MANAGEMENT SCIENCE
Vol. 18, No. 5, January, Part 2, 1972
Printed in U.S.A.

A GENERALIZED NASH SOLUTION FOR TWO-PERSON BARGAINING GAMES WITH INCOMPLETE INFORMATION*

JOHN C. HARSANYI† AND REINHARD SELTEN‡

The paper extends Nash's theory of two-person bargaining games with fixed threats to bargaining situations with incomplete information. After defining such bargaining situations, a formal bargaining model (bargaining game) will be proposed for them. This bargaining game, regarded as a noncooperative game, will be analyzed in terms of a certain class of equilibrium points with special stability properties, to be called "strict" equilibrium points. Finally an axiomatic theory will be developed in order to select a unique solution from the set X of payoff vectors corresponding to such strict equilibrium points (as well as to probability mixtures of the latter). It will be shown that the solution satisfying the axioms proposed in this paper is the point where a certain generalized Nash product is maximized over this set X.

Part I. The Problem

1. *Introduction*

The purpose of this paper[1] is to analyze two-player bargaining situations of the following kind. The two players may adopt any one of several possible *agreements*—or may get into a *conflict* by being unable to accept any agreement by mutual consent. However, in contrast to bargaining situations with complete information, investigated by Nash (see [5] and [6] and also [1] and [2]), one or both players have only *incomplete information* about some important parameters of the bargaining situation, and in particular about the utility payoffs that each player would receive under various possible agreements and/or under a conflict.

Bargaining situations, whether they involve complete or incomplete information, may be divided into those with variable threats and those with fixed threats. A bargaining situation is said to allow *variable threats* if each player can choose any one of several retaliatory strategies available to him, and can commit himself to use this strategy—called his *threat strategy*—against the other player if they cannot reach an agreement. Therefore, the payoffs the players would receive in such a conflict situation could depend on the threat strategies to which they had chosen to commit themselves.

On the other hand, we say that a given bargaining situation allows only *fixed threats* if the payoffs which the players would receive in the absence of an agreement are determined by the nature of the bargaining situation itself, instead of being determined by the players' choice of threat strategies or by any other actions the players may take.[2]

* Received March 1970; revised April 1971.

† University of California, Berkeley.
‡ Free University of Berlin.

[1] This paper grew out of discussions between the two authors and some other participants at the International Game Theory Workshop held at the Hebrew University, Jerusalem, in 1965. Further development of this work was supported by the U. S. Arms Control and Disarmament Agency through a contract with *Mathematica*, Princeton, N. J., and Washington, D. C. Distributions of this paper were supported by a grant from the National Science Foundation to one of the authors through the Center for Research in Management Science, University of California, Berkeley. The two authors wish to express their gratitude to all these persons and organizations.

[2] As to bargaining situations with complete information, Nash discussed the fixed-threats case in [5], and discussed the variable-threats case in [6]. However, the terms "fixed threats" and "variable threats" are due to later writers.

The fixed-threats case is clearly appropriate for many purely economic situations, such as selling or hiring; in case of no agreement, there is simply "no deal." It is equally clear that the variable-threats case is appropriate for most military situations, where a "failure to reach agreement" can have a wide range of possible outcomes. Intermediate cases also exist, where it is not at all clear, *prima facie*, whether or not the threats are fixed; arms control negotiations and labor negotiations are examples.

In this paper we shall be concerned only with the *fixed-threats* case under incomplete information. The variable-threats case under incomplete information appears to be much more subtle, and there are deep theoretical problems which we have only partially resolved.

Our analysis will be based on a model that Harsanyi has recently proposed for games with incomplete information [3]. Under this model, incomplete information on the part of either player is always interpreted as ignorance about certain attributes of the other player, e.g., as ignorance about the other player's utility function, about the physical, technological, and social resources available to him, or about his information and his beliefs concerning the bargaining situation.

More formally, it is assumed that either player can belong to any one of several possible *types*, and that different types of the same player may have different utility functions, different amounts of various resources available to them, different degrees of information, and different beliefs about the bargaining situation. Each player will know his own type but in general will not know the other player's actual type. In this paper we shall assume that the number of possible types to which either player may belong is *finite*, there being K possible alternative types for player 1 and M possible alternative types for player 2.

Furthermore, it will be assumed that as a result of his uncertainty about the other player's actual type, each player will formulate a *subjective probability distribution* over all possible types of his opponent. As has been shown in Part III of [3], in many cases there will exist a probability matrix $r = (r_{km})$ which can be used to represent *both* players' subjective probability distributions over each other's possible types at the same time. Each element r_{km} of this matrix can be interpreted as the joint probability of player 1 belonging to type k ($k = 1, \cdots, K$) while player 2 belongs to type m ($m = 1, \cdots, M$). If such a probability matrix r exists, it is called the *basic probability matrix*, and the two players' subjective probability distributions are said to be mutually *consistent*. If such a probability matrix cannot be found, the subjective distributions are said to be mutually *inconsistent*. In this paper we shall restrict our analysis to the consistent case, though a method has been proposed whereby our results can be extended also to the inconsistent case.[3] In the consistent case, where the basic probability matrix r exists, it is assumed to be known to both players.

Of course, the probability matrix r must satisfy the usual conditions

$$r_{km} \geqq 0 \quad \text{for all values of} \quad k \text{ and } m \tag{1.1}$$

and

$$\sum_{k=1}^{K} \sum_{m=1}^{M} r_{km} = 1. \tag{1.2}$$

The corresponding marginal probabilities

$$p_k = \sum_{m=1}^{M} r_{km} \quad \text{and} \quad q_m = \sum_{k=1}^{K} r_{km} \tag{1.3}$$

can be interpreted, respectively, as the probability that player 1 will be of type

[3] This method, due to Reinhard Selten, is described in [4, pp. 134–137].

k ($k = 1, \cdots, K$), and the probability that player 2 will be of type m ($m = 1, \cdots, M$). For convenience we shall assume that all these marginal probabilities p_k and q_m are *larger than zero*. (This assumption involves no loss of generality because any player type whose probability of occurrence is zero can be simply eliminated from our model.)

As an example, consider a situation where two countries called A ("America") and R ("Russia") are conducting negotiations about various possible arms control treaties. Such negotiations are virtually always carried on under conditions of incomplete information. Thus, in general, neither side will have precise knowledge of the other side's armament levels, economic capability, and technology, nor of the utility values that the other side would assign to alternative arms control arrangements. Under our model, all these uncertainties will be represented by the assumption that there are several possible American player types, which may be called $A_1, A_2, \cdots$, and there are also several possible Russian player types, which may be called $R_1, R_2, \cdots$. Different types of either player would be in general characterized by different combinations of armament levels, economic capabilities, technological know-how, utility scales, etc.

However, while each side will know his own type, neither side will know the opponent's type. Thus the American player will know that he is in actual fact, say, player type A_2 and is not A_1 or A_3; but he will not know whether the Russian player is in actual fact player type R_1 or player type R_2. The converse will be true about the Russian player. On the other hand, we are essentially forced to assume that both players know the probabilities associated with all possible player-type combinations. Thus both of them will know that the joint probability of the two players' belonging to types A_2 and R_3 has the numerical value of, say, $\frac{1}{6}$, etc.

2. *Definition of a Bargaining Situation with Incomplete Information*

Under this model, the utility payoffs that the two players would derive from any specific course of action (whether it results from a successful agreement or from the conflict situation) will depend on both players' actual types, i.e., on the two type variables k and m. For any given agreement, this information can be summarized in the form of a $K \times M$ bimatrix

$$u = (u_{ikm}), \qquad i = 1, 2; \quad k = 1, \cdots, K; \quad m = 1, \cdots, M; \tag{2.1}$$

where a typical element u_{ikm} specifies the payoff that player i will receive under this agreement if player 1 is of type k while player 2 is of type m. [By a $K \times M$ bimatrix u we mean a $K \times M$ matrix of which each element is not a *single* number u_{km} but a *pair* of numbers of the form (u_{1km}, u_{2km}).]

Any such agreement will require the players to perform certain specified physical actions. The particular actions that a given agreement prescribes for either player may or may not depend on his own type and on the other player's type (as disclosed to him by this other player). However, in general, agreements involving such type-dependent actions will be feasible only if the players can find some mutually acceptable methods for enforcing agreements of this kind—and in particular for enforcing truthful disclosure of information between the players (e.g., by providing opportunities for verifying the information transmitted) to the extent required for implementing such agreements.

For instance, suppose that in our previous arms control example the different types $A_1, A_2, \cdots$ of player A, and again the different types $R_1, R_2, \cdots$ of player R, would differ mainly in the number of missiles of a given kind in the possession of player A and of player R, respectively. (In other words, we are assuming that the main piece of information that each player lacks about the other player is the actual number of missiles available to him.) The one kind of agreement open to the players would involve the destruction of a certain specified *number* of missiles by each side (irrespective of the actual number of missiles in the possession of either side). This kind of agreement would be the easiest to enforce because its enforcement would not require any exchange of information about the two sides' missile stocks. Another kind of agreement would require each side to destroy some specified *percentage* of his total stockpile of missiles. Such an agreement could be enforced only if both sides (or at least some mutually trusted third party) were given an opportunity to verify the total number of missiles in the possession of the other side—or if the two sides felt they could trust each other's willingness to tell the truth, even without such verification. Finally, the two sides may even agree to make the total number of missiles to be destroyed by either side a specified mathematical function of *both* sides' total missile stocks. Such an agreement of course would involve the same enforcement problem in an even more accentuated form.

However, for the purposes of our formal theory, the set of all feasible agreements will be simply regarded as *given*. In other words, we shall assume that agreements that either side would find unacceptable because of enforcement difficulties (or for other reasons) have already been eliminated from the list of feasible agreements.

Moreover, any feasible agreement will be formally identified with the corresponding payoff bimatrix $u = (u_{ikm})$. Thus two agreements giving rise to the same payoff bimatrix u will be regarded as identical.[4] The set of all feasible agreements will be called $U = \{u\}$. To simplify our mathematical analysis we shall here assume that U is a finite set. (In most cases this is not a very restrictive assumption as the set U often can be made finite simply by assuming that money payments, commodity deliveries, or decreases in armament stockpiles, etc., must represent integer multiples of certain basic units.)

The bimatrix

$$(2.2) \qquad c = (c_{ikm}), \qquad i = 1, 2; \quad k = 1, \cdots, K; \quad m = 1, \cdots, M;$$

stating the payoffs c_{ikm} that each player i would receive in case of a conflict, under all possible assumptions about the actual types k and m of players 1 and 2, will be called the *conflict point*. As our discussion of bargaining situations with incomplete information is restricted to the fixed-threats case, the bimatrix c is a fundamental and basic *datum* of our analysis, and is completely independent of any action that the players can take. Formally, c will be considered as an element of the set U. That is, it will be assumed that the players are free to agree to bring about the conflict situation if they so desire.

