BAYESIAN LEARNING
IN NORMAL FORM GAMES

by

James S. Jordan

Discussion Paper No. 257, July 1990

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minn 55455

Bayesian Learning

in

Normal Form Games

by

J.S. Jordan*

July 1990

Abstract

This paper studies the asymptotic behavior of Bayesian learning processes for general finite-player, finite-strategy normal form games. Initially, each player is presumed to know his own payoff function but not the payoff functions of the other players. Strategies are initially determined as a Bayesian Nash equilibrium of the incomplete information game in which each player's private characteristic is his payoff function. These strategies are then observed by all players, causing a revision of beliefs. The new beliefs determine a new Bayesian Nash equilibrium and so on. Assuming that the common prior distribution of payoff functions satisfies independence across players, it is proved that the conditional distributions on strategies converge to the set of Nash equilibria with probability one. Under a further assumption that the prior distributions are sufficiently uniform, convergence to the set of Nash equilibria is proved for every profile of payoff functions, that is, for every normal form game.

1. Introduction

In a Nash equilibrium of a normal form game, each player must correctly anticipate the strategies played by the others. In the case of complete information, each player can deduce the others' strategies by computing the equilibrium, provided that some selection convention is adopted when there are multiple equilibria. However, in most strategic situations of economic interest, individual payoffs are private information, so deductive reasoning cannot produce the knowledge of equilibrium strategies. This leads to the question of whether players can learn the equilibrium from experience, that is, from repeated plays of the game.

There are considerable literatures on iterative learning procedures in individual decision models (e.g., Berry and Fristedt (1985), Easely and Kiefer (1986), McLennan (1987), and references therein) and rational expectations equilibrium models (e.g., Blume, Bray, and Easely (1982), Feldman (1987), Marcet and Sargent (1989) and references therein). By comparison, very little is known about learning in games, although recent papers by Fudenberg and Kreps (1988), Kalai and Lehrer (1989), and Milgrom and Roberts (1989) indicate increasing interest in this subject. Fudenberg and Kreps study learning processes for iterated extensive form games in which players use a learning process somewhat similar to fictitious play (e.g. Robinson (1951)) except that players occasionally experiment by choosing strategies at random. They show that if the learning process converges, the limit must be a Nash equilibrium of the one-shot extensive form game. Milgrom and Roberts show that for iterated normal form games a large class of strategy adjustment processes result in strategy sequences which are eventually contained in the set of serially undominated strategies. The Milgrom and Roberts result applies to the Cournot response process, fictitious play, and the non-experimental strategies generated by an interesting variation of a Fudenberg-Kreps process. Kalai and Lehrer (1990) study two-player repeated games in which the players use Bayesian inference to learn each other's strategies. They show that if each player's prior distribution gives positive probability to the true strategy of the

1

other player, then the two strategies converge to a Nash equilibrium of the repeated game.

The present paper studies a class of Bayesian learning processes for iterated normal form games with a finite number of players and a finite number of pure strategies. The main result is that if the common prior beliefs satisfy a certain uniformity condition then for every normal form game, the learning process converges to a subset of the set of one-shot Nash equilibria. As far as I am aware, this is the first general convergence result for normal form games.

We assume that each player knows his own payoff function but not the payoff functions of the other players. This uncertainty is summarized by a common prior distribution over payoff functions which satisfies independence across players. At the initial stage, strategies are selected according to a Bayesian Nash equilibrium determined by the prior distribution over payoff functions. That is, each possibly payoff function π_p for player p is associated with a strategy s_p. This association, together with the prior distribution μ_p over π_p, determines a probability distribution over s_p. In a Bayesian Nash equilibrium, the strategy s_p associated with π_p maximizes the expected payoff determined by π_p and the anticipated probability distribution over s_q for each $q \neq p$. After an initial strategy s_q is observed, player p knows that π_q lies in the set of payoff functions associated with s_q, and updates the prior μ_q accordingly. The updated beliefs determine a new Bayesian Nash equilibrium, and so on. At each step, players choose their strategies myopically, seeking to maximize their current expected payoff without regard to future iterations. The resulting stochastic process of strategy n-tuples is termed a *Bayesian strategy process* (BSP). In particular, a BSP describes a joint probability distribution of games $(\pi_p)_p$ and infinite sequences of strategy n-tuples $s_1, s_2, \ldots$. The focus of this paper is the relation between the Nash equilibria of the game $(\pi_p)_p$ and the sequence of probability distributions over strategies in period $t + 1$ conditional on the observed histories $s_1, \ldots, s_t$. These conditional probabilities represent each player p's beliefs about the strategies to be chosen in period $t + 1$ by the other players, so, together with π_p,

they determine player p's strategy in period $t + 1$. Thus player p's strategy in period $t + 1$ is chosen to maximize the expectation of π_p against the conditional probability distribution over other players' strategies. Hence the desired result is the convergence of the conditional probability distributions to the set of Nash equilibria of the game $(\pi_p)_p$.

This paper provides two such convergence theorems. Theorem 3.2, below, asserts that convergence occurs with probability one under any Bayesian strategy process. However, this leaves open the possibility that convergence fails to occur for some games in a set of prior probability zero. Thus Theorem 3.2 does not justify the interpretation of a BSP as a learning process for general normal form games. General convergence requires a further assumption on the prior beliefs about payoff functions. Since the set of strategy n-tuples is finite, payoff functions lie in a finite-dimensional Euclidean space, and we can rescale payoff functions to lie in the unit ball without loss of generality. Since the unit ball is compact, the uniform distribution, that is, normalized Lebesgue measure, is well-defined. Theorem 3.8, more transparently stated as Corollary 3.10, states that if the prior beliefs are sufficiently uniform then convergence occurs for every game $(\pi_p)_p$.

Bayesian strategy processes are defined formally in Section 2 below. Section 3 is devoted to the convergence results, and some concluding remarks are given in Section 4. Example 2.7 in Section 2 describes a BSP for the two-player, two-strategy case with uniform priors. This example shows that for certain two-by-two games with three Nash equilibria, a BSP can generate three subsequences of conditional probabilities converging, respectively, to each of the three Nash equilibria. Thus Bayesian learning does not appear to resolve the multiplicity of Nash equilibria in normal form games. This example also serves to emphasize the distinction between the learning mechanism represented by a BSP and mechanisms, such as the *tracing procedure* (Harsanyi (1975)), which players can use to select a particular Nash equilibrium strategy if all players know all of the payoff functions $(\pi_p)_p$. A BSP has the advantage that players are not required to know each others' payoff functions, and the disadvantage that multiple Nash equilibria can result in multiple cluster

3

points of the expectations sequence.

2. Bayesian Strategy Processes

2.1 Definitions: There are n players, $n \geq 2$, indexed by the subscript p. For each $1 \leq p \leq n$, let S_p be a finite set, and let $S = \Pi_{p=1}^n S_p$. For each p, let $S_{-p} = \Pi_{q \neq p} S_q$, with generic element $s_{-p} = (s_1, \ldots, s_{p-1}, s_{p+1}, \ldots, s_n)$. Each player p is characterized by a payoff function $\pi_p : S \to R$. The space of possible payoff functions for each player is the unit ball in R^S, that is, $B = \{\pi \in R^S : (\Sigma_{s \in S} \pi(s)^2)^{1/2} \leq 1\}$. Given the strategy sets S_p, a normal form game is completely specified by an n-tuple of payoff functions $(\pi_p)_p \in B^n$, so let $G = B^n$ denote the space of all games, with generic element $(\pi_p)_p$.

A probability distribution σ on S is a *product distribution* if $\sigma = \sigma_1 \times \cdots \times \sigma_n$, where each σ_p is the marginal distribution on S_p. That is, $\sigma(s) = \sigma_1(s_1) \cdots \sigma_n(s_n)$ for each $s \in S$. A *Nash equilibrium* for a game $(\pi_p)_p \in G$ is a product distribution σ^* on S such that for each p, σ_p^* maximizes $\Sigma_{s_p} \sigma_p(s_p) \Sigma_{s_{-p}} \pi_p(s_p, s_{-1}) \sigma_{-p}^*(s_{-p})$ over the set of probability measures σ_p on S_p, where $\pi_p(s_p, s_{-p}) = \pi_p(s_1, \ldots, s_{p-1}, s_p, s_{p+1}, \ldots, s_n)$, and σ_{-p}^* is the distribution $\Pi_{q \neq p} \sigma_q^*$ on S_{-p}. For each $(\pi_p)_p \in G$, let $N((\pi_p)_p)$ denote the set of Nash equilibria.

2.2 Remarks: It is clear that the normalization of payoff functions to the unit ball in R^S involves no loss of generality. Payoff functions could be normalized further to the unit sphere, but the convexity of the unit ball is more convenient for the proof of Theorem 3.8 below.

In order to define the process which generates successive strategies and expectations, we will first introduce some notation for conditional probabilities. Fortunately, the usual measure theoretic subtleties involved with conditional probability are avoided here because

of the finiteness of the strategy sets. Thus, any finite sequence of strategy n-tuples either has positive probability or lies outside the support of the relevant distribution.

2.3 Definitions: Suppose that X and Y are metric spaces and Z is the product space, $Z = X \times Y$. Let φ be a Borel probability measure on Z. The support of φ, written "supp φ", is defined to be the smallest closed subset $C \subset Z$ with $\varphi(C) = 1$. The support of the marginal probability distribution on X is the smallest closed subset $C \subset X$ such that $\varphi(C \times Y) = 1$. If $x \in X$, the statement "$x \in$ supp φ" will be understood to mean that x is in the support of the marginal distribution on X. If E is a Borel subset of X, $\varphi(E \times Y)$ will often be written simply $\varphi(E)$, and if $x \in X$, $\varphi(x)$ denotes $\varphi(\{x\})$, that is, $\varphi(\{x\} \times Y)$.

