Working Paper

Learning Dynamics in Games with Stochastic Perturbations

Yuri M. Kaniovski
H. Peyton Young

WP-94-30
December 1994

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 807 □ Fax: +43 2236 71313 □ E-Mail: info@iiasa.ac.at

Learning Dynamics in Games with Stochastic Perturbations

Yuri M. Kaniovski
H. Peyton Young

WP-94-30
December 1994

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 807 □ Fax: +43 2236 71313 □ E-Mail: info@iiasa.ac.at

Preface

This new research project at IIASA is concerned with modeling technological and organisational change; the broader economic developments that are associated with technological change, both as cause and effect; the processes by which economic agents – first of all, business firms – acquire and develop the capabilities to generate, imitate and adopt technological and organisational innovations; and the aggregate dynamics – at the levels of single industries and whole economies – engendered by the interactions among agents which are heterogeneous in their innovative abilities, behavioural rules and expectations. The central purpose is to develop stronger theory and better modeling techniques. However, the basic philosophy is that such theoretical and modeling work is most fruitful when attention is paid to the known empirical details of the phenomena the work aims to address: therefore, a considerable effort is put into a better understanding of the 'stylized facts' concerning corporate organisation routines and strategy; industrial evolution and the 'demography' of firms; patterns of macroeconomic growth and trade.

From a modeling perspective, over the last decade considerable progress has been made on various techniques of dynamic modeling. Some of this work has employed ordinary differential and difference equations, and some of it stochastic equations. A number of efforts have taken advantage of the growing power of simulation techniques. Others have employed more traditional mathematics. As a result of this theoretical work, the toolkit for modeling technological and economic dynamics is significantly richer than it was a decade ago.

During the same period, there have been major advances in the empirical understanding. There are now many more detailed technological histories available. Much more is known about the similarities and differencers of technical advance in different fields and industries and there is some understanding of the key variables that lie behind those differences. A number of studies have provided rich information about how industry structure co-evolves with technology. In addition to empirical work at the technology or sector level, the last decade has also seen a great deal of empirical research on productivity growth and measured technical advance at the level of whole economies. A considerable body of empirical research now exists on the facts that seem associated with different rates of productivity growth across the range of nations, with the dynamics of convergence and divergence in the levels and rates of growth of income in different countries, with the diverse national institutional arrangements in which technological change is embedded.

As a result of this recent empirical work, the questions that successful theory and useful modeling techniques ought to address now are much more clearly defined. The theoretical work described above often has been undertaken in appreciation of certain stylized facts that needed to be explained. The list of these 'facts' is indeed very long, ranging from the microeconomic evidence concerning for example dynamic increasing returns in learning activities or the persistence of particular sets of problem-solving routines within business firms; the industry-level evidence on entry, exit and size-distributions – approximately log-normal; all the way to the evidence regarding the time-series properties of major economic aggregates. However, the connection between the theoretical work and the empirical phenomena has so far not been very close. The philosophy of this project is that the chances of developing powerful new theory and useful new analytical techniques can be greatly enhanced by performing the work in an environment where scholars who understand the empirical phenomena provide questions and challenges for the theorists and their work.

In particular, the project is meant to pursue an 'evolutionary' interpretation of technological and economic dynamics modeling, first, the processes by which individual agents and organisations learn, search, adapt; second, the economic analogues of 'natural selection' by which interactive environments – often markets – winnow out a population whose members have different attributes and behavioural traits; and, third, the collective emergence of statistical patterns, regularities and higher-level structures as the aggregate outcomes of the two former processes.

Together with a group of researchers located permanently at IIASA, the project coordinates multiple research efforts undertaken in several institutions around the world, organises workshops and provides a venue of scientific discussion among scholars working on evolutionary modeling, computer simulation and non-linear dynamical systems.

The research will focus upon the following three major areas:

1. Learning Processes and Organisational Competence.

2. Technological and Industrial Dynamics

3. Innovation, Competition and Macrodynamics

Summary

Consider a game that is played repeatedly by two populations of agents. In fictitious play, agents learn by choosing best replies to the frequency distribution of actions taken by the other side. We consider a more general class of learning processes in which agents' choices are perturbed by incomplete information about what the other side has done, variability in their payoffs, and unexplained trembles. These perturbed best reply dynamics define a non-stationary Markov process on an infinite state space. We show that for 2×2 games it converges with probability one to a neighborhood of the stable Nash equilibria, whether pure or mixed. This generalizes a result of Fudenberg and Kreps (1993), who demonstrate convergence when the game has a unique mixed equilibrium.

Key words: evolutionary game theory, perturbed best reply dynamic, convergence with probability one, non-stationary Markov processes, stable Nash equilibria, rate of convergence.

Contents

Learning Dynamics in Games with Stochastic Perturbations*

By Yuri M. Kaniovski and H. Peyton Young

1 A model of technological adoption

Consider two classes of agents who are deciding whether to adopt complementary technologies. Suppose, for example, that a new kind kind of gasoline (technology X) comes on the market. Filling stations must decide whether to stock X, and they will base their choice on an estimate of how many consumers have cars that run on X. Similarly, a consumer faced with the choice of whether to buy a car that runs on X will consider how many filling stations already offer X. In both cases the individual's decision depends on the proportion of people in the other class who have already adopted X, but these proportions are not precisely known. A consumer, in driving around, will notice that some stations offer X and some do not. Similarly, a filling station owner observes that some cars use X some do not. From these casual and somewhat random observations they infer the proportions that are relevant to their decisions, but the information on which they base their decisions is incomplete. A similar story can be told for any technological innovation that complements other innovations.

