MATHEMATICS OF OPERATIONS RESEARCH
Vol. 10, No. 1, February 1985
Printed in U.S.A.

NONZERO-SUM TWO-PERSON REPEATED GAMES WITH INCOMPLETE INFORMATION*†

SERGIU HART

Tel-Aviv University

Characterization of all equilibria of nonzero-sum two-person repeated games with incomplete information, in the standard one-sided information case. Informally, each such equilibrium is described by a sequence of communications between the players (consisting of information transmission and coordination), leading to some individually rational agreement. Formally, the concept of a *bi-martingale* is introduced.

1. Introduction. An *incomplete information* environment is one where at least some of the participants do not possess all the relevant data. Much interest has been devoted in recent years to the analysis of such situations. In the economic theory literature, for example: the principal-agent problem; the theory of actions; signalling (e.g., in insurance markets); rational expectations equilibria; and so on.

What are the main difficulties in such problems? First, consider the "informed" persons—those who know more than others. On one hand, it is to their advantage to make use of their additional information (in order to improve their own final outcome). On the other hand, by doing so they actually reveal this information—and their relative advantage vanishes. Thus, what is the good of being more informed, if one cannot profit from it? This type of conflict is an essential issue in the analysis of incomplete information environments.

As an idealized example, assume someone has "inside information" that a certain small company has just succeeded in developing a new product, for which a very profitable market exists. He thus expects that the value of the shares of this company on the Stock Exchange will rise dramatically. Should he immediately buy a large quantity of these shares? By doing so, he will implicitly signal to the others the success of the company—and everyone will want to buy its shares, raising their value immediately and lowering the profits of the initially informed person. The answer clearly lies in his buying the "right" quantity of shares—not too large to draw attention, and not too small to make his profit insignificant.

The results of the analysis of such models of incomplete information usually indicate that some transmission of information does occur (possibly, in an implicit way only; namely, deducing information from actions taken by those possessing it). Thus, some sort of communication and cooperation may arise (e.g., "trading information")—even though everything is based on purely selfish (noncooperative) motives.

There is yet another conflict—this time, for the "uninformed" participants. Should they trust the information transmitted by the informed ones? In the Stock Exchange

*Received October 4, 1982; revised October 31, 1983.
AMS 1980 subject classification. Primary: 90D15. Secondary: 90D15, 93E05.
OR/MS Index 1978 subject classification. Primary: 231 Games/group decisions.
Key words. Games with incomplete information, repeated games, nonzero-sum games, bi-martingale.
†This work was partially done at the Institute for Mathematical Studies in the Social Sciences, Stanford University, and at CORE and the Mathematics Department, Université Catholique de Louvain. It was supported at the first institution by National Science Foundation Grant SES80-06654-A1, 2.

I wish to acknowledge many useful discussions with Robert J. Aumann, Jean-François Mertens and Shmuel Zamir; their suggestions and comments were of great help. Thanks are also due to an anonymous referee who pointed out an inaccuracy in the proof of Proposition 4.31.

0364-765X/85/1001/0117$01.25

example, maybe the purpose of buying a large quantity of shares is just to convince everyone that a technological breakthrough indeed occurred, leading to a big buying activity, which may finally make a good profit for the one that started it all—whether or not a new product has been developed by the company at all!

Game theory is a tool for studying conflict situations—by definition, *inter-personal* conflicts. However, one obtains as an outcome resolution of *intra-personal* conflicts (like the ones mentioned above) as well—based on individual rational behavior. This is true in particular for games with incomplete information—a class of which forms the subject of this paper.

An important development in game theory in recent years has been in the study of multi-stage games—especially, the so-called *repeated games*, where the same game is played repeatedly. This suggests itself as a good framework for incomplete information games, for two main reasons.

The first one is that, by its very nature, a repeated game has enough structure to allow the kinds of complex behaviour we described above (and many others as well). There is enough "time" to enable players to "generate" certain beliefs in other people, or to make deductions, statistical inferences, and so on. There is also place for threats, for punishments—and for rewards too.

The second reason is more formal—although closely related to the first one. Consider an infinitely repeated game with *complete* information. A well-known result (called the "Folk Theorem" since its authorship is not clear) states that the noncooperative equilibria in the repeated game precisely correspond to the individually rational and jointly feasible points in the one-shot game. The importance of this result is that one obtains *cooperative* outcomes in the one-shot game from noncooperative behaviour in the infinite game. Thus, the cooperation we usually observe is explained here not as an outcome of altruistic motives—but of purely selfish noncooperative ones (which many feel are the only rational ones).

One is therefore led in a natural way to the study of *repeated games of incomplete information*. The first research on these was done in the Mathematica (1966–68) reports, in particular by Aumann, Maschler and Stearns. It turned out that the very complex structure of these games—which, as we pointed out above, is one of the reasons for studying them—creates many difficulties. Up to date, essentially only *two-person zero-sum* games have been extensively analyzed (see the forthcoming book of Mertens and Zamir 1980, or the notes of Sorin 1980 for details).

As for the *nonzero-sum* case (still, only *two players*), a first study has been done by Aumann, Maschler and Stearns (1968). They characterized a special class of equilibria, in the so-called *standard one-sided information* case, where one player has more information than the other one, and both observe during the play all the actions taken. These equilibria—called "enforceable joint plans"—essentially consist of a transmission of information from the informed to the uninformed player ("signalling"), followed by a completely nonrevealing play from then on (similar to the Folk Theorem). Moreover, they showed that this does not exhaust all equilibria—one could have joint randomizations of enforceable joint plans, and so on.

Our main result in this paper is the *complete characterization of all equilibria* in such games. We will show that every equilibrium is equivalent to a collection of nonrevealing "plans", one of which is chosen at random. This choice is done via a sequence of communications, which are of two types: signalling (i.e., implicit transmissions of information), and jointly controlled randomizations (i.e., "lotteries" in which no single player can unilaterally change the probabilities). The reader is referred to §3 for a more detailed description.

Thus, we are able to characterize in a formal way all the kinds of cooperation and communication that arise out of noncooperative behavior in these games; moreover, we obtain a precise structure that guarantees it does not pay any player to do anything else (e.g., revealing less or more, or double-crossing, cheating, and so on). We would like to point out that the model is not the most general possible (in particular, in terms of the information structure); this paper is to be regarded as a first step in the analysis of nonzero-sum repeated games with incomplete information.

The formal model is described in §2, together with various notions of equilibrium. The main results are stated in §3, which also includes additional discussion and intuitive interpretations. §§4 and 5 are devoted to the two parts of the proof, and in §6 we present some results on enforceable joint plans. (We would like to point out that Sorin 1983 has recently proved the existence of such equilibria whenever the number of possible games is two.) The paper is concluded with an example analyzed in §7.

Some notation: R is the real line, and R^n the n-dimensional Euclidean space. For vectors $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in R^n, $x \geq y$ means $x_i \geq y_i$ for all $i = 1, 2, \ldots, n$, and $x \cdot y$ is the scalar product $\sum_{i=1}^n x_i y_i$. For a finite set L, $|L|$ is the number of elements of L, and R^L the $|L|$-dimensional Euclidean space with coordinates indexed by the members of L (thus, we write $x = (x_l)_{l \in L} = (x(l))_{l \in L}$ for x in R^L). The unit simplex in R^L will be denoted by Δ^L:

$$\Delta^L = \left\{ x \in R^L : x_l \geq 0 \text{ for all } l \text{ in } L, \sum_{l \in L} x_l = 1 \right\}.$$

Finally, N is the set of positive integers $\{1, 2, \ldots\}$.

2. The model. The class of games we study is given by the following:

(i) Two players, player 1 and player 2.

(ii) A finite set I of choices for player 1 and a finite set J of choices for players 2; I and J contain each at least two elements.

(iii) A finite set K of games; to each k in K there corresponds a pair of $I \times J$ matrices (A^k, B^k), with $A^k = (A^k(i,j))_{i \in I;\, j \in J}$, $B^k = (B^k(i,j))_{i \in I;\, j \in J}$.

(iv) A probability vector $p = (p^k)_{k \in K}$ on the set K (i.e., $p \in \Delta^K$); without loss of generality, we assume $p^k > 0$ for all k in K; otherwise, we may discard those k that have zero probability.

Based on (i)–(iv), a *game of incomplete information* $\Gamma_\infty(p)$ is given as follows:

(v) An element κ of K is chosen according to the probability vector p; κ is told to player 1 but not to player 2.

(vi) At each stage $t = 1, 2, \ldots,$ player 1 chooses an element i_t in I and player 2 chooses an element j_t in J; the choices are made simultaneously (or, without either player knowing what the other did).

(vii) Both players are then told the pair (i_t, j_t), and they get the payoffs $A^\kappa(i_t, j_t)$ and $B^\kappa(i_t, j_t)$, respectively (but they do not observe these payoffs).

(viii) Both players have perfect recall (i.e., they do not forget what they were told at all previous stages).

(ix) All of (i)–(viii) is common knowledge to both players (see Aumann 1976 for a precise definition).

Usually, (v) is called *one-sided information* (see also the discussion below), and (vii) and (viii)—*standard information*. Note that the players observe only the actual choices i_t and j_t, and not the randomizations used.

Following Harsanyi (1967–1968), this can be equivalently viewed as a game with complete but imperfect information (namely, where the uncertainty players have is not about the "rules of the game"—e.g., payoffs—but only about moves previously made, by the players or by chance). This is done by adding a stage $t = 0$, at which "nature" chooses an element κ of K according to the probability p. At each stage $t = 1, 2, \ldots,$ the information player 2 has consists of the sequence of previous choices by both players: $(i_1, j_1), (i_2, j_2), \ldots, (i_{t-1}, j_{t-1})$. As for player 1, he in addition knows κ.

This completes the description of $\Gamma_\infty(p)$. It should be pointed out that more general games can be made to fit into this model. In particular, consider the case where player 1 does not have full information on κ, but player 2 knows even less (see Mertens and Zamir 1980, Chapter III). Formally, a partition of the set K is given for each player, which is informed only what element of his partition contains the chosen κ. For example, let $K = \{1,2,3,4,5\}$, the partition of player 1 is $\{1,2\},\{3\},\{4\},\{5\}$, and that of player 2 is $\{1,2,3\},\{4,5\}$. The (common) prior p is $(1/5,1/5,1/5,1/5,1/5)$. First, we observe that *both* players distinguish between $\{1,2,3\}$ and $\{4,5\}$—thus, there are two completely disjoint games. In the first, *both* do not distinguish between 1 and 2; therefore, this corresponds to $K' = \{\{1,2\},\{3\}\}$ and $p' = (2/3,1/3)$, where the payoff matrix for $\{1,2\}$ is $A^{\{1,2\}} = (1/2)A^1 + (1/2)A^2$, and similarly for B. Note that 2/3 is the conditional probability that $\kappa \in \{1,2\}$ given $\kappa \in \{1,2,3\}$, 1/3 is $P(\kappa = 3 \mid \kappa \in \{1,2,3\})$, and $1/2 = P(\kappa = 1 \mid \kappa \in \{1,2\}) = P(\kappa = 2 \mid \kappa \in \{1,2\})$. In the second game, $K'' = \{4,5\}$ and $p'' = (1/2,1/2)$. Thus, the original game has been decomposed into two games, each fitting our model. It should be clear how to generalize the construction given for this example.

Next we describe the sets of strategies of the players in $\Gamma_\infty(p)$. For each $t = 1, 2, \ldots,$ let H_t be the set of histories up to (but not including) stage t, namely,[1] $H_t = (I \times J)^{t-1}$. A *pure strategy* σ of player 1 is collection $\sigma = \{\sigma_t\}_{t=1}^{\infty}$, where

$$\sigma_t : H_t \times K \to I \tag{2.1}$$

for all $t = 1,2,\ldots$. Thus, for every history h_t in H_t and every k in K (the "true" game κ chosen), $\sigma_t(h_t;k)$ is the choice i_t made by player 1. In a similar way, a *pure strategy* τ *of player* 2 is $\tau = \{\tau_t\}_{t=1}^{\infty}$, where

$$\tau_t : H_t \to J \tag{2.2}$$

for all $t = 1,2,\ldots$.

A *mixed strategy* is, as usual, a probability distribution over the set of pure strategies. Since $\Gamma_\infty(p)$ is a game with perfect recall, one can restrict the study to behaviour strategies (cf. Kuhn 1953 and Aumann 1964), where players make independent randomizations at each move. A *behaviour strategy* is thus defined in the same way as a pure strategy, with (2.1) replaced by

$$\sigma_t : H_t \times K \to \Delta^I, \tag{2.3}$$

and (2.2) replaced by

$$\tau_t : H_t \to \Delta^J. \tag{2.4}$$

Since we never use pure strategies specifically, the term "strategy" will henceforth mean behaviour or mixed strategy.

We have not yet defined payoffs in $\Gamma_\infty(p)$, only sequences of payoffs. Given a pair of strategies (σ,τ) of the two players, we denote

$$a_T^k = \frac{1}{T}\sum_{t=1}^{T} A^k(i_t, j_t), \tag{2.5}$$

$$\beta_T = \frac{1}{T}\sum_{t=1}^{T} B^\kappa(i_t, j_t), \tag{2.6}$$

for all $T = 1,2,\ldots$ and all k in K. Thus, a_T^k is the average payoff up to (and including) stage T to player 1, if the true game is $\kappa = k$; this depends on the choices of

[1] The set H_1, being an empty product, is defined to consist of one element only.

i_t's and j_t's, made according to σ and τ (actually, only $\sigma(\cdot;k)$ and τ matter). Let $E^k_{\sigma,\tau}(a^k_T)$ denote its expectation. For player 2, β_T is his average payoff up to T; it depends on σ, τ and also on the choice of κ (according to p). Let $E_{\sigma,\tau,p}(\beta_T)$ be its expectation.

A pair of (σ,τ) of strategies is a (Nash) *equilibrium point* in $\Gamma_\infty(p)$ if

$$\liminf_{T\to\infty} E^k_{\sigma,\tau}(a^k_T) \geqslant \limsup_{T\to\infty} E^k_{\sigma',\tau}(a^k_T) \tag{2.7}$$

for all strategies σ' of player 1 and all k in K, and

$$\liminf_{T\to\infty} E_{\sigma,\tau,p}(\beta_T) \geqslant \limsup_{T\to\infty} E_{\sigma,\tau',p}(\beta_T) \tag{2.8}$$

for all strategies τ' of player 2. If we take $\sigma' = \sigma$ in (2.7), we get a vector $a = (a^k)_{k\in K}$ such that

$$\lim_{T\to\infty} E^k_{\sigma,\tau}(a^k_T) = a^k \tag{2.9}$$

for all k in K. Similarly, $\tau' = \tau$ in (2.8) gives β with

$$\lim_{T\to\infty} E_{\sigma,\tau,p}(\beta_T) = \beta. \tag{2.10}$$

We will call a and β the *payoffs* of the equilibrium point (σ,τ). Note that they are computed *ex post*—namely, after the choice of κ was made and player 1 was informed of it. Therefore, player 1 considers his payoffs in each possible state $\kappa = k$, whereas for player 2 only his expectation over κ matters. It can be easily checked that the definition does not change if we replace (2.7) by *ex ante* optimality, namely:

$$\liminf_{T\to\infty} E_{\sigma,\tau,p}(\alpha_T) \geqslant \limsup_{T\to\infty} E_{\sigma',\tau,p}(\alpha_T),$$

where α_T is defined in the same way as β_T (thus, $\alpha_T = a^\kappa_T$). Indeed, since the value k of κ is in any case part of the information player 1 has at every stage, he can choose his best response against τ independently for each k (this is $\sigma(\cdot;k)$). In the imperfect information version of $\Gamma_\infty(p)$, the adequate payoff is indeed this expectation over κ; in the incomplete information one, the vector of payoffs should be considered instead, since given any "type" k—in Harsanyi's terminology—it does not care about the payoffs to all the other possible types.

A strengthening of the definition of equilibrium is suggested by the results obtained in the zero-sum case (i.e., where $A^k + B^k \equiv 0$ for all k in K). A pair of strategies (σ,τ) is a *uniform equilibrium point* in $\Gamma_\infty(p)$ if

$$\liminf_{T\to\infty} E^k_{\sigma,\tau}(a^k_T) \geqslant \limsup_{T\to\infty}\Big(\sup_{\sigma'} E^k_{\sigma',\tau}(a^k_T)\Big) \tag{2.11}$$

for all k in K, and

$$\liminf_{T\to\infty} E_{\sigma,\tau,p}(\beta_T) \geqslant \limsup_{T\to\infty}\Big(\sup_{\tau'} E_{\sigma,\tau',p}(\beta)\Big). \tag{2.12}$$

Clearly, every uniform equilibrium point is also an equilibrium point (if we change the order of lim sup and sup in (2.11) and (2.12), we obtain (2.7) and (2.8), respectively). The *payoffs* (a, β) are given by (2.9) and (2.10).

To emphasize the difference between the two definitions, we translate them into the "ϵ-language". A uniform equilibrium satisfies the following: for every $\epsilon > 0$ there exists $T_0 \equiv T_0(\epsilon)$ large enough such that for all $T > T_0$,

$$E^k_{\sigma',\tau}(a^k_T) < a^k + \epsilon \quad\text{and}\quad E_{\sigma,\tau',p}(\beta_T) < \beta + \epsilon \tag{2.13}$$

for all k in K and *all* strategies σ' of player 1 and τ' of player 2. For a regular equilibrium (according to (2.7) and (2.8)), T_0 may also depend on σ' and τ'. The importance of (2.13) uniformly in σ' and τ' is that it implies that (σ, τ) generates an ϵ-equilibrium in all long enough but *finite* games $\Gamma_T(p)$ (which are defined in the same way as $\Gamma_\infty(p)$, but they only last T stages). Since $\Gamma_\infty(p)$ may be viewed as an "idealization" of such games, the uniform definition may seem more appropriate.

However, we will prove the following result:

PROPOSITION 2.14. *The sets of payoffs of equilibrium points and of uniform equilibrium points in $\Gamma_\infty(p)$ coincide.*

Thus, although it is clear that there exist equilibrium points that are not uniform, they are always payoff-equivalent to uniform ones.

Other definitions of equilibrium are also possible. For example, one could use Abel instead of Cesaro summability; namely, limits as $\rho > 0$ converge to 0 of

$$E\left(\rho \sum_{t=1}^{\infty} \frac{x_t}{(1+\rho)^t}\right)$$

where $\{x_t\}_{t=1}^{\infty}$ is the corresponding sequence of payoffs (this is interpreted as the limit, as the interest rate goes to zero, of the current value). Banach limits (see §4) can also be used. However, in all cases the set of equilibrium payoffs will be the same as in Proposition 2.14.