These considerations suggest the following definition. A *bargaining situation with incomplete information* is formally defined as a triplet S, such that

$$(2.3) \qquad S = (U, c, r),$$

where U is the set of all feasible agreements; c is the conflict point (which is itself a special element of the set U); and r is the basic probability matrix.

[4] That is, they are identical as far as the present formal theory is concerned. Many psychological aspects must then either be ignored (with some loss of realism) or incorporated into the payoffs (with great increase in complexity).

3. *Comparison with Nash's Bargaining Theory*

As our own theory will be a generalization of Nash's [1950] theory of bargaining under complete information, we shall now briefly restate Nash's theory in such a way as to make it more easily comparable with our own theory to be described below. We shall restrict ourselves to those aspects of his theory which are important for our present purposes, but shall include some ideas that he made explicit only in his second paper on bargaining games [6]. We shall use our own terminology and notation.

Nash's analysis may be divided into four steps.

Step 1. *Definition of a bargaining situation with complete information.* Any such bargaining situation S can be characterized by specifying the set $U = \{u\}$ of all payoff vectors $u = (u_1, u_2)$ that the players can achieve by mutual agreement, as well as the payoff vector $c = (c_1, c_2)$ that the players will obtain in a conflict situation, i.e., in the absence of any agreement. Thus formally we can define

$$S = (U, c). \tag{3.1}$$

Equation (3.1) is an obvious analog of equation (2.3), except that it does not exhibit the probability matrix r, because complete information means that there is only one possible type for each player. (That is, $K = M = 1$.)

Step 2. *A bargaining model,* specifying the nature of the bargaining moves that the players can make, as well as the outcome (viz., "agreement" or "conflict") that will result in each case. More particularly, Nash assumes that each player i ($i = 1, 2$), independently of the other player, has to choose some payoff u_i as his payoff demand. If the two players' demands are compatible, i.e., if $u = (u_1, u_2) \in U$, then we say that the players have reached an *agreement* on choosing this payoff vector u, and so they will receive the corresponding payoffs u_1 and u_2. On the other hand, if $u \notin U$ then we say that the players have reached a *conflict situation,* and they will receive the conflict payoffs c_1 and c_2.

Step 3. *Analysis of the bargaining model (bargaining game) as a noncooperative game, in terms of its equilibrium points.* Nash has shown that the equilibrium points of the bargaining game postulated by him are given by all payoff vectors $u = (u_1, u_2)$ lying on the upper right boundary of the feasible set U and satisfying the inequalities

$$u_1 \geqq c_1, \quad u_2 \geqq c_2. \tag{3.2}$$

Thus, typically the set V of all equilibrium points is an infinite set (though in certain degenerate cases V can shrink to a unique point).

Step 4. *An axiomatic theory for choosing one particular equilibrium point as the solution of the game.* Nash argues that the solution of the game should satisfy certain intuitively plausible axioms, and then shows that a solution satisfying these axioms always exists and is always unique; and it is given by the point $u = (u_1, u_2)$ at which the Nash product

$$\pi = (u_1 - c_1) \cdot (u_2 - c_2) \tag{3.3}$$

is maximized over the feasible set U (or equivalently over the set of all equilibrium points, V), subject to the inequalities (3.2).

Our own theory will likewise involve four steps, analogous to those of Nash's theory. We have already discussed Step 1 (the formal definition of a bargaining situation S) at the end of §2. We shall now propose a bargaining model for bargaining situations with incomplete information (Step 2).

Then, we shall analyze the resulting bargaining game as a noncooperative game. However, for reasons to be stated below, our analysis will not be based on the concept of *equilibrium point* as such. Instead, it will be based on the concept of *strict* equilibrium point (which will be defined as an equilibrium point satisfying a certain special stability requirement). We shall also make use of *probability mixtures* of strict equilibrium points, involving joint randomization by the players over strategy combinations representing such strict equilibrium points. We shall argue that the solution of the game should be chosen from the set X of payoff vectors corresponding to *probability mixtures* of strict equilibrium points—rather than from the set V of payoff vectors corresponding to *ordinary* equilibrium points, or even from the set Y of payoff vectors corresponding to *strict* equilibrium points without the use of probability mixtures of the latter (Step 3).

Finally, we shall propose an axiomatic theory to select a unique solution from this set X. The axioms we shall use will be closely related to those used by Nash. We shall show that the solution defined by these axioms always exists and is always unique, and can be mathematically characterized in terms of maximization of a certain generalized Nash product (Step 4).

Of the remaining steps of our analysis, Steps 2 and 3 will be discussed in Part II of this paper, while Step 4 will be discussed in Part III.

Part II. The Bargaining Model and Its Analysis as a Noncooperative Game

4. *Description of Our Bargaining Model*

The purpose of a bargaining model is to make the bargaining process between the two players accessible to formal game theoretical analysis, by representing it as a mathematically well-defined "bargaining game" with precise rules about permissible bargaining moves, commitments, agreements, etc.

To analyze bargaining situations with incomplete information, we shall need a more complicated bargaining model than Nash needed in the complete-information case. But we feel that the model we are going to propose is the simplest bargaining model which is still rich enough to represent the most important features of a bargaining process under incomplete information.

One might argue that it is not really necessary to introduce a bargaining model since one could develop a theory which simply selected one of the feasible agreements as the solution. But this approach would not be satisfactory. An appropriate bargaining model will often permit different pairs of types of the two players to choose different agreements. For example, type k of player 1 and type m of player 2 may wish to choose an agreement which is very favorable to them but would be wholly unacceptable to types k' and m' of the two players. In general, it will depend on the situation and on the rules of the bargaining game whether the players actually can make their choice of an agreement type-dependent in this way.

On the one hand, the players can usually benefit by making the outcome type-dependent. On the other hand, they may have incentives to act as if their types were different from what they really are, which may make it impossible for them to arrive at different agreements for different pairs of types. These are very important features of the bargaining process, and they cannot be captured without analysing the players' behavior at the equilibrium points of a formal bargaining game.

A multi-stage acceptance model of bargaining. Under our model the bargaining process will consist of a finite number of consecutive stages. For this reason our model may be called a multi-stage model. The authors have also explored one-stage models, which do not seem to be able to provide adequate representation of the players' strategic possibilities in bargaining under incomplete information.

We shall assume that at any stage of the bargaining process each player can irrevocably commit himself to *accept* any particular agreement (should the latter be proposed at any time by the other player for mutual acceptance); but he cannot irrevocably commit himself to reject any particular agreement. In this sense our model is an *acceptance* model and not a *rejection* model.

Under an acceptance model, as bargaining proceeds, each player will declare more and more possible agreements to be acceptable to him, while under a rejection model each player will declare more and more possible agreements to be *un*acceptable from his point of view. Thus, acceptance models give clearer expression to the empirical fact that during the bargaining process the two sides often make increasing concessions to each other, rather than make increasingly more extreme demands.[5]

More specifically, under our model, bargaining will proceed as follows. At every stage each player will make an offer, by naming some agreement $u \in U$ as being acceptable to him. Each time the two players will choose their offers simultaneously, and independently of each other. Neither player can withdraw any offer once he has made it. When the two players make their offers, they will always know all offers made at previous stages, and so will know what agreements each player has already declared to be acceptable to him.

At any given stage a player may make a *new offer*, by proposing an agreement which he has not proposed yet at any earlier stage; or he may simply repeat an *old offer* he has already made. It is assumed that, if at some stage no new offer is made by either player but only old offers are repeated by both, then the negotiations will break down and a conflict will result. In this case the players' payoffs will be specified by the conflict point c. Thus, if the players want to avoid a conflict, they must keep the negotiations going by at least one of them making a new concession at every stage until an agreement acceptable to both players is found.

If at some stage j ($j = 1, 2, \cdots$) a given player has made an offer u which has also been proposed by the other player (at the same time or at some earlier time), then we say that u is an *agreement accepted at stage* j. As soon as at least one agreement u has been accepted at some stage j, no further offers will be made by the two players.

If at some stage j exactly one agreement u is accepted then this agreement u will be the outcome of the bargaining game, and the players will receive the corresponding payoffs. The rules of the bargaining game do not exclude the possibility that at some stage j *two* different agreements will be accepted, and so we have to define the outcome of the game also in such cases. This possibility arises because it may happen that each of the two players at the same time will propose an offer already proposed by the other player at some earlier stage, so that under our previous definition both of these offers will now become "accepted" agreements. In such cases our model always empowers one of the two players to break this deadlock. More specifically, we assume that a random move will be made to select one of the two players for this purpose, with equal

[5] However, one may argue that, even if the two sides do become more concessive in their overt behavior, their inner attitudes may very well become in some sense more intransigent, e.g., because they may come to learn more and more about each other's weak points during the negotiations. Perhaps rejection models could be used to represent such shifts to more intransigent inner attitudes by the two players.

chances for the two players. Then the player so selected has to choose one of the two accepted agreements. The agreement chosen by him will be regarded as the outcome of the bargaining game, and the players will receive the corresponding payoffs.[6]

5. *Bargaining Strategies and the N-Player Form of the Bargaining Game*

Bargaining strategies. A strategy of player 1 or player 2 in the bargaining game defined by the above bargaining model is called a bargaining strategy. In order that we can give a more precise formal characterization of the two players' bargaining strategies, we introduce the concept of an *offer sequence.* Consider a sequence

$$(5.1) \qquad b = \begin{pmatrix} u^{11}, u^{12}, \cdots, u^{1J} \\ u^{21}, u^{22}, \cdots, u^{2J} \end{pmatrix}$$

which consists of a finite number of pairs of the form u^{1j}, u^{2j}, where u^{ij} is interpreted as the offer $u = u^{ij}$ of player i at stage j. Such a sequence is called an offer sequence if it is compatible with the rules of the bargaining game. This will be the case if at every stage $j = 1, \cdots, J$ at least one of the two offers u^{1j} and u^{2j} is a new offer, and if none of the offers u^{1j} and u^{2j} with $j = 1, \cdots, J - 1$ is an accepted offer. (Otherwise the bargaining game would not have reached stage J.)

An offer sequence b is called a *conflict sequence* if both u^{1J} and u^{2J} are old offers. If exactly one of the two offers u^{1J} and u^{2J} is an accepted offer, or if $u^{1J} = u^{2J}$, then b is called a *one-agreement sequence.* A *complete sequence* is an offer sequence which is either a conflict sequence or a one-agreement sequence.