Such a dynamical adjustment process exhibits several features that are found in many different learning situations, whether the learning is by individuals or groups of individuals. First, the decision of each agent hinges on the actions taken by other agents. In other words it has the structure of a game. Second, an agent may know *some* of the previous actions taken by others, but there is no reason to suppose that she actually knows *all* of them. Third, while a well-informed and highly sophisticated individual might, in theory, be able to forecast how such a process is going to evolve over time, we do not want to assume that individuals are especially well-informed or highly sophisticated. We prefer to assume that they do more or less sensible things given a limited knowledge of the world around them. Finally, no matter how carefully we try to specify individuals' decision making processes, there will inevitably be some random variation in their responses that arise from unmodeled factors.

In this paper we examine a class of learning dynamics that incorporate these features. Specifically, we consider a stochastic version of fictitious play in which agents' information is incom-

*An earlier version of the paper was entitled "Dynamic Equilibrium Selection Under Incomplete Information".

plete, their payoffs functions wobble, and their choices are sometimes random. We then analyze the behavior of such process for 2×2 games. We show that, under suitable regularity conditions on the perturbations, the learning process converges with probability one to a neighborhood of a stable Nash equilibrium. In particular, if the game has a unique Nash equilibrium (pure or mixed), then it is stable and the process converges to a neighborhood of it with probability one. If on the other hand the game has exactly three Nash equilibria (two pure and one mixed), then the process converges to a neighborhood of one or both of the pure equilibria (which are stable), and with probability zero to the mixed equilibrium (which is unstable). The size of the neighborhood shrinks to zero as the probability of making random errors becomes vanishingly small.

This result is related to other recent work in evolutionary game theory and learning, particularly Fudenberg and Kreps (1993)[1]. They showed that, when agents play a 2×2 game repeatedly with slightly perturbed payoffs, then the frequency distribution of play converges with probability one to a neighborhood of the mixed strategy equilibrium *provided* that the game has a unique, completely mixed equilibrium. We show in a more general setting that convergence obtains for all 2×2 non-degenerate games whether they have pure or mixed equilibria. Moreover only the stable equilibria are attained with positive probability. In particular, if the game has exactly three equilibria – two pure and one mixed – then the former are attained with probability one and the mixed one with probability zero.

The paper proceeds as follows. In section 2 we define a stochastic version of fictitious play in which the only noise arises from incomplete information (i.e. sampling variability). This stripped-down version of the model exhibits many of the key features mentioned above, and is easy to grasp intuitively. Section 3 shows how to analyze the long-run behavior of such processes using stochastic approximation techniques (see, for example, Nevelson and Hasminskii (1976)). Unlike most other work in this area, we do not rely on Lyapunov functions to prove convergence (indeed we do not know how to construct the relevant Lyapunov functions for some classes of games). Rather we derive the relevant stability conditions for the system of differential equations using a geometric argument. Numerical simulations of fictitious play with sampling are given in section 4. We then broaden the framework in sections 5 and 7 to include other sources of noise such as random perturbations in the players' choices and in the payoff functions. In section 6 we introduce the concept of a "perturbed best reply dynamic," which covers all the above sources of noise, as well as many others. We then prove a general result concerning the almost

[1] For stochastic evolutionary game theory models of this kind see also Foster and Young (1990), Kandori, Mailath and Rob (1993), Young (1993a, 1993b), Ellison and Fudenberg (1994), Dosi and Kaniovski (1994), Posch (1994).

sure convergence of such a process to a Nash equilibrium. The rate of convergence is studied in section 8.

2 Fictitious play with sampling

Fix a two-person game G with payoff matrix

$$\begin{pmatrix} \alpha_{11},\ \beta_{11} & \alpha_{12},\ \beta_{12} \\ \alpha_{21},\ \beta_{21} & \alpha_{22},\ \beta_{22} \end{pmatrix}.$$

Assume there are two populations of agents: row players (R) and column players (C). Each of these populations consists of one or more players. In every time period $t = 1, 2, \ldots$ one pair is drawn from $R \times C$ to play the game. The *state* at t is a vector of nonnegative integers $(a_1^t, a_2^t, b_1^t, b_2^t)$, where a_1^t, a_2^t are the numbers of row players who have chosen strategies 1 and 2 respectively up to and including time t, and b_1^t, b_2^t are the numbers of column players who have chosen 1 and 2 respectively. We assume the agents selected to play the game in period $t + 1$ have incomplete information about the current state, which they gather by randomly sampling from previous actions (as in Young (1993a)). For notational simplicity we assume that all players have the same sample size s (a positive integer), though in fact our results extend to the case where players have different sample sizes. The sample size measures the extent of an agent's information, but we do not view it as the result of an optimal search. Rather, it reflects the extent to which the agent "gets around", i.e., is networked with other members of the population. We take this as exogenously given.