In view of this result, we can unambiguously define the *set of equilibrium payoffs of* $\Gamma_\infty(p)$. Our main result will be a characterization of this set.

Can one further strengthen the definition of equilibrium by changing the order of limit and expectation? The answer is no—as an example by J.-F. Mertens and the author shows already in the zero-sum case.

3. Statement and interpretation of the main result. In this section we state our main result—the characterization of all equilibria in $\Gamma_\infty(p)$.

The Folk Theorem in the complete information case states that the set of equilibrium payoffs coincides with the set of feasible and individually rational payoffs. We consider first the notion of individual rationality; it is to be understood in the sense of what each player cannot be prevented from obtaining (i.e., the "minmax"). The study of the zero-sum case (Aumann and Maschler 1966) enables us to characterize individual rationality in $\Gamma_\infty(p)$.

We need some notation first. Let p be a probability vector in Δ^K; let $p \cdot A$ be the matrix $\sum_{k\in K} p^k A^k$ (i.e., whose (i,j)th element is $\sum_{k\in K} p^k A^k(i,j)$). Consider the two-person zero-sum game with payoffs to player 1 given by $p \cdot A$ and let $(\text{val}_1 A)(p)$ denote its value (when played just once). Thus,

$$(\text{val}_1 A)(p) = \max_{x\in\Delta^I} \min_{y\in\Delta^J} (p\cdot A)(x,y) = \min_{y\in\Delta^J} \max_{x\in\Delta^I} (p\cdot A)(x,y) \tag{3.1}$$

where $x = (x_i)_{i\in I}$, $y = (y_j)_{j\in J}$, and

$$(p\cdot A)(x,y) = \sum_{i\in I}\sum_{j\in J} x_i y_j \sum_{k\in K} p^k A^k(i,j).$$

Similarly, let $(\text{val}_2 B)(p)$ be the value of player 2 of the two-person zero-sum game with payoff matrix $p \cdot B$ to player 2. Clearly,

$$(\text{val}_2 B)(p) = -(\text{val}_1(-B))(p). \tag{3.2}$$

For a function f on Δ^K, let $\text{vex} f$ denote its convexification; namely, $\text{vex} f$ is the largest convex function on Δ^K that does not exceed f. We will write $(\text{vex}\,\text{val}_2 B)(p)$ for the evaluation of the function $\text{vex}(\text{val}_2 B)$ at the point p.

We can now define: a vector $a = (a^k)_{k\in K}$ in R^K is an *individually rational payoff vector to player* 1 *in* $\Gamma_\infty(p)$ if

$$q \cdot a \geqslant (\text{val}_1 A)(q), \qquad \text{for all} \quad q \text{ in } \Delta^K. \tag{3.3}$$

A scalar β in R is an *individually rational payoff to player* 2 *in* $\Gamma_\infty(p)$ if

$$\beta \geqslant (\text{vex}\,\text{val}_2 B)(p). \tag{3.4}$$

These definitions are the correct ones, in view of the following results. A set Q in R^K is *approachable by player* 2 (cf. Blackwell 1956) if there exists a strategy τ of player 2 such that

$$\lim_{T\to\infty}\left(\sup_{\sigma} E_{\sigma,\tau}(d(Q, a_T))\right) = 0,$$

where $a_T = (a_T^k)_{k\in K}$ (recall (2.5)), d is the Euclidean distance in R^K, and the supremum is over all strategies σ of player 1.

PROPOSITION 3.5. *Let a be a vector in R^K. Then* (3.3) *is a necessary and sufficient condition for the set $Q = \{x \in R^K : x \leqslant a\}$ to be approachable by player* 2.

PROOF. Blackwell (1956); see Aumann and Maschler (1966). ∎

Thus, if (3.3) is satisfied, then player 2 can guarantee that the payoffs to player 1 will not, in the limit, exceed a^k for all k in K simultaneously. If (3.3) is *not* satisfied, then *given* any strategy of player 2, player 1 has a strategy such that, for at least one k in K, he will get more than a^k.

For player 2, we have

PROPOSITION 3.6. *Let β be a scalar in R. Then* (3.4) *is a necessary and sufficient condition for player* 1 *to have a strategy σ such that*

$$\limsup_{T\to\infty}\left(\sup_{\tau} E_{\sigma,\tau,p}(\beta_T)\right) \leqslant \beta,$$

where β_T is given by (2.6) *and the supremum is over all strategies τ of player* 2.

PROOF. Aumann and Maschler (1966); see (3.2). ∎

Again, this means that player 1 can hold player 2 down to $(\text{vex}\,\text{val}_2 B)(p)$ in $\Gamma_\infty(p)$, but to no less than that.

Having completed the study of individual rationality, we come next to feasibility. Let us consider a simple case first. Fix i in I and j in J: is there an equilibrium resulting in the pair (i, j) being chosen at every stage and for all k in K? Clearly, the answer depends on the actions the players will take "outside of equilibrium"—namely, when (i, j) is no longer been played. Again as in the Folk Theorem, it is easy to see that the necessary and sufficient condition is precisely individual rationality for both players (each will use the corresponding strategy given by Propositions 3.5 and 3.6, respectively, immediately after the other deviates from (i, j)). Therefore, the payoffs $a = (A^k(i, j))_{k\in K}$ and $\beta = \sum_{k\in K} p^k B^k(i, j)$ will be equilibrium payoffs in $\Gamma_\infty(p)$ if and only if (3.3) and (3.4) are satisfied.

This reasoning can now be extended to any convex combination by using the corresponding frequencies. It generates a class of equilibria in $\Gamma_\infty(p)$, which result in player 1 actually playing the same for all k in K (i.e., independent of κ). Note that this is true only "in equilibrium" (i.e., so long as there are no defections); "out of equilibrium", the strategy given by Proposition 3.6 may depend on κ. We will thus call these equilibria *nonrevealing*.

To define formally the corresponding payoffs, we denote by $(A,B)(i,j)$ the vector

$$(A,B)(i,j)=\left(\left(A^k(i,j)\right)_{k\in K},\left(B^k(i,j)\right)_{k\in K}\right)\in R^K\times R^K,$$

for all i in I and j in J. Then, let

$$F=\mathrm{co}\{(A,B)(i,j):i\in I,j\in J\},\tag{3.7}$$

where "co" denotes the convex hull of a set. F can be viewed as the set of feasible vector payoffs (in the one-shot game).

Let M be the maximum absolute value of any possible payoff:

$$M=\max\{|A^k(i,j)|,|B^k(i,j)|:i\in I,j\in J,k\in K\}.\tag{3.8}$$

We then write R_M^K for the set of all vectors in R^K, all of whose coordinates are bounded by M. We also put R_M for the real interval $[-M,M]$ (thus $R_M^K=(R_M)^K$). Clearly, F is a subset of $R_M^K\times R_M^K$.

Finally, we define the set G as follows: it consists of all triples (a,β,p), with a in R_M^K, β in R_M and p in Δ^K, such that (3.3) and (3.4) are satisfied, and there exist c and d in R^K with

$$(c,d)\in F,\tag{3.9}$$

$$a\geqslant c\quad\text{and}\quad p\cdot a=p\cdot c,\tag{3.10}$$

$$\beta=p\cdot d.\tag{3.11}$$

As in the zero-sum case, we will find it necessary to consider all the games $\Gamma_\infty(p)$, as p ranges over Δ^K, at the same time; a triple (a,β,p) is understood as (a,β) being payoffs in $\Gamma_\infty(p)$.

In view of our previous discussion, G is essentially the set of payoffs corresponding to nonrevealing equilibria (note that (3.10) can be restated as: $a^k=c^k$ if $p^k>0$, $a^k\geqslant c^k$ otherwise—therefore, a and c are identical for all relevant games).

Our main result states that, based on the set G, we can characterize all equilibrium payoffs. We thus define the concept of a G-process, as follows.

Let $g=(a,\beta,p)\in R_M^K\times R_M\times\Delta^K$. A sequence $\{g_n\}_{n=1}^\infty=\{(a_n,\beta_n,p_n)\}_{n=1}^\infty$ of $(R_M^K\times R_M\times\Delta^K)$-valued random variables (on some probability space) is called a *G-process starting at g* if:

$$g_1\equiv g\qquad\text{a.s.}\tag{3.12}$$

There exists a nondecreasing sequence $\{\mathscr{G}_n\}_{n=1}^\infty$ of finite fields[2] with respect to which $\{g_n\}_{n=1}^\infty$ is a martingale. (3.13)

Let g_∞ be an a.s. limit of g_n (as $n\to\infty$); then $g_\infty\in G$ a.s. (3.14)

For each $n=1,2,\ldots,$ either $a_{n+1}=a_n$ a.s., or $p_{n+1}=p_n$ a.s. (3.15)

The martingale condition in (3.13) means that g_n is $\mathscr{G}_n$-measurable and $E(g_{n+1}|\mathscr{G}_n)=g_n$ a.s. for all n. Together with (3.12), it implies $E(g_n)=g$ for all n. Since the sequence is uniformly bounded, the Martingale Convergence Theorem implies that it has an a.s. limit—thus (3.14) is well defined. It then means that $g_\infty=(a_\infty,\beta_\infty,p_\infty)$ satisfies a.s. individual rationality for both players (i.e., (3.3) and (3.4)), and also (3.9)–(3.11).

[2] A finite field means a field with finitely many elements; such a field is equivalent to a finite partition of the space (the atoms of the field being the elements of the partition).

The last condition (3.15) is slightly unusual: at every step, either a_n or p_n remain constant (while the other may change—but in such a way that the conditional expectation does not, by (3.13)). If we disregard the β_n coordinate, such a process may be called a *bi-martingale* (see Proposition 3.18 below). These objects are studied in Aumann and Hart (1983).

Finally, we define G^* as the set of all points $g = (a, \beta, p)$ in $R_M^K \times R_M \times \Delta^K$ such that there exists a G-process starting at g. We note here that (3.12) and (3.15) are essential conditions; without either one, G^* will just be the convex hull of G.

We are now ready to state our main result.

MAIN THEOREM. *Let $a \in R^K$ and $\beta \in R$. Then (a, β) are equilibrium payoffs in $\Gamma_\infty(p)$ if and only if $(a, \beta, p) \in G^*$.*

Thus, the set G^* is the graph of the equilibrium payoffs correspondence (as p ranges over Δ^K). It can be easily checked that G is a nonempty set; in particular, any $(a^k, b^k) \in R \times R$ that is feasible and individually rational in the game (A^k, B^k) (cf. the "Folk Theorem") generates a point (a, β, p) in G, with $a = (M, \ldots, M, a^k, M, \ldots, M)$, $\beta = b^k$ and p the kth unit vector in Δ^K. Hence, G^* is also a nonempty set; however, this does *not* imply that for *every* p in Δ^K there exist (a, β) with $(a, \beta, p) \in G^*$, i.e., that every $\Gamma_\infty(p)$ has at least one equilibrium point (recently, Sorin 1983 has shown this to be the case when $|K| = 2$).

The Main Theorem and Proposition 2.14 will be proved together (we know of no direct proof of the latter alone). This will be done by showing first (in §4) that all equilibrium payoffs, according to the regular definition (2.7)–(2.10), belong to G^*. And second, by constructing (in §5) a uniform equilibrium (cf. (2.11) and (2.12)) corresponding to any point in G^*.

The second part of the proof leads us to an important additional result; namely, that all equilibrium points in $\Gamma_\infty(p)$ are equivalent to a special class of equilibria (those we construct in §5). Informally, such an equilibrium consists of a "*master plan*", which is followed by each player so long as the other does it too; and of "*punishments*", which come into effect after a deviation from the master plan has been detected.

The master plan is a sequence of "*communications*" between the two players, the purpose of which is to eventually settle on a point in G which is played forever from then on (using frequencies), and leads to the desired "*payoffs*". The communications are of two sorts: "*signalling*", where the informed player 1 plays dependent on κ (and thus reveals some of this information to player 2, who can update his posterior probabilities); and joint decisions, more precisely "*jointly controlled lotteries*", where the two players make together a randomization on how to continue the play.[3] Signalling has already been obtained in the zero-sum case; however, the jointly controlled lotteries (in which the uninformed player plays a no lesser role than the informed one) are a feature of the nonzero-sum case only.

At the end of the communication period (which we assume for the moment to consist of finitely many stages only), player 1 will play independent of κ (otherwise, he will reveal additional information)—and thus a nonrevealing equilibrium results from then on (a point in G). In the general case, the sequence of communications may be infinite. However, after a long enough time, almost everything that was ever going to

[3]The standard example is the well-known children's way of choosing among two alternatives with equal probability ("two-finger Morra"): they each show, simultaneously, either one or two fingers. If they match (i.e., both show the same number), the first alternative is chosen; if not, the second one. If both choose the number of fingers at random (i.e., with probabilities 1/2, 1/2), the two alternatives each have probability 1/2, even when one of the participants uses any other strategy! (This is better than tossing a coin, which may be counterfeit—a fact known to one but not to the other.) This idea of jointly controlled randomizations is due to Aumann, Maschler and Stearns (1968).

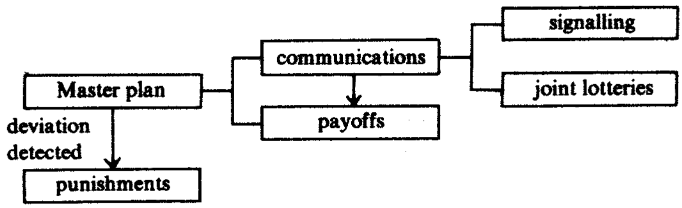

FIGURE 1

be revealed (by player 1) or decided (by a joint randomization) has already occurred—and we are essentially at a nonrevealing point again (i.e., in G). To generate the right payoffs, "payoff accumulation" periods are then introduced between communications—at which both players choose prescribed moves (again—with the correct frequencies).

Finally, punishments are always in accordance to the strategies given by Propositions 3.5 and 3.6, respectively[4] (see Proposition 3.16).

The structure of such equilibria is summarized in Figure 1.

The G-process is thus "followed" during the play. At each stage, the corresponding $g_n = (a_n, \beta_n, p_n)$ will serve as a "state variable", with a_n in R_M^K being the vector payoff player 1 will get from then on, β_n in R_M the same for player 2 (averaging over κ), and p_n in Δ^K the vector of posterior probabilities for[5] κ.

Why is an equilibrium thus obtained? Deviations during communication stages are not helpful: jointly controlled lotteries are so designed as to have each one of the players generate the right probabilities even if the other does not; as for signalling by player 1, it occurs precisely when $a_{n+1} = a_n$ in the G-process, which makes him indifferent among the various alternatives. In all other cases, the punishments keep the players in line. This is due to the following:

PROPOSITION 3.16. *Let $\{(a_n, \beta_n, p_n)\}_{n=1}^{\infty}$ be a $(R_M^K \times R_M \times \Delta^K)$-valued martingale, converging a.s. to $(a_\infty, \beta_\infty, p_\infty)$. Then*

(i) *a_∞ satisfies* (3.3) *a.s. if and only if for all $n = 1, 2, \ldots$, a_n satisfies* (3.3) *a.s.*

(ii) *(β_∞, p_∞) satisfies* (3.4) *a.s. if and only if for all $n = 1, 2, \ldots$, (β_n, p_n) satisfies* (3.4) *a.s.*

PROOF. The "if" part is obtained by taking the limit as $n \to \infty$ (in (ii), we use the continuity of the function $\operatorname{vex} \operatorname{val}_2 B$—e.g., see Mertens and Zamir 1980, Theorem 3.14).

Let $\{\mathcal{G}_n\}_{n=1}^{\infty}$ be the corresponding sequence of σ-fields, then we have $a_n = E(a_\infty \mid \mathcal{G}_n)$ by the martingale theorem. The "only if" part in (i) is obtained by taking conditional expectations over $\mathcal{G}_n$. As for (ii),

$$\beta_n = E(\beta_\infty \mid \mathcal{G}_n) \geq E((\operatorname{vex} \operatorname{val}_2 B)(p_\infty) \mid \mathcal{G}_n)$$
$$\geq (\operatorname{vex} \operatorname{val}_2 B)(E(p_\infty \mid \mathcal{G}_n)) = (\operatorname{vex} \operatorname{val}_2 B)(p_n),$$

where we used the convexity of the function $\operatorname{vex} \operatorname{val}_2 B$. ∎

[4] In the complete information case, the so-called "Perfect Folk Theorem" (cf. Aumann and Shapley 1976 and Rubinstein 1977) shows that any equilibrium can be made *perfect* (i.e. without "unbelievable threats", that hurt the "punisher" as well as the "punished"). However, this is not the case in the games of incomplete information we consider here. Indeed, player 1 may have to reveal additional information in order to keep player 2 at his individual rational level, after which a return to the master plan may not be possible. (Such an example may be easily constructed.)

[5] Our result shows that every randomization of player 2 can be replaced by a jointly controlled lottery. Therefore, one can essentially ignore the equilibrium conditions for player 2, and obtain a sequential structure (with the move of player 2 at every stage following that of player 1)—which can be then described by a state variable as above.

This last result leads to an additional interpretation of G^* as outcomes of bargaining processes—see Aumann (1981).

COROLLARY 3.17. *Let (α, β) be equilibrium payoffs in $\Gamma_\infty(p)$. Then α and β are individually rational for player 1 and player 2, respectively.*

PROOF. Proposition 3.16 for $n = 1$ (recall (3.12)); or, directly from Propositions 3.5 and 3.6. ∎

Another property of a G-process (which led to the name "bi-martingale") is as follows.

PROPOSITION 3.18. *Let $\{(a_n, p_n)\}_{n=1}^\infty$ be a $(R_M^K \times \Delta^K)$-valued martingale with respect to a nondecreasing sequence of σ-fields $\{\mathcal{G}_n\}_{n=1}^\infty$. If (3.15) is satisfied, then $\{a_n \cdot p_n\}_{n=1}^\infty$ is also a martingale with respect to $\{\mathcal{G}_n\}_{n=1}^\infty$.*

PROOF. Let n be such that $a_{n+1} = a_n$ a.s.; then $E(a_{n+1} \cdot p_{n+1} | \mathcal{G}_n) = E(a_n \cdot p_{n+1} | \mathcal{G}_n) = a_n \cdot E(p_{n+1} | \mathcal{G}_n) = a_n \cdot p_n$, since a_n is $\mathcal{G}_n$-measurable. The same when $p_{n+1} = p_n$. ∎

4. From equilibrium to martingale. This section contains the proof of the first part of our result; namely, given an equilibrium point we construct the corresponding G-process (see Proposition 4.43 at the end of the section for a precise statement).