If both u^{1J} and u^{2J} are accepted offers, yet if $u^{1J} \neq u^{2J}$ then b is called a *semicomplete sequence.* Finally, an *incomplete sequence* is an offer sequence that is neither complete nor semicomplete. An *empty sequence,* which corresponds to the situation before the bargaining has started, and which contains no offers at all, is also called *incomplete.*

After these definitions, we can give the following formal description of a pure bargaining strategy: *A pure bargaining strategy f_i of player i is a function which assigns some agreement $f_i(b) \in U$ to every incomplete offer sequence b, and also assigns one of the two last-stage offers $f_i(b) = u^{1J}$ or $f_i(b) = u^{2J}$ to every semicomplete offer sequence b.*

We shall now show that the number of pure bargaining strategies for either player is *finite.*

Let $|U|$ be the number of elements in U, i.e., the number of feasible agreements. In any incomplete or semicomplete offer sequence there must be at least one new offer at each stage, other than the last stage. Yet a given player cannot make more than $|U|$ new offers during the game. Therefore $2 \cdot |U| + 1$ is an upper bound for the number of stages in an incomplete or semicomplete offer sequence, and so the number of such offer sequences is finite. Consequently, the number of pure bargaining strategies is also finite, as desired.

Another implication is that the bargaining game will always end after a *finite number of stages.*

The payoff expectations resulting from a given pair of bargaining strategies f_1 and f_2 used by the two players. The set F of bargaining strategies is the same for both players. For any given pair of bargaining strategies (f_1, f_2) with $f_1 \in F$ and $f_2 \in F$, we can compute a

[6] Our experience with numerical examples seems to suggest that the solution we shall propose is not significantly affected by the rule actually chosen to define the outcome of the bargaining game in the case where two different agreements are "accepted."

$K \times M$ bimatrix $h = h(f_1, f_2)$ of which a typical element $h_{ikm} = h_{ikm}(f_1, f_2)$ specifies the expected payoff to player i, when players 1 and 2 are of types k and m, respectively, and when they use the bargaining strategies f_1 and f_2 in the bargaining game. This bimatrix can be computed as follows:

Let $b(f_1, f_2)$ be the complete or semicomplete sequence we shall eventually obtain if we repeatedly apply the formulas

$$u^{ij} = f_1(b^{j-1}) \quad \text{for} \quad i = 1, 2 \quad \text{and} \quad j = 1, \cdots \tag{5.2}$$

and

$$b^j = \begin{pmatrix} u^{11}, & \cdots, & u^{1j} \\ u^{21}, & \cdots, & u^{2j} \end{pmatrix} \tag{5.3}$$

and go on doing so as long as b^{j-1} remains an incomplete sequence.

Moreover, let $u(f_1, f_2)$ be the unique accepted offer in sequence $b(f_1, f_2)$ if the latter is a one-offer sequence. Then we can define the bimatrix $h = h(f_1, f_2)$ as follows.

$$\begin{aligned} h(f_1, f_2) &= u(f_1, f_2) \text{ if } b(f_1, f_2) \text{ is a one-agreement sequence,} \\ h(f_1, f_2) &= c \text{ if } b(f_1, f_2) \text{ is a conflict sequence, and} \\ h(f_1, f_2) &= \tfrac{1}{2}f_1(b(f_1, f_2)) + \tfrac{1}{2}f_2(b(f_1, f_2)) \text{ if } b(f_1, f_2) \text{ is a semicomplete sequence.} \end{aligned} \tag{5.4}$$

Each expected payoff $h_{ikm} = h_{ikm}(f_1, f_2)$ will be simply the appropriate element of this bimatrix $h = h(f_1, f_2)$ just defined.

Mixed bargaining strategies. We do not want to exclude the possibility that the two players may use mixed strategies in the bargaining game. A mixed strategy g_i of player i will be a probability distribution over the set F of pure bargaining strategies. Obviously the set G of mixed bargaining strategies will be the same for both players. We shall write $h(g_1, g_2)$ to denote the expected value of the bimatrix $h(f_1, f_2)$ when player 1 uses the mixed strategy g_1 while player 2 uses the mixed strategy g_2. Since $h(g_1, g_2)$ is the expected value of the $K \times M$ bimatrix $h(f_1, f_2)$, it will be itself also a bimatrix of the same size. A typical element $h_{ikm} = h_{ikm}(g_1, g_2)$ of the bimatrix $h(g_1, g_2)$ will again specify the expected payoff to player i, when players 1 and 2 are of types k and m respectively, and when they use the mixed bargaining strategies g_1 and g_2.

The N-player form of the bargaining game. For many purposes it is useful to look upon the K alternative types of player 1, and the M alternative types of player 2, as being the $K + M = N$ players of an N-player game. When we want to refer to this interpretation of the bargaining game it will be convenient to call these N types of players 1 and 2 *subplayers,* and to number them consecutively from 1 to $N = K + M$. In particular, the types $1, \cdots, k, \cdots, K$ of player 1 will be called subplayers $1, \cdots, k, \cdots, K$; whereas the types $1, \cdots, m, \cdots, M$ of player 2 will be called subplayers $K + 1, \cdots, K + m, \cdots, K + M$.

In the N-player bargaining game played by these N subplayers, each subplayer i will be assumed to choose some bargaining strategy, which may be a *pure* bargaining strategy $s_i \in F$, or may be a *mixed* bargaining strategy $t_i \in G$.[7] In terms of its intuitive

[7] Technically, the N-player game played by the N subplayers is the *Selten game* corresponding to the two-player bargaining game between players 1 and 2, and is based on the *Selten model* for games with incomplete information. For a more extensive discussion of Selten games and their intuitive interpretations, see Part I of [3, pp. 177–180].

interpretation, for any subplayer $i = k = 1, \cdots, K$, this bargaining strategy s_i or t_i is simply the *pure* bargaining strategy $f_1 = f_1^k \in F$ or the *mixed* bargaining strategy $g_1 = g_1^k \in G$ that player 1 would use in the two-player bargaining game between players 1 and 2, *if he were* of type $i = k$. Likewise, for any subplayer $i = K + m = K + 1, \cdots, K + M$, this bargaining strategy s_i or t_i is simply the *pure* bargaining strategy $f_2 = f_2^m \in F$ or the *mixed* bargaining strategy $g_2 = g_2^m \in G$ that player 2 would use *if he were* of type $m = i - K$.

As pure strategies are special cases of mixed strategies, no generality will be lost if we assume that the N subplayers will choose the mixed-strategy combination

$$t = (t_1, \cdots, t_N). \tag{5.5}$$

To compute the expected payoffs of the first K subplayers when this strategy combination t is chosen, we must use the fact that any given subplayer k simply represents type k of player 1. Therefore, his expected payoff $x_k(t)$ will be given by player 1's conditional payoff expectation under the assumption that player 1 is known to be of type k. Thus we can write

$$x_k(t) = \frac{1}{p_k} \sum_{m=1}^{M} r_{km} h_{1km}(t_k, t_{K+m}) \quad \text{for} \quad k = 1, \cdots, K. \tag{5.6}$$

On the other hand, to compute the expected payoffs of the remaining M subplayers when the strategy combination t is chosen, we must use the fact that any given subplayer $K + m$ simply represents type m of player 2. Therefore, his expected payoff $x_{K+m}(t)$ will be given by player 2's *conditional payoff expectation* under the assumption that player 2 is known to be of type m. Thus we can write

$$x_{K+m}(t) = \frac{1}{q_m} \sum_{k=1}^{K} r_{km} h_{2km}(t_k, t_{K+m}) \quad \text{for} \quad m = 1, \cdots, M. \tag{5.7}$$

The $K + M = N$ quantities defined by equations (5.6) and (5.7) together form the vector

$$x(t) = (x_1(t), \cdots, x_N(t)). \tag{5.8}$$

As we have seen, each component $x_i(t)$ of this vector is the payoff expectation of some subplayer i. But, more fundamentally, it also represents a certain conditional payoff expectation of player 1 or 2. In view of this latter interpretation, the vector $x(t)$ will be called the *conditional payoff vector* corresponding to the strategy combination $t = (t_1, \cdots, t_N)$.

6. *Strict Equilibrium Points*

Strict equilibrium points in the N-player bargaining game. We assume that the bargaining game will be played noncooperatively. Therefore, rational players may be expected to use bargaining strategies corresponding to an equilibrium point. Indeed, we want to argue that only equilibrium points satisfying a special stability requirement—to be called *strict* equilibrium points—can represent strategy combinations acceptable to rational players.

Equilibrium points in mixed strategies often have the undesirable property that the payoffs of certain players j will be changed if a given player i deviates from his mixed equilibrium strategy t_i, and shifts to some pure strategy s_i occurring with a positive probability in this mixed strategy t_i. Yet, there is no reason to expect that deviations of this kind will not happen, because such a pure strategy s_i will always be a best reply to the other players' equilibrium strategies, just as much as the equilibrium strategy t_i itself is. Consequently, these players j will be unable to form definite expectations about the payoffs they would receive if they stuck to their own equilibrium strategies t_j, which again could easily induce these latter players, also, to deviate from these equilibrium strategies, etc. Thus equilibrium points of this sort will not represent stable strategy combinations.

To avoid this difficulty, we introduce the concept of *strict* equilibrium points.[8] Let $t = (t_1, \cdots, t_N)$ be some combination of bargaining strategies in the N-player bargaining game, and let t'_i be some bargaining strategy of subplayer i. We shall write t/t'_i to denote the strategy combination we shall obtain from t, if we replace the strategy t_i by t'_i but leave all other components of t unchanged.

A *best reply* of subplayer i against a given strategy combination $t' = (t'_1, \cdots, t'_N)$ is defined as a strategy t_i^0 such that

$$x(t'/t_i^0) = \max_{t_i \in G} x(t'/t_i). \tag{6.1}$$

(For our purposes it is convenient to use a somewhat unorthodox terminology, and to call this strategy t_i^0 a best reply against the *whole* strategy N-tuple $t' = (t'_1, \cdots, t'_N)$, rather than call it a best reply merely against the strategy $(n - 1)$-tuple $(t'_1, \cdots, t'_{i-1}, t'_{i+1}, \cdots, t'_N)$, comprising only the strategies of the *other* $(N - 1)$ players with the exclusion of player i's own strategy t'_i. This terminology is perhaps less natural than the more customary one, but it simplifies the phrasing of some of the statements we are going to make.)

An *equilibrium point* can be characterized as a strategy combination $t^* = (t_1^*, \cdots, t_N^*)$ in which every component t_i^* is a best reply against t^*.