The process unfolds as follows. At time $t+1$ one new row player and one new column player come forward. The row player draws a subset of s actions taken so far by the column players. The total number of such actions is $b^t = b_1^t + b_2^t$. For convenience we shall assume that all samples of size s are equally likely to be drawn.

Let the random variables B_1^t, B_2^t denote the actual numbers of previous actions by column players that Row draws at time $t+1$. Row then adopts strategy 1 or 2 according as the following criterion is positive or non-positive

$$\alpha_{11} B_1^t + \alpha_{12} B_2^t - \alpha_{21} B_1^t - \alpha_{22} B_2^t. \tag{1}$$

Independently and simultaneously Column draws a subset of s previous actions by Row, the total number of such actions being $a^t = a_1^t + a_2^t$. The random variables A_1^t, A_2^t denote the number of actions of each type in the column player's sample. She then adopts strategy 1 or 2 according as the following expression is positive or non-positive

$$\beta_{11} A_1^t + \beta_{12} A_2^t - \beta_{21} A_1^t - \beta_{22} A_2^t. \tag{2}$$

These definitions yield a stochastic process of form

$$(a_1^{t+1}, a_2^{t+1}, b_1^{t+1}, b_2^{t+1}) = (a_1^t, a_2^t, b_1^t, b_2^t) + \vec{I}^t(a_1^t, a_2^t, b_1^t, b_2^t),\ t \geq 1, \tag{3}$$

where $\vec{I}^t(\cdot,\cdot,\cdot,\cdot)$ are independent random vectors that take the values $(1,0,1,0)$, $(0,1,1,0)$, $(0,1,0,1)$ and $(1,0,0,1)$ with probabilities that depend on the current state.

To analyze this process, we project it into the space of proportions of the two populations. Let $X^t = a_1^t/a^t$ and $Y^t = b_1^t/b^t$. Then there exist Bernoulli random variables $\xi^t(y) = 0$ or 1 and $\psi^t(x) = 0$ or 1 such that

$$\begin{aligned} X^{t+1} &= X^t + (1/a^{t+1})[\xi^t(Y^t) - X^t],\ t \geq 1,\ X^1 = a_1^1/a^1, \\ Y^{t+1} &= Y^t + (1/b^{t+1})[\psi^t(X^t) - Y^t],\ t \geq 1,\ Y^1 = b_1^1/b^1. \end{aligned} \tag{4}$$

These equations define two parallel or *co-evolving* process on the space $[0,1] \times [0,1]$. The two-dimensional process (X^t, Y^t) is Markovian but non-stationary because the denominators a^{t+1} and b^{t+1} depend on t. In fact we have the simple relations $a^{t+1} = t + a^1$ and $b^{t+1} = t + b^1$ because the number of actions already taken grows by one in each period. Note that the distributions of $\xi^t(\cdot)$ and $\psi^t(\cdot)$ depend on the *number* of agents in the other class (not just their *proportions*) because the sampling is without replacement. We call this process *fictitious play with sampling.*

The process can also be represented as an urn scheme. Imagine two urns R and C of infinite capacity. Each contains two colors of balls – red for strategy 1 and white for strategy 2. Initially there are a_1^1 red balls and a_2^1 white balls in the first urn. Similarly, there are b_1^1 red balls and b_2^1 white balls in the second urn. In the first period, a representative row player reaches into the second urn and pulls out s balls at random. Then he adds a red ball to the first urn if the criterion (1) is positive and adds a white ball if it is non-positive. Simultaneously and independently a representative column player reaches into the first urn and pulls out s balls at random. He then applies criterion (2) to determine what color of ball to add to the second urn. We call this a *co-evolving urn scheme.*

It can also be represented (in a more complicated way) by a single urn containing four colors of balls. At each stage $t = 1, 2, \ldots$ two balls of various colors are added according to a probability distribution that depends on t and the proportions of balls currently in the urn. Let us identify a ball of the first color with a red ball in the first urn, a ball of the second color with a white ball in the first urn, a ball of the third color with a red ball in the second urn, and a ball of the fourth color with a white ball in the second urn. Designate by x_i^t the current proportion of balls of the i-th color, $i = 1, 2, 3$. (The value x_4^t is determined by these.) Then

$$X^t = \frac{x_1^t}{x_1^t + x_2^t} \quad \text{and} \quad Y^t = \frac{x_3^t}{1 - x_1^t + x_2^t},\ t \geq 1.$$

Now we can characterize the process as follows. Add one ball of the first color and one ball of the third color if both (1) and (2) are positive. If (1) is non-positive but (2) is positive, add one ball of the second color and one of the third color, and so on. This an example of a generalized single-urn scheme with multiple additions (Arthur, Ermoliev, and Kaniovski (1987)). Unfortunately, proving convergence for such processes using the approach of Arthur, Ermoliev and Kaniovski (1987) requires the construction of a Lyapunov function, which poses difficulties in this case. Instead, we shall develop a new approach that exploits the geometry of the situation together with the qualitative theory of ordinary differential equations.

3 Asymptotic behavior of fictitious play with sampling

We begin by analyzing the situation when the only source of noise is sampling variability. Players always choose best replies given the information in their samples; there is no variability in their payoffs and no trembling. This model is easy to grasp and contains almost all of the essential features of the more general case.