We start with an informal discussion of the proof. Let (σ, τ) be an equilibrium point; to simplify the arguments, let us assume that the frequencies with which the various pairs (i, j) in $I \times J$ are played always converge. Let $c_\infty = (c_\infty^k)_{k\in K}$ and $d_\infty = (d_\infty^k)_{k\in K}$ be the limit payoffs; clearly, $(c_\infty, d_\infty) \in F$. For every history h_t up to stage t, we then define the following: for each k in K, $a^k(h_t)$ is the expected payoff to player 1 if $\kappa = k$ (thus, $a^k(h_t)$ is just the expectation of c_∞^k given h_t); $\delta(h_t)$ is the expected payoff to player 2 (the expectation of d_∞^κ given h_t); and $p^k(h_t)$ is the (posterior) probability that $\kappa = k$ (again, given h_t). We next introduce "half-steps", i.e., we define the above conditional expectations when given both h_t and the next move i_t of player 1; we will thus write $a^k(h_t, i_t)$, and so on.

Assume h_t has positive probability of occurring when $\kappa = k$. Then all possible moves i_t of player 1 (i.e., those with $\sigma_t(h_t; k)(i_t) > 0$) must have the same expected payoff $a^k(h_t, i_t)$. Otherwise, player 1 could give probability 1 to that i_t leading to the highest payoff; this would be "undetected" by player 2 (since this i_t is possible according to σ), and giving an expected payoff of $a^k(h_t, i_t)$, which is higher than that given by σ. This contradicts the equilibrium conditions, therefore $a^k(h_t, i_t)$ must be constant—hence, equal to $a^k(h_t)$—for all possible i_t's. A similar argument shows that for the other i_t's an inequality is obtained (if they are not chosen, then their corresponding payoff cannot be higher); this will eventually lead to the condition (3.10) in the limit.

Next, consider the half-step from (h_t, i_t) to $h_{t+1} = (h_t, i_t, j_t)$. Since player 2 does not know κ, j_t is independent of it, and the posterior probabilities cannot change. We thus have[6] $p^k(h_t, i_t) = p^k(h_{t+1})$.

It is easy to check that in all other cases, the martingale conditions are satisfied; e.g., $E(a^k(h_{t+1}) | h_t, i_t) = a^k(h_t, i_t)$, and so on. We have therefore obtained a martingale (with the index set being that of half-steps), which furthermore satisfies (3.15). The individual rationality conditions (3.3) and (3.4) also hold (since otherwise (σ, τ) will not be an equilibrium), and one can show that, in the limit (which exists by the martingale convergence theorem), a point in G is a.s. reached.

The actual proof will be quite complicated. Since we have no convergence of the

[6] Note that we ignore the equilibrium conditions for player 2—since we can replace his randomizations by jointly controlled lotteries (see previous footnote).

payoffs, we will need to use Banach limits.[7] To facilitate following the arguments, we divided the proof into a sequence of subsections.

4.1. *The probability space.* For each $t \in N$ (the set of positive integers), we defined $H_t = (I \times J)^{t-1}$, the set of histories before stage t. We also define the set of infinite histories $H_\infty = \prod_{t=1}^{\infty}(I \times J)$, an element of H_∞ being a sequence $\{(i_t, j_t)\}_{t=1}^{\infty}$ of moves made by the two players at all stages.

On H_∞ we define for each $t \in N$ the finite field generated by H_t, and call it $\mathscr{H}_t$; thus, two infinite histories belong to the same atom in $\mathscr{H}_t$ if and only if they coincide up to (but not including) t. Let $\mathscr{H}_\infty$ be the σ-field generated by all the $\mathscr{H}_t$'s (usually called the cylindrical or the product σ-field on the space H_∞).

The basic probablity space will also include the choice of κ in K by chance. Thus, let $\Omega = H_\infty \times K$ be endowed with the σ-field $\mathscr{H}_\infty \otimes 2^K$. Each pair of strategies (σ, τ) and each probability vector $p \in \Delta^K$ for the initial chance move determine a probability distribution on this space. We denote it by $P_{\sigma,\tau,p}$; note that $E_{\sigma,\tau,p}$ used in §2 is precisely the expectation with respect to $P_{\sigma,\tau,p}$, and $E^k_{\sigma,\tau}$ is the conditional expectation given $\kappa = k$.

We will use some additional fields on H_∞. For each $t \in N$, let $H_{t+1/2} = (I \times J)^{t-1} \times I = H_t \times I$, and denote by $\mathscr{H}_{t+1/2}$ the finite field it generates. We have now defined H_s and $\mathscr{H}_s$ for all *half-integers* s, namely all[8] $s \in N_2 \equiv \{1, 1\frac{1}{2}, 2, 2\frac{1}{2}, \dots\}$. Note that $\{\mathscr{H}_s\}_{s \in N_2}$ is an increasing sequence of finite subfields of $\mathscr{H}_\infty$, converging to $\mathscr{H}_\infty$ as $s \to \infty$.

Since our probability space is actually $\Omega = H_\infty \times K$ and not H_∞, we will denote the field generated by H_s on Ω also by $\mathscr{H}_s$; this will generate no confusion.

4.2. *Banach limit.* In order to deal with the nonsummability of the sequences of payoffs, we introduce the concept of a "Banach limit" (e.g., see Dunford and Schwartz 1958, p. 73).

As usual, let l_∞ be the (Banach) space of all real bounded sequences $x = \{x_n\}_{n=1}^{\infty}$. A *Banach limit* is a real operator L on l_∞ with the following properties[9] (holding for all $x = \{x_n\}_n$ and $y = \{y_n\}_n$ in l_∞, and λ, μ in R):

$$L(\{\lambda x_n + \mu y_n\}) = \lambda L(\{x_n\}) + \mu L(\{y_n\}), \tag{4.1}$$

$$L(\{x_{n+1}\}_{n=1}^{\infty}) = L(\{x_n\}_{n=1}^{\infty}), \tag{4.2}$$

$$\liminf_{n\to\infty} x_n \leq L(\{x_n\}) \leq \limsup_{n\to\infty} x_n . \tag{4.3}$$

In particular, note that (4.3) implies:

$$L(\{x_n\}) = \lim_{n\to\infty} x_n, \qquad \text{if } \{x_n\} \text{ is a convergent sequence.} \tag{4.4}$$

Therefore the Banach limit is an extension of the notion of limit (to *all* bounded sequences). To slightly simplify the notation, we will henceforth write $L[x_n]$ for $L(\{x_n\}_{n=1}^{\infty})$.

Three further properties of Banach limits will be needed.

LEMMA 4.5. *Let L be a Banach limit, and let $\{x_n\}_n, \{y_n\}_n \in l_\infty$. Then*

$$|L[x_n] - L[y_n]| \leq \limsup_{n\to\infty} |x_n - y_n|.$$

[7] The use of Banach limits has been suggested by J.-F. Mertens.

[8] In this section we will always use t for integers in N, and s for half-integers in N_2.

[9] We list here only those we will need in our proofs; the existence of such L is guaranteed by the Hahn–Banach Theorem (see the reference above).

PROOF. Immediate by (4.1) and (4.3). ∎

LEMMA 4.6. *Let L be a Banach limit, and $X = \{X_n\}_{n=1}^{\infty}$ an l_∞-valued random variable (i.e., X is a measurable function from some probability space into l_∞). If X has only finitely many values, then $L[E(X_n)] = E(L[X_n])$.*

PROOF. Immediate by (4.1). ∎

In particular, this result will be useful for conditional expectations over *finite* fields. One could actually define a stronger concept of Banach limit, which commutes with the expectation operator for any uniformly bounded (or, even uniformly integrable) sequence of random variables—without the finiteness assumption. Although it will simplify some of the arguments below, the construction of such a so-called "medial limit" requires however the use of the continuum hypothesis—and it is not needed in our proof (cf. Mokobodzki, see Meyer 1973).

LEMMA 4.7. *Let L be a Banach limit, and C a compact and convex subset of some Euclidean space R^m. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in C, with $x_n = (\xi_n^{(1)}, \xi_n^{(2)}, \ldots, \xi_n^{(m)})$. Let $\eta^{(r)} = L[\xi_n^{(r)}]$ for $r = 1, 2, \ldots, m$. Then $y = (\eta^{(1)}, \eta^{(2)}, \ldots, \eta^{(m)}) \in C$.*

PROOF. Let q be any vector in R^m, then by (4.1), (4.3) and $x_n \in C$,

$$q \cdot y = L[q \cdot x_n] \leqslant \limsup_{n\to\infty} q \cdot x_n \leqslant \sup\{q \cdot c : c \in C\}.$$

This holds for all q; since C is a compact convex set, it implies $y \in C$. ∎

Given a Banach limit L, we can now define the concept of an *L-equilibrium point* in $\Gamma_\infty(p)$, by replacing (2.7) with

$$L\left[E^k_{\sigma,\tau}(a^k_T)\right] \geqslant L\left[E^k_{\sigma',\tau}(a^k_T)\right], \tag{4.8}$$

and (2.8) with

$$L\left[E_{\sigma,\tau,p}(\beta_T)\right] \geqslant L\left[E_{\sigma,\tau',p}(\beta_T)\right], \tag{4.9}$$

where the limit L is taken with respect to the index $T = 1, 2, \ldots$; this convention will be kept throughout this section. The corresponding *payoffs* will then be

$$L\left[E^k_{\sigma,\tau}(a^k_T)\right] = a^k \tag{4.10}$$

for each k in K, and

$$L\left[E_{\sigma,\tau,p}(\beta_T)\right] = \beta. \tag{4.11}$$

We put $a = (a^k)_{k\in K}$.

In view of (4.3), every equilibrium point is also an L-equilibrium point for any Banach limit L.

Throughout this section, we fix the following: a Banach limit L, a probability vector p in Δ^K, and an L-equilibrium point (σ, τ) in $\Gamma_\infty(p)$ with payoffs $(a, \beta) \in R^K_M \times R_M$. Unless stated otherwise, the probability measure $P \equiv P_{\sigma,\tau,p}$ is assumed (on the space Ω), with $E \equiv E_{\sigma,\tau,p}$ the corresponding expectation operator. Thus, all statements "a.s.", "martingale", and so on, will be with respect to P. Also, we will use E^k for the conditional expectation $E(\cdot \mid \kappa = k)$.

Our purpose is to construct a G-process starting at (a, β, p). The probability space on which it will be defined is Ω, and the sequence of fields is $\{\mathcal{H}_s\}_{s\in N_2}$.

4.3. *The martingale* $\{p_s\}$. For each $k \in K$, $s \in N_2$ and an history $h_s \in H_s$, let $p^k_s = p^k_s(h_s)$ be the conditional probability of the "true" game κ of being k, given σ, τ, p and h_s (namely, if $s = t \in N$, the first $t-1$ moves of each player; if $s = t + \frac{1}{2}$, $t \in N$, the first t moves of player 1 and $t-1$ moves of player 2). We can thus write $p_s^k = P_{\sigma,\tau,p}(\kappa = k \mid \mathscr{H}_s) = P(k \mid \mathscr{H}_s)$ (on each atom $h_s \in H_s$ of $\mathscr{H}_s$, p_s^k is a.s. constant, thus a.s. equal to $p_s^k(h_s)$). We put $p_s = (p_s^k)_{k\in K}$.

PROPOSITION 4.12. *The sequence* $\{p_s\}_{s\in N_2}$ *is a* Δ^K*-valued martingale with respect to* $\{\mathscr{H}_s\}_{s\in N_2}$, *satisfying*

$$p_1 \equiv p. \tag{4.13}$$

$$p_{t+1/2} = p_{t+1} \quad \text{for all} \quad t \in N. \tag{4.14}$$

There exists a Δ^K*-valued random variable* p_∞ *such that* $p_s \to p_\infty$ *a.s. as* $s \to \infty$. (4.15)

PROOF. The fact that $\{p_s^k\}_s$ forms a martingale is immediate from its definition. Since it is bounded, it must converge a.s., say to p_∞^k; then $p_\infty = (p_\infty^k)_{k\in K}$. (4.14) follows from the fact that given $h_{t+1/2}$ (actually, only h_t suffices), the tth move j_t of player 2 is independent of κ; as for (4.13), at $t = 1$ there is no history yet, hence posteriors and priors coincide. ∎

4.4. *The martingales* $\{\gamma_s\}$ *and* $\{\delta_s\}$. In §2, we defined the average payoffs of the two players up to time T (see (2.5) and (2.6)). We will find it useful to define also

$$\alpha_T = \frac{1}{T}\sum_{t=1}^{T} A^{\kappa}(i_t, j_t) \tag{4.16}$$

(i.e., $\alpha_T = a_T^\kappa$). For each $s \in N_2$, let $\gamma_s = L[E(\alpha_T \mid \mathscr{H}_s)]$, $\delta_s = L[E(\beta_T \mid \mathscr{H}_s)]$. Thus, γ_s and δ_s are the (Banach) limits of the expected average payoffs to player 1 and player 2, respectively, given a history h_s.

PROPOSITION 4.17. *The sequences* $\{\gamma_s\}_{s\in N_2}$ *and* $\{\delta_s\}_{s\in N_2}$ *are* R_M*-valued martingales with respect to* $\{\mathscr{H}_s\}_{s\in N_2}$, *satisfying*:

$$\gamma_1 \equiv p\cdot a \quad \text{and} \quad \delta_1 \equiv \beta. \tag{4.18}$$

There exist R_M*-valued random variables* γ_∞ *and* δ_∞
such that $\gamma_s \to \gamma_\infty$ *and* $\delta \to \delta_\infty$ *a.s. as* $s \to \infty$. (4.19)

PROOF. We can use Lemma 4.6: the field $\mathscr{H}_s$ being finite, γ_s has finitely many values:

$$\begin{aligned} E(\gamma_{s+1/2} \mid \mathscr{H}_s) &= E\big(L\big[E(\alpha_T \mid \mathscr{H}_{s+1/2})\big] \mid \mathscr{H}_s\big) \\ &= L\big[E(E(\alpha_T \mid \mathscr{H}_{s+1/2}) \mid \mathscr{H}_s)\big] \\ &= L\big[E(\alpha_T \mid \mathscr{H}_s)\big] = \gamma_s. \end{aligned}$$

Thus $\{\gamma_s\}_{s\in N_2}$ forms a martingale. It is bounded by M (which bounds all possible payoffs by (3.8), hence also averages, expectations and limits—by (4.3)—of those). Therefore it converges to some limit γ_∞. For $s = 1$, we have

$$E(\alpha_T \mid \mathscr{H}_1) = E(\alpha_T) = \sum_{k\in K} p^k E^k(a_T^k),$$

hence (4.1) and (4.10) give $\gamma_1 \equiv p\cdot a$. The sequence $\{\delta_s\}$ is dealt with in a similar way ($\delta_1 \equiv \beta$ is just (4.11)). ∎

4.5. *The martingales* $\{c_s\}$ *and* $\{d_s\}$. We now associate vector payoffs to each infinite history. We define, for each k in K and T in N,

$$b_T^k = \frac{1}{T} \sum_{t=1}^{T} B^k(i_t, j_t),$$

in a similar way to the definition (2.5) of a_T^k. Note that these are random variables, $\mathscr{H}_{T+1}$-measurable (k is fixed; in contrast, α_T and β_T in (4.16) and (2.6) are $(\mathscr{H}_{T+1} \otimes 2^K)$-measurable). We further remark that a_T^k and b_T^k are defined for all histories—even those which may be incompatible with $\kappa = k$ according to (σ, τ).

If the limit of a_T^k (as $T \to \infty$) would always exist, it would imply $E(\lim a_T^k) = \lim E(a_T^k)$. However, this is not the case, and the Banach limit L commutes with the expectation operator if there are only finitely many values (see Lemma 4.6 and the discussion thereafter). We define, for each $s \in N_2$: $c_s^k = L[E(a_T^k \mid \mathscr{H}_s)]$, $d_s^k = L[E(b_T^k \mid \mathscr{H}_s)]$ (again, for each k in K). Note that the expectations are *not* conditional on $\kappa = k$; thus, the probability of any history is its total probability, summed over all k in K.

One can interpret c_s^k and d_s^k as follows. Let $h_s \in H_s$ have positive probability, then $p_s = p_s(h_s)$ is the vector of posterior probabilities for the various games k. Assume that after h_s occurred, player 1 replaces his strategy σ by his average nonrevealing strategy there; namely, for *all* k in K, he uses $\sum_{k \in K} p_s^k \sigma_t(h_t; k)$ instead of $\sigma_t(h_t; k)$ whenever $t > s$ and h_t coincides with h_s up to s. The expected average payoffs up to T in game k will then be $E(a_T^k \mid h_s)$ and $E(b_T^k \mid h_s)$, respectively. As we shall see later, the difference in payoffs due to this change in strategy becomes negligible as $s \to \infty$ (Proposition 4.23). Intuitively, this is due to the fact that after sufficiently many stages, player 1 has already revealed (almost) everything he is ever going to reveal about the true game κ; thus, he must thereafter play (almost) nonrevealing, or (almost) independent of κ. In technical terms, this occurs whenever the martingales are close to their limits.

As usual, we write c_s for $(c_s^k)_{k \in K}$ and d_s for $(d_s^k)_{k \in K}$. The set F was defined in (3.7) as the set of all "feasible" vector payoffs to both players (in the one-shot game).

Proposition 4.20. *The sequences $\{c_s\}_{s \in N_2}$ and $\{d_s\}_{s \in N_2}$ are R_M^K-valued martingales with respect to $\{\mathscr{H}_s\}_{s \in N_2}$, satisfying:*

There exist R_M^K-valued random variables c_∞ and d_∞ such that $c_s \to c_\infty$ and $d_s \to d_\infty$ a.s. as $s \to \infty$. (4.21)

$$(c_\infty, d_\infty) \in F \text{ a.s.} \tag{4.22}$$

Proof. The martingale property and (4.21) are proved in a similar way to Proposition 4.17. For every T in N, the vector $((a_T^k)_{k \in K}, (b_T^k)_{k \in K})$ belongs to the compact convex set F, as an average of such vectors. The same holds for its expectations, and by Lemma 4.7 for its Banach limits (c_s, d_s) too. (4.22) now follows by letting $s \to \infty$. ∎

The next proposition makes precise the statement that, as $s \to \infty$, player 1 plays "almost" nonrevealing after s (see the discussion following the definition of c_s^k and d_s^k).