An equilibrium point t^* is called *strict* if, for *every* subplayer i, and for *every* best reply t_i^0 that this subplayer i has against t^*, we can write

$$x(t^*/t_i^0) = x(t^*). \tag{6.2}$$

Thus, if t^* is a strict equilibrium point then any player j using his equilibrium strategy t_j^* will have some protection against any deviation by another player i from his equilibrium strategy t_i^* to some alternative best-reply strategy t_i^0, since such a deviation would not affect player j's payoff.

The concept of strict equilibrium points has useful applications, not only in the analysis of bargaining games, but also in the analysis of general n-person noncooperative games. (Of course, in the latter case strict equilibrium points must be defined with reference to "players," instead of "subplayers.")[9]

Guaranteed equilibrium points will be defined in terms of a much stronger stability

[8] For another approach to the problem resulting from the fact that at many equilibrium points $t=(t_1, \ldots, t_N)$ the equilibrium strategy t_i of a given player i is not his *only* best reply to the other players' equilibrium strategies, see [2]. The approach proposed there is based on the concepts of *strong* equilibrium points and of *centroid* equilibrium points. But for our present purposes the concept of strict equilibrium points seems to be a preferable analytical tool (cf. Lemmas 6.1, 6.2 and 6.3 below).

[9] However, as one of the editorial referees has pointed out, one can construct examples for games where all strict equilibrium points are rather inefficient. It is not clear whether this problem can arise in bargaining games. But nevertheless one may wish to develop a solution based on the set of all equilibrium points rather than on the set of strict equilibrium points alone. As the reader can easily verify, very little has to be changed in our theory in order to do this.

requirement than used in the definition of strict equilibrium points. However, as we shall see, owing to certain special characteristics of bargaining games, any strict equilibrium point of a bargaining game is also a guaranteed equilibrium point.

Let $t^* = (t_1^*, \cdots, t_N^*)$ be an equilibrium point, and let $t^0 = (t_1^0, \cdots, t_N^0)$ be some combination of strategies, such that every component t_i^0 of t^0 is a best reply against t^*. The equilibrium point t^* is called a *guaranteed* equilibrium point, if for every such combination t^0 of best replies against t^* we can write

$$x_j(t^0/t_j^*) = x_j(t^*) \quad \text{for} \quad j = 1, \cdots, N. \tag{6.3}$$

Thus, if t^* is a guaranteed equilibrium point then every player j using his equilibrium strategy t_j^* will be protected, not only against deviations by single players, but also against simultaneous deviations by several players—or even by all other players—to alternative best-reply strategies.

Lemma 6.1. *Every strict equilibrium point* $t^* = (t_1^*, \cdots, t_N^*)$ *of a bargaining game is also a guaranteed equilibrium point.*

Proof. Suppose that $t^0 = (t_1^0, \cdots, t_N^0)$ is a combination of best replies against t^*. Consider subplayers j and $K + m$, where subplayer j $(j = 1, \cdots, K)$ represents type j of player 1, whereas subplayer $K + m$ $(m = 1, \cdots, M)$ represents type m of player 2. Suppose that subplayer j shifts from strategy t_j^* to strategy t_j^0 while all other subplayers $i \neq j$ stick to the strategies t_i^*. That is, suppose that strategy N-tuple t^* is replaced by strategy N-tuple t^*/t_j^0. What effect will this have on the payoff of subplayer $K + m$?

By equation (5.7), his original payoff $x_{K+m}(t^*)$ is a sum of K terms. When t_j^* is replaced by t_j^0, only one term will be affected, viz. the term corresponding to $k = j$. Consequently, we can write

$$x_{K+m}(t^*/t_j^0) = x_{K+m}(t^*) - \frac{r_{jm}}{q_m} h_{2jm}(t_j^*, t_{K+m}^*) + \frac{r_{jm}}{q_m} h_{2jm}(t_j^0, t_{K+m}^*). \tag{6.4}$$

On the other hand, as t^* is a strict equilibrium point, by equation (6.2) we have $x_{K+m}(t^*/t_j^0) = x_{K+m}(t^*)$. Therefore, we must have

$$\frac{r_{jm}}{q_m} h_{2jm}(t_j^0, t_{K+m}^*) = \frac{r_{jm}}{q_m} h_{2jm}(t_j^*, t_{K+m}^*) \tag{6.5}$$

for $j = 1, \cdots, K$ and for $m = 1, \cdots, M$.

Next, consider the case where some subplayer $K + j$, representing type j of player 2 $(j = 1, \cdots, M)$, shifts from strategy t_{K+j}^* to strategy t_{K+j}^0, so that strategy N-tuple t^* is replaced by strategy N-tuple t^*/t_{K+j}^0. By similar reasoning to that used in deriving equation (6.5), one can verify that

$$\frac{r_{kj}}{p_k} h_{1kj}(t_k^*, t_{K+j}^0) = \frac{r_{kj}}{p_k} h_{1kj}(t_k^*, t_{K+j}^*) \tag{6.6}$$

for $k = 1, \cdots, K$ and for $j = 1, \cdots, M$. However, equations (5.6), (5.7), (6.5) and (6.6) together immediately imply that t^* has property (6.3). This completes the proof.

Let $Y = \{y\}$ be the set of all payoff vectors $y = (y_1, \cdots, y_N)$ which are conditional payoff vectors of the form $y = x(t^*)$, corresponding to some strict equilibrium point t^* of the N-player bargaining game.

The number of strict equilibrium payoff vectors y will now be characterized by two lemmas.

Lemma 6.2. *The set Y of strict equilibrium payoff vectors y is always nonempty.*

Proof. Consider the strategy combination $s = (s_1, \cdots, s_n)$ under which every subplayer i selects a pure bargaining strategy s_i assigning the conflict point c to every incomplete offer sequence. This strategy combination s will be a strict equilibrium point. Hence the set Y is always *nonempty.*

Lemma 6.3. *The set Y of strict equilibrium payoff vectors y is always a finite set.*

Proof. To prove the lemma, we introduce the concept of a total equilibrium point. A *total* equilibrium point is a guaranteed equilibrium point t^* which has the additional property that every pure strategy of each player is a best reply against t^*. If a game has several total equilibrium points, then all of them must have the same equilibrium payoff vector, because the use of the equilibrium strategy guarantees the equilibrium payoff regardless of the strategy choices of the other players.

Let t^* be a strict equilibrium point of the N-player bargaining game. Let F_i^* be the set of pure strategies to which the equilibrium strategy t_i^* assigns positive probabilities $(i = 1, \cdots, N)$. If the set of pure strategies of every subplayer i is narrowed down to F_i^*, a *restricted* bargaining game results from the original one. As we have seen, every strict equilibrium point of the bargaining game is a guaranteed equilibrium point. Therefore, in the restricted bargaining game, t^* (or, more precisely, the strategy combination corresponding to t^* in this restricted game) is a total equilibrium point. Therefore, every payoff vector $y \in Y$ is the payoff vector of a total equilibrium point in a restricted bargaining game. Since the number of pure bargaining strategies is finite, only a finite number of restricted games can be derived from strict equilibrium points of the bargaining game. Each of these restricted games has only one total equilibrium payoff vector. Therefore *Y is a finite set.*

7. *The Equilibrium Set*

Let X be the convex hull of the set Y. We call X the equilibrium set of the bargaining game. It represents the set of all conditional payoff vectors $x = (x_1, \cdots, x_N)$ that can be obtained by means of jointly-randomized strategies corresponding to mixtures of strict equilibrium points.

Under our model the bargaining game is a noncooperative game. As we have argued in a noncooperative game rational players will always use some strategy combination that is an equilibrium point, and indeed is a *strict* equilibrium point. To restate our previous argument in somewhat different terms, the basic reason is that only strict equilibrium points can give rise to *stable conformistic expectations* on the part of intelligent players. By a *conformistic expectation* for a given strict equilibrium point $t = (t_1, \cdots, t_N)$ we mean the expectation that every player i will use his equilibrium strategy t_i, or at worst will engage in some deviation not affecting the payoffs of the other players (as will be the case with deviations to alternative best-reply strategies t'_i). Now suppose that for some reason or another the players come to form conformistic expectations about one another's behavior with respect to some strict equilibrium point t. Then these expectations will be *stable* because they will be fully consistent with the incentives confronting the players, and of course the players will know this.

In contrast, if a given strategy combination $t = (t_1, \cdots, t_N)$ is *not* a strict equilibrium point then it *cannot* give rise to stable conformistic expectations. For suppose the players would adopt conformistic expectations about one another, i.e., would adopt the expectation that all other players would use the strategies t_i (or would adopt strategies equivalent to the latter in terms of the resulting payoffs). Then *these expecta-*

tions themselves would give some of the players the incentive to engage in nonconformistic behavior, contrary to these expectations; and of course the players will again know this. Therefore, such expectations could not be stable with rational players.

On the other hand, it is our contention that *probability mixtures* of strict equilibrium points do have the ability to generate stable conformistic expectations in the same way as single strict equilibrium points do.

We are assuming that the players are free to *communicate* before the actual exchange of offers as prescribed by our bargaining model. Given such preplay communication, they may reach an informal understanding to make their choice among alternative strict equilibrium points, $t, t^*, \cdots$, of the bargaining game depend on the outcome of some chance event (which may occur spontaneously, or may be produced by the players themselves for the very purpose of obtaining the desired probability mixture of strict equilibrium points). Even if the rules of the bargaining game do not make such an informal understanding legally enforceable, it will be self-enforcing and therefore will result in *stable conformistic expectations.*

For example, suppose that the chance mechanism used by the players actually decides in favor of some particular strict equilibrium point t; and suppose that the players expect one another to conform to this assumed informal understanding. This means that now the players will expect one another to use strategies conforming to this equilibrium point t chosen by the chance mechanism. However, as t is assumed to be a strict equilibrium point, once the players adopt these conformistic expectations with respect to t these expectations will be *stable.* In other words, if the players decide to use some probability mixture of strict equilibrium points, this decision will generate stable conformistic expectations—simply because any strict equilibrium point, chosen by a chance mechanism or by any other method, will itself generate stable conformistic expectations on the part of the players.

This leads us to the conclusion that the conditional payoff vectors the players can achieve in the bargaining game not only include all points $y = (y_1, \cdots, y_N)$ in the set Y of strict equilibrium payoff vectors, but also include all points $x = (x_1, \cdots, x_N)$ in the *convex hull* X of this set Y (which we have called the equilibrium set X).