Consider the following example

$$\begin{pmatrix} 2,\ 2 & 4,\ 0 \\ 1,\ 3 & 6,\ 4 \end{pmatrix}. \tag{5}$$

This game has three equilibria: $((0,1),(0,1))$, $((1,0),(1,0))$, and $((1/3,2/3),(2/3,1/3))$. To simplify notation we shall refer to these equilibria as $(0,0)$, $(1,1)$, and $(1/3,2/3)$ respectively. The direction of motion of ordinary fictitious play (which is deterministic) are shown in Figure 1. Note that each of the pure equilibria is *dynamically stable* in the sense that it is the unique limit of the deterministic process whenever the process starts in a sufficiently small neighborhood of that equilibrium. The mixed strategy equilibrium, by contrast, is dynamically *unstable.*

Consider now the stochastic process defined by (4) when agents have sample information with sample size s. Let the process begin in an arbitrary state $(a_1^1, a_2^1, b_1^1, b_2^1)$. (We have to assume here that $a_1^1 + a_2^1 \geq s$ and $b_1^1 + b_2^1 \geq s$ to be sure that the samples are feasible.) Sampling changes fictitious play in two ways: i) it creates variability around the best-reply path; ii) it creates bias in the replies because of the finiteness of the sample.

To state our result precisely, let us say that a 2×2 game G is *non-degenerate* if it has exactly three Nash equilibria (two pure and one mixed) or exactly one mixed Nash equilibrium.

Theorem 1 *Let G be a non-degenerate 2×2 game. For all sufficiently large sample sizes, fictitious play with sampling converges with probability one to a random vector (X^*, Y^*) which lies close to a stable Nash equilibrium of G. That is, for every $\epsilon > 0$ there exists a positive integer s_ϵ such that whenever $s \geq s_\epsilon$ $\lim_{t \to \infty}(X^t, Y^t) = (X^*, Y^*)$ exists with probability one, and its support lies within an ϵ-neighborhood of the stable Nash equilibria.*

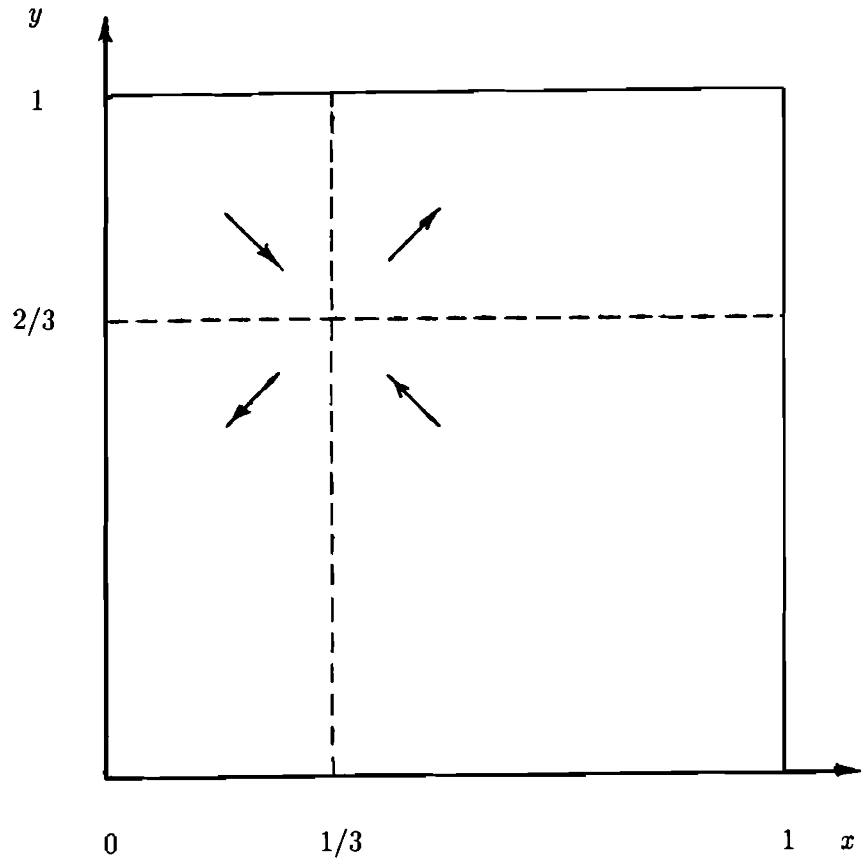

Figure 1: Direction of motion of fictitious play.

We now give the intuition behind this theorem. (A formal proof of a more general result from which this one follows is given in Appendix.) We can think of the process as fictitious play with a small noise or "wobble". The source of the wobble is the stochastic variation in the choices that agents make each period. As time runs on, each new choice counts for less and less relative to the total number of choices that have already been made. The incremental changes in the population proportions decreases as $1/t$, and so does the variability in these increments. Thus we have an annealing process in which the level of noise damps down over time. The result says that the state of the system – projected into the space of proportions – converges with probability one. Moreover, the limit of the process is precisely a fixed point of the expected motion, which is equal to (or close to) a Nash equilibrium of the game when s is large. If this equilibrium is dynamically unstable, however, then because of the perpetual wobble the process will not converge to it (except with probability zero).

All of this make sense intuitively. What remains to be shown is that the process converges almost surely, and that the only things to which it can converge (with positive probability) do in fact lie close to the stable Nash equilibria of the game. Here we shall sketch the outlines of the argument.