For each $1 \le t < \infty$, let $S^t = \Pi_{\tau=1}^t S$, with generic element $s^t = (s_1, \ldots, s_t)$, and let $S^\infty = \Pi_{\tau=1}^\infty S$. We topologize each S^t and S^∞ as products of the discrete space S. Let G have the Euclidean topology, and let φ be a Borel probability measure on $G \times S^\infty$. For each p, t, each $s^t \in$ supp φ, and each $s_p \in S_p$, let $\varphi_p(s_p|s^t)$ be the probability that $s_{pt+1} = s_p$ conditional on s^t, that is, $\varphi_p(s_p|s^t) = \varphi(s^t, s_p)/\varphi(s^t)$. For each $s_{-p} \in S_{-p}$, let $\varphi_{-p}(s_{-p}|s^t)$ be the probability that $s_{-pt+1} = s_{-p}$ conditional on s^t, that is, $\varphi_{-p}(s_{-p}|s^t) = \varphi(s^t, s_{-p})/\varphi(s^t)$. If $t = 0$, we adopt the convention that $(s^t, s_p) = s_p$, $(s^t, s_{-p}) = s_{-p}$, and $\varphi_p(s_p|s^t)$ (resp. $\varphi_{-p}(s_{-p}|s^t)$) represents the unconditional probability that $s_{p1} = s_p$ (resp. $s_{-p1} = s_{-p}$). Finally, for each $t \ge 1$, and each $s^t \in$ supp φ, $\varphi(s|s^t)$ denotes the probability that $s_{t+1} = s$ conditional on s^t, with the analogous convention for $t = 0$.

2.4 Bayesian Strategy Processes: For each p, let μ_p be a Borel probability distribution on B, and let μ denote the product measure. $\mu = \mu_1 \times \cdots \times \mu_n$ on G. A Borel probability measure φ on $G \times S^\infty$ is a *Bayesian strategy process* (BSP) for μ if

i) the marginal distribution on G agrees with μ;

ii) for each p, each $t \geq 0$, and each $(\pi_p; s^t, s_p) \in$ supp φ, s_p maximizes

$$\Sigma_{s_{-p}} \pi_p(\,\cdot\,, s_{-p})\varphi_{-p}(s_{-p}|s^t) \text{ on } S_p; \text{ and}$$

iii) for each $s^t \in$ supp φ, $\varphi(\,\cdot\,|s^t)$ is a product distribution, that is,

$$\varphi(\,\cdot\,|s^t) = \varphi_1(\,\cdot\,|s^t) \times \cdots \times \varphi_p(\,\cdot\,|s^t) \times \cdots \times \varphi_n(\,\cdot\,|s^t).$$

2.5 Remarks: The essential property of a BSP is the maximization condition 2.4 *(ii)*, which states that in period $t+1$, s_p is chosen to maximize player p's expected payoff conditioned on the observed strategy n-tuples $s_1, \ldots, s_t$. Condition 2.4 *(ii)* does not completely rule out correlated strategies because of the possibility that, if different players each have several maximal strategies, they could correlate their choices among these strategies and still satisfy 2.4 *(ii)*. Condition 2.4 *(iii)* rules out this possibility by requiring the strategies of different players to be conditionally independent each period. Condition 2.4 *(iii)* does not quite ensure the conditional independence of s_p and π_q for $q \neq p$, but we will not need a stronger independence condition. We will observe in Section 3.7, below, that if each μ_p is absolutely continuous with respect to Lebesgue measure on B, so that multiple maxima occur only for a null set of payoff functions, then 2.4 *(iii)* is implied by 2.4 *(i* and *ii)*.

Although the successive updating of the prior probability distribution μ is not explicitly specified in the above definition, it can be used to construct a BSP as a sequence of Bayesian Nash equilibria. Given μ, with the players' payoff functions interpreted as their characteristics, the assumption that $\mu = \mu_1 \times \cdots \times \mu_n$ immediately implies the hypothesis of the Existence Theorem of Milgrom and Weber (1985, p. 626). Hence, for each p, there is a Borel measurable function $f_p : B \times S_p \to R$ such that for each $\pi_p \in B$, $f_p(\pi_p, \cdot)$ is a probability distribution on S_p which maximizes

$$\Sigma_{s_p} \sigma_p(s_p) \int_{B^{n-1}} \{\Sigma_{s_{-p}} \pi_p(s_p, s_{-p}) f_{-p}(\pi_{-p}, s_{-p})\} \, d\mu_{-p}$$

on the set of all probability distributions σ_p on S_p, where

6

$f_{-p}(\pi_{-p}, \ s_{-p}) = \Pi_{q \neq p} f_q(\pi_q, \ s_q)$ and $\mu_{-p} = \Pi_{q \neq p} \mu_q$. Then μ and the functions f_p determine a distribution φ^1, on $G \times S$, so for each $s \in \text{supp } \varphi^1$, we have a conditional distribution $\mu^1(s)$ on G. By the construction of φ^1, $\mu^1(s) = \mu_1^1(s) \times \cdots \times \mu_n^1(s)$, so the hypothesis of the Milgrom and Weber Existence Theorem is again satisfied. Hence, there is a Bayesian Nash equilibrium for the updated prior distribution $\mu^1(s)$, for every $s \in \text{supp } \varphi^1$. Proceeding iteratively in this fashion, we can obtain for every $t < \infty$ a distribution φ^t on $G \times S^t$ which satisfies conditions 2.4 (i - iii) by construction. By the Kolmogoroff Extension Theorem, these distributions define a distribution φ on $G \times S^\infty$ which is a BSP for μ. Thus, we can state the following Proposition as a direct implication of the Milgrom and Weber Existence Theorem.[†]

2.6 Proposition: For each p, let μ_p be a Borel probability distribution on B, and let $\mu = \mu_1 \times \cdots \times \mu_n$. Then there exists a BSP for μ.

2.7 Example: The most transparent example is the 2×2 case with uniform priors. Let player 1 be the row player, let player 2 be the column player, and let $S_1 = \{T, B\}$, $S_2 = \{L, R\}$. Then each player p's payoff function is a 2×2 matrix with the entries $\pi_p(T, L)$, $\pi_p(T, R)$, $\pi_p(B, L)$, $\pi_p(B, R)$. However, it will be convenient to employ the following normalization which reduces each player p's space of possible payoff functions to the unit circle in R^2. To motivate this normalization, suppose that player 1 anticipates that player 2 will play L with probability $\varphi_2(L)$. Then player 1's optimal strategy is T, B, or both, as the quantity

$$\varphi_2(L) \left[\pi_1(T, \ L) - \pi_1(B, \ L) \right] + (1 - \varphi_2(L))[\pi_1(T, \ R) - \pi_1(B, \ R)]$$

is positive, negative, or zero; respectively. Therefore we can subtract the second row of player 1's payoff matrix from each row, so that the top row is now $(\pi_1(T, \ L) - \pi_1(B, \ L), \ \pi_1(T, \ R) - \pi_1(B, \ R))$ and the bottom row is $(0, \ 0)$. Applying the same normalization to the columns of player 2's payoff matrix reduces the payoff bimatrix to

the form

	L	R
T	$a,\ \alpha$	$b,\ 0$
B	$0,\ \beta$	$0,\ 0$

.

If we ignore the measure zero possibility that $a = b = 0$, that is, $\pi_1(T, L) = \pi_1(T, R)$ and $\pi_1(B, L) = \pi_1(B, R)$, then we can further normalize (a, b) to the unit circle without affecting player 1's response to any mixed strategy played by player 2. Thus we can assume that $a^2 + b^2 = 1$, and for Player 2, that $\alpha^2 + \beta^2 = 1$. Under this normalization, each player's payoff function is a point on the unit circle, and each 2×2 game is a point on the torus (we will continue to exclude the degenerate cases $a = b = 0$ and $\alpha = \beta = 0$).

One immediate benefit of this normalization is a simple graphical representation of the Nash equilibrium correspondence for 2×2 games. Figure 1 pictures the Nash equilibrium correspondence when points on the unit circle are measured in radians, with $(1, 0) = 0 = 2\pi$ and $(0, 1) = \pi/2$. Thus, the game $(-\pi/4, -\pi/4)$ at the lower left corner of Figure 1 is the bimatrix game

	L	R
T	$1/\sqrt{2},\ 1/\sqrt{2}$	$-1/\sqrt{2},\ 0$
B	$0,\ -1/\sqrt{2}$	$0,\ 0$

.

Since $-\pi/4 = 7\pi/4$, the two sides of the figure are identified, and the top and bottom of the figure are identified, so Figure 1 is the usual square depiction of the torus (e.g., Guillemin and Pollak (1974), Figure 1-10, p. 17). In the white areas, the Nash equilibrium is a unique pure strategy equilibrium, so each white area is labelled with the respective equilibrium pure strategies. The gray areas are intersections of white areas. Therefore, each game in the gray square $[\pi/2, \pi] \times [\pi/2, \pi]$, in the center of Figure 1, for example, has the two pure strategy equilibria (T, R) and (B, L). Games in the interior, $(\pi/2, \pi) \times (\pi/2, \pi)$, also have a single mixed strategy equilibrium. These games behave as "battle of the sexes" games. Note that because of the identifications along the sides and the top