The extended bargaining game. Instead of assuming that the players use jointly randomized strategies to achieve probability mixtures of strict equilibrium points, we may obtain these probability mixtures also by a formal extension of our bargaining model. The *extended bargaining model* we shall use for this purpose will assume that before the beginning of the actual bargaining a random move will select a random number λ which is uniformly distributed over the interval $0 \leqq \lambda \leqq 1$. This random number is then announced to both players. Only after they have been informed about λ, do they begin to play the bargaining game described above. We call the bargaining game together with the preplay random move the *extended bargaining game.* It is easy to see that every strict equilibrium point of this extended bargaining game can be regarded as a function which assigns a strict equilibrium point of the original bargaining game to almost every λ in the interval $0 \leqq \lambda \leqq 1$. Therefore, *the set of conditional payoff vectors corresponding to strict equilibrium points of the extended bargaining game is the equilibrium set X.*

The conflict payoff vector $w = (w_1, \cdots, w_N)$ is defined as the conditional payoff vector corresponding to the conflict point c, so that

$$w_k = \frac{1}{p_k} \sum_{m=1}^{M} r_{km} c_{1km}, \quad \text{for} \quad k = 1, \cdots, K, \tag{7.1}$$

and

$$w_{K+m} = \frac{1}{q_m} \sum_{k=1}^{K} r_{km} c_{2km}, \quad \text{for} \quad m = 1, \cdots, M. \tag{7.2}$$

In other words, w_k is player 1's *conditional payoff expectation* in case of a conflict, when his type is known to be k; whereas w_{K+m} is player 2's *conditional payoff expectation* in case of a conflict, when his type is known to be m.

If every subplayer i selects a pure bargaining strategy s_i which assigns the conflict point c to every incomplete offer sequence, then the resulting strategy combination $s = (s_1, \cdots, s_N)$ will be a strict equilibrium point with $w = x(s)$ as the corresponding equilibrium payoff vector. Hence

$$w \in X. \tag{7.3}$$

By using the strategy s_i, each subplayer i can guarantee at least the payoff w_i for himself, regardless of the strategy choices of all other subplayers. Therefore, we have

$$w_i \leqq x_i \quad \text{for all} \quad x = (x_1, \cdots, x_N) \in X. \tag{7.4}$$

In view of (7.3) and (7.4), *the conflict payoff vector w can be characterized as the minimal element of the equilibrium set X.*

The bargaining basis. It is a fundamental assumption of our theory that the equilibrium set X and the basic probability matrix r are the crucial parameters determining the outcome of bargaining between rational players. Thus, two bargaining situations will be regarded as equivalent if they have the same basic probability matrix r and also lead to the same equilibrium set X. Therefore we shall call the ordered pair

$$B = (X, r) \tag{7.5}$$

the *bargaining basis* for a given bargaining game.

More generally, we shall call any pair $B = (X, r)$ a bargaining basis, if $r = (r_{km})$ is a probability matrix of size $K \times M$, satisfying conditions (1.1) and (1.2) as well as the positivity requirement for all marginal probabilities p_k and q_m ($k = 1, \cdots, K$ and $m = 1, \cdots, M$); and if X is a compact and convex set of N-vectors (with $N = K + M$), such that X has a minimal element w satisfying condition (7.4).

However, bargaining bases derived from bargaining games have some special properties in addition to the general properties listed in the last paragraph. One such special property is that the equilibrium set X of such a bargaining basis is always the convex hull of a *finite* set of payoff vectors. (This follows from Lemma 6.3, according to which the set Y of strict equilibrium payoff vectors is always a finite set: for X has been defined as the convex hull of this set Y.)

It could be argued that our theory should be developed in terms of bargaining bases possessing the special properties which make them derivable from bargaining situations. We shall not adopt this point of view because it would lead to unnecessary mathematical complications. Instead, we shall develop our theory in terms of the more general definition of bargaining bases stated above. The mathematical justification of this approach lies in two facts. On the one hand, every bargaining basis $B = (X, r)$ can be approximated by a bargaining basis $B^* = (X^*, r)$ derived from a bargaining game, up to a Hausdorff distance between X and X^* as small as desired. On the other hand, as will be shown in §11, in terms of the same distance, the solution we shall define is a continuous function of the equilibrium set X.

Polyhedral bargaining bases. A bargaining basis $B = (X, r)$ will be called polyhedral if it has the following properties:

(a) the equilibrium set X is a convex hull of a *finite* set Y; and

(b) for every element $y \neq w$ in set Y we have

$$w_i < y_i \quad \text{for} \quad i = 1, \cdots, N. \tag{7.6}$$

(Here w again denotes the minimal element of X.)

LEMMA 7.1. *Any bargaining basis $B = (X, r)$ can be approximated in the sense of the Hausdorff metric, as closely as desired, by a polyhedral bargaining basis $B^* = (X^*, r)$.*

PROOF. The lemma follows from the fact that any compact and convex set X of a finite-dimensional Euclidean space can be approximated by a polyhedral set X^*, which is the convex hull of a finite number of points, and has an arbitrarily small Hausdorff distance from X.

LEMMA 7.2. *Every polyhedral bargaining basis can be derived from some appropriately chosen bargaining situation $S = (U, c, r)$.*

PROOF. Let Y be the finite set of which X is the convex hull. We shall construct a bargaining situation $S = (U, c, r)$ which has B as its bargaining basis.

We define the first component of S, the agreement set U, as the set of all bimatrices $u = (u_{ikm}) = u(y), y \in Y$, which can be derived from any element $y = (y_1, \cdots, y_N)$ of set Y according to the formula

$$\begin{aligned} u_{ikm} &= y_k \quad &\text{if} \quad i = 1 \\ &= y_{K+m} \quad &\text{if} \quad i = 2. \end{aligned} \tag{7.7}$$

The second component of S, the conflict point c, is defined as the minimal element of set X. Finally, as the third component of S we choose the basic probability matrix r occurring in the definition of $B = (X, r)$.

In view of these definitions, any combination of pure bargaining strategies for the two players must lead to some agreement $u \in U$. Therefore, all conditional payoff vectors corresponding to combinations of pure or mixed strategies in the N-person bargaining game derived from $S = (U, c, r)$ are convex combinations of some payoff vectors y in set Y, and are consequently elements of set X. It follows that the equilibrium set of the bargaining game derived from S is a subset of X.

In order to prove that this equilibrium set is actually identical with X, we must show that every point y in set Y is a conditional payoff vector corresponding to some strict equilibrium point of the bargaining game derived from S. If every subplayer i chooses a pure bargaining strategy $s_i = s_i(y)$ which to each incomplete offer sequence always assigns the agreement $u = u(y)$ corresponding to some specific point y in set Y, then the resulting strategy combination $s(y) = (s_1(y), \cdots, s_N(y))$ will be a strict equilibrium point with the equilibrium payoff vector $x(s(y)) = y$. The strategy combination $s(y)$ will be a strict equilibrium point because any deviation by some subplayer i from his equilibrium strategy $s_i(y)$ can only have the result that in some cases the agreement $u = u(y)$ will be replaced by the conflict point c as the outcome of the bargaining game. Therefore, any deviation by subplayer i which influences the payoffs of the other subplayers $j \neq i$ will reduce his own payoff from y_i to $w_i < y_i$; thus $s(y)$ satisfies the definition of strict equilibrium points. Consequently, X is in fact the equilibrium set of the bargaining game derived from S, and so $B = (X, r)$ is the bargaining basis of this bargaining game. This completes the proof.

Part III. An Axiomatic Theory for Defining the Solution of the Bargaining Game

8. *Some Definitions*

Solution functions. We now propose to present an axiomatic theory for a rational selection of a unique conditional payoff vector $x^* = (x_1^*, \cdots, x_N^*) \in X$ when the bargaining basis $B = (X, r)$ is given. Formally the purpose of our axioms will be to select a unique solution function L as defined below.

A *bargaining basis* $B = (X, r)$ is called *regular* if the equilibrium set X contains at least one element $x = (x_1, \cdots, x_N)$ with $x_i > w_i$ for every player $(i = 1, \cdots, N)$. (Thus, informally speaking, a bargaining basis is regular if there is some possible stable outcome of the bargaining game which would yield each player i a payoff expectation x_i greater than his conflict payoff w_i.) In order to avoid unnecessary mathematical complications we shall develop the axiomatization in terms of regular bargaining bases. Later it will be easy to show how the theory can be extended to irregular bargaining bases.

A *solution function* is a function L which assigns an element $x^* = L(B)$ of the equilibrium set X to every regular bargaining basis $B = (X, r)$; we call $x^* = L(B)$ the *solution* of B.

The generalized Nash product. The solution function L, which we shall axiomatize, selects that $x^* = L(B)$ which maximizes a generalized Nash product. In order to have a convenient way of writing this product we shall adopt the notation

$$p_{K+m} = q_m \tag{8.1}$$

for the marginal probabilities $q_1, \cdots, q_M$ associated with types $1, \cdots, M$ of player 2. Then we can write the generalized Nash product in the form

$$\pi = \Pi_{i=1}^{N} (x_i - w_i)^{p_i}. \tag{8.2}$$

If $B = (X, r)$ is a regular bargaining basis, then the maximum of the generalized Nash product π is attained at a uniquely determined conditional payoff vector $x = x^*$. This follows from the convexity of X and from the existence of conditional payoff vectors $x \in X$ with $x_i > w_i$ for $i = 1, \cdots, N$.

If the bargaining basis $B = (X, r)$ is not regular, then $x_i = w_i$ holds for at least one component of every vector $x \in X$. In that case the product π vanishes for all $x \in X$. Therefore, the rule of maximizing the generalized Nash product π must be slightly modified for irregular bargaining bases. This question will be discussed later.

From now on we shall use the symbol L^* for the solution function which maximizes the generalized Nash product. L will denote an unspecified solution function.

Operations. The axioms which we shall formulate in order to characterize the solution function described in the previous section will refer to some operations which can be applied to bargaining bases. It will always be clear that the result of the application of one of these operations to a regular bargaining basis is also a regular bargaining basis.

Operation 1. *Interchanging the players.* Let $x = (x_1, \cdots, x_N)$ be a conditional payoff vector in the equilibrium set X of a bargaining basis $B = (X, r)$. Then the payoff vector $x' = (x'_1, \cdots, x'_N)$ with

$$x'_m = x_{K+m} \quad \text{for} \quad m = 1, \cdots, M \tag{8.3}$$

and

$$x'_{M+k} = x_k \quad \text{for} \quad k - 1, \cdots, K \tag{8.4}$$

is called the payoff vector *derived from x by interchanging the players*. Let X' be the set of all payoff vectors which can be derived from the vectors $x \in X$ by interchanging the players and let r' be the transpose of r. Then $B' = (X', r')$ is called the bargaining basis *derived from B by interchanging the players.*

Operation 2. *Interchanging two types of the same player.* Let h and j be two integers with $h \neq j$ and with $1 \leqq j \leqq K$ and $1 \leqq h \leqq K$. We define the payoff vector $x' = (x'_1, \cdots, x'_N)$ *derived from* $x = (x_1, \cdots, x_N)$ *by interchanging the types h and j of player* 1. This payoff vector has the components

$$x'_h = x_j, \tag{8.5}$$

$$x'_j = x_h, \tag{8.6}$$

$$x'_i = x_i \quad \text{for all} \quad i \quad \text{with} \quad i \neq h \quad \text{and} \quad i \neq j. \tag{8.7}$$

The bargaining basis $B' = (X', r')$ *derived from* $B = (X, r)$ *by interchanging the types h and j of player* 1 is defined as follows: X' is the set of all vectors x' which can be derived from vectors $x \in X$ by interchanging the types h and j of player 1 and r' results from r by interchanging the jth and hth rows.