Let G be a non-degenerate game as above. The situation in which G has exactly three equilibria will be called "case 1" and the situation where it has a unique equilibrium will be termed “case 2”.

In both cases the formula for the mixed equilibrium (β, α) is (see, for example, Vorob'ev (1977), p.p. 99-103)

$$\beta = \frac{\beta_{22} - \beta_{21}}{\beta_{11} - \beta_{21} - \beta_{12} + \beta_{22}}, \quad \alpha = \frac{\alpha_{22} - \alpha_{12}}{\alpha_{11} - \alpha_{21} - \alpha_{12} + \alpha_{22}}.$$

Without loss of generality in case 1 we have

$$\alpha_{11} - \alpha_{21} - \alpha_{12} + \alpha_{22} > 0 \text{ and } \beta_{11} - \beta_{21} - \beta_{12} + \beta_{22} > 0,$$

while in case 2 we may assume that

$$\alpha_{11} - \alpha_{21} - \alpha_{12} + \alpha_{22} > 0 \text{ and } \beta_{11} - \beta_{21} - \beta_{12} + \beta_{22} < 0.$$

Now let us derive analytic expressions for the distributions of the Bernoulli random variables involved in (4). Assume we are in case 1. At time t , as above, X^t stands for the proportion of strategy 1 chosen so far by the row players, and Y^t the proportion of strategy 1 chosen so far by the column players. Let $\xi^t(Y^t)$ be the indicator of the random event that Row plays strategy 1 in period t and let $\psi^t(X^t)$ be the indicator of the random event that Column plays strategy 1 in period t. Let s be the sample size. The random variable B_1^t denotes the number of 1's that appear in Row's sample, while $B_2^t = s - B_1^t$ denotes the number of 2's in Row's sample. Define A_1^t and A_2^t similarly. Then we have the relations[2]

$$P\{\xi^t(Y^t) = 1\} = P\{B_1^t > \alpha s\}, \tag{6}$$

$$P\{\psi^t(X^t) = 1\} = P\{A_1^t > \beta s\}. \tag{7}$$

In case 2, (6) still applies but (7) must be replaced by

$$P\{\psi^t(X^t) = 1\} = 1 - P\{A_1^t > \beta s\}. \tag{8}$$

(This follows from the fact that in case 2, $\alpha_{11} - \alpha_{21} - \alpha_{12} + \alpha_{22} > 0$ and $\beta_{11} - \beta_{21} - \beta_{12} + \beta_{22} < 0$.)

Consider the inequality $B_1^t > \alpha s$. The probability of this event is equal to

$$\sum_{i > \alpha s}^{s} H(i; b^t, s, b_1^t),$$

where $H(i; b^t, s, b_1^t)$ is the hypergeometric distribution

$$H(i; b^t, s, b_1^t) = \frac{\begin{pmatrix} b_1^t \\ i \end{pmatrix} \begin{pmatrix} b^t - b_1^t \\ s - i \end{pmatrix}}{\begin{pmatrix} b^t \\ s \end{pmatrix}}. \tag{9}$$

[2] From these formulae it is clear why we require G to be non-degenerate. If, for example, $\alpha \notin (0,1)$ then the event $\{B_1^t > \alpha s\}$ becomes deterministic and the dynamic for X^t is also deterministic. Hence sampling does not create anything interesting for degenerate games.

To avoid the trivial case when this probability is identically 0 or 1, we have to assume that

$$b_1^1 > [\alpha s] \text{ and } b_2^1 \geq s - [\alpha s] - 1,$$

where $[\alpha s]$ designates the integer part of αs. The analogous condition for the row players is

$$a_1^1 > [\beta s] \text{ and } a_2^1 > s - [\beta s] - 1.$$

We say that the initial state is *rich* if the above inequalities hold. We shall henceforth assume this condition to avoid the less interesting case where the process is deterministic.

Since $b^t = b^1 + t - 1$ and $Y^t = b_1^t / b^t$, we can write (9) in the following form

$$\begin{pmatrix} s \\ i \end{pmatrix} \frac{Y^t(Y^t - \frac{1}{b^1+t-1})\dots(Y^t - \frac{i-1}{b^1+t-1})(1-Y^t)(1-Y^t - \frac{1}{b^1+t-1})\dots(1-Y^t-\frac{s-i-1}{b^1+t-1})}{1(1-\frac{1}{b^1+t-1})\dots(1-\frac{s-1}{b^1+t-1})}.$$

Consequently we have

$$\sum_{i>\alpha s}^{s} H(i; b^t, s, b_1^t) = f_\alpha^s(Y^t) + \delta_\alpha^s(b^t, Y^t),$$

where

$$f_\alpha^s(y) = \sum_{i>\alpha s}^{s} \begin{pmatrix} s \\ i \end{pmatrix} y^i (1-y)^{s-i}, \tag{10}$$

and

$$\sup_{y\in[0,1]} |\delta_\alpha^s(n, y)| \leq c_{\alpha,s}/n. \tag{11}$$

We can sum up these observations in the following lemma.