Proposition 4.23. *$\gamma_s - p_s \cdot c_s \to 0$ and $\delta_s - p_s \cdot d_s \to 0$ a.s. as $s \to \infty$.*

Proof. We prove here the first part. Fix $s \in N_2$, and let $t > s$, $t \in N$. Conditioning over $\mathscr{H}_{t+1}$ and κ gives (recall that $p_{t+1}^k = P(\kappa = k \mid \mathscr{H}_{t+1})$):

$$\begin{aligned} E(A^\kappa(i_t, j_t) \mid \mathscr{H}_s) &= E\Big(\sum_{k \in K} p_{t+1}^k A^k(i_t, j_t) \mid \mathscr{H}_s\Big) \\ &= \sum_{k \in K} p_s^k E\big(A^k(i_t, j_t) \mid \mathscr{H}_s\big) + \sum_{k \in K} E\big((p_{t+1}^k - p_s^k) A^k(i_t, j_t) \mid \mathscr{H}_s\big). \end{aligned}$$

We sum this for all t in the range $s \leq t < T$, and note that total payoffs up to s are bounded by sM, to obtain

$$\left|E(\alpha_T \mid \mathscr{H}_s) - \sum_{k\in K} p_s^k E(a_T^k \mid \mathscr{H}_s)\right| \leq 2\frac{sM}{T} + \frac{1}{T}M \sum_{\substack{t\in N\\ s\leq t<T}} \sum_{k\in K} E(|p_{t+1}^k - p_s^k| \mid \mathscr{H}_s).$$

We denote (for each $s \in N_2$)

$$\pi_s = \sum_{k\in K} \sup_{\substack{t\in N\\ t\geq s}} |p_{t+1}^k - p_s^k|,$$

and let $T \to \infty$. By Lemma 4.5 and (4.1), $|\gamma_s - p_s \cdot c_s| \leq ME(\pi_s \mid \mathscr{H}_s)$. Since $\{p_s\}_s$ converges a.s. as $s \to \infty$ by (4.15), it follows that $E(\pi_s \mid \mathscr{H}_s) \to 0$ a.s. as $s \to \infty$; this assertion is proved in the next lemma. ∎

LEMMA 4.24. *Let $\{X_n\}_{n=1}^{\infty}$ be a bounded sequence of real random variables, converging a.s. as $n \to \infty$, and let $\{\mathscr{F}_n\}_{n=1}^{\infty}$ be a nondecreasing sequence of σ-fields. Define $Y_n = \sup_{m\geq n}|X_m - X_n|$, then $E(Y_n \mid \mathscr{F}_n) \to 0$ a.s. as $n \to \infty$.*

PROOF.[10] Let $X_\infty = \lim X_n$, and put $Z_n = \sup_{m\geq n}|X_m - X_\infty|$. Then $\{Z_n\}_{n=1}^{\infty}$ is a nondecreasing sequence as $n \to \infty$, converging a.s. to zero. Therefore $\{E(Z_n \mid \mathscr{F}_n)\}_{n=1}^{\infty}$ is a bounded super-martingale[11] with respect to $\{\mathscr{F}_n\}_{n=1}^{\infty}$, hence converges a.s. to some Z_∞. Now $E(E(Z_n \mid \mathscr{F}_n)) = E(Z_n) \to 0$, thus $E(Z_\infty) = 0$ and $Z_\infty = 0$ a.s. Noting that $Y_n \leq 2Z_n$ completes the proof. ∎

Finally we have

COROLLARY 4.25. *$\gamma_\infty = p_\infty \cdot c_\infty$ a.s. and $\delta_\infty = p_\infty \cdot c_\infty$ a.s.*

PROOF. (4.15), (4.19), (4.21) and Proposition 4.23. ∎

4.6. *The martingales $\{e_s\}$ and $\{f_s\}$.* For each k in K and s in N_2, we define

$$e_s^k = \sup_{\sigma'} L\left[E_{\sigma',\tau}^k(a_T^k \mid \mathscr{H}_s)\right],$$

where σ' ranges over all strategies of player 1 (note that the expectation now is conditional on $\kappa = k$). Thus, for every history $h_s \in H_s$, e_s^k is the most player 1 can obtain if the true game is k and player 2 uses τ—given that h_s has already occurred.

PROPOSITION 4.26. *For every k in K, s in N_2 and t in N:*

$$e_1^k \equiv a^k. \tag{4.27}$$

$$e_s^k \geq c_s^k. \tag{4.28}$$

$$e_{t+1/2}^k = E(e_{t+1}^k \mid \mathscr{H}_{t+1/2}). \tag{4.29}$$

$$e_t^k(h_t) = \max_{i_t\in I} e_{t+1/2}^k(h_t, i_t) \qquad \textit{for all } h_t \textit{ in } H_t. \tag{4.30}$$

PROOF. (4.27) is just (4.8) and (4.10). To obtain (4.28), we consider the following σ': if $\kappa = k$ and h_s occurred, play the average nonrevealing strategy given by σ; namely, $\sigma_t'(h_t; k) = \sum_{k'\in K} p_s^{k'} \sigma_t(h_t; k')$ for all $t \geq s$ and h_t in H_t that coincide with h_s up to s (see the discussion following the definition of c_s^k and d_s^k in §4.5).

To prove (4.29), note that the additional information from $t + \frac{1}{2}$ to $t + 1$ is j_t, whose

[10] Suggested by J.-F. Mertens.

[11] I.e., $E[E(Z_{n+1} \mid \mathscr{F}_{n+1}) \mid \mathscr{F}_n] \leq E(Z_n \mid \mathscr{F}_n)$.

distribution depends on τ and h_t only, hence is the same in $E^k_{\sigma',\tau}$ as in E. Therefore,

$$\sup_{\sigma'} L\left[E^k_{\sigma',\tau}\left(a^k_T \mid h_{t+1/2}\right)\right] = \sup_{\sigma'} L\left[E\left(E^k_{\sigma',\tau}\left(a^k_T \mid h_{t+1/2}, j_t\right) \mid h_{t+1/2}\right)\right]$$
$$= \sup_{\sigma'} E\left(L\left[E^k_{\sigma',\tau}\left(a^k_T \mid h_{t+1}\right)\right] \mid h_{t+1/2}\right)$$

(we used Lemma 4.6). Given $h_{t+1/2}$, the first stage player 1 has to choose a move is $t+1$, and by that time he will already know j_t. Thus, the best he can do given $h_{t+1/2}$ is just to do his best given $(h_{t+1/2}, j_t) = h_{t+1}$, for *each* possible j_t. Therefore, the last expression is

$$= E\left(\sup_{\sigma'} L\left[E^k_{\sigma',\tau}\left(a^k_T \mid h_{t+1}\right)\right] \mid h_{t+1/2}\right),$$

proving (4.29).

Next, let $h_t \in H_t$ be given. For any σ', its relevant part for $L[E^k_{\sigma',\tau}(a^k_T \mid h_t)]$ consists of a probability distribution $\pi = \sigma'_t(h_t; k)$ in Δ^I for choosing i_t, and some strategy afterwards $\sigma'' \equiv \sigma''(i_t) = \sigma'((h_t, i_t, \cdot); k)$, for each possible i_t. Therefore (again using Lemma 4.6)

$$e^k_t(h_t) = \sup_{\pi \in \Delta^I} \sup_{\sigma''(i_t)} E'^k\left(L\left[E'^k\left(a^k_T \mid h_t, i_t\right)\right] \mid h_t\right),$$

where E'^k is just $E^k_{\sigma',\tau}$. The choice of $\sigma''(i_t)$ can be done separately for each i_t, therefore we can interchange the first E'^k with the supremum over $\sigma''(i_t)$, to obtain

$$e^k_t(h_t) = \sup_{\pi \in \Delta^I} \sum_{i_t \in I} \pi(i_t) e^k_{t+1/2}(h_t, i_t).$$

The supremum is attained by giving positive probability $\pi(i_t)$ only to those i_t for which $e^k_{t+1/2}(h_t, i_t)$ is maximal; this proves (4.30). ∎

It is easy to see that (4.29) and (4.30) imply that $e_s = (e^k_s)_{k \in K}$ forms a supermartingale: $e_s \geq E(e_{s+1/2} \mid \mathscr{H}_{s+1/2})$ for all $s \in N_2$. We want to obtain a martingale with values in R^K_M too. For this purpose, fix $k \in K$ and let $0 \leq \lambda^k_{t+1/2} \leq 1$ be such that $M - e^k_t = \lambda^k_{t+1/2}(M - e^k_{t+1/2})$ for all $t \in N$ (recall (4.30)). Now define f^k_s for all $s \in N_2$ by

$$M - f^k_s = \left(\prod_{\substack{r \in N \\ r < s}} \lambda^k_{r+1/2}\right)(M - e^k_s);$$

put $f_s = (f^k_s)_{k \in K}$.

PROPOSITION 4.31. *The sequence $\{f_s\}_{s \in N_2}$ is an R^K_M-valued martingale with respect to $\{\mathscr{H}_s\}_{s \in N_2}$, satisfying:*

$$f_1 \equiv a. \tag{4.32}$$

$$f_t = f_{t+1/2} \quad \textit{for all } t \textit{ in } N. \tag{4.33}$$

There exists an R^K_M-valued random variable f_∞ such that $f_s \to f_\infty$ a.s. as $s \to \infty$. (4.34)

$$f_s \geq e_s \geq c_s \quad \textit{for all } s \textit{ in } N_2. \tag{4.35}$$

$$f_\infty \geq c_\infty \quad \textit{and} \quad p_\infty \cdot f_\infty = p_\infty \cdot c_\infty \textit{ a.s.} \tag{4.36}$$

PROOF. All the λ's are in [0, 1], therefore $e^k_s \leq f^k_s \leq M$. (4.32) is immediate from (4.27) and the definition of f_1. Let $t \in N$, then

$$M - f^k_{t+1/2} = \left(\prod_{\substack{r \in N \\ r < t}} \lambda^k_{r+1/2}\right)\lambda^k_{t+1/2}(M - e^k_{t+1/2}) = \left(\prod_{\substack{r \in N \\ r < t}} \lambda^k_{r+1/2}\right)(M - e^k_t) = M - f^k_t,$$

proving (4.33). Moreover, $\lambda^k_{r+1/2}$ is $\mathscr{H}_{t+1/2}$-measurable for all $r < t$, and by (4.29) we obtain

$$M - E\left(f^k_{t+1} \mid \mathscr{H}_{t+1/2}\right) = \left(\prod_{\substack{r \in N \\ r < t+1}} \lambda^k_{r+1/2}\right)\left(M - E\left(e^k_{t+1} \mid \mathscr{H}_{t+1/2}\right)\right)$$

$$= \left(\prod_{\substack{r \in N \\ r < t+1/2}} \lambda^k_{r+1/2}\right)\left(M - e^k_{t+1/2}\right) = M - f^k_{t+1/2}.$$

This completes the proof that $\{f_s\}$ indeed forms a martingale in R^K_M (note: with respect to the probability measure $P \equiv P_{\sigma,\tau,p}$). Since it is bounded, (4.34) follows, and $f_s \geq e_s \geq c_s$ implies $f_\infty \geq c_\infty$ a.s. (recall (4.28) and (4.21)).

Therefore $p_\infty \cdot f_\infty \geq p_\infty \cdot c_\infty = \gamma_\infty$ (by Corollary 4.25). To obtain the opposite inequality, we note that $\{p_s\}$ and $\{f_s\}$ form a bi-martingale, hence $\{p_s \cdot f_s\}$ is a martingale (Proposition 3.18), and we have by (4.13), (4.32) and (4.18)

$$E(p_\infty \cdot f_\infty) = E(p_1 \cdot f_1) = p \cdot a = E(\gamma_1) = E(\gamma_\infty),$$

proving that $p_\infty \cdot f_\infty = \gamma_\infty = p_\infty \cdot c_\infty$ a.s. ∎

4.7. *Individual rationality.* We start with player 1.

PROPOSITION 4.37. $q \cdot f_s \geq q \cdot e_s \geq (\mathrm{val}_1 A)(q)$ *for all vectors* q *in* Δ^K *and all* s *in* N_2.

PROOF. Let $q \in \Delta^K$, and consider the one-shot zero-sum game $A(q)$. By definition (3.1), player 1 has a strategy $u \in \Delta^I$ such that for any strategy $v \in \Delta^J$ of player 2,

$$(\mathrm{val}_1 A)(q) \leq \sum_{i \in I} \sum_{j \in J} u_i v_j \sum_{k \in K} q^k A^k(i, j). \tag{4.38}$$

Let $h_s \in H_s$ have positive probability (under (σ, τ)). Define a new strategy σ' of player 1 as follows: $\sigma'((h_s, \cdot); k) = u$ for all k, and σ' equals σ otherwise (thus, after h_s has occurred, player 1 makes independent randomizations with distribution u at all stages for all k). By (4.38), we have for all t in N, $t \geq s$

$$(\mathrm{val}_1 A)(q) \leq E_{\sigma',\tau}\left(\sum_{k \in K} q^k A^k(i_t, j_t) \mid h_s\right).$$

As $T \to \infty$, payoffs before s become negligible in a^k_T, and we have (by (4.2), (4.3) and then (4.1)):

$$(\mathrm{val}_1 A)(q) \leq L\left[E_{\sigma',\tau}\left(\sum_{k \in K} q^k a^k_T \mid h_s\right)\right] = \sum_{k \in K} q^k L\left[E_{\sigma',\tau}\left(a^k_T \mid h_s\right)\right].$$

Recalling the definition of e^k_s (note that given h_s, $E^k_{\sigma',\tau}$ is independent of k),

$$(\mathrm{val}_1 A)(q) \leq \sum_{k \in K} q^k e^k_s(h_s) = q \cdot e_s(h_s),$$

and $q \cdot e_s \leq q \cdot f_s$ follows from (4.35). ∎

COROLLARY 4.39. $q \cdot f_\infty \geq (\mathrm{val}_1 A)(q)$ *a.s. for all* q *in* Δ^K.

PROOF. (4.34) and Propositions 3.16(i) and 4.37. ∎

We consider now player 2.

PROPOSITION 4.40. $\delta_s \geq (\mathrm{vex\,val}_2 B)(p_s)$ *a.s. for all* s *in* N_2.

PROOF. For each q in Δ^K, let $\Gamma^2_\infty(q)$ be defined in the same way as $\Gamma_\infty(q)$, but with payoff matrices $(-B^k)_{k \in K}$ instead of $(A^k)_{k \in K}$ for player 1. This is a zero-sum repeated game; therefore player 2 (the uninformed player) has a strategy $\hat{\tau} \equiv \hat{\tau}(q)$ such that

$$\limsup_{T\to\infty} E_{\sigma',\hat{\tau},q}\left(\frac{1}{T}\sum_{t=1}^{T}(-B^{\kappa}(i_t,j_t))\right) \leq (\operatorname{cav}(\operatorname{val}_1(-B)))(q)$$
$$= -(\operatorname{vex}\operatorname{val}_2 B)(q)$$

for all σ' (cf. Aumann and Maschler 1966—$\hat{\tau}$ may be taken to be the corresponding Blackwell strategy; "cav" is the concavification of a function, and we use (3.2)). Thus,

$$\liminf_{T\to\infty} E_{\sigma',\hat{\tau},q}(\beta_T) \geq (\operatorname{vex}\operatorname{val}_2 B)(q). \tag{4.41}$$

Let $h_s \in H_s$ have positive probability under (σ,τ), and consider the following strategy τ' of player 2: after h_s has occurred, τ' is $\hat{\tau}(p_s)$, where $p_s = p_s(h_s)$ is the vector of posterior probabilities given h_s; otherwise, τ' equals τ. Let $T \geq s$, then we condition on $\mathscr{H}_s$ to obtain

$$E(\beta_T) - E'(\beta_T) = P(h_s)(E(\beta_T \mid h_s) - E'(\beta_T \mid h_s)),$$

where $E' = E_{\sigma,\tau',p}$ (up to stage s—no difference between E and E'; aftwards—only if h_s has occurred). Apply the Banach limit L as $T\to\infty$; since (σ,τ) is an equilibrium (see (4.9)), we get by (4.1)

$$0 \leq P(h_s)(L[E(\beta_T \mid h_s)] - L[E'(\beta_T \mid h_s)]),$$

hence

$$\delta_s(h_s) = L[E(\beta_T \mid h_s)] \geq L[E'(\beta_T \mid h_s)]$$

(since $P(h_s) > 0$). By (4.3) and (4.41) (with $\sigma' = \sigma$: note that payoffs up to s do not matter as $T\to\infty$), the proof is completed. ∎

COROLLARY 4.42. $\delta_\infty \geq (\operatorname{vex}\operatorname{val}_2 B)(p_\infty)$ *a.s.*

PROOF. (4.15), (4.19) and Propositions 3.16(ii) and 4.40. ∎

4.8. *The G-process.* We have thus completed the proof of

PROPOSITION 4.43. *Let (a, β) be the payoffs of an L-equilibrium point (σ,τ) in $\Gamma_\infty(p)$. Then there exists a G-process starting at (a, β, p).*

PROOF. The probability space is $(\Omega, \mathscr{H}_\infty, P_{\sigma,\tau,p})$; the sequence of fields is $\{\mathscr{H}_s\}_{s\in N_2}$, and the G-process $\{g_s\}_{s\in N_2}$ is given by $g_s = (f_s, \delta_s, p_s)$. All the required properties are indeed satisfied: $g_1 \equiv (a, \beta, p)$ by (4.13), (4.18) and (4.32); the limit $g_\infty = (f_\infty, \delta_\infty, p_\infty)$ (see (4.15), (4.19) and (4.34)) belongs to the set G a.s. by (4.22), (4.36) and Corollaries 4.25, 4.39 and 4.42; and finally the "bi" property (3.15) is given in (4.14) and (4.33). ∎

5. From martingale to equilibrium. This section is devoted to the proof of the second half of our result; namely, given a G-process the corresponding uniform equilibrium point is constructed.