The operation of interchanging two types of player 2 is analogous to that of interchanging two types of player 1. It can be defined as follows: the bargaining basis $B'' = (X'', r'')$ *derived from* $B = (X, r)$ *by interchanging the types h and j of player* 2 (here h and j are integers with $1 \leqq h \leqq M$ and $1 \leqq j \leqq M$) is that bargaining basis which results from B by first interchanging the players, then interchanging the types h and j of player 1, and finally interchanging the players once more.

Operation 3. *Order-preserving linear utility transformation.* Consider a system of order-preserving linear transformations

$$T_i(x_i) = \alpha_i x_i + \beta_i; \quad \alpha_i > 0; \quad i = 1, \cdots, N. \tag{8.8}$$

Let $T(x)$ be the vector

$$T(x) = (T_1(x_1), \cdots, T_N(x_N)). \tag{8.9}$$

We shall call T itself also an order-preserving linear transformation. Let $T(X)$ be the set of all $T(x)$ with $x \in X$. Then $B' = (T(X), r)$ is the bargaining basis derived from $B = (X, r)$ *by the order-preserving linear utility transformation* $T = (T_1, \cdots, T_N)$.

Operation 4. *Splitting a type.* Let $B = (X, r)$ be a bargaining basis and let j be an integer with $1 \leqq j \leqq K$. For every payoff vector $x = (x_1, \cdots, x_N)$ we define a payoff vector $x' = (x'_1, \cdots, X'_{N+1})$ *derived from x by splitting type j of player* 1 *into two types.* This payoff vector has the components

$$x'_i = x_i \quad \text{for} \quad i = 1, \cdots, j, \tag{8.10}$$

$$x'_{j+1} = x_j, \tag{8.11}$$

$$x'_i = x_{i-1} \quad \text{for} \quad i = j + 2, \cdots, N + 1. \tag{8.12}$$

The basic probability matrix r', *derived from r by splitting type j of player* 1 *into two types with the probabilities* v and $1 - v$, is defined as a $(K + 1) \times M$ matrix, related to r

as follows:

$$r'_{km} = r_{km} \quad \text{for} \quad k = 1, \cdots, j-1, \tag{8.13}$$

$$r'_{jm} = vr_{jm}, \tag{8.14}$$

$$r'_{j+1,m} = (1-v)r_{jm}, \tag{8.15}$$

$$r'_{km} = r_{k-1,m} \quad \text{for} \quad k = j+2, \cdots, K+1, \tag{8.16}$$

each of these four equations holding for $m = 1, \cdots, M$. Here v is a probability with $0 < v < 1$. Let X' be the set of all payoff vectors x' which can be derived from the payoff vectors x in X by splitting type j of player 1 into two types. Then $B' = (X', r')$ is the bargaining basis *derived from* $B = (X, r)$ *by splitting type* j *of player* 1 *into two types with the probabilities* v *and* $1 - v$.

The operation of splitting a type j of player 2 is analogous to the operation of splitting a type j of player 1. Let $B'' = (X'', r'')$ denote the bargaining basis derived from $B = (X, r)$ by splitting type j of player 2 into two types with the probabilities v and $1 - v$. We can define B'' as the bargaining basis we obtain if we first interchange the two players, then split type j of player 1 into two types with the probabilities v and $1 - v$, and finally interchange the two players once more.

Operation 5. *Dividing a type*. Consider the bargaining basis $B' = (X', r')$ which is obtained from $B = (X, r)$ by splitting type j of player 1 into two types with the probabilities v and $1 - v$. Let r'' be a $(K + 1) \times M$ matrix, related to r as follows:

$$r''_{km} = r_{km} \quad \text{for} \quad k = 1, \cdots, j-1, \tag{8.17}$$

$$r''_{jm} = v_m r_{jm}, \tag{8.18}$$

$$r''_{j+1,m} = (1-v_m)r_{jm}, \tag{8.19}$$

$$r''_{km} = r_{k-1,m} \quad \text{for} \quad k = j+2, \cdots, K+1, \tag{8.20}$$

each of these four equations holding for $m = 1, \cdots, M$. Here the quantities $v_1, \cdots, v_M$ are probabilities which may or may not be equal to one another. If now in $B' = (X', r')$ the matrix r' is replaced by r'' then we obtain the bargaining basis $B'' = (X', r'')$. Any bargaining basis B'' obtained in this way will be called a bargaining basis *derived from* $B = (X, r)$ *by dividing type* j *of player* 1 *into two types*. Obviously, the operation of "splitting" a given type j, already defined, is simply a special case of the operation of "dividing" this type j: it corresponds to that special case where $v_1 = \cdots = v_M = v$.

Again the operation of dividing type j of player 2 into two types is analogous to the operation of doing this for type j of player 1. In order to obtain the bargaining basis $B'' = (X'', r'')$ derived from $B = (X, r)$ by dividing type j of player 2, we can proceed as follows. We first interchange the two players, then we divide type j of player 1 into two types, and finally interchange the two players once more.

9. *The Axioms*

We now formulate a set of axioms for solution functions L. Later it will be shown that there is one and only one solution function satisfying these axioms. This is the solution function L^* which is generated by the maximization of the generalized Nash product.

AXIOM 1 (PROFITABILITY). The solution $x^* = (x_1^*, \cdots, x_N^*) = L(B)$ of a regular bargaining basis $B = (X, r)$, where $w = (w_1, \cdots, w_N)$ is the minimal element of X, satisfies the inequalities $x_i^* > w_i$ for $i = 1, \cdots, N$.

Axiom 2 (Player Symmetry). If $B' = (X', r')$ is derived from $B = (X, r)$ by interchanging the players, then $x'^* = L(B')$ is derived from $x^* = L(B)$ by interchanging the players.

Axiom 3 (Type Symmetry). If $B' = (X', R')$ is derived from $B = (X, r)$ by interchanging the types h and j of player 1, then $x'^* = L(B')$ is derived from $x^* = L(B)$ by interchanging the types h and j of player 1.

Axiom 4 (Efficiency). The solution $x^* = L(B)$ of a bargaining basis $B = (X, r)$ has the property that there is no point x in X with $x \neq x^*$ and with $x_i \geqq x_i^*$ for $i = 1, \cdots, N$.

Axiom 5 (Linear Invariance). If $B' = (T(X), r)$ is derived from $B = (X, r)$ by application of the system $T = (T_1, \cdots, T_N)$ of order-preserving linear utility transformations, then $L(B') = T(L(B))$.

Axiom 6 (Irrelevant Alternatives). If $B = (X, r)$ and $B' = (X', r)$ are two bargaining bases with $X' \subseteq X$ and $L(B) \in X'$, then $L(B') = L(B)$.

Axiom 7 (Splitting Types). If $B' = (X', r')$ is derived from $B = (X, r)$ by splitting type j of player 1 into two types with probabilities v and $1 - v$, then $x'^* = L(B')$ is derived from $x^* = L(B)$ by splitting type j of player 1 into two types.

Axiom 8 (Mixing Basic Probability Matrices). If $B = (X, r)$ and $B' = (X, r')$ have the same solution vector $L(B) = L(B')$ and if r' has as many rows and as many columns as r, then for every $B'' = (X, r'')$ with $r'' = vr + (1 - v)r'$, where v is a probability with $0 \leqq v \leqq 1$, we have $L(B'') = L(B') = L(B)$.

The solution function L^* satisfies the axioms. In the following we shall show that the solution function L^*, which maximizes the generalized Nash product, satisfies Axioms 1 through 8.

Obviously $x^* > w$ must be true for the conditional payoff vector which maximizes the generalized Nash product for a regular bargaining basis. Thus L^* satisfies Axiom 1. The generalized Nash product does not change if the players or two types of one player are interchanged. Therefore, L^* satisfies the symmetry Axioms 2 and 3. It is also clear that Axiom 4 is satisfied, because any $x \in X$ with $x \neq x^*$ and $x \geqq x^*$ would be associated with a higher Nash product than x^*. The solution function L^* satisfies Axiom 5 because the order-preserving linear transformation $T = (T_1, \cdots, T_N)$, defined by equations (8.8) and (8.9), transforms the generalized Nash product π of equation (8.2) into the quantity

$$\pi' = \Pi_{i=1}^{N} (\alpha_i x_i - \alpha_i w_i)^{p_i} = \pi \cdot \Pi_{i=1}^{N} \alpha_i, \tag{9.1}$$

and consequently maps the solution of $B = (X, r)$ into the solution of $B' = (T(x), r)$.

Axiom 6 is satisfied because the payoff vector $x^* = L^*(B)$, maximizing the generalized Nash product π over the set X, maximizes the same product π also over any subset X' of X, *if* $x^* \in X'$.