Lemma 1 *Let G be a non-degenerate 2×2 game with mixed Nash equilibrium (β,α). For all $(x,y)\in[0,1]\times[0,1]$,*

$$\text{case 1}: \; P\{\xi^t(y)=1\} = f_\alpha^s(y) + \delta_\alpha^s(b^t, y), \quad P\{\psi^t(x)=1\} = f_\beta^s(x) + \delta_\beta^s(a^t, x),$$

$$\text{case 2}: \; P\{\xi^t(y)=1\} = f_\alpha^s(y) + \delta_\alpha^s(b^t, y), \quad P\{\psi^t(x)=1\} = 1 - f_\beta^s(x) - \delta_\beta^s(a^t, x),$$

where the functions involved are given by (10) and (11).

Define $\Xi^t(y) = \xi^t(y) - E\xi^t(y)$ and $\Psi^t(y) = \psi^t(y) - E\psi^t(y)$ (here E designates mathematical expectation), and rewrite (4) in the following way:

case 1 :

$$X^{t+1} = X^t + (1/a^{t+1})\{[f_\alpha^s(Y^t) + \delta_\alpha^s(b^t, Y^t) - X^t] + \Xi^t(Y^t)\}, \; t\geq 1, \; X^1 = a_1^1/a^1,$$

$$Y^{t+1} = Y^t + (1/b^{t+1})\{[f_\beta^s(X^t) + \delta_\beta^s(a^t, X^t) - Y^t] + \Psi^t(X^t)\}, \; t\geq 1, \; Y^1 = b_1^1/b^1;$$

case 2 :

$$X^{t+1} = X^t + (1/a^{t+1})\{[f^s_\alpha(Y^t) + \delta^s_\alpha(b^t, Y^t) - X^t] + \Xi^t(Y^t)\},\ t \geq 1,\ X^1 = a^1_1/a^1,$$

$$Y^{t+1} = Y^t + (1/b^{t+1})\{[1 - f^s_\beta(X^t) - \delta^s_\beta(a^t, X^t) - Y^t] + \Psi^t(X^t)\},\ t \geq 1,\ Y^1 = b^1_1/b^1.$$

These equations define a two-dimensional stochastic approximation procedure (see, for example, Nevelson and Hasminskii (1976)).

Suppose that at time t the process is at the point (x, y). Then the *expected motion* is as follows

$$\text{case 1}:\ (x,y) \mapsto (x + (1/a^{t+1})[f^s_\alpha(y) + \delta^s_\alpha(b^t, y) - x], y + (1/b^{t+1})[f^s_\beta(x) + \delta^s_\beta(a^t, x) - y]);$$

$$\text{case 2}:\ (x,y) \mapsto (x + (1/a^{t+1})[f^s_\alpha(y) + \delta^s_\alpha(b^t, y) - x], y + (1/b^{t+1})[1 - f^s_\beta(x) - \delta^s_\beta(a^t, x) - y]).$$

Since $a^{t+1} = a^1 + t$ and $b^{t+1} = b^1 + t$, we might reasonably conjecture that, as $t \to \infty$, this process behaves like the system of ordinary differential equations

$$\begin{aligned} &\text{case 1}:\ \dot{x} = f^s_\alpha(y) - x, \quad \dot{y} = f^s_\beta(x) - y; \\ &\text{case 2}:\ \dot{x} = f^s_\alpha(y) - x, \quad \dot{y} = 1 - f^s_\beta(x) - y. \end{aligned} \tag{12}$$

(By (11) we have neglected terms of order t^{-1}.) This conjecture turns out to be correct. (The argument is given in Appendix.)

Assuming this holds, we can now see that the stationary points of (12) determine the possible limits of these systems, and these are precisely the solutions of

$$\begin{aligned} &\text{case 1}:\ x = f^s_\alpha(y), \quad y = f^s_\beta(f^s_\alpha(y)); \\ &\text{case 2}:\ x = f^s_\alpha(y), \quad y = 1 - f^s_\beta(f^s_\alpha(y)). \end{aligned} \tag{13}$$

The next step is to show that the solutions of (13) are close to the Nash equilibria of the game. Consider $f^s_\alpha(\cdot)$ for a fixed α and variable s. As s increases, $f^s_\alpha(\cdot)$ becomes more and more S-shaped and approaches the step function (see Figure 2)

$$f^\infty_\alpha(x) = \begin{cases} 1 & \text{if } x \in (\alpha, 1), \\ 1/2 & \text{if } x = \alpha, \\ 0 & \text{if } x \in (0, \alpha). \end{cases}$$

We may state this result more exactly as follows. Say that a function $f(\cdot) : [0,1] \mapsto [0,1]$ is *convex-concave* if for some $z \in (0,1)$, $f(\cdot)$ is convex on $[0, z)$ and concave on $(z, 1]$.