Let $g = (a, \beta, p)$ belong to G^*. Thus, we are given a probability space[12] $(Z, \mathscr{Q}, Q)$, a nondecreasing sequence $\{\mathscr{Q}_n\}_{n=1}^{\infty}$ of finite subfields of $\mathscr{Q}$, and a G-process $\{g_n\}_{n=1}^{\infty} = \{(f_n, \delta_n, p_n)\}_{n=1}^{\infty}$ with respect to $\{\mathscr{Q}_n\}_{n=1}^{\infty}$, starting at g; i.e.

$$(f_1, \delta_1, p_1) \equiv (a, \beta, p) \quad Q\text{-a.s.} \tag{5.1}$$

[12] Z is a set, $\mathscr{Q}$ a σ-field on Z, and Q a probability measure on $\mathscr{Q}$.

Without loss of generality, we can thus assume that $\mathscr{G}_1$ is the trivial field $\{Z, \phi\}$. Let $g_\infty = (f_\infty, \delta_\infty, p_\infty)$ be a Q-a.s. limit of g_n as $n \to \infty$: then $g_\infty \in G$ a.s. We will find it useful to weaken the bi-property (3.15) to the following:

$$\|f_{n+1} - f_n\| \cdot \|p_{n+1} - p_n\| = 0 \qquad \text{a.s. for all} \quad n = 1, 2, \ldots . \tag{5.2}$$

This means that on each atom of $\mathscr{G}_n$, either f_{n+1} is constant (and thus equals f_n), or p_{n+1} is constant (and equals p_n); however, which one of the two is true may differ from one atom to another. It is easy to see that G^* does not change (to obtain (3.15) from (5.2), insert between each $\mathscr{G}_n$ and $\mathscr{G}_{n+1}$ an additional field $\mathscr{G}_{n+1/2}$, and put $g_{n+1/2} = g_{n+1}$ if $f_{n+1} = f_n$ and $g_{n+1/2} = g_n$ otherwise).

5.1. *Standard G-process.* To simplify the construction of the equilibrium point, we will work with a G-process having the following additional property:

$$\begin{aligned}&\text{For every atom } z_n \text{ of } \mathscr{G}_n \text{ there are exactly two}\\ &\text{atoms } z'_{n+1} \text{ and } z''_{n+1} \text{ of } \mathscr{G}_{n+1} \text{ contained in } z_n,\\ &\text{and } Q(z'_{n+1} \mid z_n) = Q(z''_{n+1} \mid z_n) = 1/2.\end{aligned} \tag{5.3}$$

Such a G-process will be called *standard.*

PROPOSITION 5.4. *For every g in G^* there exists a standard G-process starting at g.*

PROOF. We will show how to "transform" a G-process into a standard one.

Given a G-process $\{g_n\}_{n=1}^\infty$ with respect to $\{\mathscr{G}_n\}_{n=1}^\infty$, we can describe the sequence of fields as a "probability tree" as follows. The nodes in the nth layer are the atoms of $\mathscr{G}_n$; the root (i.e., the first layer) can be taken to be Z (by (5.1)). A (directed) arc leads from an atom z_n of $\mathscr{G}_n$ to an atom z_m of $\mathscr{G}_m$ if and only if $m = n + 1$ and $z_m \equiv z_{n+1} \subset z_n$. We associate the probability $Q(z_{n+1} \mid z_n)$ to this arc and define the probability of a finite path starting at the root to be the product of the probabilities of all its arcs. This clearly equals $Q(z_n)$, where z_n is the endpoint of the path. This probability distribution is then uniquely extended in a standard way to all infinite paths in the tree starting at the root; we will denote this probability measure also by Q. This completes the description of our probability tree.

The G-process $\{g_n\}_{n=1}^\infty$ can now be regarded as being defined on the nodes of the tree; we will write $g_n(z_n)$ for the value of g_n on the atom z_n of $\mathscr{G}_n$. The properties (3.12)–(3.14) and (5.2) defining a G-process become:

(i) $g_1(z_1) = g$.

(ii) $E[g_{n+1}(z_{n+1}) \mid z_{n+1} \text{ succeeds } z_n] = g_n(z_n)$ for all z_n.

(iii) The sequence $\{g_n(z_n)\}_{n=1}^\infty$ converges for almost all infinite paths, and the limit g_∞ belongs to G a.s.

(iv) For each node z_n, either $f_{n+1}(z_{n+1}) = f_n(z_n)$ for all successors z_{n+1} of z_n, or $p_{n+1}(z_{n+1}) = p_n(z_n)$ for all successors z_{n+1} of z_n.

In order to obtain property (5.3), we need two kinds of modifications of the tree—and thus, of the G-process. First, we make the number of (immediate) successors of each node exactly two; and second, we make the probability of every arc precisely 1/2.

For the former, we have two cases. If there is only one successor z_{n+1} of z_n, we can add an additional copy of the whole subtree starting at z_{n+1}, and thus obtain two successors z_{n+1} and z'_{n+1} (which is identical to z_{n+1})—and moreover with probability 1/2 each (from z_n). Now assume z_n has more than two successors, say $\{z^r_{n+1}\}_{r=1}^m$. We then introduce additional nodes in between; e.g., at level $n + 1$ we will have z^1_{n+1} and the union of $z^2_{n+1}, \ldots, z^m_{n+1}$; from the latter, at level $n + 2$ we will have z^2_{n+1} and the union of $z^3_{n+1}, \ldots, z^m_{n+1}$; and so on. The probabilities of the new arcs will be defined

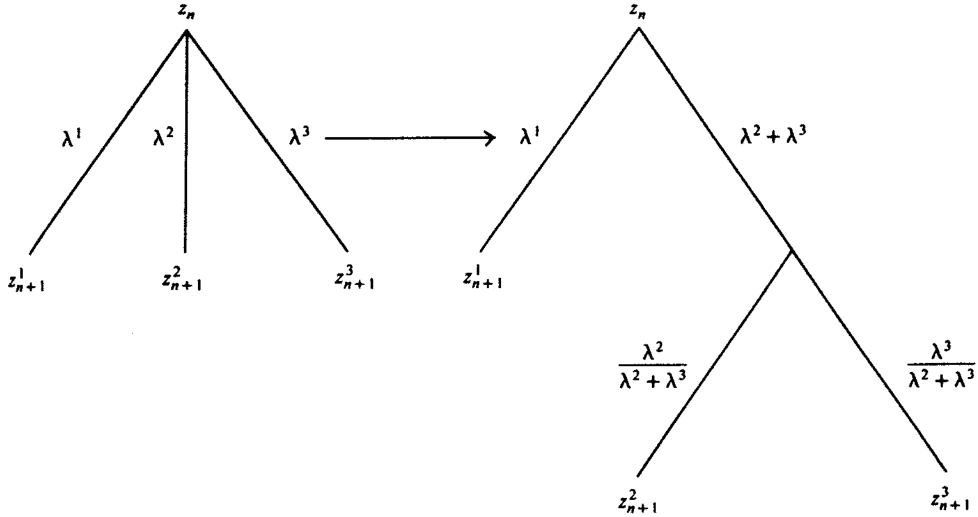

FIGURE 2

as the corresponding conditional probabilities; the value of the G-process at the new nodes, as the conditional expectation. As an example, see Figure 2; the value of the G-process at the new node will be

$$\frac{\lambda^2}{\lambda^2+\lambda^3}\, g_{n+1}(z^2_{n+1}) + \frac{\lambda^3}{\lambda^2+\lambda^3}\, g_{n+1}(z^3_{n+1}).$$

Clearly, all four properties (i)–(iv) continue to hold after such modifications.

Next we have to make the probabilities of all arcs precisely 1/2. Let z'_{n+1} and z''_{n+1} be the two successors of z_n, and let λ' and $\lambda'' = 1 - \lambda'$ be the corresponding probabilities. We want to obtain z'_{n+1} with probability λ' and z''_{n+1} with probability λ'' by using the probability 1/2 only. This is done as follows: we express λ' as a binary fraction

$$\lambda' = \sum_{m=1}^{\infty} \frac{1}{2^m} \lambda_m,$$

with $\lambda_m = 0$ or 1 for all m. We then consider an infinite sequence of independent Bernoulli trials, with "success" and "failure" having probability 1/2 each, up to the first occurrence of "success". If this happens after m trials, then z'_{n+1} "results" if $\lambda_m = 1$ and z''_{n+1} "results" if $\lambda_m = 0$. Thus, the total probability of z'_{n+1} is precisely λ' (since the first "success" occurs at the mth trial with probability $1/2^m$), and that of z''_{n+1} is λ''. This structure now replaces the original randomization between z'_{n+1} and z''_{n+1} in the tree. As an example, see Figure 3 (note that $\lambda' = 2/3$ gives $\lambda_m = 1$ for m odd and $\lambda_m = 0$ for m even). Again, the value of the G-process at a new node is the corresponding expectation.

With probability one, either z'_{n+1} or z''_{n+1} will be reached. If we do this modification at all nodes, the properties (i)–(iv) will not be affected (there are only countable many nodes, hence the probability of "success" not occurring in even one case is still zero). ∎

Henceforth we will assume that the G-process we start with is already standard.

5.2. *The sequence* $\{\theta_n\}$. The limit g_∞ of the G-process belongs to G a.s.; by (3.9), a corresponding point in F is thus obtained—and from it, a point in the set $\Delta^{I\times J}$ of "feasible joint actions".

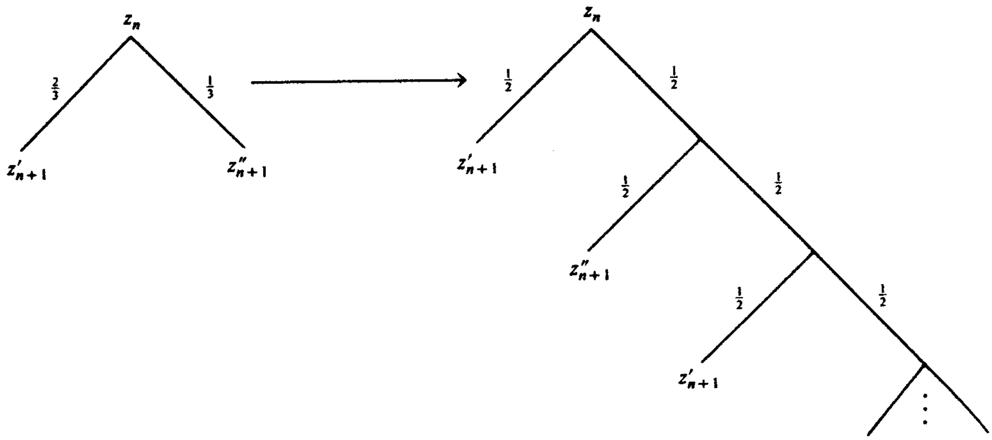

FIGURE 3

For $\theta = (\theta(i,j))_{i \in I, j \in J}$ in $\Delta^{I \times J}$ and k in K, we will denote

$$A^k(\theta) = \sum_{i \in I} \sum_{j \in J} \theta(i,j) A^k(i,j),$$

and $A(\theta) = (A^k(\theta))_{k \in K}$; similarly for B.

PROPOSITION 5.5. *There exists a $\Delta^{I \times J}$-valued random variable θ_∞ satisfying Q-a.s.*

$$f_\infty \geqslant A(\theta_\infty) \quad \text{and} \quad p_\infty \cdot f_\infty = p_\infty \cdot A(\theta_\infty), \tag{5.6}$$

$$\delta_\infty = p_\infty \cdot B(\theta_\infty). \tag{5.7}$$

PROOF. By definition, $g_\infty \in G$ implies the existence of $(c_\infty, d_\infty) \in F$ satisfying $f_\infty \geqslant c_\infty$, $p_\infty \cdot f_\infty = p_\infty \cdot c_\infty$ and $\delta_\infty = p_\infty \cdot d_\infty$. Since F is precisely the set of $(A(\theta), B(\theta))$ for all θ in $\Delta^{I \times J}$, there is θ_∞ in $\Delta^{I \times J}$ such that $c_\infty = A(\theta_\infty)$ and $d_\infty = B(\theta_\infty)$. The measurability is obtained by the Measurable Selection Theorem (e.g., see Hildenbrand (1974)). ∎

PROPOSITION 5.8. *There exists a sequence $\{\theta_n\}_{n=1}^\infty$ of $\Delta^{I \times J}$-valued random variables, satisfying Q-a.s. for all n in N and (i,j) in $I \times J$:*

$$\theta_n \text{ is } \mathcal{Y}_n\text{-measurable}. \tag{5.9}$$

$$|\theta_n(i,j) - E(\theta_\infty(i,j) \mid \mathcal{Y}_n)| < 1/n. \tag{5.10}$$

$$\theta_n \to \theta_\infty \quad \text{as} \quad n \to \infty. \tag{5.11}$$

$$n\theta_n(i,j) \text{ is an integer}. \tag{5.12}$$

PROOF. Define $\bar{\theta}_n = E(\theta_\infty \mid \mathcal{Y}_n)$, then $\{\bar{\theta}_n\}_{n=1}^\infty$ forms a martingale converging to θ_∞. Choose θ_n to be a rational approximation to $\bar{\theta}_n$ with denominators n (e.g., let $\theta'_n(i,j) = [n\bar{\theta}_n(i,j)]/n$, where $[x]$ denotes the largest integer not exceeding x, then $\theta_n(i,j)$ is either $\theta'_n(i,j)$ or $\theta'_n(i,j) + 1/n$, so as to have the sum equal 1). ∎

5.3. *The strategies σ and τ.* We can now define the pair of strategies (σ, τ). In a similar way to the so-called "Folk Theorem" for repeated games with *complete* information (for a detailed proof, see the Lecture Notes of Hart 1980, Section IV), they are based on a *master plan* and *punishments*. Each player follows the master plan as long as the other one does it too (at least, as long as no deviation is detected), and uses the corresponding punishment otherwise.

The master plan consists of two parts. Stages $t = n!$, for all $n = 1, 2, \ldots$, are *communication* stages; the moves made serve as a mean of transmitting information (from the informed to the uninformed player), or of making a joint decision. All the other stages are *payoff* periods; well-determined moves (namely, pure) are used in order for both players to accumulate the "right" payoffs. The sequence $n!$ was chosen since $(n-1)!$ is negligible relative to $n!$ as n goes to infinity—thus only the last period[13] really counts. Any other sequence with the same property could be used just as well.

The master plan is derived from the G-process. The moves at stage $t = n!$ correspond to the arcs from z_n to z_{n+1} in the tree (see §5.1), whereas at stages $(n-1)! < t < n!$, one "stays" at z_n. Thus, a function ζ is defined inductively from the set of finite histories in the game for which no deviations occurred, to the set of atoms of the fields $\{\mathcal{G}_n\}$—or, equivalently, to the set of nodes in the tree.

Let $i' \neq i''$ be two elements of I, and $j' \neq j''$ two elements of J, fixed throughout the remainder of this section. These two (pure) moves for each player will actually be their communication alphabet (thus, they essentially "talk" in a binary language).[14]

Let $t = n!$, let h_t be a history with no deviation from the master plan, and let $z_n = \zeta(h_t)$ be the corresponding node in the tree. We will define now the behaviour of each player at stage t, and also the resulting $\zeta(h_{t+1})$. If $p_{n+1} = p_n$ and $f_{n+1} = f_n$ at both nodes z'_{n+1} and z''_{n+1} succeeding z_n, the two players play arbitrarily at $t = n!$, and for all i_t and j_t, $\zeta(h_{t+1}) = z'_{n+1}$, say. Otherwise, we distinguish two cases:

(i) $p_{n+1} \neq p_n$ (and thus $f_{n+1} = f_n$),

(ii) $f_{n+1} \neq f_n$ (and thus $p_{n+1} = p_n$).

In case (i), we define for each k in K

$$\sigma_t(h_t; k)(i_t) = \begin{cases} \dfrac{p^k_{n+1}(z'_{n+1})}{2p^k_n(z_n)}, & \text{if } i_t = i'. \\[2ex] \dfrac{p^k_{n+1}(z''_{n+1})}{2p^k_n(z_n)}, & \text{if } i_t = i'', \\[2ex] 0, & \text{otherwise.} \end{cases}$$

Since $p_n(z_n) = (p_{n+1}(z'_{n+1}) + p_{n+1}(z''_{n+1}))/2$ by (5.3), $\sigma_t(h_t; k)$ is indeed a probability distribution over I. As for player 2, we let $\tau_t(h_t)$ be arbitrary in this case, and then for all j_t in J, we put $\zeta(h_t, i', j_t) = z'_{n+1}$ and $\zeta(h_t, i'', j_t) = z''_{n+1}$.

LEMMA 5.13. *Assume that* $P[\kappa = k \mid h_t] = p^k_n(\zeta(h_t))$ *for all* $k \in K$. *Then*

$$P\big[\zeta(h_{t+1}) = z'_{n+1} \mid h_t\big] = P\big[\zeta(h_{t+1}) = z''_{n+1} \mid h_t\big] = \tfrac{1}{2} \quad \textit{and}$$

$$P\big[\kappa = k \mid h_{t+1}\big] = p^k_n(\zeta(h_{t+1})).$$

PROOF. Assume $i_t = i'$, then we have

$$P\big[i_t = i' \mid h_t\big] = \sum_{k \in K} P\big[i_t = i' \mid h_t, \kappa = k\big] P\big[\kappa = k \mid h_t\big]$$

$$= \sum_{k \in K} \frac{p^k_{n+1}(z'_{n+1})}{2p^k_n(z_n)} p^k_n(z_n) = \frac{1}{2} \sum_{k \in K} p^k_{n+1}(z'_{n+1}) = \frac{1}{2}.$$

[13] By "period" we will usually mean the stages from $(n-1)!$ to $n!$ for some n.

[14] If there are more than two strategies, the communications may be "shortened" (i.e., less stages). This is not important in our model, since payoffs in finitely many periods do not matter, but will be so if a fixed discount rate is assumed.

Therefore

$$P[\kappa = k \mid h_{t+1}] = P[\kappa = k \mid h_t, i_t = i']$$
$$= \frac{P[i_t = i' \mid h_t, \kappa = k]P[\kappa = k \mid h_t]}{P[i_t = i' \mid h_t]}$$
$$= \frac{\dfrac{p_{n+1}^k(z'_{n+1})}{2p_n^k(z_n)} p_n^k(z_n)}{\frac{1}{2}} = p_{n+1}^k(z'_{n+1}).$$

Similarly for $i_t = i''$. ∎

Thus, if the posterior probabilities for the various values of k at stage $t = n!$ are p_n^k, then the new posteriors generated after the moves at time t are precisely p_{n+1}^k. Case (i) therefore corresponds to a transmission of information (about the value k of κ) from player 1 to player 2; we will henceforth call this *signalling* (by player 1).