8

and bottom of Figure 1, the four small gray squares in the corners adjoin to form the square $[3\pi/2, 0] \times [3\pi/2, 0]$ on the torus. Each game in the interior of a black area has a mixed strategy equilibrium as its unique Nash equilibrium. The games on the boundaries of regions are "singular games" which typically have a continuum of Nash equilibria.
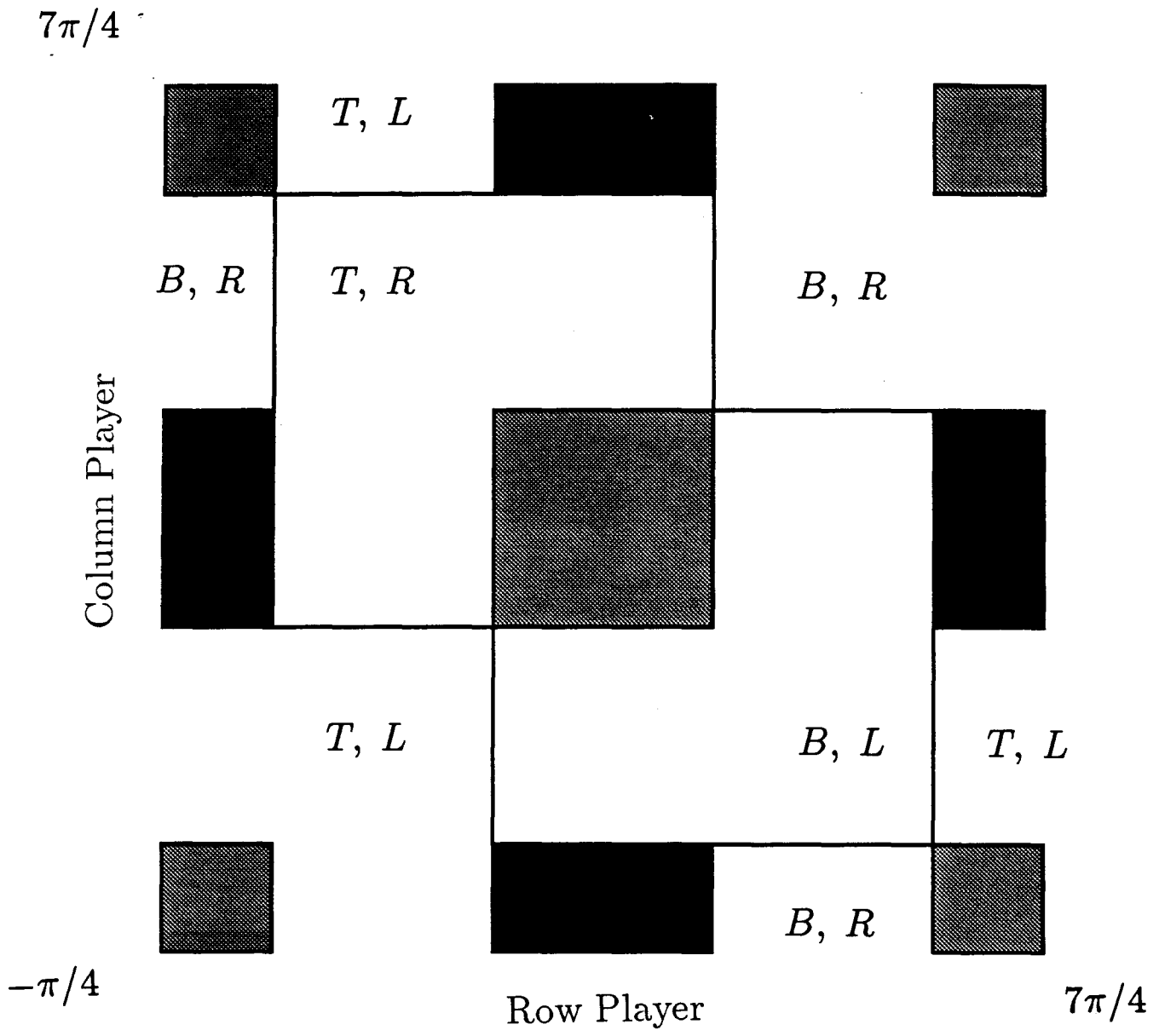
FIGURE 1

The first element of a Bayesian strategy process is a prior probability distribution over each players' space of possible payoff functions, in this case, over the unit circle for each player. A natural choice is the uniform distribution for each player. For this prior distribution, we will study a Bayesian strategy process along the two-period history $((T, R), (B, L))$. More precisely, we will compute the first period expectations $\varphi(\cdot)$ and second period expectations $\varphi(\cdot \mid T, L)$, which are uniquely determined, and show that there are three possible choices for the third period expectations $\varphi(\cdot \mid (T, L), (B, R))$. First, let φ_2 denote the initial probability distribution on $S_2 = \{L, R\}$ facing player 1. That is, $\varphi_2(L)$ is the probability that player 2 will play L in the first period. Then player 1 will play T in period 1 if player 1's payoff function, represented by (a, b) satisfies $a\varphi_2(L) + b(1 - \varphi_2(L)) > 0$. That is, the set of player 1 "types" that play T in period 1 is the set $T_1 = \{(a, b) : a\varphi_2(L) + b(1 - \varphi_2(L)) > 0\}$. The set $\{(a, b) : a\varphi_2(L) + b(1 - \varphi_2(L)) = 0\}$ has prior probability zero and thus can be ignored. Hence, for any expectation $\varphi_2(L)$, the probability that player 1 will play T, which is simply the prior probability of the set T_1, equals 1/2. Since this reasoning applies to both players symmetrically, the unique first period Bayesian Nash equilibrium is $\varphi_1(T) = \varphi_1(B) = 1/2$ and $\varphi_2(L) = \varphi_2(R) = 1/2$. Figure 2 illustrates this reasoning for player 1. The circle is the space of payoff functions for player 1, and the unit simplex represents the possible first period expectations $(\varphi_2(L), \varphi_2(R))$. The semi-circle above the line perpendicular to the expectation vector is the set T_1 of player 1 "types" that play T in period 1.
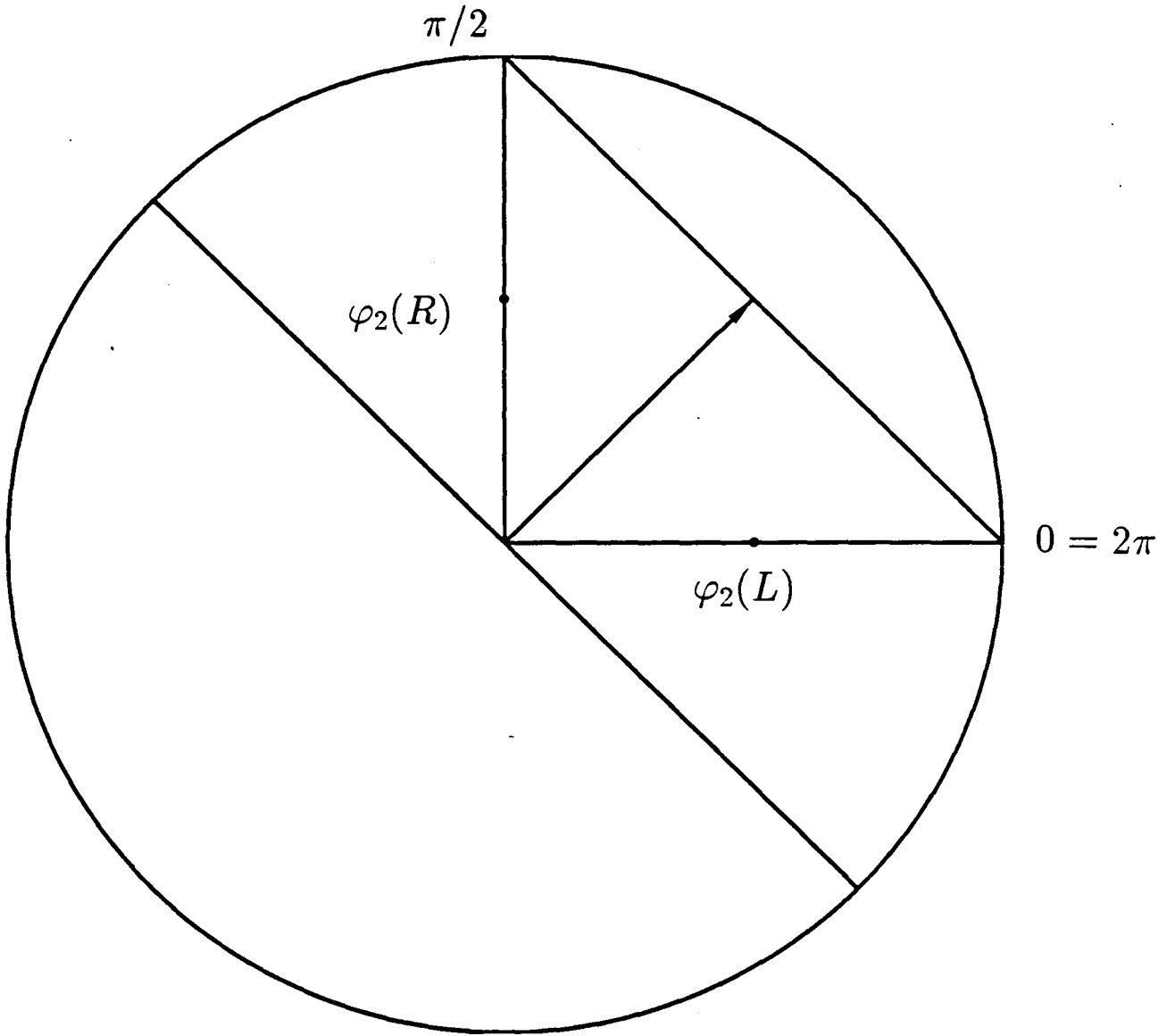
FIGURE 2