Lemma 2 *For every* $\alpha \in (0,1)$

i) $\lim_{s\to\infty} f^s_\alpha(x) = f^\infty_\alpha(x)$ *for all* $x \in [0,1]$*;*

ii) $f^s_\alpha(x)$ *is strictly increasing in* x*, continuously differentiable, and for large* s *it is convex-concave.*

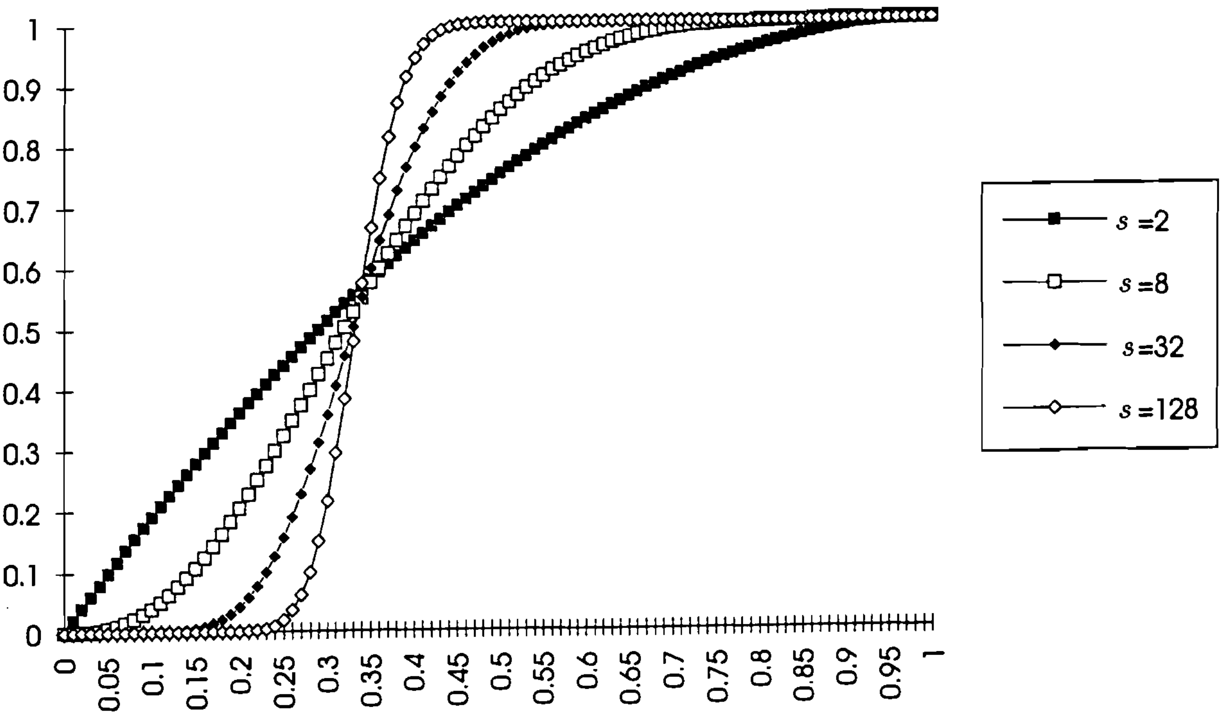

Figure 2: The function $f^s_{1/3}(\cdot)$ for $s = 2, 8, 32, 128$.

Proof. The value $f^s_\alpha(x)$ can be thought as

$$P\{\sum_{k=1}^{s} a_k \geq \alpha s\},$$

where a_k are independent Bernoulli random variables satisfying

$$a_k = \begin{cases} 1 & \text{with probability} \quad x, \\ 0 & \text{with probability} \quad 1 - x. \end{cases}$$

Set $b_k = a_k - x$. Then the above probability equals

$$P\{\sum_{k=1}^{s} b_k \geq (\alpha - x)s\}.$$

From the law of large numbers it follows that, in probability,

$$\lim_{s\to\infty} \frac{1}{s} \sum_{k=1}^{s} b_k = 0.$$

Consequently,

$$\lim_{s\to\infty} f^s_\alpha(x) = \lim_{s\to\infty} P\{\frac{1}{s}\sum_{k=1}^{s} b_k \geq \alpha - x\} = 0 \text{ for } \alpha > x,$$

and

$$\lim_{s\to\infty} f^s_\alpha(x) = \lim_{s\to\infty} P\{\frac{1}{s}\sum_{k=1}^{s} b_k \geq \alpha - x\} = 1 \text{ for } \alpha < x.$$

Finally, $f^\infty_\alpha(\alpha) = 1/2$ by symmetry. This proves statement i).

To prove statement ii) let us observe that $f^s_\alpha(x)$ can be expressed in another way. Namely, $f^s_\alpha(x)$ is the probability that $k = [\alpha s]$ of the a_i, $i = 1, 2, \ldots, s$, are less than or equal to x. Thus $f^s_\alpha(x)$ is the distribution function of the k-th order statistic of the $a_i, i = 1, 2, \ldots, s$. It follows from standard arguments (see, for example, David (1981), p. 9) that the corresponding density is given by

$$\frac{d}{dx} f^s_\alpha(x) = \frac{s!}{(k-1)!(s-k)!} x^{k-1}(1-x)^{s-k}.$$

It is easily checked that its derivative is positive on $[0, (k-1)/(s-1))$ and negative on $((k-1)/(s-1), 1]$. Hence $f^s_\alpha(\cdot)$ is convex-concave; moreover it is continuously differentiable and increasing. This completes the proof of Lemma 2.

Lemma 3 *Let G be non-degenerate with unique mixed equilibrium (β, α).*

Case 1: for all sufficiently large s, (13) has exactly three solutions: $(0,0)$, $(1,1)$, (x_s, y_s), *and* $\lim_{s\to\infty}(x_s, y_s) = (\beta, \alpha)$;

case 2: for each s, (13) has exactly one solution (x'_s, y'_s), *and* $\lim_{s\to\infty}(x'_s, y'_s) = (\beta, \alpha)$.