In case (ii), $f_{n+1} \neq f_n$ and $p_{n+1} = p_n$; we define for all k in K:

$$\sigma_t(h_t\,; k)(i_t) = \begin{cases} \frac{1}{2} & \text{if } i_t = i', \\ \frac{1}{2} & \text{if } i_t = i'', \\ 0 & \text{otherwise,} \end{cases}$$

$$\tau_t(h_t)(j_t) = \begin{cases} \frac{1}{2} & \text{if } j_t = j', \\ \frac{1}{2} & \text{if } j_t = j'', \\ 0 & \text{otherwise,} \end{cases}$$

and then $\zeta(h_t, i', j') = \zeta(h_t, i'', j'') = z'_{n+1}$, $\zeta(h_t, i', j'') = \zeta(h_t, i'', j') = z''_{n+1}$.

LEMMA 5.14.

$$P[\zeta(h_{t+1}) = z_{n+1} \mid h_t] = P[\zeta(h_{t+1}) = z_{n+1} \mid h_t, i_t]$$
$$= P[\zeta(h_{t+1}) = z_{n+1} \mid h_t, j_t] = \tfrac{1}{2}$$

where z_{n+1} stands for either z'_{n+1} or z''_{n+1}, i_t for i' or i'', and j_t for j' or j''.

PROOF. The choices of i_t and j_t are made independently. ∎

Thus, in case (ii) a lottery with probabilities 1/2, 1/2 is performed among z'_{n+1} and z''_{n+1}. Moreover, no player has any control over the outcome—whichever of his two possible moves he chooses, the probabilities are the same (1/2, 1/2). Therefore, this is called (following Aumann, Maschler and Stearns 1968) a *jointly controlled lottery*.

This completes the definition of the master plan for $t = n!$ (the communication stages). It corresponds to advancing one step in the tree (from z_n to z_{n+1}).

We next consider the *payoff* periods. Let $z_n = \zeta(h_{(n-1)!+1})$ (thus, we are just after z_n was determined at stage $(n-1)!$). Let $\theta_n = \theta_n(z_n)$ in $\Delta^{I \times J}$ be given by Proposition 5.8 (see (5.9)). At stages $(n-1)! + 1$ through $n! - 1$, the players will play θ_n by frequencies; namely, the pair (i, j) will be played $\theta_n(i, j)$ of the time. Since all the denominators are n by (5.12), this can be done in cycles of length n each. For example, assume $\theta_n(i', j') = 1/n$, $\theta_n(i'', j'') = (n-1)/n$ and $\theta_n(i, j) = 0$ otherwise, then player 1 plays i' once (at $t = (n-1)! + 1$), then $n - 1$ times i'' (at $t = (n-1)! + 2, \ldots, (n-1)! + n$), repeating this n-stage cycle up to (and including) $t = n! - 1$; as for player 2, he chooses j' at $t = (n-1)! + 1$ and j'' at $t = (n-1)! + 2, \ldots, (n-1)! + n$, and so on. Clearly, we put $\zeta(h_t) = z_n$ for all $(n-1)! < t < n!$, when the two players play as described.

We introduce the following notation: for every i in I, j in J and u, v in N with $u < v$, let

$$\phi_u^v(i,j) = \frac{1}{v-u+1}\left|\{t \in N : u \leq t \leq v, i_t = i, j_t = j\}\right|.$$

Thus, $\phi_u^v(i,j)$ is the frequency that the pair of moves (i,j) was used at stages $u, u+1, \ldots, v$. Note that it is $\mathscr{H}_{v+1}$-measurable.

LEMMA 5.16. *Let $t \in N$, $(n-1)! < t < n!$. Then, for all i in I and j in J*

$$|\phi^t_{(n-1)!+1}(i,j) - \theta_n(i,j)| \leq \frac{n-1}{t-(n-1)!}.$$

PROOF. Every n stages, the frequency θ_n is precisely obtained. The inequality follows by ignoring the (at most) $n-1$ stages following the last complete n cycle. ∎

Finally, we have to define the *punishments*—what each player does after detecting a deviation from the master plan by the other player. Two results are needed from the theory of zero-sum games (see Propositions 3.5 and 3.6; the more precise statements here are needed to obtain a uniform equilibrium).

PROPOSITION 5.17. *Assume the vector y in R^k satisfies* (3.3). *Then player* 2 *has a strategy $\bar{\tau} \equiv \bar{\tau}(y)$ such that $E^k_{\sigma',\bar{\tau}}(a^k_T) \leq y^k + 2M/\sqrt{T}$ for all strategies σ' of player* 1, *all k in K and all T in N.*

PROOF. The precise bound is obtained from the proof of the approachability theorem (cf. Blackwell 1956, or Mertens and Zamir 1980, Chapter 1). ∎

PROPOSITION 5.18. *For every q in Δ^k, player* 1 *has a strategy $\bar{\sigma} \equiv \bar{\sigma}(q)$ such that $E_{\bar{\sigma},\tau',q}(\beta_T) \leq (\text{vex}\,\text{val}_2 B)(q)$ for all strategies τ' of player* 2 *and all T in N.*

PROOF. The above inequality actually holds with $B^\kappa(i_t, j_t)$ instead of β_T, for all t (e.g., see Aumann and Maschler 1966 or Mertens and Zamir 1980, Theorem 3.15). ∎

The definitions of σ and τ can now be completed. Assume first that player 1 deviated from the master plan, either by playing $i_t \neq i', i''$ at some $t = n!$ or by not playing the "right" i_t at some $(n-1)! < t < n!$. Let D be the stage at which this deviation of player 1 occurred. Thus, all moves in h_D are according to the master plan, and i_D is the deviation move (which is observed by player 2 before stage $D+1$). Let $z_n = \zeta(h_D)$ be the corresponding node just before the deviation; the strategy τ prescribes then that after h_{D+1} (i.e., from stage $D+1$ on), player 2 should use $\bar{\tau}(y)$ with $y = f_n(z_n)$ (see Proposition 5.17, and note that (3.3) is satisfied in view of Proposition 3.16(i)).

Next, assume player 2 deviated from the master plan at stage D (and was detected). From stage $D+1$ on, player 1 then uses $\bar{\sigma}(q)$ with $q = p_n(z_n)$ as defined in Proposition 5.18 (again, $z_n = \zeta(h_D)$).

This ends the definition of the pair of strategies σ and τ.

5.4. *Payoffs and probabilities for* (σ, τ). In this subsection we assume that both players use σ and τ, respectively. Thus, only the master plan matters; there are no deviations and no punishments.

We first analyze the payoffs. Let $T \in N$, $(n-1)! < T \leq n!$, and let h_T in H_T be a history possible under (σ, τ); i.e., $P_{\sigma,\tau,p}(h_T) > 0$. We will write θ_n for $\theta_n(\zeta(h_T))$—the value of θ_n on the atom $\zeta(h_T)$ of $\mathscr{G}_n$; similarly for the other random variables defined on Z. Recalling definition (5.15) of the frequencies ϕ, we have

PROPOSITION 5.19. *Let $T \in N$, $(n-1)! < T \leq n!$, $h_T \in H_T$ with $P_{\sigma,\tau,p}(h_T) > 0$.*

Then, for all i in I and j in J,

$$\left|\phi_1^{T-1}(i,j) - \left[\left(1 - \frac{(n-1)!}{T-1}\right)\theta_n(i,j) + \frac{(n-1)!}{T-1}\theta_{n-1}(i,j)\right]\right| \leq \frac{4}{n}.$$

PROOF. If $4/n \geq 1$, there is nothing to prove (both ϕ and the expression [. . .] lie in the interval [0, 1]). Let $n \geq 5$, then

$$\phi_1^{T-1} = \frac{(n-1)!}{T-1}\phi_1^{(n-1)!} + \frac{(T-1)-(n-1)!}{T-1}\phi_{(n-1)!+1}^{T-1}.$$

The frequency θ_{n-1} is "played" at stages $(n-2)! < t < (n-1)!$, therefore

$$|\phi_1^{(n-1)!}(i,j) - \theta_{n-1}(i,j)| \leq \frac{(n-2)!+1}{(n-1)!}$$

(the difference is due to $1 \leq t \leq (n-2)!$ and $t = (n-1)!$; in total, $(n-2)!+1$ stages; note that θ_{n-1} requires a cycle of length $n-1$, which divides $(n-1)!$). Together with the inequality in Lemma 5.16 for the second term, we obtain an overall difference of at most

$$\frac{(n-2)!+1+n-1}{T-1} \leq \frac{(n-2)!+n}{(n-1)!} < \frac{4}{n},$$

the last inequality being easily checked (for $n \geq 5$). ∎

This result shows that the frequencies obtained are close to those given by the sequence $\{\theta_n\}$. The next two corollaries will be needed in the sequel; we define $\overline{M} = M|I||J|$.

COROLLARY 5.20. *For all k in K,*

$$a_{T-1}^k \leq f_n^k + \frac{(n-1)!}{T-1}(f_{n-1}^k - f_n^k) + \frac{5\overline{M}}{n}.$$

PROOF. By Proposition 5.19,

$$a_{T-1}^k = A^k(\phi_1^{T-1}) \leq \lambda_T A^k(\theta_n) + \lambda'_T A^k(\theta_{n-1}) + 4\overline{M}/n,$$

with $\lambda'_T = (n-1)!/(T-1)$ and $\lambda_T = 1 - \lambda'_T$.

Recalling (5.10) gives $A^k(\theta_n) \leq E(A^k(\theta_\infty)|\mathcal{G}_n) + \overline{M}/n$. By (5.6) $A^k(\theta_\infty) \leq f_\infty^k$, hence $A^k(\theta_n) \leq f_n^k + \overline{M}/n$. Similarly for θ_{n-1}, and we obtain

$$A^k(\phi_1^{T-1}) \leq \lambda_T f_n^k + \lambda'_T f_{n-1}^k + 5\overline{M}/n,$$

from which the result follows. ∎

COROLLARY 5.21. *For all q in Δ^K,*

$$q \cdot B(\phi_1^{T-1}) \leq \max\{q \cdot B(\theta_n), q \cdot B(\theta_{n-1})\} + 4\overline{M}/n.$$

PROOF. Immediate from Proposition 5.19. ∎

Next, we deal with the probability $P_{\sigma,\tau,p}$ and the induced posteriors.

PROPOSITION 5.22. *Let $n \in N$, z_n an atom of $\mathcal{G}_n$ and $T \in N$, $(n-1) < T \leq n!$.*

Then:

$$P_{\sigma,\tau,p}\left[\zeta(h_T) = z_n\right] = Q(z_n), \tag{5.23}$$

$$P_{\sigma,\tau,p}\left[\kappa = k \mid h_T\right] = p_n^k(z_n) \tag{5.24}$$

for all h_T *in* H_T *with* $\zeta(h_T) = z_n$ *and all* k *in* K.

PROOF. Induction on n. For $n = 1$, there is only one history h_1 (the "empty" history), thus (5.23) is just $1 = 1$ and (5.24) is $p = p_1$ (recall (5.1)). The induction now proceeds as follows.

At stages $(n-1)! < t < n!$ (payoff stages), neither ζ nor any probabilities change (both players make pure choices). At $t = n!$, the probabilities for $\zeta(h_{t+1}) = z'_{n+1}$ or z''_{n+1} are $1/2$ each by Lemmata 5.13 and 5.14, thus equal to $Q(z'_{n+1} \mid z_n) = Q(z''_{n+1} \mid z_n)$ by (5.3) (recall that our G-process is now assumed standard). As for the posterior probabilities (of $\kappa = k$), they change only when there is signalling ($t = n!$ and case (i)) —and we use again Lemma 5.13. ∎

We will now show that the payoffs of (σ, τ) are (α, β). We need first the following result ($E^k_{\sigma,\tau}$ is the conditional expectation given $\kappa = k$, and $f^k_n = f^k_n(\zeta(h_T))$).

PROPOSITION 5.25. *Let* $T \in N$, $(n-1)! < T \leq n!$ *and* $k \in K$. *Then* $E^k_{\sigma,\tau}(f^k_n) = a^k$.

PROOF. The probability distribution of $z_n = \zeta(h_T)$ induced by $P_{\sigma,\tau,p}$ is precisely Q (by (5.23)); therefore $E_{\sigma,\tau,p}(f^k_n(\zeta(h_t))) = E_{(Z)}(f^k_n)$, where $E_{(Z)}$ denotes expectation on the space Z (with respect to Q; note that $E_{\sigma,\tau,p}$ and $E^k_{\sigma,\tau}$ are on Ω). Since $\{f^k_n\}_n$ is a martingale and $f_1 = a$ by (5.1), the above equals a^k.

We claim that the same expectation is obtained when using $P^k_{\sigma,\tau}$ instead of $P_{\sigma,\tau,p}$. Indeed, the induced probability distributions over the tree differ only in case of signalling (in a jointly controlled lottery, it is $1/2$ in both cases by Lemma 5.14); however, in that case $f_{m+1} = f_m$, so that the expectations are the same. ∎

REMARK. Actually, the conditional expectations (with respect to P and P^k) are also the same—thus, $\{f^k_n\}_n$ is a martingale also with respect to the probability distribution induced by P^k. Moreover, any strategy σ' of player 1 that differs from σ only in the probabilities used for signalling has this property (as we shall see in the next subsection).

PROPOSITION 5.26. $\lim_{T\to\infty} E^k_{\sigma,\tau}(a^k_T) = a^k$ *for all* k *in* K, *and* $\lim_{T\to\infty} E_{\sigma,\tau,p}(\beta_T) = \beta$.

PROOF. We start with player 2. Let $(n-1)! < T \leq n!$; conditioning on $\mathscr{H}_T$, we obtain ($E = E_{\sigma,\tau,p}$):

$$E(\beta_{T-1} \mid h_T) = \sum_{k\in K} p^k_n B^k(\phi_1^{T-1}),$$

with $p^k_n = p^k_n(\zeta(h_T))$ as usual. By Proposition 5.19,

$$\left|B^k(\phi_1^{T-1}) - \left[\lambda_T B^k(\theta_n) + \lambda'_T B^k(\theta_{n-1})\right]\right| \leq 4\overline{M}/n,$$

where $\lambda'_T = (n-1)!/(T-1)$ and $\lambda_T = 1 - \lambda'_T$. As in the proof of Proposition 5.25, the distribution of $z_n = \zeta(h_T)$ induced by $P = P_{\sigma,\tau,p}$ is Q (see (5.23)). Therefore,

$$\left|E(\beta_{T-1}) - E_{(Z)}\left(p_n \cdot \left[\lambda_T B^k(\theta_n) + \lambda'_T B^k(\theta_{n-1})\right]\right)\right| \leq 4\overline{M}/n.$$

As $n \to \infty$, $\theta_n \to \theta_\infty$ and $\theta_{n-1} \to \theta_\infty$ Q-a.s. by (5.11); also $p_n \to p_\infty$, hence ($\lambda_T, \lambda'_T \geq 0$, $\lambda_T + \lambda'_T = 1$ and everything is bounded):

$$\lim_{T\to\infty} E(\beta_{T-1}) = E_{(Z)}(p_\infty \cdot B(\theta_\infty)).$$

By (5.7), this is $E_{(Z)}(\delta_\infty) = \delta_1 = \beta$ (recall (5.1)).

For player 1, the same argument gives (α_T was defined in (4.16)):

$$\lim_{T\to\infty} E(\alpha_{T-1}) = E_{(Z)}(p_\infty \cdot A(\theta_\infty)) = E_{(Z)}(p_\infty \cdot f_\infty) = p_1 \cdot f_1 = p \cdot a$$

(see (5.6), Proposition 3.18 and (5.1)). If we condition on κ, we have $E(\alpha_{T-1}) = \sum_{k\in K} p^k E^k(a^k_{T-1})$, where $E^k = E^k_{\sigma,\tau} = E_{\sigma,\tau,p}(\cdot \mid \kappa = k)$. As in the proof of Corollary 5.20, we obtain $a^k_{T-1} \leq \lambda_T f^k_n + \lambda'_T f^k_{n-1} + 5\overline{M}/n$. By Proposition 5.25, $E^k(a^k_{T-1}) \leq \lambda_T a^k + \lambda'_T a^k + 5\overline{M}/n$, hence

$$\limsup_{T\to\infty} E^k(a^k_T) \leq a^k,$$

which together with

$$\lim_{T\to\infty} \sum_{k\in K} p^k E^k(a^k_T) = \sum_{k\in K} p^k a^k$$

and $p^k > 0$ for all k completes the proof. ∎

REMARK. The above proof actually shows that, if both players use (σ,τ), then the average payoffs converge a.s. to the corresponding $\theta_\infty = \lim_{n\to\infty}\theta_n$, where $\theta_n = \theta_n(\zeta(h_T))$. Therefore one can interchange the order of limit (as $T\to\infty$) and expectation (and there is no need for Banach limits!).

At this point we can show intuitively that (σ,τ) is indeed an equilibrium point (at least—in the weak sense (2.7)–(2.8); the complicated inequalities in the next two subsections are in part due to the fact that we want to prove the uniform property (2.11)–(2.12)).

Consider player 1, and fix k in K (the true game). As we noted in the last Remark, the average payoff to player 1 for infinite histories without deviations from the master plan is the corresponding $A^k(\theta_\infty)$, which is $\leq f^k_\infty$. If the game has proceeded up to a point corresponding to the node z_n in the tree, his expected payoff will thus be at most $E^k(f^k_\infty \mid z_n) = f^k_n$ (see Proposition 5.25 and the Remark following it). This will also be the expected payoff if player 1 decides to deviate now—by Proposition 5.17. This shows that he cannot gain by detectable deviations. How about undetected ones? He can only make those at communication stages (at payoff stages, the moves are pure). If a jointly controlled lottery is performed, he cannot influence the outcome—the alternatives have probability 1/2 each no matter what player 1 chooses (since player 2 randomizes truly according to τ). If he is in a signalling case, then $f^k_{n+1} = f^k_n$, thus f^k_{n+1} is constant, and any "signal" he uses gives him the same expectation. Therefore undetectable deviations do not help either, and σ is optimal against τ.

Consider now player 2. Since player 1 uses σ, the posterior probabilities are given by p_n. Therefore the expected average payoff of player 2 at a node z_n—if he does not deviate—is precisely $p_n \cdot E(B(\theta_\infty) \mid z_n)$, which for n large enough is close to $E(p_\infty \cdot B(\theta_\infty) \mid z_n) = E(\delta_\infty \mid z_n) = \delta_n$ (since $p_n \to p_\infty$). If he makes a detectable deviation, he will get thereafter at most $(\text{vex}\,\text{val}_2 B)(p_n) \leq \delta_n$—thus he cannot gain by doing so. The only other possible change in strategy is in a jointly controlled lottery; again, if player 1 uses σ, player 2 cannot influence the resulting probabilities. Thus τ is a best response against σ.