Now suppose that the strategies (T, R) are played in period 1. This reveals that player 1's payoff function lies on the upper semi-circle depicted in Figure 3, and that player 2's payoff function lies on the lower semi-circle depicted in Figure 4. To solve for the Bayesian Nash equilibrium conditional expectations $\varphi(\cdot \mid (T, R))$, let $x = \varphi_1(T \mid (T, R))$, the probability that player 1 plays T in period 2, and let $y = \varphi_2(L \mid (T, R))$. Given y, x is simply $\mu_1(\{(a, b) \in \mathcal{T}_1 : ay + b(1 - y) > 0\})/\mu_1(\mathcal{T}_1)$, where μ_1 denotes the uniform distribution on the unit circle. Thus x is simply the relative arc length given by the formula

$$(1) \qquad x = (\pi - |\theta(y)|)/\pi,$$

where $\theta(y)$ is the angle between the expectations vectors $(1/2, 1/2)$ and $(y, 1 - y)$, measured in radians. This is depicted in Figure 3. Figure 4 is the analogous diagram for player 2, and the analogous formula for y as a function of x is

$$(2) \qquad 1 - y = (\pi - |\theta(x)|)/\pi.$$

The equations (1) and (2) have a unique solution:

$$(3) \qquad x^* \approx 0.82, \quad y^* \approx 0.18,$$

so the unique second period Bayesian Nash equilibrium expectations are $\varphi_1(T \mid (T, R)) = x^*$ and $\varphi_2(L \mid (T, R)) = y^*$.
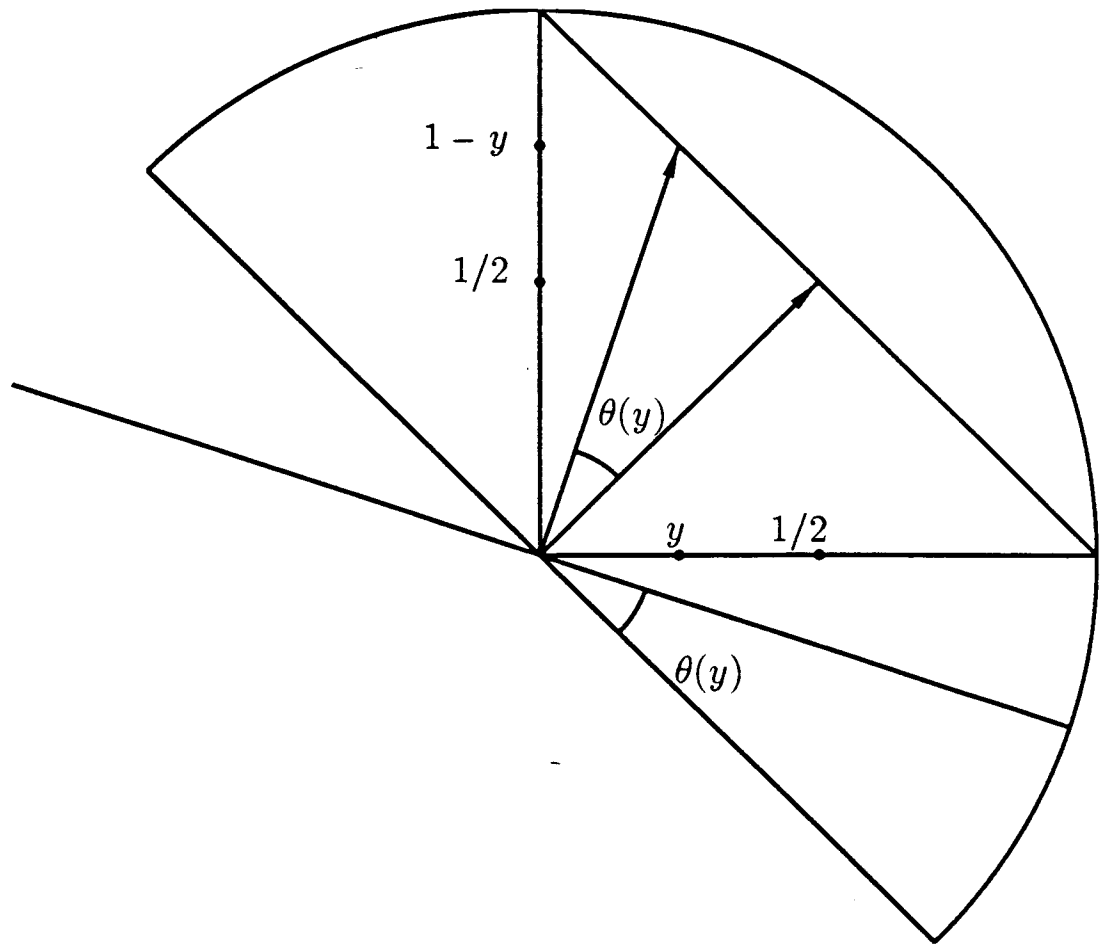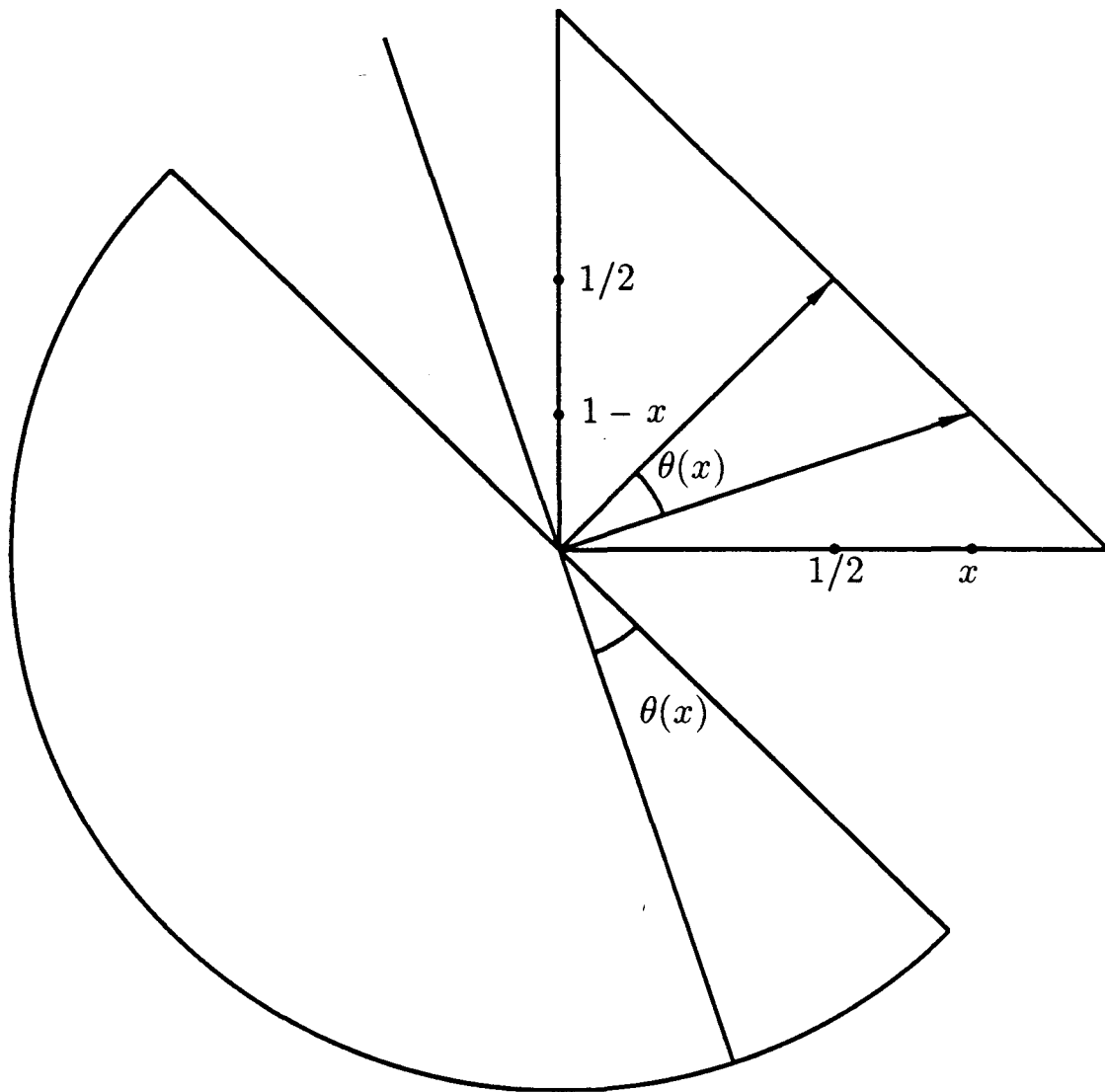
FIGURE 3

FIGURE 4

Now suppose that the strategies (B, L) are played in period 2. This reveals that player 1's payoff function lies in the arc between the line perpendicular to the expectation vector $(1/2, 1/2)$ and the line perpendicular to the expectation vector $(y^*, 1 - y^*)$, and the analogous inference can be drawn concerning player 2. Algebraically, (a, b) and (α, β) have been revealed to satisfy

(4) $(1/2)a + (1/2)b > 0,$

$y^*a + (1 - y^*)b < 0;$

$(1/2)\alpha + (1/2)\beta < 0;$ and

$x^*\alpha + (1 - x^*)\beta > 0.$

Since $y^* < 1/2$ and $x^* > 1/2$, (4) implies

(5) $a > 0, b < 0; \quad \alpha > 0, \beta < 0.$

It follows from (5), or from Figure 1, that (T, L) and (B, R) are pure strategy Nash equilibria for every game (π_1, π_2) such that $((\pi_1, \pi_2), (T, R), (B, L)) \in$ supp φ. It also follows that there is a mixed strategy Nash equilibrium, but the equilibrium mixed strategies are not yet revealed.

Therefore, in period 3, there are three possible Bayesian Nash equilibrium expectations:

(6) $\varphi_1(T \,|\, (T, R), (B, L)) = \varphi_2(L \,|\, (T, R), (B, L)) = 1;$

$\varphi_1(T \,|\, (T, R), (B, L)) = \varphi_2(L \,|\, (T, R), (B, L)) = 0;$ and

$\varphi_1(T \,|\, (T, R), (B, L)) \approx 0.61, \quad \varphi_2(L \,|\, (T, R), (B, L)) \approx 0.39.$

The first two are "pure strategy" Bayesian Nash equilibria which correspond to the pure strategy Nash equilibria and reveal no further information about each player's payoff function. The third is a "quasi-mixed" Bayesian Nash equilibrium which further partitions the two arcs revealed by the history $((T, R), (B, L))$. This multiplicity of Bayesian Nash

12

equilibria continues for all future periods. If the "quasi-mixed" equilibrium is selected infinitely often, it will converge to the mixed strategy Nash equilibrium determined by the true payoff functions (a, b), (α, β).