Axiom 7 is satisfied because if $B' = (X', r')$ is derived from $B = (X, r)$ by splitting type j of player 1 into two types with the probabilities v and $1 - v$, then the generalized Nash product π associated with any payoff vector x in B will be identically equal to the generalized Nash product π' associated with the corresponding payoff vector x' in B', since

$$\pi' = \Pi_{i=1}^{N+1} (x'_i - w'_i)^{p'_i} = \Pi_{i=1}^{N} (x_i - w_i)^{p_i} = \pi. \tag{9.2}$$

This follows from equations (8.10), (8.11), (8.12) and from the fact that, by equations (8.13), (8.14), (8.15), and (8.16), the marginal probabilities $p'_1, \cdots, p'_{N+1}$ associated with r' and the marginal probabilities $p_1, \cdots, p_N$ associated with r are related as follows:

$$p'_i = p_i \quad \text{for} \quad i = 1, \cdots, j-1, \tag{9.3}$$

$$p'_j + p'_{j+1} = p_j, \tag{9.4}$$

$$p'_i = p_{i-1} \quad \text{for} \quad i = j+2, \cdots, N+1. \tag{9.5}$$

In order to prove that L^* also satisfies Axiom 8, we shall look at the logarithm of the generalized Nash product π. If x^* maximizes π for $B = (X, r)$ then we must have

$$\textstyle\sum_{i=1}^{N} p_i \log (x_i^* - w_i) > \sum_{i=1}^{N} p_i \log (x_i - w_i) \tag{9.6}$$

for every $x \in X$ with $x \neq x^*$. As $x^* = L^*(B) = L^*(B')$, a similar inequality holds also for the bargaining basis $B' = (X, r')$, so that

$$\textstyle\sum_{i=1}^{N} p'_i \log (x_i^* - w_i) > \sum_{i=1}^{N} p'_i \log (x_i - w_i) \tag{9.7}$$

for every $x \in X$ with $x \neq x^*$. Here the quantities $p'_1, \cdots, p'_N$ are the marginal probabilities associated with r'. The marginal probabilities for r'' of Axiom 8 are computed as follows:

$$p''_i = vp_i + (1 - v)p_i \quad \text{for} \quad i = 1, \cdots, N. \tag{9.8}$$

Now, multiplying (9.6) by v, then multiplying (9.7) by $1 - v$, and then finally adding up the resulting two inequalities and using (9.8), we obtain:

$$\textstyle\sum_{i=1}^{N} p''_i \log (x_i^* - w_i) > \sum_{i=1}^{N} p''_i \log (x_i - w_i) \tag{9.9}$$

or every $x \in X$ with $x \neq x^*$. Therefore, vector x^* maximizes the Nash product π also in $B'' = (X, r'')$, in accordance with Axiom 8.

10. *The Main Theorem. Characterization, Existence, and Uniqueness of the Solution*

Characterization of L^* by the Axioms 1 *through* 8. In the preceding section we have already proved the first half of the following Theorem:

THEOREM. *There is one and only one solution function which satisfies Axioms* 1 *through* 8, *namely the solution function L^*, which maximizes the generalized Nash product.*

We must show that L^* is the only solution function satisfying Axioms 1 through 8. In order to do this, we shall prove several lemmas in which we shall make use of the following definitions. A bargaining basis $B = (X, r)$ is called *linear*, if the equilibrium set X can be described as the set of vectors $x = (x_1, \cdots, x_N)$ satisfying the inequalities

$$x_i \geqq w_i \quad \text{for} \quad i = 1, \cdots, N \tag{10.1}$$

and

$$\textstyle\sum_{i=1}^{N} a_i x_i \leqq b, \tag{10.2}$$

where $a_1, \cdots, a_N$ are positive constants. A special case is the set X^N defined by the inequalities

$$x_i \geqq 0 \quad \text{for} \quad i = 1, \cdots, N \tag{10.3}$$

and

$$\textstyle\sum_{i=1}^{N} x_i \leqq 2. \tag{10.4}$$

Given any basic probability matrix r, the bargaining basis $B^r = (X^N, r)$ will be called the *norm basis* of r. These norm bases will be important for the proof of the Theorem. Obviously any regular linear bargaining basis can be obtained from some norm basis by using an order-preserving linear transformation T.

LEMMA 10.1. *Let r be a fixed basic probability matrix and let R be the set of all regular bargaining bases $B = (X, r)$ with this basic probability matrix r. Let L be a solution function which satisfies Axioms* 1, 4, 5, *and* 6, *and let $e = L(B^r)$ be the solution of the norm basis $B^r = (X^N, r)$ of r. The components of this vector $e = (e_1, \cdots, e_N)$ will satisfy the conditions*:

$$e_i > 0 \quad \textit{for} \quad i = 1, \cdots, N \tag{10.5}$$

and

$$\textstyle\sum_{i=1}^{N} e_i = 2. \tag{10.6}$$

Furthermore, if $x^* = L^*(B)$ is the solution of a bargaining basis $B = (X, r)$ in R, then x^* is the uniquely determined conditional payoff vector which maximizes the product

$$\pi_e = \Pi_{i=1}^{N} (x_i - w_i)^{e_i} \tag{10.7}$$

over the set X. Here $w = (w_1, \cdots, w_N)$ denotes the minimal element of X.

PROOF OF THE LEMMA. The inequality $e_i > 0$ follows from Axiom 1 because e is the solution of B^r. Equation (10.6) is a consequence of Axiom 4.

The product π_e assumes its maximum at the same point as does the function

$$\log \pi_e = \textstyle\sum_{i=1}^{N} e_i \log (x_i - w_i). \tag{10.8}$$

In the case of $B^r = (X^N, r)$ we have

$$\log \pi_e = \textstyle\sum_{i=1}^{N} e_i \log x_i. \tag{10.9}$$

Maximizing this function under the constraint

$$\textstyle\sum_{i=1}^{N} x_i = 2 \tag{10.10}$$

yields the necessary conditions

$$e_i/x_i = \lambda \quad \text{for} \quad i = 1, \cdots N, \tag{10.11}$$

where λ is the Lagrange multiplier. The solution of the set of equations (10.10) and (10.11) is $x = e$. Since $\log \pi_e$ is strictly concave in x, the product π_e assumes its maximum over X^N at exactly one point, namely at $x = e$. This shows that $e = L(B^r)$ maximizes π_e for B^r.

Any regular linear bargaining basis in R has the form $B = (T(X^N), r)$ where $T = (T_1, \cdots, T_N)$ is an order-preserving linear utility transformation. Therefore, the product π_e for B is nothing else than the corresponding product for B^r multiplied by a constant positive factor. Consequently, the product π_e for B is maximized by $x^* = T(e)$. Thus, in view of Axiom 5, the conditional payoff vector $x^* = T(e)$ is the solution of B. This shows that $x^* = L(B)$ is maximal with respect to π_e for every regular linear bargaining basis B.

Consider an arbitrary bargaining basis $B = (X, r)$ in R. Let w be the minimal element of X and let x^* be the point where the product π_e assumes its maximum over X. Let π_e^* be the value of π_e for $x = x^*$. The hypersurface H consisting of all points x for which we have $\pi_e = \pi_e^*$ can be represented by the equation

$$\textstyle\sum_{i=1}^{N} e_i \log (x_i - w_i) = \log \pi_e^*. \tag{10.12}$$

The equation of the hyperplane $H^{\#}$ tangential to this hypersurface H at the point $x = x^*$ is

$$\sum_{i=1}^{N} (x_i - x_i^*)e_i/(x_i^* - w_i) = 0. \tag{10.13}$$

Let $X^{\#}$ be the set of all points x satisfying the conditions

$$x_i \geqq w_i \quad \text{for} \quad i = 1, \cdots, N \tag{10.14}$$

and

$$\sum_{i=1}^{N} x_i e_i/(x_i^* - w_i) \leqq \sum_{i=1}^{N} x_i^* e_i/(x_i^* - w_i). \tag{10.15}$$

Since the quantities $e_i/(x_i^* - w_i)$ are all positive, $B^{\#} = (X^{\#}, r)$ is a linear bargaining basis in R. Obviously $x = x^*$ is the point where π_e assumes its maximum for $B^{\#}$. Therefore, $B^{\#}$ has the solution $x^* = L(B^{\#})$.

Now X is a subset of $X^{\#}$. This can be seen as follows. Assume that X contains a point y outside $X^{\#}$. The straight line which connects y with x^* contains a point $x = x^0$ above the hypersurface H and for this point x^0 the product π_e is greater than π_e^*. This follows from the strict concavity of the left side of the equation (10.12) defining H. Hence, because of the convexity of X the point x^0 must be in X. Since π_e^* is the maximum of π_e over X, this is a contradiction.

Since X is a subset of $X^{\#}$ it follows from Axiom 6 that $B = (X, r)$ and $B^{\#} = (X^{\#}, r)$ have the same solution x^*. This completes the proof.

Lemma 10.2. *Let L be a solution function satisfying Axioms 3, 7, and 8. Then the following is true: If $B' = (X', r')$ is derived from $B = (X, r)$ by dividing type j of player 1 into two types, then $L(B')$ is derived from $L(B)$ by splitting type j of player 1.*

Proof of Lemma 10.2. We must have $x_j^* = x_{j+1}'^*$ for $x'^* = L(B')$ because $x'_j = x'_{j+1}$ holds for every $x' \in X'$ and x'^* must be an element of X'. Consider the bargaining basis $B'' = (X'', r'')$ which is derived from $B' = (X', r')$ by interchanging the types j and $j + 1$ of player 1. In view of Axiom 3 we must have $x_{j+1}''^* = x_j'^*$ and $x_j''^* = x_{j+1}'^*$ for $x'^* = L(B')$ and $x''^* = L(B'')$. We know already that $x_j'^* = x_{j+1}''^*$. Consequently, we have $x_j''^* = x_{j+1}''^* = x_j'^* = x_{j+1}'^*$.

By the manner in which $B' = (X', r')$ is derived from $B = (X, r)$ it is clear that we have $X'' = X'$. Consider the basic probability matrix $r^0 = \frac{1}{2}r' + \frac{1}{2}r''$ and the bargaining basis $B^0 = (X', r^0)$. In r'' the rows j and $j + 1$ of r' are interchanged but apart from that there is no difference between r' and r''. Obviously r^0 is the basic probability matrix which results if type j of player 1 in $B = (X, r)$ is split into two types with probabilities $\frac{1}{2}$ and $\frac{1}{2}$. Therefore, $B^0 = (X', r^0)$ is the bargaining basis which is derived from $B = (X, r)$ by splitting type j of player 1 into two types with probabilities $\frac{1}{2}$ and $\frac{1}{2}$. Consequently, by Axiom 7 the solution $L(B'')$ is derived from $L(B)$ by splitting type j of player 1 into two types.

We know that $X'' = X'$ is true and we have proved $L(B'') = L(B')$. Therefore, it is possible to apply Axiom 8 to B', B'', and B^0. This yields $L(B^0) = L(B')$. Consequently, $L(B')$ is derived from $L(B)$ by splitting type j of player 1 into two types. This completes the proof of Lemma 10.2.