5.5. *σ is optimal against τ.* We will show here that σ is a best response of player 1 against the strategy τ of player 2. Moreover, the uniform condition (2.11) will be proved.

Thus, let $\epsilon > 0$; we have to find $T_0 \equiv T_0(\epsilon)$ such that for all $T > T_0$ and all σ' $E^k_{\sigma',\tau}(a^k_T) \leq a^k + \epsilon$ for all k in K (see (2.13) and Proposition 5.26).

As usual, P, P^k, E and E^k refer to (σ,τ,p), whereas P', P'^k, E' and E'^k to (σ',τ,p). If both players use (σ,τ), there are no deviations from the master plan, and ζ is defined for all possible histories (i.e., those with positive P). However, when we consider alternative strategies, it will be useful to define ζ for *all* histories (i.e., even those that are not possible under (σ,τ)); $\zeta(h_t)$ will be ζ of the part of h_t up to the first stage a (detectable) deviation occurred. Thus, we define

$$D = \sup\{t \in N : P(h_t) > 0\}. \tag{5.27}$$

D is a random variable on Ω with values in $N \cup \{\infty\}$, and is $\mathscr{H}_\infty$-measurable. For every infinite history, D is the stage of the first detectable deviation, if any; $D = \infty$ otherwise. Note that $P(h_t) > 0$ just means that the sequence of moves used by both players at stages $1, 2, \ldots, t-1$ is possible under (σ,τ); more precisely, that h_t is possible under (σ,τ) when $\kappa = k$ for some k in K.

Let $D \wedge t \equiv \min\{D, t\}$, then we define for all t in N and h_t in H_t

$$\zeta(h_t) = \zeta(h_{D\wedge t}). \tag{5.28}$$

The right-hand side was defined in §5.3; we thus extend the definition of ζ to all histories.

Next, we "translate" the G-process to the space Ω as follows:

$$\hat{g}_t(h_t) \equiv g_m(z_m), \tag{5.29}$$

where $z_m = \zeta(h_t)$ (and thus, by (5.28), we have $(m-1)! < D \wedge t \leqslant m!$). As usual, $\hat{g}_t = (\hat{f}_t, \hat{\delta}_t, \hat{p}_t)$, with $\hat{f}_t = (\hat{f}_t^k)_{k\in K}$ in R^K, $\hat{\delta}_t$ in R and $\hat{p}_t = (\hat{p}_t^k)_{k \in K}$ in Δ^K.

PROPOSITION 5.30. *Let $k \in K$. The sequence $\{\hat{f}_t^k\}_{t=1}^\infty$ is a martingale on $(\Omega, \mathscr{H}_\infty, P'^k)$ with respect to $\{\mathscr{H}_t\}_{t=1}^\infty$. Moreover, for all t in N, h_t in H_t and all i_t in I,*

$$E'^k(\hat{f}_{t+1}^k \mid h_t, i_t) = \hat{f}_t^k \qquad P'^k\text{-a.s.} \tag{5.31}$$

This proposition is a crucial assertion in our proof. (5.31) means that the strategies σ and τ have been constructed in such a way that player 1 is indifferent among his various choices of i_t at all histories h_t—and this includes both detectable and undetectable deviations from σ (see also Proposition 5.25 and the subsequent Remark).

PROOF. The measurability of $\hat{f}_t^k$ with respect to $\mathscr{H}_t$ is immediate by definition. As for the martingale property, namely $E'^k(\hat{f}_{t+1}^k \mid h_t) = \hat{f}_t$, it will follow from (5.31) (which is stronger, since it holds for all i_t in I, not only in the average).

We now prove (5.31). It is easy to see by (5.29) that $\hat{f}_{t+1}^k \neq \hat{f}_t^k$ only when $D \wedge t = m!$ and $D \wedge (t+1) = m! + 1$, and thus $D \geqslant t+1$ and $t = m!$. Let $z_m = \zeta(h_t)$; since $f_{m+1}^k \neq f_m^k$, case (ii)—a jointly controlled lottery—occurs at z_m. But i_t must be either i' or i'' (otherwise, player 1 deviated and $D = t$); in both instances, $\zeta(h_{t+1})$ is z'_{m+1} or z''_{m+1} with probability $1/2$ each by Lemma 5.14, and (5.31) reduces to

$$f_m^k(z_m) = \tfrac{1}{2} f_{m+1}^k(z'_{m+1}) + \tfrac{1}{2} f_{m+1}^k(z''_{m+1}),$$

which holds by (5.3). ∎

For each T in N, we define $\mathscr{H}_{D\wedge T}$ to be the finite field generated by all events of the form[15] $\{h_t \text{ and } D \wedge T \geqslant t\}$ for t in N, $t \leqslant T$ and h_t in H_T. This is the field of events *prior* to the first detectable deviation. Note that $D+1$ (but not D) is a *stopping time* relative to the sequence $\{\mathscr{H}_t\}_{t=1}^\infty$, and so is $(D+1) \wedge (T+1) = (D \wedge T) + 1$. The field of events *strictly before* $(D \wedge T) + 1$ is precisely $\mathscr{H}_{D\wedge T}$. It is easy to see that $\mathscr{H}_{D\wedge T} \subset \mathscr{H}_T$, and an atom in $\mathscr{H}_{D\wedge T}$, which we denote by $h_{D\wedge T}$, is of the form $h_{D\wedge T} = \{h_t \text{ and } D \wedge T = t\}$ for some t in N, $t \leqslant T$ and h_t in H_t.

[15] This is the set of infinite histories which coincide with h_t up to time t, and for which $D \wedge T$ is no less than t.

From now on, we fix a strategy σ' of player 1, an element k in K, and T in N. To shorten notation, we will write D for $D \wedge T$ and $\mathscr{H}_D$ for $\mathscr{H}_{D\wedge T}$.

Consider $E'^k(a^k_{T-1})$; we separate it into three parts: before D, at D, and after D (note that only the first one is always nonempty). Thus,

$$a^k_{T-1} = \frac{D-1}{T-1} a^k_{D-1} + \frac{1}{T-1} A^k(i_D, j_D) + \frac{1}{T-1} \sum_{t=D+1}^{T-1} A^k(i_t, j_t).$$

The middle term is at most $M/(T-1)$, hence

$$E'^k(a^k_{T-1}) \leqslant E'^k\left(\frac{D-1}{T-1} a^k_{D-1}\right) + \frac{M}{T-1} + E'^k\left(\frac{1}{T-1} \sum_{t=D+1}^{T-1} A^k(i_t, j_t)\right). \quad (5.32)$$

For the first term, we have

LEMMA 5.33.

$$E'^k\left(\frac{D-1}{T-1} a^k_{D-1}\right) \leqslant E'^k\left(\frac{D-1}{T-1} \hat{f}^k_T\right) + \frac{5\overline{M} + 2M}{n-1}.$$

PROOF. We use Corollary 5.20:

$$E'^k\left(\frac{D-1}{T-1} a^k_{D-1}\right) \leqslant E'^k\left(\frac{D-1}{T-1} f^k_m\right) + E'^k\left(\frac{D-1}{T-1} \frac{(m-1)!}{D-1} (f^k_{m-1} - f^k_m)\right)$$
$$+ 5\overline{M}E'^k\left(\frac{D-1}{T-1} \frac{1}{m}\right) \quad (5.34)$$

where f^k_m and f^k_{m-1} are evaluated at the corresponding $\zeta(h_D)$ (note that m here is a random variable, $(m-1)! < D \leqslant m!$).

By definitions (5.28) and (5.29), the first term is precisely $E'^k(((D-1)/(T-1))\hat{f}^k_T)$. The second term is separated into two parts. If $D \leqslant (n-1)!$, then $m \leqslant n-1$, hence

$$\frac{(m-1)!}{T-1} \leqslant \frac{(n-2)!}{(n-1)!} = \frac{1}{n-1},$$

giving a bound of $2M/(n-1)$. Next, we claim that

$$E'^k\left(\frac{(m-1)!}{T-1} (f^k_{m-1} - f^k_m)\chi_{\{D>(n-1)!\}}\right) = 0,$$

where $\chi_{\{\cdots\}}$ denotes the indicator function of the event $\{\cdots\}$. Let $u \equiv (n-1)!$, then $D > u$ if and only if $P(h_{u+1}) > 0$ (see the definition (5.27) of D). But player 2 does not deviate from τ, therefore we have P'^k-a.s.: $P(h_{u+1}) > 0$ if and only if $P(h_u, i_u) > 0$. Conditioning on h_u and i_u gives

$$E'^k\left(\frac{u}{T-1} \chi_{\{P(h_u,i_u)>0\}} E'^k(f^k_{n-1} - f^k_n \mid h_u, i_u)\right).$$

By (5.29), $f^k_{n-1} = \hat{f}^k_u$ and $f^k_n = \hat{f}^k_{u+1}$; recalling (5.31) shows that the whole expression is zero.

The last term in (5.34) is also separated into two: for $D \leqslant (n-2)!$,

$$\frac{D-1}{T-1} \frac{1}{m} \leqslant \frac{(n-2)!}{(n-1)!} = \frac{1}{n-1};$$

for $D > (n-2)!$, $m \geqslant n-1$ and

$$\frac{D-1}{T-1} \frac{1}{m} \leqslant \frac{1}{n-1}.$$

This gives a bound of $5\overline{M}/(n-1)$; together with $2M/(n-1)$ from the second term, the proof is completed. ∎

For the last term in (5.32), we condition on $\mathscr{H}_D$.

LEMMA 5.35.

$$E'^k\left(\frac{1}{T-1}\sum_{t=D+1}^{T-1} A^k(i_t, j_t)\,\middle|\,\mathscr{H}_D\right) < \frac{T-D-1}{T-1}\hat{f}_T^k + \frac{2M}{\sqrt{T-1}}.$$

PROOF. Given h_D, player 2 uses the punishment strategy $\bar{\tau}(y)$ starting at $t = D+1$, with $y = f_m(\zeta(h_D)) = \hat{f}_T$ (see (5.29)). The inequality is obtained from Proposition 5.17, applied to

$$\frac{1}{T-D-1}\sum_{t=D+1}^{T-1} A^k(i_t, j_t).\quad ∎$$

PROPOSITION 5.36. *For every $\epsilon > 0$ there exists $T_0 \equiv T_0(\epsilon)$ such that for all $T > T_0$ and all σ', $E'^k(a_T^k) \leq a^k + \epsilon$ for all k in K.*

PROOF. Combining the inequalities in Lemmata 5.33 and 5.35, we obtain from (5.32)

$$E'^k(a_{T-1}^k) \leq E'^k\left(\frac{T-2}{T-1}\hat{f}_T^k\right) + \frac{M}{T-1} + \frac{5\overline{M}+2M}{n-1} + \frac{2M}{\sqrt{T-1}}.$$

The first term differs from $E'^k(\hat{f}_T^k) = f_1^k = a^k$ (see Proposition 5.30 and (5.1)) by at most another $M/(T-1)$. All additional terms are independent of σ', and converge to zero as $T \to \infty$ (hence, $n \to \infty$ too). ∎

5.6. *τ is optimal against σ.* Here we prove that τ is a best response of player 2 against σ of player 1; as in §5.5, we obtain the uniform property (2.12): Given $\epsilon > 0$, we show that there is $T_0 \equiv T_0(\epsilon)$ such that for all $T > T_0$ and all τ', $E_{\sigma,\tau',p}(\beta_T) \leq \beta + \epsilon$ (recall Proposition 5.26). We will use the notations E' and P' for σ, τ', p.

In §5.5, the time of the first detectable deviation was defined (see (5.27)); also, the G-process was translated to the space of histories by (5.28) and (5.29). Thus, $\hat{\delta}_t(h_t)$ is the value of the sequence $\{\delta_m(z_m)\}_m$ just before the deviation (if any, up to stage t). We have

PROPOSITION 5.37. *The sequence $\{\hat{\delta}_t\}_{t=1}^{\infty}$ is martingale on $(\Omega, \mathscr{H}_\infty, P')$ with respect to $\{\mathscr{H}_t\}_{t=1}^{\infty}$.*

PROOF. Similar to that of Proposition 5.30. Again we actually prove a stronger assertion (but which will not be needed in the sequel), namely $E'(\hat{\delta}_{t+1} \mid h_t, j_t) = \hat{\delta}_t$, P'-a.s., for all t in N, h_t in H_t and j_t in J. The only case to check is $t = m!$ and $D > t+1$. The probabilities of z'_{m+1} and z''_{m+1} are $1/2$ each no matter what player 2 chooses at stage t; this is so by definition of σ in case (i) (signalling), and by Lemma 5.14 in case (ii) (a jointly controlled lottery; since $D > t+1$, $j_t = j'$ or j'' then). ∎

An important property of σ is that player 2 cannot *increase* the probability of reaching any node z_n in the tree (he may be able to decrease it by making detectable deviations in previous stages).

PROPOSITION 5.38. *Let $t \in N$, $(n-1)! < t \leq n!$, and let z_n be an atom of $\mathscr{G}_n$. Then $P'[\zeta(h_t) = z_n] \leq Q(z_n)$.*

PROOF. Induction on n. For $n = 1$ we clearly have equality. Let $u = (n-1)!$, then definition (5.28) gives

$$P'[\zeta(h_t) = z_n] = P'[\zeta(h_{u+1}) = z_n \text{ and } D \geq u+1]$$
$$= P'[\zeta(h_{u+1}) = z_n \mid \zeta(h_u) = z_{n-1} \text{ and } D \geq u+1]$$
$$\cdot P'[D \geq u+1 \mid \zeta(h_u) = z_{n-1}] \cdot P'[\zeta(h_u) = z_{n-1}].$$

The same argument as in the proof of Proposition 5.37 shows that the first factor is $1/2$ (since $D \geq u+1$, in both cases (i) and (ii) the probabilities do not change); the second is at most 1, and the third at most $Q(z_{n-1})$ by induction. This completes the proof, since $Q(z_n) = (1/2) \cdot Q(z_{n-1})$ by (5.3). ∎

Let τ' be a fixed strategy of player 2, and fix T in N, $(n-1)! < T \leq n!$. As in §5.5, we divide $E'(\beta_{T-1})$ into three parts, as follows (D stands for $D \wedge T$):

$$E'(\beta_{T-1}) \leq E'\left(\frac{D-1}{T-1}\beta_{D-1}\right) + \frac{M}{T-1} + E'\left(\frac{1}{T-1}\sum_{t=D+1}^{T-1} B^{\kappa}(i_t, j_t)\right). \quad (5.39)$$

In the first term we condition on $\mathscr{H}_D$, to obtain

$$E'\left(\frac{D-1}{T-1}\beta_{T-1}\right) = E'\left(\frac{D-1}{T-1}\sum_{k\in K} P'(\kappa = k \mid \mathscr{H}_D) B^k(\phi_1^{D-1})\right). \quad (5.40)$$

LEMMA 5.41. *Let h_D be an atom of $\mathscr{H}_D$, and $(m-1)! < D \leq m!$. Then $P'[\kappa = k \mid h_D] = p_m^k(\zeta(h_D))$ for all k in K.*

PROOF. Given h_D, player 1 did not detect any deviations by player 2 up to D. Therefore he played according to the master plan, and the result follows by (5.24). ∎

LEMMA 5.42. *For every $\epsilon > 0$ there exists $T_1 \equiv T_1(\epsilon)$ independent of τ' such that*

$$E'\left(\frac{D-1}{T-1}\beta_{T-1}\right) \leq E'\left(\frac{D-1}{T-1}\hat{\delta}_T\right) + \epsilon.$$

PROOF. On the space Z, we define for all n in N, $\eta_n = \max\{p_n \cdot B(\theta_n), p_n \cdot B(\theta_{n-1})\}$. Clearly η_n is $\mathscr{G}_n$-measurable, and $\eta_n \to \delta_\infty$ Q-a.s. as $n \to \infty$ (indeed: $p_n \to p_\infty$, $\theta_n \to \theta_\infty$ by (5.11), and $\delta_\infty = p_\infty \cdot B(\theta_\infty)$ by (5.7)). We also have $\delta_n \to \delta_\infty$ as $n \to \infty$, therefore $\lim_{n\to\infty} E_{(Z)}(|\eta_n - \delta_n|) = 0$ (everything is bounded by M). Thus, for every $\epsilon > 0$ there is $n_1 \equiv n_1(\epsilon)$ large enough such that $E_{(Z)}(|\eta_n - \delta_n|) < \epsilon$ for all $n > n_1$.

Using Lemma 5.41 and then Corollary 5.21 (with $q = p_m$), we obtain by (5.40)

$$E'\left(\frac{D-1}{T-1}\beta_{T-1}\right) \leq E'\left(\frac{D-1}{T-1}\eta_m\right) + 4\overline{M}E'\left(\frac{D-1}{T-1}\frac{1}{m}\right),$$

where $(m-1)! < D \leq m!$, and $\eta_m = \eta_m(\zeta(h_D))$. As in the proof of Lemma 5.33, the last term is no more than $4\overline{M}/(n-1)$, which can be made arbitrarily small for large enough T (independent of τ').

Therefore it remains to bound $E'(((D-1)/(T-1))|\eta_m - \delta_m|)$ (recall that $\hat{\delta}_T = \delta_m(\zeta(h_D))$). We separate into three parts: $m \leq n-2$, $m = n-1$, and $m = n$. The first one is bounded by $2M/(n-1)$ (since $D \leq (n-2)!$). Let z_n be an atom of $\mathscr{G}_n$, then

$$P'[\zeta(h_D) = z_n] = P'[\zeta(h_T) = z_n] \leq Q(z_n)$$

by (5.28) and Proposition 5.38. This implies that

$$E'\left(\frac{D-1}{T-1}|\eta_m - \delta_m|\chi_{\{m=n\}}\right) \leq E_{(Z)}(|\eta_n - \delta_n|).$$

Similarly, let z_{n-1} be an atom of $\mathcal{G}_{n-1}$, then

$$P'[\zeta(h_D) = z_{n-1}] = P'[\zeta(h_{n!}) = z_{n-1} \text{ and } D \leq n!] \leq Q(z_{n-1}) \quad \text{and}$$

$$E'\left(\frac{D-1}{T-1}|\eta_m - \delta_m|\chi_{\{m=n-1\}}\right) \leq E_{(Z)}(|\eta_{n-1} - \delta_{n-1}|).$$

If $n > n_1(\epsilon)$, both expectations are bounded by ϵ, which completes the proof. ∎

For the last term in (5.39), we condition on $\mathcal{H}_D$.