Figure 5 depicts the possible limits of expectations for any Bayesian strategy process with uniform priors. The white, black, and gray regions have the same interpretation as in Figure 1. Where the Nash equilibrium is unique, it is also the unique limit of any Bayesian Nash equilibrium expectations sequence. Where there are multiple Nash equilibria, a comparison of Figure 1 and Figure 5 indicates that for some games the expectations limit is unique, and for some games there can be multiple cluster points. The games discussed above, for which (T, R) and (B, L) are played in periods 1 and 2, respectively, constitute the large gray square $\approx [-\pi/4, -\pi/15] \times [47\pi/30, 7\pi/4]$. The contiguous smaller gray square is associated with the history $((T, R), (T, R), (B, L), \ldots)$, and so on, with each successively smaller square associated with an additional play of (T, R) followed by a switch to (B, L). As the picture suggests, the gray squares decrease exponentially to the point $(0, 3\pi/2)$. More generally, all four cascades of gray squares correspond to successively longer runs of a given strategy pair followed by a simultaneous switch to the opposite strategies by both players.
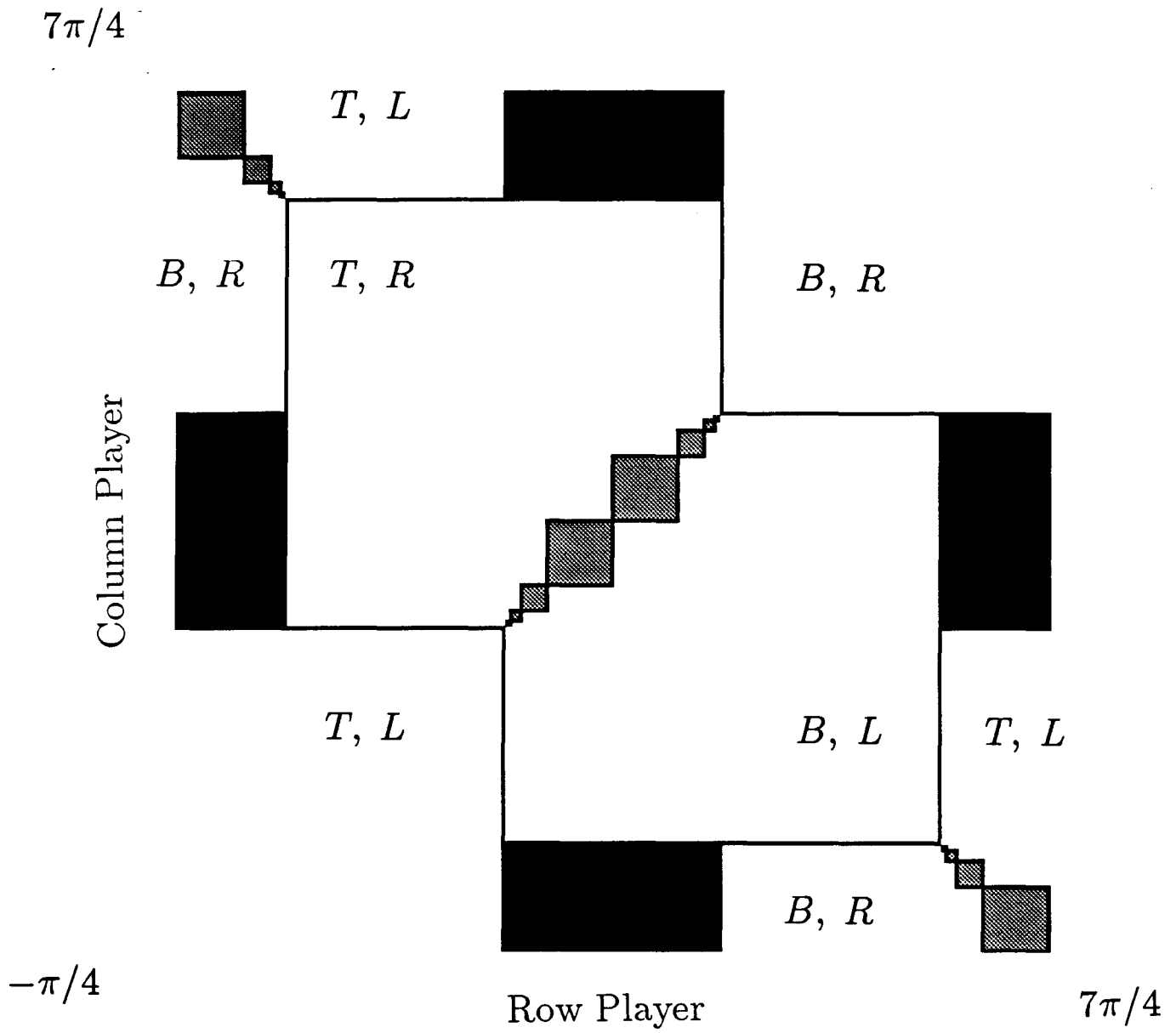
$7\pi/4$

T, L

B, R T, R B, R

Column Player

T, L B, L T, L

B, R

$-\pi/4$ Row Player $7\pi/4$

FIGURE 5

2.8 Strategic issues: The above example also illustrates the assumption of myopic behavior. Each period, each player chooses s_p to maximize the current expected payoff given the current Bayesian Nash equilibrium expectations. The choice of s_p subsequently reveals information about player p's payoff function to the other players, leading to next periods' Bayesian Nash equilibrium, and so on. In some cases it is possible for player p to choose s_p in a more sophisticated fashion so as to manipulate the expectations of the other players. The assumption of myopic behavior precludes any consideration of the strategic aspects of information revelation. Of course, loosely speaking, the potential gains from more sophisticated strategy choices diminish as the number of players increases. In the ideal case in which each player in the above model represents a continuum of players of the same type, individual players have no incentive to deviate from myopic behavior. This assumption is used by Feldman (1987) in proving the convergence of a myopic Bayesian learning process for firms in a market equilibrium model (see also Townsend (1978)).

If, in the small numbers case, players anticipate future iterations when they select current strategies, repeated game equilibria other than the one shot Nash equilibria may be more appropriate limits of learning processes. We have not attempted to extend the above model to the more general problem of learning repeated game equilibria, but some positive results have been obtained recently by Kalai and Lehrer (1990) for Bayesian learning in two-person games.

3. Convergence to Nash Equilibrium

This Section contains two results on the convergence of the conditional expectations $\varphi(\,\cdot\,|\,s^t)$ to Nash equilibrium. Example 2.7 indicates that the sequence $\varphi(\,\cdot\,|\,s^t)$ for a given game $(\pi_p)_p$ may have multiple cluster-points. However, Theorem 3.2, below, asserts that with φ-probability one, every cluster-point is a Nash equilibrium. The intuition behind Theorem 3.2 is fairly straightforward. Let $((\pi_p)_p,\, s^\infty) \in \operatorname{supp} \varphi$ and suppose that some product distribution σ is a cluster-point of the sequence $\varphi(\,\cdot\,|\,s^t)$. The only way σ can

fail to be a Nash equilibrium for $(\pi_p)_p$ is if, for some p, there is some strategy s_p^- which does not maximize the expectation of π_p against σ_{-p}, but $\sigma_p(s_p^-) > 0$. However, since $\varphi(\cdot \mid s^t)$ is near σ infinitely often, there are infinitely many periods t such that $\varphi(\cdot \mid s^t)$ is near σ_{-p}, and $s_{pt+1} = s_p^-$ with probability $\sigma_p(s_p^- \mid s^t)$, which is near $\sigma_p(s_p^-) > 0$. A player p with the payoff function π_p would never play s_p^- in these circumstances, so the likelihood that π_p is associated with s^t is reduced by a fraction near $\sigma_p(s_p^-)$ in each such period t. Thus the probability of disequilibrium situations can be driven to zero.

3.1 Definitions: If $t > \tau$, $s^t \in S^t$, and $s^\tau \in S^\tau$, we will write $s^t > s^\tau$ if s^τ consists of the first τ terms of s^t, that is $s_u^\tau = s_u^t$ for all $u \le \tau$. A sequence $\{s^{t_k}\}_{k=1}^\infty$ is said to be *increasing* if $t_k \to \infty$, and for each k, $t_{k+1} > t_k$ and $s^{t_{k+1}} > s^{t_k}$. If some $s^\infty \in S^\infty$ has been quantified, we will sometimes refer to s^t without explicitly writing $s^t < s^\infty$, where there is no risk of confusion.

Given probability distributions σ and σ' on S, define $\|\sigma - \sigma'\| = \max\{|\sigma(s) - \sigma'(s)| : s \in S\}$. Given σ and a set A of probability distributions on S, define $\|\sigma - A\| = \inf\{\|\sigma - \sigma'\| : \sigma' \in A\}$.

3.2 Theorem: Let φ be a BSP. Then $\varphi\{((\pi_p)_p, s^\infty) : \lim_{t\to\infty} \|\varphi(\cdot \mid s^t) - N((\pi_p)_p)\| = 0\} = 1$.

Proof: Let $D = \{((\pi_p)_p, s^\infty) \in \text{supp } \varphi : \|\varphi(\cdot \mid s^t) - N((\pi_p)_p)\| \not\to 0\}$. We will show that D is a countable union of sets of probability zero. Let $((\pi_p)_p, s^\infty) \in D$. Since the set of probability measures on S is compact there is an increasing sequence $\{s^{t_k}\}_{k=1}^\infty$, with $s^{t_k} < s^\infty$ for each k and $\varphi(\cdot \mid s^{t_k}) \to \sigma^*$ for some $\sigma^* \notin N((\pi_p)_p)$. Since φ is a BSP, 2.4 (*iii*) implies that $\sigma^* = \Pi_{p=1}^n \sigma_p^*$, where σ_p^* is the marginal distribution of σ^* on S_p, for each p. Since $\sigma^* \notin N((\pi_p)_p)$, there is some q, some $s_q^+, s_q^- \in S_q$, and some $\varepsilon > 0$ such that

(7) $\quad \varphi_q(s_q^- \mid s^{t_k}) > \varepsilon$; and

15

(8) $\quad \Sigma_{s_{-q}} \left[\pi_q(s_q^+, s_{-q}) - \pi_q(s_q^-, s_{-q}) \right] \varphi_{-q}(s_{-q}|s^{t_k}) > \varepsilon,$

for all k sufficiently large.

