of Economic Studies 44: 199-209.
- Deschamps, R. and Gevers, L. (1978) Leximin and utilitarian rules: A joint characterization. Journal of Economic Theory 17: 143-163.
- Debreu, G. (1954) Representation of a preference ordering by a numerical function, in Thrall, R.M., Coombs, C.H., and Davis, R.L. (Eds) Decision Processes (New York: Wiley).
- Debreu, G. (1960) Topological methods in cardinal utility theory, in Arrow, K.J., Karlin, S. and Suppes, P. (Eds) Mathematical Methods in the Social Sciences (Stanford: Stanford University Press).

- Debreu, G. and Koopmans, T.C. (1982) Additively decomposed quasi-convex functions. Mathematical Programming 24: 1-38.
- Eichhorn, W. (1978) Functional Equations in Economics (Reading, Mass.: Addison-Wesley).
- Fleming, M. (1952) A cardinal concept of welfare. Quarterly Journal of Economics 66: 366-384.
- Gale, D. (1960) A note on revealed preference. Economica 27: 348-354.
- Harsanyi, J.C. (1955) Cardinal welfare, individual ethics and inter-personal comparisons of utility. Journal of Political Economy 63: 309-321.
- Harsanyi, J.C. (1959) A bargaining model for the cooperative n-person game, in Tucker, A.W. and R.D. Luce (Eds.), Contributions to the Theory of Games IV, Annals of Mathematical Studies, No. 40, (Princeton: Princeton University Press).
- Harsanyi, J.C. (1963) A simplified bargaining model for the n-person cooperative game. International Economic Review 4: 194-220.
- Harsanyi, J.C. (1977) Rational Behavior and Bargaining Equilibrium in Games and Social Situations. (Cambridge: Cambridge University Press).

- Harsanyi, J.C. and Selten, R. (1972) A generalized Nash solution for two-person bargaining games with incomplete information. Management Science, 18(5)II: 80-106.
- Houthakker, H.S. (1950) Revealed preference and the utility function. Economica 7: 159-174.
- Hurwicz, L. (1971) On the problem of integrability of demand functions, in Chipman et al.
- Hurwicz, L. and Richter, M.K. (1971) Revealed preference without demand continuity assumptions, in Chipman et al.
- Hurwicz, L. and Richter, M.K. (1979) Ville axioms and consumer theory. Econometrica 47: 603-619.
- Imai, H. (1983) Individual monotonicity and lexicographic maximin solution. Econometrica 51: 398-401.
- Kalai, E. (1977a) Proportional solutions to bargaining situations: Interpersonal utility comparisons. Econometrica 45: 1623-1630.
- Kalai, E. (1977b) Nonsymmetric Nash solutions and replications of two-person bargaining. International Journal of Game Theory 6: 129-133.
- Kalai, E. and M. Smorodinsky (1975) Other solutions to Nash's bargaining problem. Econometrica 43: 513-518.

Kaneko, M. and Nakamura, K. (1979) The Nash social welfare function. Econometrica 47(2): 423-435.

Kihlstrom, R., Mas-Colell, A. and Sonnenschein, H. (1976) The demand theory of the weak axiom of revealed preference. Econometrica 44: 971-978.

Luce, D. and H. Raiffa (1957) Games and Decisions (New York: Wiley).

Myerson, R.B. (1977) Two-person bargaining problems and comparable utility. Econometrica 45(7): 1631-1637.

Myerson, R.B. (1981) Utilitarianism, egalitarianism and the timing effect in social choice problems. Econometrica 49: 883-897.

Nash, J.F. (1950) The bargaining problem. Econometrica 18: 155-162.

Nash, J.F. (1953) Two-person cooperative games. Econometrica 21: 128-140.

Nydegger, R.V. and Owen, G. (1974) Two-person bargaining: An experimental test of the Nash axioms. International Journal of Game Theory 3: 239-249.

Raiffa, H. (1953) Arbitration schemes for generalized two-person games, in Kuhn, H.W. and A.W. Tucker (Eds.) Contributions to the Theory of Games II, Annals of Mathematical Studies, No. 28 (Princeton: Princeton University Press).

- Rawls, J. (1971) A Theory of Justice (Cambr. Mass.: Harvard University Press).
- Richter, M.K. (1966) Revealed preference theory. Econometrica 34: 634-645.
- Richter, M.K. (1971) Rational choice, in Chipman et al.
- Roberts, K.W.S. (1980) Interpersonal comparability and social choice theory. Review of Economic Studies 48: 421-439.
- Roth, A.E. (1977) Independence of irrelevant alternatives and solutions to Nash's bargaining problem. Journal of Economic Theory 16: 247-251.
- Roth, A.E. (1979a) Proportional solutions to the bargaining problem. Econometrica 47(3): 775-780.
- Roth, A.E. (1979b) Axiomatic Models of Bargaining, (New York: Springer-Verlag).
- Roth, A.E. and Malouf, W.K. (1979) Game-theoretic models and the role of information in bargaining. Psychological Review 86: 574-594.
- Roth, A.E. and Malouf, W.K. (1982) Scale changes and shared information in bargaining: An experimental study. Mathematical Social Sciences 3: 157-177.

Samuelson, P.A. (1938) A note on the pure theory of consumer's behavior.

Economica 5: 61-71.

Samuelson, P.A. (1947) Foundations of Economic Analysis (Cambr. Mass.: Harvard University Press).

Sen, A. (1970) Collective Choice and Social Welfare (San Fransisco: Holden-Day).

Thomson, W. (1981a) A class of solutions to bargaining problems. Journal of Economic Theory 25(3): 431-441.

Thomson, W. (1981b) Independence of irrelevant expansions. International Journal of Game Theory 10: 107-114.

Thomson, W. (1983a) The fair division of a fixed supply among a growing population. Mathematics of Operations Research 8(3): 319-326.

Thomson, W. (1983b) Problems of fair division and the egalitarian solution. Journal of Economic Theory 31(2): 211-226.

Thomson, W. and R.B. Myerson (1980) Monotonicity and independence axioms. International Journal of Game Theory 9: 37-49.

Uzawa, H. (1971) Preference and rational choice in the theory of consumption, in Chipman et al.

Varian, H.R. (1981) Dynamical systems with applications to economics, in Arrow, K.J. and Intriligator, M.D. (Eds) Handbook of Mathematical Economics, vol. I (Amsterdam: North Holland).

Ville, J. (1951) The existence of a total utility function. Review of Economic Studies 19: 123-128.