LEMMA 5.43.

$$E'\left(\frac{1}{T-1}\sum_{t=D+1}^{T-1} B^{\kappa}(i_t, j_t)\middle|\mathcal{H}_D\right) \leq \frac{T-D-1}{T-1}\hat{\delta}_T.$$

PROOF. By Lemma 5.41, the posteriors at h_D are given by $p_m = p_m(\zeta(h_D))$. From stage $D+1$ on, player 1 uses his punishment strategy; by Proposition 5.18, the expression we consider is thus no more than

$$\frac{T-D-1}{T-1}(\text{vex}\,\text{val}_2 B)(p_m) \leq \frac{T-D-1}{T-1}\delta_m$$

(we used Proposition 3.16(ii)). But $\hat{\delta}_T = \delta_m(\zeta(h_D))$, completing the proof. ∎

PROPOSITION 5.44. *For every $\epsilon > 0$ there exists $T_0 \equiv T_0(\epsilon)$ such that for all $T > T_0$ and all τ', $E'(\beta_T) \leq \beta + \epsilon$.*

PROOF. Similarly to Proposition 5.36, we combine (5.39), Lemmata 5.42 and 5.43, Proposition 5.37, and (5.1). ∎

We have completed the proof of the second half of our main result.

PROPOSITION 5.45. *Let $(a, \beta, p) \in G^*$. Then there exists a uniform equilibrium point (σ, τ) in $\Gamma_\infty(p)$ with payoffs (a, β).*

PROOF. Propositions 5.26, 5.36 and 5.44. ∎

6. Enforceable joint plans. Let us consider now equilibria that require *finite* sequences of communications. For every positive integer m, let $G^m = \{g \in G^*$: there exists a G-process $\{g_n\}_{n=1}^{\infty}$ starting at g such that $g_n = g_m$ for all $n \geq m\}$. Thus, G^m corresponds to those G-processes for which the limit g_∞ is reached already at stage m. Clearly, $G^1 = G$ (recall (3.12)). Therefore, the first such set to study is G^2.

The following is easily obtained: A point $g = (a, \beta, p)$ belongs to G^2 if and only if g can be expressed as a convex combination of points in G, all of which have the same a coordinate or the same p coordinate. Thus, there is a finite set S such that $g = \sum_{s\in S}\rho(s)g(s)$, with $\rho = (\rho(s))_{s\in S}$ in Δ^S, $g(s) = (a(s), \beta(s), p(s))$ in G for all s in S, and either $a(s) = a$ for all s or $p(s) = p$ for all s.

The latter case ($p(s) = p$ for all s) leads to no additional points outside G; this is due to the fact that, for a fixed p, the set of (a, β) such that (a, β, p) belongs to G is a convex set (indeed, all conditions (3.3), (3.4), (3.9)–(3.11) are invariant under convex combinations—again, when p is constant).

Therefore, the only interesting case is $a(s) = a$ for all s (and $p(s)$ not constant). This generates points in G^2 that do not necessarily belong to G, and that correspond to equilibria with one communication only[16] (signalling), followed by payoff accumula-

[16] If the G-process is standard (cf. §5.1), this would require one stage in the game; in general, this may take longer (e.g., if player 1 uses only i' and i'' as in §5.3, then at least $\log_2 l$ stages are needed, where l is the number of different values of g_2).

tion henceforth (using frequencies). Following Aumann, Maschler and Stearns (1968), this is called an *enforceable joint plan*.[17]

An interesting question is: how many different signals are needed? Since the only information player 1 has (that player 2 has not) is the value of κ, it seems reasonable that no more than $|K|$ signals should be required. Namely, the most player 1 can transmit to player 2 is just κ, which has $|K|$ possible values. However, it turns out that this is not the case, and the correct bound is $|K| + 1$ rather than $|K|$; i.e., no more than $|K| + 1$ signals are needed, and there are examples which do indeed require $|K| + 1$.

For every integer l, let $G^2(l)$ be the set of all $g = (a, \beta, p)$ in G^2 such that

$$g = \sum_{s=1}^{l} \rho(s) g(s), \qquad \sum_{s=1}^{l} \rho(s) = 1,$$

$$g(s) = (a, \beta(s), p(s)) \in G \quad \text{and} \quad \rho(s) > 0 \qquad \text{for all} \quad s = 1, 2, \ldots, l.$$

PROPOSITION 6.1. $G^2 = G^2(|K| + 1)$.

PROOF. For fixed a, the vector (β, p) lies in $R \times \Delta^K$, which is a $|K|$-dimensional Euclidean space; we now apply Caratheodory's Theorem. ∎

We will next present an example where $G^2 \neq G^2(|K|)$, showing that $|K| + 1$ is the best bound.

EXAMPLE 6.2. Let $K = \{1,2\}$, $I = \{1,2\}$, $J = \{1,2,3,4,5,6,7\}$. The two games are (player 1 chooses the row, player 2 the column):

	i \ j	1	2	3	4	5	6	7
$k = 1$	1	0, 0	0, 4	0, −5	−1, −9	−1, −3	−1, 3	−1, 6
	2	0, 0	0, 4	0, −5	−1, −9	−1, −3	−1, 3	−1, 6

	i \ j	1	2	3	4	5	6	7
$k = 2$	1	0, 0	0, −5	0, 4	−1, 6	−1, 3	−1, −3	−1, −9
	2	0, 0	0, −5	0, 4	−1, 6	−1, 3	−1, −3	−1, −9

It is easy to see that $(\text{val}_1 A)(p) = -1$ for all p in Δ^K, and $(\text{vex}\,\text{val}_2 B)(p) = (\text{val}_2 B)(p) = \max\{-9p^1 + 6p^2, -3p^1 + 3p^2, 3p^1 - 3p^2, 6p^1 - 9p^2\}$, where $p = (p^1, p^2)$. Therefore, the intersection of G with the hyperplane $a = (0,0)$ consists of exactly three points:

$$g(1) = ((0,0), 0, (\tfrac{1}{2}, \tfrac{1}{2})), \qquad g(2) = ((0,0), 1, (\tfrac{2}{3}, \tfrac{1}{3})), \qquad g(3) = ((0,0), 1, (\tfrac{1}{3}, \tfrac{2}{3})),$$

where we write as usual $g = ((a^1, a^2), \beta, (p^1, p^2))$; these three points correspond to $j = 1$, $j = 2$ and $j = 3$, respectively (i does not matter). Indeed, since $a = (0,0)$, $j = 4$, 5, 6 and 7 are not possible; individual rationality for player 2 (namely (3.4)) then implies that $j = 1$ can be used only at $p = (1/2, 1/2)$, $j = 2$ only at $p = (2/3, 1/3)$, and $j = 3$ only at $p = (1/3, 2/3)$.

Therefore G^2 will contain the convex hull of $g(1)$, $g(2)$ and $g(3)$; however, no interior point of this triangle can be expressed using only two of its vertices.

[17] They only define "joint plans"—and then find conditions under which these can be "enforced" by equilibria. As Sorin (1983) pointed out, one of their conditions should be slightly weakened—and then it corresponds to our characterization of G^2.

It is easily seen in this example that an additional condition may reduce the number of signals to $2 = |K|$. In general, we have

PROPOSITION 6.3. *Let $(a, \beta, p) \in G^2$. Then there exists β' in R such that $\beta' > \beta$ and $(a, \beta', p) \in G^2(|K|)$.*

PROOF. By Proposition 6.3,

$$g = (a, \beta, p) = \sum_{s=1}^{l} \rho(s) g(s), \qquad \sum_{s=1}^{l} \rho(s) = 1,$$

$g(s) = (a, \beta(s), p(s)) \in G$ and $\rho(s) > 0$ for all $s = 1, 2, \ldots, l$, where $l < |K| + 1$. Assume $l = |K| + 1$, and consider the l vectors $\{(p(s), 1)\}_{s=1}^{l}$ in $\Delta^K \times R$. They must be linearly dependent; let $\{\pi(s)\}_{s=1}^{l}$ be not all zero and such that $\sum_{s=1}^{l} \pi(s) p(s) = 0$ and $\sum_{s=1}^{l} \pi(s) = 0$. Without loss of generality, we assume that $\sum_{s=1}^{l} \pi(s) \beta(s) > 0$ (otherwise, replace all $\pi(s)$ by $-\pi(s)$). Let $\eta = \min\{-\rho(s)/\pi(s) : \pi(s) < 0\}$, and put $\rho'(s) = \rho(s) + \eta\pi(s)$. Then $\rho'(s) > 0$ for all $s = 1, 2, \ldots, l$ and at least one $\rho'(s)$ is zero; moreover, $\sum_{s=1}^{l} \rho'(s) = 1$, $\sum_{s=1}^{l} \rho'(s) p(s) = p$ and $\beta' \equiv \sum_{s=1}^{l} \rho'(s) \beta(s) > \beta$. ∎

The following is now immediate.

COROLLARY 6.5. *Let (a, β) be the payoffs of a Pareto optimal*[18] *enforceable joint plan equilibrium in $\Gamma_\infty(p)$. Then no more than $|K|$ signals are needed; namely, $(a, \beta, p) \in G^2(|K|)$.*

What about G^m for larger values of m? It is easy to see that each G^m is obtained from the previous G^{m-1} by taking convex combinations—with either a fixed (when m is even) or p fixed (when m is odd). New points are usually obtained; Aumann, Maschler and Stearns (1968) provide examples where $G^3 \neq G^2$ and $G^4 \neq G^3$. It is probably not difficult to use the same ideas in order to generate examples where $G^m \neq G^{m-1}$ for arbitrary m.

An open question still remains. Is G^* the union of all the G^m? If one ignores the game structure, and just considers the notions of bi-convexification (G^m) and bi-martingale (G^*), the answer is negative: G^* may contain points that do not belong to any G^m (and it is not just a matter of closure either—G^* may be a very different set). For details on these problems, the reader is referred to Aumann and Hart (1983).

7. Example. In this section we will analyze an example and find its equilibria.

EXAMPLE 7.1.[19] $|I| = |J| = |K| = 2$. The games are as follows (player 1 chooses the row and player 2 the column):

$k = 1$

$i \backslash j$	1	2
1	1,0	−1,1
2	1,0	0,1

$k = 2$

$i \backslash j$	1	2
1	1,1	0,0
2	1,1	−1,0

It is straightforward to obtain $(\text{val}_1 A)(p) = \max\{-p^1, -p^2\}$ and $(\text{vex}\,\text{val}_2 B)(p) = (\text{val}_2 B)(p) = \max\{p^1, p^2\}$ for all $p = (p^1, p^2)$ in Δ^K. Therefore $a = (a^1, a^2)$ is individually rational for player 1 (i.e., satisfies (3.3)) if and only if $a^1, a^2 > 0$.

[18] I.e., such that there is no other enforceable joint plan equilibrium in $\Gamma_\infty(p)$ with payoffs (a', β') satisfying $(a', \beta') > (a, \beta)$ and $(a', \beta') \neq (a, \beta)$.

[19] Suggested by S. Zamir.

We first find the set G. Let $g = (a, \beta, p) \in G$ and let $(c, d) \in F$ correspond to it (recall (3.9)–(3.11)). We distinguish five cases.

(i) $p^1 = 0$. Individual rationality for player 2 (i.e., (3.4)) implies that $d^2 = \beta \geq 1$, therefore only $j = 1$ may be used, in which case $a = (1, 1)$ and $g = ((1,1), 1, (1,0)) \in G$.

(ii) $p^1 = 1$. By (3.4) we have $d^1 = \beta \geq 1$, therefore $j = 2$. But $c^1 = a^1 \geq 0$ by (3.10) and (3.3), hence $i = 2$ and $g = ((0, a^2), 1, (1,0)) \in G$ for all $a^2 \geq 0$.

(iii) $0 < p^1 < \frac{1}{2}$. Here $\beta \geq p^2$ by (3.4), hence $j = 1$ (for $j = 2$, $\beta = p^1 < p^2$), and $g = ((1,1), \beta, p) \in G$ with $\beta = p^2$.

(iv) $\frac{1}{2} < p^1 < 1$. Here $j = 2$ by (3.4), but then $a^1, a^2 < 0$, contradicting (3.3). Hence no points in G in this case.

(v) $p^1 = \frac{1}{2}$. Condition (3.4) imposes no restriction, hence $g = (a, \frac{1}{2}, (\frac{1}{2}, \frac{1}{2})) \in G$ for all $a \in \mathrm{co}\{(1,1), (-1,0), (0,-1)\} \cap \{a \geq 0\} = \mathrm{co}\{(1,1), (\frac{1}{2}, 0), (0, \frac{1}{2}), (0,0)\} \equiv V$.

In conclusion, G consists of the following points (a, β, p):

$$\begin{array}{lll} a = (1,1), & \beta = p^2, & 0 \leq p^1 < \frac{1}{2}, \\ a \in V, & \beta = \frac{1}{2}, & p^1 = \frac{1}{2}, \\ a^1 = 0, a^2 \geq 0, & \beta = 1, & p^1 = 1. \end{array}$$

The set G^2 (see the previous section) is now easily obtained (convex combinations of points in G with the same a). It includes G together with all the points (a, β, p) satisfying

$$a^1 = 0, \qquad 0 \leq a^2 \leq \tfrac{1}{2}, \qquad \beta = p^1, \qquad \tfrac{1}{2} < p^1 < 1.$$

Indeed, each such point is a convex combination of $(a, \frac{1}{2}, (\frac{1}{2}, \frac{1}{2}))$ and $(a, 1, (1,0))$ when a satisfies both $a \in V$ and $a^1 = 0$, $a^2 \geq 0$.

Using the methods developed in Aumann and Hart (1983), we can next show that G^2 is actually G^*. Consider the following functions on $R_M^K \times \Delta^K$ (the β coordinate is ignored; $M = 1$):

$$f_1(a, p) = \left[\tfrac{1}{2} - p^1\right]_+ d(a, (1,1)), \qquad f_2(a, p) = \left[1 - p^1\right]_+ d(a, V),$$

$$f_3(a, p) = \left[p^1 - \tfrac{1}{2}\right]_+ a^1,$$

where $[x]_+ \equiv \max\{x, 0\}$ for real x and d is the (Euclidean) distance. All these functions are easily seen to be bi-convex, bounded, continuous, nonnegative, and vanishing on G. They must therefore vanish on G^* too (cf. Proposition 4.8 in Aumann and Hart 1983). Let $(a, \beta, p) \in G^*$; if $p^1 < \frac{1}{2}$ then $a = (1,1)$ since $f_1(a, p) = 0$; if $p^1 < 1$ then $a \in V$ since $f_2(a, p) = 0$; and finally if $p^1 > \frac{1}{2}$ then $a^1 = 0$ since $f_3(a, p) = 0$. All these conditions however imply $(a, \beta, p) \in G^2$, therefore $G^* = G^2$.

To summarize: all equilibria of $\Gamma_\infty(p)$ are equivalent to nonrevealing equilibria when $p^1 < \frac{1}{2}$, and to enforceable joint plans when $p^1 > \frac{1}{2}$.

References

Aumann, R. J. (1964). Mixed and Behaviour Strategies in Infinite Extensive Games. in: *Advances in Game Theory*. Ann. Math. Studies 52, M. Dresher et al. (eds.), Princeton University Press, Princeton, N.J., 627–650.

———. (1976). Agreeing to Disagree. *Ann. Statist*. 4 1236–1239.

———. (1981). Repetition as a Paradigm for Cooperation in Games of Incomplete Information. mimeo.

——— and Hart, S. (1983). Bi-Convexity and Bi-Martingales. RM-56, Center for Research In Mathematical Economics and Game Theory, The Hebrew University, Jerusalem. mimeo.

——— and Maschler, M. (1966). Game Theoretic Aspects of Gradual Disarmament. *Mathematica* V.1–V.55.

———, ——— and Stearns, R. E. (1968). Repeated Games of Incomplete Information: An Approach to the Non-Zero-Sum Case. *Mathematica* 117–216.

——— and Shapley, L. S. (1976). Long Term Competition—A Game Theoretic Analysis. mimeo.

Blackwell, D. (1956). An Analogue for the Minimax Theorem for Vector Payoffs. *Pacific J. Math.* **6** 1–8.

Dunford, N. and Schwartz, J. T. (1958). *Linear Operators*. Vol. I, Interscience, New York.

Harsanyi, J. C. (1967–1968). Games with Incomplete Information Played by Bayesian Players. Part I, II, III. Management Sci. **14** 159–182, 320–334, 486–502.

Hart, S. (1979). Lecture Notes on Special Topics in Game Theory. IMSSS-Economics, Stanford University. mimeo.

Hildenbrand, W. (1974). *Core and Equilibria in a Large Economy*. Princeton University Press, Princeton, N.J.

Kuhn, H. W. (1953), Extensive Games and the Problem of Information. in *Contributions to the Theory of Games II*. Ann. Math. Studies **28**, H. W. Kuhn and A. W. Tucker (Eds.), Princeton University Press, Princeton, N.J., 193–216.

Mathematica (1966–1968), Reports to the U.S. Arms Control and Disarmament Agency, prepared by *Mathematica, Inc*., Princeton, N.J.: ST-80 (1966), ST-116 (1967), ST-140 (1968).

Mertens, J.-F. and Zamir, S. (1980). Repeated Games. mimeo.

Meyer, P. A. (1973). Limites Médiales d'après Mokobodzki. in *Sém. de Probabilités VIII*, Université de Strasbourg (1971–1972), Springer-Verlag, 198–204.

Rubinstein, A. (1977). Equilibrium in Supergames. RM-25, Center for Research in Mathematical Economics and Game Theory, The Hebrew University, Jerusalem. mimeo.

Sorin, S. (1980). An Introduction to Two-Person-Zero-Sum Repeated Games with Incomplete Information. IMSSS—Economics TR-312, Stanford University. mimeo.

———. (1983). Some Results on the Existence of Nash Equilibria for Non-Zero-Sum Games with Incomplete Information. *Internat. J. Game Theory* **12** 193–205.

DEPARTMENT OF STATISTICS, SCHOOL OF MATHEMATICAL SCIENCES, TEL-AVIV UNIVERSITY, 69978 TEL-AVIV, ISRAEL.