Accordingly, fix $s_q^+, s_q^- \in S_q$, let $0 < \varepsilon < 1$, and let σ° be a product distribution on S with $\sigma_q^\circ(s_q^-) > \varepsilon$. Let $\delta > 0$ such that for any product distribution σ on S which satisfies $\|\sigma - \sigma^\circ\| < \delta$, we have $\sigma_q(s_q^-) > \varepsilon$. Let $U(\sigma^\circ, \delta) = \{\sigma : \sigma$ is a product distribution and $\|\sigma - \sigma^\circ\| < \delta\}$. Let $P_q(\sigma^\circ, \delta) = \{\pi_q \in B : \Sigma_{s_{-q}}[\pi_q(s_q^+, s_{-q}) - \pi_q(s_q^-, s_{-q})]\sigma_{-q}(s_{-q}) > 0$ for every $\sigma \in N(\sigma^\circ, \delta)\}$. Finally, define $D^\circ = \{((\pi_p)_p, s^\infty) \in \text{supp } \varphi : \pi_q \in P_q(\sigma^\circ, \delta)$ and $\varphi(\cdot|s^t) \in U(\sigma^\circ, \delta)$ for infinitely many $t\}$. It is apparent from (7) and (8) that D is the union of sets D° over all $1 \leq q \leq n$, all $s_q^+, s_q^- \in S_q$, countably many values of ε and δ, and a countable dense subset of product distributions σ° on S. Thus it suffices to prove that $\varphi(D^\circ) = 0$.

For each $s^\infty \in \text{supp } \varphi$ and each $t \geq 1$, let $c_t(s^\infty) = \#\{\tau \leq t : \varphi(\cdot|s^\tau) \in U(\sigma^\circ, \delta)\}$. Then $D^\circ = \{((\pi_p)_p, s^\infty) \in \text{supp } \varphi : \pi_q \in P_q(\sigma^\circ, \delta)$ and $c_t(s^\infty) \to \infty\}$. For each $t \geq 1$ and each integer $k \geq 1$, let

$$D_k^\circ = \{((\pi_p)_p, s^\infty) \in \text{supp } \varphi : \pi_q \in P_q(\sigma^\circ, \delta) \text{ and } c_t(s^\infty) = k \text{ for some } t\}.$$

Then $D_{k+1}^\circ \subset D_k^\circ$ for each k, and $D^\circ = \cap_k D_k^\circ$. We will obtain an upper bound on $\varphi(D_k^\circ)$. Since φ is a BSP, property 2.4 (ii) implies that if $((\pi_p)_p, s^\infty) \in \text{supp } \varphi$ with $\pi_q \in P_q(\sigma^\circ, \delta)$ and $\varphi(\cdot|s^t) \in U(\sigma^\circ, \delta)$ then $s_{qt+1} \neq s_q^-$. However if $\varphi(\cdot|s^t) \in U(\sigma^\circ, \delta)$ then $\varphi_q(s_q^-|s^t) > \varepsilon$. Therefore $\varphi\{((\pi_p)_p, s^t) \in \text{supp } \varphi : \pi_q \in P_q(\sigma^\circ, s)$ and $\varphi(\cdot|s^t) \in U(\sigma^\circ, \delta)\} < (1-\varepsilon) \varphi \{((\pi_p)_p, s^t) \in \text{supp } \varphi : \varphi(\cdot|s^t) \in U(\sigma^\circ, \delta)\} \leq (1-\varepsilon)$. Proceeding iteratively, we obtain $\varphi(D_k^\circ) < (1-\varepsilon)^k$. Hence $\varphi(D^\circ) = 0$. ∎

3.3 Remarks: Theorem 3.2 states that convergence to Nash equilibrium fails only on a set of φ-probability zero. However, the Theorem allows the possibility that for some games $(\pi_p)_p \in \text{supp } \mu$ and some histories s^∞ with $((\pi_p)_p, s^\infty) \in \text{supp } \varphi$, there is a subsequence of expectations $\varphi(\cdot|s^{t_k})$ which is bounded away from $N((\pi_p)_p)$.

Example 3.11, below, explicitly demonstrates this possibility. Thus Theorem 3.2 ensures convergence only to the extent that the prior distribution μ is interpreted literally as the distribution from which the game is drawn. If μ is misspecified, the "true" probability of nonconvergence may be positive. In any case, a much stronger convergence result is needed to justify the interpretation of Bayesian strategy processes as learning processes for general normal form games.

A general convergence result is stated in Theorem 3.8, below, under the assumption that each μ_p is very similar to the uniform distribution on B. This property is defined formally below.

3.4 Definitions: Let m denote Lebesgue measure on B. For $1 \leq p \leq n$, the notation $\mu_p \ll m$ states that μ_p is absolutely continuous with respect to Lebesgue measure. For each p, we will write $\mu_p \sim m$ if there is some number $\alpha > 0$ such that for every Borel set $E \subset B$, $\alpha m(E) \geq \mu_p(E) \geq m(E) / \alpha$.

3.5 Remarks: If $\mu_p \sim m$, then μ_p and m are mutually absolutely continuous. Moreover, if g_p is a density function for μ_p with respect to m, then for some $\alpha > 0$, $\alpha > g_p(\pi_p) > 1/\alpha$ for almost every $\pi_p \in B$. We will discuss the role of this property in Section 3.7, below.

The proof of Theorem 3.8 will use a characterization of Bayesian strategy processes which relies only on the absolute continuity of each μ_p with respect to Lebesgue measure on B. Recall that in period $t + 1$, given an observed history s^t, player p faces the distribution $\varphi_{-p}(\cdot \mid s^t)$ over his opponents' strategies. Except for a set of payoff functions π_p having Lebesgue measure zero, player p has a unique maximal response to $\varphi_{-p}(\cdot \mid s^t)$. Therefore, if $\mu_p \ll m$ for each p, one can express the probability of each history s^t as the μ-measure of games $(\pi_p)_p$ such that, for each $\tau < t$, each $s_{p\tau+1}$ uniquely maximizes the expectation of π_p against $\pi_{-p}(\cdot \mid s^\tau)$. In particular, mixed-strategies can

be ignored. This characterization is stated in the following proposition.

3.6 Proposition: Suppose that $\mu_p \ll m$ for each p and that φ is a BSP for μ. Let $s^t \in \text{supp } \varphi$ with $t \geq 1$, and for each p, define $C_p(s^t) = \{\pi_p \in B : \text{for each } \tau < t, \, s^t_{p\tau+1} \text{ maximizes } \Sigma_{s_{-p}} \pi_p(\cdot, s_{-p})\varphi_{-p}(s_{-p}|s^\tau) \text{ on } S_p, \text{ where } s^\tau < s^t\}$, and define $C^\circ_p(s^t) = \{\pi_p \in B : \text{for each } \tau < t, \, s^t_{p\tau+1} \text{ is the unique maximizer of } \Sigma_{s_{-p}} \pi_p(\cdot, s_{-p})\varphi_{-p}(s_{-p}|s^\tau), \text{ where } s^\tau < s^t\}$. Then, for each p,

    i) $C_p(s^t)$ is a compact convex set, and $C^\circ_p(s^t)$ is a convex set which is open relative to B;

    ii) $\mu_p(C_p(s^t) \setminus C^\circ_p(s^t)) = m(C_p(s^t) \setminus C^\circ_p(s^t)) = 0$;

    iii) $\{\pi_p \in B : (\pi_p, s^t) \in \text{supp } \varphi\} \subset C_p(s^t)$; and

    iv) $\varphi(s^t) = \mu_1(C_1(s^t)) \times \cdots \times \mu_n(C_n(s^t))$.

Proof: Assertion *(i)* follows immediately from the definitions of $C_p(s^t)$ and $C^\circ_p(s^t)$. To prove *(ii)*, note that $m(\{\pi_p \in B : \text{for some } \tau < t \text{ and some}$
$s_p \neq s'_p \in S_p, \Sigma_{s_{-p}} [\pi_p(s_p, s_{-p}) - \pi_p(s'_p, s_{-p})] \varphi_{-p}(s_{-p}|s^\tau) = 0\}) = 0$, because this event is a finite union of linear subspaces of B. Then *(ii)* follows from the assumption that $\mu_p \ll m$. Assertion *(iii)* follows iteratively from the property 2.4 *(ii)* of the BSP φ. To prove *(iv)*, note that *(iii)* implies that for each $1 \leq p \leq n$,

$$C^\circ_p(s^t) \cap \{\pi_p : (\pi_p, s'^t) \in \text{supp } \varphi\} = \emptyset$$

for every s'^t such that $s'^t_\tau \neq s^t_{p\tau}$ for some $\tau \leq t$. Since the marginal distribution of φ on G agrees with μ, this implies that $\Pi^n_{p=1}\mu_p(C^\circ_p(s^t)) \leq \varphi(s^t)$. Since *(iii)* also implies that $\varphi(s^t) \leq \Pi^n_{p=1} \mu_p(C_p(s^t))$, *(iv)* now follows from *(ii)*. ∎

3.7 Remarks: The proof of Proposition 3.6 does not use the independence property, 2.4 *(iii)*, of Bayesian strategy processes. Since 2.4 *(iii)* is a direct implication of 3.6 *(iv)*, it is

apparent that if $\mu_p \ll m$ for each p, then 2.4 (iii) follows from 2.4 (i and ii). Moreover, 3.6 (iv) implies that, conditional on each $s^t \in \text{supp } \varphi$, the variables (π_p, s_{pt+1}) and (π_q, s_{qt+1}) are independent for $p \neq q$. Thus 3.6 (iv) is a slightly stronger independence condition than 2.4 (iii).

Each set $C_p(s^t)$ is the intersection of B and finitely many closed half-spaces: one half-space for each revealed preference for $s_{p\tau+1}$ over s_p given $\varphi_{-p}(\cdot \mid s^\tau)$, for each $\tau < t$. If we define $C_p(s^t, s_p) = \{\pi_p \in C_p(s^t) : s_p$ is expected payoff maximizing against $\varphi_{-p}(\cdot \mid s^t)\}$, then 3.6 ($iv$) implies that $\varphi_p(s_p \mid s^t) = \mu_p(C_p(s^t, s_p)) / \mu_p(C_p(s^t))$. As π_p varies over the set $C_p(s^t)$, the expected utility differences

$$\Sigma_{s_{-p}} [\pi_p(s_p^+, s_{-p}) - \pi_p(s_p^-, s_{-p})] \varphi_{-p}(s_{-p} \mid s^t),$$

for various strategies s_p^+, $s_p^- \in S_p$, are related to the relative diameters of the sets $C_p(s^t, s_p)$. If $\mu_p \sim m$ for each p, the relative diameters are related to the relative μ_p-probabilities, and thus to the conditional probabilities $\varphi_p(s_p \mid s^t)$. The proof of Theorem 3.8 uses this relation to show that if $\pi_p \in C_p(s^t)$ for all t and s_p^- is an asymptotically inferior strategy for π_p against $\varphi_{-p}(\cdot \mid s^t)$, then $\varphi_p(s_p^- \mid s^t) \to 0$.