Lemma 10.3. *Let L be a solution function satisfying Axioms 1 and 3 through 8 and let $e(r) = L(B^r)$ be the solution of the norm basis B^r of r. The vector function $e(r) = (e_1(r), \cdots, e_N(r))$ has the following property:*

If r' is derived from r by dividing type j of player 1 or by dividing type $j - K$ of player 2 *then the following are true:*

(10.16) $$e_i(r') = e_i(r) \quad \textit{for} \quad i = 1, \cdots, j-1,$$

(10.17) $$e_j(r') + e_{j+1}(r') = e_j(r),$$

(10.18) $$e_i(r') = e_{i-1}(r) \quad \textit{for} \quad i = j+2, \cdots, N+1.$$

Proof of Lemma 10.3. We shall assume that $j \leqq K$. The case where j is a subplayer of player 2 can be treated analogously. We can derive a bargaining basis $B' = (X', r')$ from the norm basis $B^r = (X^N, r)$ by dividing type j of player 1. In view of Lemma 2 the solution $e' = L(B')$ must be derived from $e(r)$ by splitting type j of player 1. On the other hand, $e' = L(B')$ must be equal to the conditional payoff vector x' which maximizes the product

(10.19) $$\pi_e = \sum_{i=1}^{N+1} (x'_i)^{e_i(r')}$$

subject to the constraints

(10.20) $$x'_j = x'_{j+1}$$

and

(10.21) $$\sum_{i=1;\, i\neq j+1}^{N+1} x'_i = 2.$$

This follows from the definition of B' and from Lemma 10.1. It can easily be seen that this product π_e of equation (10.19) assumes its maximum at the point $x' = (x'_1, \cdots, x'_N)$ with

(10.22) $$x'_1 = e_i(r') \quad \text{for} \quad i = 1, \cdots, j-1,$$

(10.23) $$x'_j = x'_{j+1} = e_j(r') + e_{j+1}(r'),$$

(10.24) $$x'_i = e_{i-1}(r') \quad \text{for} \quad i = j+2, \cdots, N+1.$$

But we already know that $x' = L(B') = e'$, where e' is the payoff vector obtained from $e(r)$ by splitting type j of player 1. This fact together with equations (10.22) to (10.24) shows that the lemma is true.

Lemma 10.4. *Let L be a solution function satisfying Axiom* 1 *as well as Axioms* 3 *to* 8. *Let*

(10.25) $$r = \begin{pmatrix} p^1 \\ \vdots \\ p_K \end{pmatrix}$$

be a basic probability matrix with K rows but only one column. Then each component $e_i(r)$ *of the vector* $e(r)$ *will be a function of the one quantity* p_i *alone so that we can write*

(10.26) $$e_i(r) = \varphi(p_i) \quad \textit{for} \quad i = 1, \cdots, K,$$

and the function φ *will be independent of i.*

Proof of Lemma 10.4. The matrix r can be derived from the 1×2 matrix

(10.27) $$r' = \begin{pmatrix} p_1 \\ 1 - p_1 \end{pmatrix}$$

by successive applications of the operation of dividing a type j of player 1 with $j > 1$. Therefore, in view of equation (10.16) in Lemma 10.3, the first component $e_1(r)$ of vector $e(r)$ will always remain unchanged under these operations. Hence, $e_1(r)$ is a function of p_1 alone so that we can write $e_1(r) = \varphi(p_1)$.

If we interchange the types 1 and j in the norm basis $B^r = (X^N, r)$, we obtain the norm basis $B^{r'} = (X^N, r')$. Since B^r and $B^{r'}$ have the solutions $e(r)$ and $e(r')$, Axiom 3 yields

$$e_j(r) = e_1(r') = \varphi(p_j). \tag{10.28}$$

This shows that the lemma is true.

LEMMA 10.5. *Let L be a solution function satisfying Axioms* 1 *through* 8. *Then we have* $\varphi(p) = p$.

PROOF OF LEMMA 10.5. Consider the basic probability matrix r' with one row and one column and the single element 1. The norm basis of r' remains unchanged if the players are interchanged. Therefore, it follows from (10.7) in Lemma 1 and from Axiom 2 that we have

$$L(B^{r'}) = e(r') = (1, 1). \tag{10.29}$$

This yields

$$\varphi(1) = 1. \tag{10.30}$$

The matrix

$$r' = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} \tag{10.31}$$

can be derived from

$$r = \begin{pmatrix} p_1 + p_2 \\ 1 - p_1 - p_2 \end{pmatrix} \tag{10.32}$$

by dividing type 1 of player 1. Therefore, it follows from equation (10.23) in Lemma 3 that φ has the property

$$\varphi(p_1 + p_2) = \varphi(p_1) + \varphi(p_2). \tag{10.33}$$

In view of equations (10.30) and (10.33), we must have $\varphi(1/Z) = 1/Z$ for every integer Z and consequently also

$$\varphi(z/Z) = z/Z \tag{10.34}$$

for every positive rational number z/Z between 0 and 1. In view of inequality (10.5) in Lemma 1 we have $\varphi(p) > 0$ for $p > 0$. Therefore, it follows from (10.33) that $\varphi(p') > \varphi(p)$ holds for $p' > p$.

There can be no p with $0 < p \leqq 1$ and $\varphi(p) \neq p$. This can be shown as follows: Let p be such a number. Consider the case $\varphi(p) < p$. We can find a rational number p' with $\varphi(p) < p' < p$. For this p' we would have $\varphi(p') > \varphi(p)$ in spite of $p' < p$. We get a similar contradiction for $\varphi(p) > p$; for a rational p' with $\varphi(p) > p' > p$ we would have $\varphi(p') < \varphi(p)$ in spite of $p' > p$.

PROOF OF THE THEOREM. In order to prove the Theorem, it is sufficient to show that $e(r)$ is equal to the vector $p = (p_1, \cdots, p_N)$ of the marginal probabilities of r. This follows from Lemma 1 and from the definition of L^*. If r is a $K \times 1$ matrix, then

$e_i(r) = p_i$ for $i = 1, \cdots, K$ follows from Lemma 4 and Lemma 5. In view of equation (10.6) in Lemma 1 we also have $e_N(r) = p_N = 1$.

Any $K \times M$ matrix can be derived from a $K \times 1$ matrix by successive applications of the operation of dividing a type of player 2. In view of equation (10.16) in Lemma 3 these operations do not influence the first K components of e. The marginal probabilities $p_1, \cdots, p_K$ do not change either. Therefore, we have

$$e_i(r) = p_i \quad \text{for} \quad i = 1, \cdots, K, \tag{10.35}$$

for every $K \times M$ matrix r.

By Axiom 2 equation (10.35) must hold also for $i = K + 1, \cdots, N$. This can be seen as follows. Consider the norm basis $B^{r'} = (X^N, r')$ which is derived from the norm basis $B^r = (X^N, r)$ by interchanging the players. The solution $e(r)$ of B^r is derived from the solution $e(r')$ of $B^{r'}$ by interchanging the players. The first M components of $e(r')$ must be equal to the M marginal probabilities of the columns of r. This proves the Theorem.

11. *A Continuity Property of the Solution*

We now propose to show that the solution has the continuity property mentioned in §7.

LEMMA 11.1. *Under the Hausdorff metric, the solution $L^*(B)$ depends continuously on the equilibrium set X of the bargaining basis $B = (X, r)$.*

PROOF. Let V_ϵ be the ϵ neighborhood of $L^*(B)$. Let $\Pi^*(B)$ be the value that the generalized Nash product π takes at $L^*(B)$. Let γ be a small positive number to be specified later. Let H, H_γ and $H_{-\gamma}$ be the hypersurfaces consisting of all points x for which the generalized Nash product π takes the value $\Pi^*(B)$, $\Pi^*(B) + \gamma$, and $\Pi^*(B) - \gamma$ respectively. Let $H^\#$ be the hyperplane tangent to H at the solution point $L^*(B)$. Finally, let $H_\gamma^\#$ be the hyperplane parallel to $H^\#$ and tangential to H_γ.

We can always choose γ in such a way that the set A_γ of all points x lying above the hypersurface $H_{-\gamma}$ but below the hyperplane $H_\gamma^\#$ is fully contained in V_ϵ.

Let M_δ be the family of all sets X' having the general properties of an equilibrium set and a Hausdorff distance less than δ from X. We have to show that for every $\epsilon > 0$ we can find some $\delta > 0$, with the property that for any bargaining basis $B' = (X', r)$ with $X' \in M_\delta$, the solution $L^*(B')$ lies in V_ϵ.

We can always choose δ in such a way that any set $X' \in M_\delta$ will have some points above $H_{-\gamma}$ and no points above $H_\gamma^\#$. The latter fact follows from the convexity of X. Therefore, for any $B' = (X', r)$ with $X' \in M_\delta$, the solution $L^*(B')$ must lie in the set A_γ, and consequently also in V_ϵ.

12. *Irregular Bargaining Bases*

Extension of the solution function L^* to irregular bargaining bases. We have defined a solution function as a function which assigns a solution to every regular bargaining basis. Irregular bargaining bases have been excluded from the region on which a solution function is defined, because it is convenient to develop the axiomatic theory in terms of this definition. This does not mean that the limiting case of irregularity is without interest. Therefore, in this section we shall extend the definition of L^* to irregular bargaining bases.

Let $B = (X, r)$ be an irregular bargaining basis, where w is the minimal element of X. Consider a subplayer j in B for whom no $x \in X$ with $x_j > w_j$ can be found. Let J be the set of all subplayers j with this property. We call these subplayers *irregular* and call the other subplayers *regular*. Because of the convexity of X there are elements in x with $x_i > w_i$ for all regular subplayers i. Therefore the product

$$\pi = \Pi_{i=1;\ i \notin J}^{N} (x_i - w_i)^{p_i} \tag{12.1}$$

will assume its maximum over X at a uniquely determined conditional payoff vector $x^* \in X$, with $x_j^* = w_j$ for $j \in J$ and $x_i^* > w_i$ for $i \notin J$. We define the extension of L^* to irregular bargaining bases by

$$L^*(B) = x^*. \tag{12.2}$$

We call a function L which assigns a conditional payoff vector $x \in X$ to every bargaining basis $B = (X, r)$ an *extended* solution function. It can easily be seen that the extended solution function L^* satisfies Axioms 1 through 8. The same reasoning which was used to show that the unextended solution function L^* satisfies Axioms 1 through 8 can be applied almost without change to the extended solution function.

References

1. HARSANYI, JOHN C., "Approaches to the Bargaining Problem Before and After the Theory of Games," *Econometrica*, Vol. 24 (1956), pp. 144–157.
2. ——, " On the Rationality Postulates Underlying the Theory of Cooperative Games," *Journal of Conflict Resolution*, Vol. 5 (1961), pp. 179–196.
3. ——, "Games with Incomplete Information Played by 'Bayesian' Players," Part I, *Management Science*, Vol. 14 (November 1967), pp. 159–182. Part II, *ibid.* (January 1968), pp. 320–334. Part III, *ibid.* (March 1968), pp. 486–502.
4. MATHEMATICA, "Models of Gradual Reduction of Arms," Final Report F-6184 on Contract ACDA/ST-116, prepared for the U. S. Arms Control and Disarmament Agency, *Mathematica*, Princeton, N. J., November 1967.
5. NASH, JOHN F., "The Bargaining Problem," *Econometrica*, Vol. 18 (1950), pp. 155–162.
6. ——, "Two-Person Cooperative Games," *Econometrica*, Vol. 21 (1953), pp. 128–140.