3.8 Theorem: Suppose that $\mu_p \sim m$ for each p, and let φ be a BSP for μ. Then for every $((\pi_p)_p, s^\infty) \in \text{supp } \varphi$, $\lim_{t \to \infty} \|\varphi(\cdot \mid s^t) - N((\pi_p)_p)\| = 0$.

Proof: Let $((\pi_p)_p, s^\infty) \in \text{supp } \varphi$. Since the set of probability distributions on S is compact, it suffices to show that every cluster-point of the sequence $\{\varphi(\cdot \mid s^t)\}_{t=1}^\infty$ lies in $N((\pi_p)_p)$. Let σ be a cluster-point, so that there is an increasing sequence $\{s^{t_k}\}_{k=1}^\infty$ with $s^{t_k} < s^\infty$ for each k and $\varphi(\cdot \mid s^{t_k}) \to \sigma$. Since φ is a BSP, 2.4 (iii) implies that σ is a product distribution. Fix q, and for each k, define

$$(9) \quad d_k = \max \{\Sigma_{s_{-q}} \pi_q(s_q, s_{-q}) \varphi_{-q}(s_{-q} \mid s^{t_k}) : s_q \in S_q\}$$

$$-\min \{\Sigma_{s_{-q}} \pi_q(s_q, s_{-q}) \varphi_{-q}(s_{-q} \mid s^{t_k}) : \sigma_q(s_q) > 0\}.$$

19

To prove that σ_q is a Nash equilibrium mixed strategy for player q with payoff function π_q, it suffices to show that $d_k \to 0$. Therefore, we can assume that $d_k > 0$ for all k. Proposition 3.6 (iii) implies that the maximum on the right hand side of (9) is achieved at $s_q = s_{qt_k+1}^{t_k+1}$. Taking a subsequence if necessary, we can assume that this is the same strategy, say s_q^+, for all k. We can also assume that the minimum is achieved at the same strategy, say s_q^-, for all k. To further condense notation, for each k and each $s_{-q} \in S_{-q}$, define $x_k(s_{-q}) = \varphi_{-q}(s_{-q}|s^{t_k})$. Then

$$(10) \quad d_k = \Sigma_{s_{-q}}[\pi_q(s_q^+, s_{-q}) - \pi_q(s_q^-, s_{-q})]\, x_k(s_{-q}),$$

and $s_q^+ = s_{qt_k+1}^{t_k+1}$ for all k. For each k, define the linear function $\delta_k : R^S \to R$ by

$$(11) \quad \delta_k(\pi_q') = \Sigma_{s_{-q}} [\pi_q'(s_q^+, s_{-q}) - \pi_q'(s_q^-, s_{-q})]\, x_k(s_{-q}).$$

In particular, $\delta_k(\pi_q) = d_k$. For each k, define

$$(12) \quad A_k^+ = \{\pi_q' \in C_q(s^{t_k}) : \delta_k(\pi_q') > 0\};$$

$$(13) \quad A_k^0 = \{\pi_q' \in C_q(s^{t_k}) : \delta_k(\pi_q') = 0\}; \text{ and}$$

$$(14) \quad A_k^- = \{\pi_q' \in C_q(s^{t_k}) : \delta_k(\pi_q') < 0\}.$$

It follows from Proposition 3.6 (ii and iv) that $\varphi_q(s_q^-|s^{t_k}) = \mu_q\{\pi_q' \in C_q(s^t) : s_q^-$ is the unique maximizer of $\Sigma_{s_{-q}} \pi_q'(\cdot, s_{-q})\, x_k(s_{-q})\}/\mu_q(C_q(s^{t_k}))$, so (14) and (11) imply

$$(15) \quad \varphi_q(s_q^-|s^{t_k}) \leq \mu_q(A_k^-)/\mu_q(C_q(s^{t_k})).$$

Therefore $1 - \varphi_q(s_q^-|s^{t_k}) \geq \mu_q(A_k^+)/\mu_q(C_q(s^{t_k}))$ and

$$(16) \quad (1 - \varphi_q(s_q^-|s^{t_k}))/\varphi_q(s_q^-|s^{t_k}) \geq \mu_q(A_k^+)/\mu_q(A_k^-).$$

Since $\varphi_q(s_q^-|s^{t_k}) \to \sigma_q(s_q^-) > 0$, we can assume that the denominators in (16) are positive for all k. For each k, define

$$(17) \quad \delta_k^- = \min\{\delta_k(\pi_q') : \pi_q' \in C_q(s^{t_k})\},$$

so $\delta_k^- < 0$. Also, define

(18) $\quad K_k^+ = \{\lambda \pi_q' + (1 - \lambda) \pi_q : 0 \le \lambda < 1 \text{ and } \pi_q' \in A_k^0\}$; and

(19) $\quad K_k = \{\lambda \pi_q' + (1 - \lambda) \pi_q : 0 \le \lambda \le 1 + (|\delta_k^-| / d_k) \text{ and } \pi_q' \in A_k^0\}$.

Then by the definitions (12), (13), and (18), $K_k^+ \subset A_k^+$. Also, for any $\pi_q' \in A_k^-$, $\delta_k(\pi_q') < 0$, so, using (17), $\gamma \pi_q + (1 - \gamma) \pi_q' \in A_k^0$ for some $0 < \gamma \le |\delta_k^-| / (|\delta_k^-| + d_k)$, which implies that $\pi_q' \in K_k$. Hence $A_k^- \subset K_k \backslash K_k^+$, so

(20) $\quad \mu_q(A_k^+) / \mu_q(A_k^-) \ge (\mu_q(K_k^+) / \mu_q(K_k \backslash K_k^+))$.

Let $\ell = \#S$, let m^0 denote Lebesgue measure on the $(\ell - 1)$-dimensional linear subspace $\{\pi_q' \in R^S : \delta_k(\pi_q') = 0\}$, and let $a_k = \max\{\delta_k(\pi_q') : \|\pi_q'\| = 1\}$. Then we have the following volume measurements:

(21) $\quad m(K_k^+) = d_k m^0(A_k^0) / \ell a_k$; and

(22) $\quad m(K_k) = [(d_k + |\delta_k^-|)^\ell / d_k^{\ell-1}] m^0(A_k^0) / \ell a_k$, so

(23) $\quad m(K_k^+) / m(K_k \backslash K_k^+) = d_k^\ell / [(d_k + |\delta_k^-|)^\ell - d_k^\ell]$.

We now show that $\delta_k^- \to 0$. For each k, $s_q^+ = s_{q t_k+1}^{t_k+1}$, so Proposition 3.6 (iii) implies that $C_q(s^{t_k+1}) \subset A_k^+ \cup A_k^0$. Hence

(24) $\quad \delta_{k+1}^- \ge \min\{\delta_{k+1}(\pi_q') : \pi_q' \in B \text{ and } \delta_k(\pi_q') \ge 0\}$.

Since $x_k(\cdot) \to \sigma_{-q}$, it follows from the definition of δ_k, (11), that

$\max\{|\delta_{k+1}(\pi_q') - \delta_k(\pi_q')| : \pi_q' \in B\} \to 0$. Then (24) implies that $\delta_k^- \to 0$.

Since $\mu_q \sim m$, by (23) and Definition 3.4, there is some $\alpha > 0$ such that

(25) $\quad \mu_q(K_k^+) / \mu_q(K_k \backslash K_k^+) \ge d_k^\ell / [(d_k + |\delta_k^-|)^\ell - d_k^\ell] \alpha^2$.

It follows from (16), (20), and (25) that

(26) $\qquad (1 - \sigma_q(s_q^-)) / \sigma_q(s_q^-) \ge \limsup_{k \to \infty} d_k^\ell / [(d_k + |\delta_k^-|)^\ell - d_k^\ell] \alpha^2$.

Since $\delta_k^- \to 0$, (26) implies that $d_k \to 0$. Since $\mu_p \sim m$ for all p, the above argument shows that σ is a Nash equilibrium for $(\pi_p)_p$. $\blacksquare$

3.9 Remarks: If $\mu_p \sim m$ for each p, then supp $\mu = G$, so every normal form game has Nash equilibria which are achieved as cluster points of $\varphi(\cdot \mid s^t)$. The following Corollary is an equivalent statement of the Theorem which emphasizes this implication.

3.10 Corollary: Suppose that $\mu_p \sim m$ for each p, and let φ be a BSP for μ. For each $(\pi_p)_p \in G$, define $L((\pi_p)_p, \varphi)$ to be the set of probability distributions σ on S such that there is an increasing sequence $\{s^{t_k}\}_{k=1}^{\infty}$ satisfying

$(i) \quad ((\pi_p)_p, s^{t_k}) \in$ supp φ for all k; and

$(ii) \quad \lim_{k \to \infty} \varphi(\cdot \mid s^{t_k}) = \sigma.$

Then for every $(\pi_p)_p \in G$, $L((\pi_p)_p, \varphi)$ is a nonempty subset of $N((\pi_p)_p)$.

3.11 Example: The assumption that $\mu_p \sim m$ for each p, or some very similar assumption, seems essential to the above result. This section provides an example showing that in the absence of such an assumption, there can be points $((\pi_p)_p, s^{\infty}) \in$ supp φ with subsequences of expectations $\{\varphi(\cdot \mid s^{t_k})\}_{k=1}^{\infty}$ converging to some $\sigma \notin N((\pi_p)_p)$.

The example will involve two players and three pure strategies for each player. We first specify a probability distribution, μ° on the unit square $[0, 1] \times [0, 1]$ which will be used to define μ_1. Points in the unit square will be denoted (x, y). The essential properties of μ° are that

$(27) \quad \mu^{\circ}(x \leq 2/3) = .6, \; \mu^{\circ}(y \leq 1/2 - 1/3 \mid x > 2/3) = .6;$

and for each integer $r > 2$,

$(28) \quad \mu^{\circ}(x \leq 1 - (r+1)^{-1} \mid x > 1 - r^{-1}, \, y > (1/2) - r^{-1}) = .6; \;$ and

$(29) \quad \mu^{\circ}(y \leq (1/2) - (r+1)^{-1} \mid x > 1 - (r+1)^{-1}, \, y > (1/2) - r^{-1}) = .6.$

Also, μ° is absolutely continuous with respect to Lebesgue measure, and the density function, f, is strictly positive on $(0, 1) \times (0, 1)$, so that supp $\mu^{\circ} = [0, 1] \times [0, 1]$. Such a μ° can be constructed iteratively, making the density function constant at the

appropriate positive level on rectangles of the form $(1-r^{-1}, 1-(r+1)^{-1}] \times ((1/2)-r^{-1}, 1]$, and $(1-(r+1)^{-1}, 1] \times ((1/2)-r^{-1}, (1/2)-(r+1)^{-1}]$. Then the density function f is bounded above, but for any $(1/2) \leq y \leq 1$, $f(x_n, y) \to 0$ as $x_n \to 1$, so f does not have a positive lower bound.

Now let $S_1 = S_2 = \{a, b, c\}$. The specification of μ_1 and μ_2 will be simplified by departing from the unit ball normalization of payoff functions. Let μ_2 be the point mass at π_2° defined by

$$(30) \qquad \pi_2^\circ(s_1, s_2) = \begin{cases} 0 & \text{if } s_1 \neq s_2; \\ 3/2 & \text{if } s_1 = s_2 = 0 \text{ or } s_1 = s_2 = b; \text{ and} \\ 1 & \text{if } s_1 = s_2 = c. \end{cases}$$

The distribution μ_1 is defined by the properties that, with probability one,

$$(31) \qquad \pi_1(s_1, s_2) = \begin{cases} 0 & \text{if } s_1 \neq s_2; \text{ and} \\ 1 & \text{if } s_1 = s_2 = c; \end{cases}$$

and the joint distribution on $(\pi_1(a, a), \pi_1(b, b))$ is given by μ°. Thus, each $(\pi_1, \pi_2) \in$ supp μ defines a "battle-of-the-sexes" game with three pure strategies.

There is a multiplicity of Bayesian strategy processes for μ, and we will exploit this multiplicity in the following construction of φ. In odd numbered periods $\varphi_p(\cdot | s^t)$ will be a mixture of strategies a and c, and in even numbered periods $\varphi_p(\cdot | s^t)$ will be a mixture of b and c. More specifically, in period 1, let

$$(32) \quad \varphi_1(s_1 = a) = .4, \quad \varphi_1(s_1 = c) = .6, \quad \varphi_2(s_2 = a) = .6, \quad \varphi_2(s_2 = c) = .4.$$

Then (27) and (30) ensure that (32) is consistent with first period Bayesian Nash equilibrium. We will be concerned with the strategy sequence generated by the game $(\pi_1^\circ, \pi_2^\circ) \in$ supp μ defined by $\pi_p^\circ(a, a) = 1$ and $\pi_1^\circ(b, b) = 1$ (or $(x, y) = (1, 1)$ in the definition of μ°, above). For this π_1°, we have, in the first period, $s_{11} = a$, which reveals to player 2 that $\pi_1(a, a) > 2/3$. Since μ_2 is a point mass at π_2°, player 1 has nothing more to learn from observing s_2 in any period. In period 2, let

(33) $\varphi_1(s_1 = b \,|\, s_{11} = a) = .4, \quad \varphi_1(s_1 = c \,|\, s_{11} = a) = .6,$

$\varphi_2(s_2 = b \,|\, s_{11} = a) = 6/7, \quad \varphi_2(s_1 = c \,|\, s_{11} = a) = 1/7.$

Then (27) and (30) ensure that (33) is consistent with second period Bayesian equilibrium. Since $\pi_p^o(b, b) = 1$, we have $s_{12} = b$. In general, let s^t be any history satisfying $s_{1\tau} = a$ if τ is odd and $s_{1\tau} = b$ if τ is even, for each $\tau \le t$. If t is even, let $r = 2 + (t/2)$, and let

(34) $\varphi_1(s_1 = a \,|\, s^t) = .4, \quad \varphi_1(s_1 = c \,|\, s^t) = .6$

$\varphi_2(s_2 = a \,|\, s^t) = (r+1)/(2r+1), \quad \varphi_2(s_2 = c \,|\, s^t) = r/(2r+1).$

If t is odd, let $r = 2 + (t-1)/2$, and let

(35) $\varphi_1(s_1 = b \,|\, s^t) = .4, \quad \varphi_1(s_1 = c \,|\, s^t) = .6$

$\varphi_2(s_2 = b \,|\, s^t) = (2r+2)/(3r+1), \quad \varphi_2(s_2 = c \,|\, s^t) = (r-1)/(3r+1).$

Then (30), together with (28) (resp. (29)) ensures that (34) (resp. (35)) is consistent with Bayesian Nash equilibrium. Also, since $\pi_1^o(a, a) = 1$ and $\pi_1^o(b, b) = 1$, (34) implies that $s_{1t+1} = a$ when t is even, and (35) implies that $s_{1t+1} = b$ when t is odd.

Finally, for each positive integer k, let $t_k = 2k - 1$. Then for each k, $((\pi_1^o, \pi_2^o), s^{t_k}) \in \text{supp } \varphi$, and $\varphi(\cdot \,|\, s^{t_k}) \to \sigma = \sigma_1 \times \sigma_2$, where

(36) $\sigma_1(s_1 = b) = .4, \quad \varphi_1(s_1 = c) = .6,$

$\sigma_2(s_2 = b) = 2/3, \quad \varphi_2(s_2 = c) = 1/3.$

However, since $\pi_1^o(b, b) = \pi_1^o(c, c) = 1$ and $\pi_1^o(b, c) = \pi_1^o(c, b) = 0$, σ_1 is not a Nash equilibrium mixed strategy for player 1.

The fact that μ^o is absolutely continuous with respect to Lebesgue measure on $[0, 1] \times [0, 1]$ suggests that a more elaborate version of this example may be possible with $\mu_p \ll m$ for each p. It appears that the relation between relative volumes and probabilities, which yields the inequality (25) in the proof of Theorem 3.8, may be essential for the general convergence result.

24

4. Conclusion

The main result of this paper is stated as Corollary 3.10, above. Given any fixed finite set of players and pure strategies, if the prior probability distribution over payoff functions is sufficiently uniform, then a Bayesian strategy process is a learning process which achieves Nash equilibria for every normal form game. In particular, a single BSP φ completely specifies a learning process for all normal form games with the given player and strategy sets.

Perhaps the most troubling feature of Bayesian strategy processes is the amount of required coordination among the players. A common prior distribution μ must be agreed upon, and selections among multiple Bayesian Nash equilibria must be agreed upon at each stage of the process. All of these choices are built into the single object φ, so they are independent of particular realizations of the game $(\pi_p)_p$. In this sense, the coordination embodied in φ is similar to the requirement in macroeconomic learning models, for example, that all agents use a particular econometric procedure (e.g. Marcet and Sargent (1989)). However, the amount of coordination embodied in φ is intuitively much greater. It would be interesting to formulate a measure of such coordination and to determine the minimal coordination required for general convergence.

Notes

[†] Strictly speaking, Milgrom and Weber (1985) define equilibrium in terms of joint distributions on $B \times S_p$, and our functions f_p are the associated conditional distributions on S_p given π_p. A more recent and more general existence theorem due to Balder (1987) is stated in terms of the conditional distributions.

References

Balder, E.J. (1987), "Generalized Equilibrium Results for Games with Incomplete Information", Mathematical Institute, University of Utrecht, mimeo.

Berry, D.A., and Fristedt, B.(1985), *Bandit Problems: Sequential Allocation of Experiments*, Chapman and Hall, London.

Blume, L., M. Bray, and D. Easely (1982), "Introduction to the Stability of Rational Expectations Equilibrium", *Journal of Economic Theory*, **26**, pp. 313-317.

Easely, D., and N. Kiefer (1986), "Controlling a Stochastic Process with Unknown Parameters", *Econometrica*, **56**, pp. 1045-1064.

Feldman, M. (1987), "An Example of Convergence to Rational Expectations with Heterogeneous Beliefs", *International Economic Review*, **28:3**, pp. 635-650.

Fudenberg, D., and D. Kreps (1988), "A Theory of Learning, Experimentation, and Equilibrium in Games", mimeo.

Guillemin, V. and A. Pollak (1974), *Differential Topology*, Prentice-Hall, Englewood Cliffs, New Jersey.

Harsanyi, J. (1975), "The Tracing Procedure: A Bayesian Approach to Defining a Solution for n-Person Noncooperative Games", *International Journal of Game Theory*, **4**, pp. 61-94.

Kalai, E. and E. Lehrer (1990), "Rational Learning Leads to Nash Equilibrium", Discussion Paper No. 858, Northwestern University.

Marcet, A., and T. Sargent (1989), "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models", *Journal of Economic Theory*, **48**, pp. 337-368.

McLennan, A. (1987), "Incomplete Learning in a Repeated Statistical Decision Problem", mimeo.

Milgrom, P. and J. Roberts (1989), "Adaptive and Sophisticated Learning in Repeated

Normal Form Games", mimeo.

Milgrom, P., and R. Weber (1985), "Distributional Strategies for Games with Incomplete Information", *Mathematics of Operations Research*, **10**, pp. 619-632.

Townsend, R. (1978), "Market Anticipations, Rational Expectations, and Bayesian Analysis", *International Economic Review*, **19:2**, pp. 481-494.