Proof. We shall give the argument for case 1; the other case is similar. Fix a small $\epsilon > 0$. By statement i) of Lemma 2 there is an integer s_ϵ such that, for all $s \geq s_\epsilon$

$$f^s_\beta(\beta - \epsilon) \leq \epsilon, \quad f^s_\beta(\beta + \epsilon) \geq 1 - \epsilon,$$

and

$$f^s_\alpha(\alpha - \epsilon) \leq \epsilon, \quad f^s_\alpha(\alpha + \epsilon) \geq 1 - \epsilon.$$

The curves $f^s_\beta(\cdot)$ and $f^s_\alpha(\cdot)$ intersect at $(0,0)$ and $(1,1)$. By choice of ϵ the only other point of intersection (if any) must be inside the box $B_\epsilon = \{(x,y) : \beta - \epsilon \leq x \leq \beta + \epsilon,\ \alpha - \epsilon \leq y \leq \alpha + \epsilon\}$. Moreover, since the functions are continuous, by statement i) of Lemma 2 they do have at least one point of intersection in B_ϵ for $s \geq s_\epsilon$. We shall show that the intersection is unique.

Let x_f be the first x such that $(x, f^s_\beta(x))$ is in the box B_ϵ. Then there is $x' \in (\beta - \epsilon, x_f)$ such that

$$f^s_\beta(x_f) - f^s_\beta(b - \epsilon) = \frac{d}{dx} f^s_\beta(x')(x_f - b + \epsilon),$$

and hence

$$\frac{d}{dx} f^s_\beta(x') \geq \frac{\alpha - 2\epsilon}{2\epsilon}.$$

(If this were not the case, the curve could not reach the box.) Similarly,

$$\frac{d}{dx} f^s_\beta(x'') \geq \frac{1 - \alpha - 2\epsilon}{2\epsilon}$$

for some $x'' \in (x_l, \beta + \epsilon)$, where x_l is the last x such that $(x, f^s_\beta(x))$ is in the box.

By statement ii) of Lemma 2, $\frac{d}{dx}f^s_\beta(\cdot)$ is first increasing, then decreasing. Hence

$$\frac{d}{dx}f^s_\beta(x) \geq \min(\frac{\alpha-2\epsilon}{2\epsilon}, \frac{1-\alpha-2\epsilon}{2\epsilon})$$

for $x \in [x_f, x_l]$. Consequently for sufficiently small ϵ and all $s \geq s_\epsilon$, we have

$$\frac{d}{dx}f^s_\beta(x) > 1 \text{ for all } (x, f^s_\beta(x)) \in B_\epsilon.$$

Similarly,

$$\frac{d}{dy}f^s_\alpha(y) > 1 \text{ for all } (y, f^s_\alpha(y)) \in B_\epsilon.$$

Suppose now that the curves of $f^s_\beta(\cdot)$ and $f^s_\alpha(\cdot)$ intersect in two distinct points $(x_1, y_1), (x_2, y_2) \in B_\epsilon$. Then there is a point $(x^0, f^s_\beta(x^0)) \in B_\epsilon$ such that $\frac{d}{dx}f^s_\beta(x^0) = (y_2 - y_1)/(x_2 - x_1)$ and a point $(f^s_\alpha(y^0), y^0) \in B_\epsilon$ such that $\frac{d}{dy}f^s_\alpha(y^0) = (x_2 - x_1)/(y_2 - y_1)$. Since at least one of these is less than or equal to 1, we have arrived at a contradiction. This proves the uniqueness of the interior intersection point (x_s, y_s) for all sufficiently large s, and that $(x_s, y_s) \to (\beta, \alpha)$ as $s \to \infty$. Thus Lemma 3 is proved.

From this it follows that if fictitious play with sampling converges, then it converges to a solution of (13), which is either a Nash equilibrium of G or close to one. The proof of convergence is given in Appendix, where we also show that for case 1 games the limit is almost surely *not* contained in every sufficiently small neighborhood of the unstable Nash equilibrium. At this point let us also note that, for case 1 games, *both* of the pure equilibria are attained in the limit with positive probability if the initial state is rich. Second, the theorem remains valid if even the sample sizes are different, so long as the smallest sample size is sufficiently large.

4 Simulations

To illustrate these ideas concretely, let G have the payoff matrix (5). This game has three equilibria: $(0,0)$, $(1,1)$ and $(1/3, 2/3)$. Let the sample size be 7 and suppose that the process begins in state $(4,4,4,4)$. Figure 3 shows the distributions of X^t and Y^t for $t = 200$, while Figure 4 gives the same distributions for $t = 600$. Compare this with Figure 5, which demonstrates the same distributions when the initial state is $(4,6,4,4)$, $s = 7$ and $t = 200$. In both cases convergence occurs to $(0,0)$ and $(1,1)$ with positive probability, but the value of the probability depends on the initial state.

As a second example, consider the following game

$$\begin{pmatrix} 2,\ 2 & 4,\ 4 \\ 1,\ 1 & 6,\ 0 \end{pmatrix}.$$

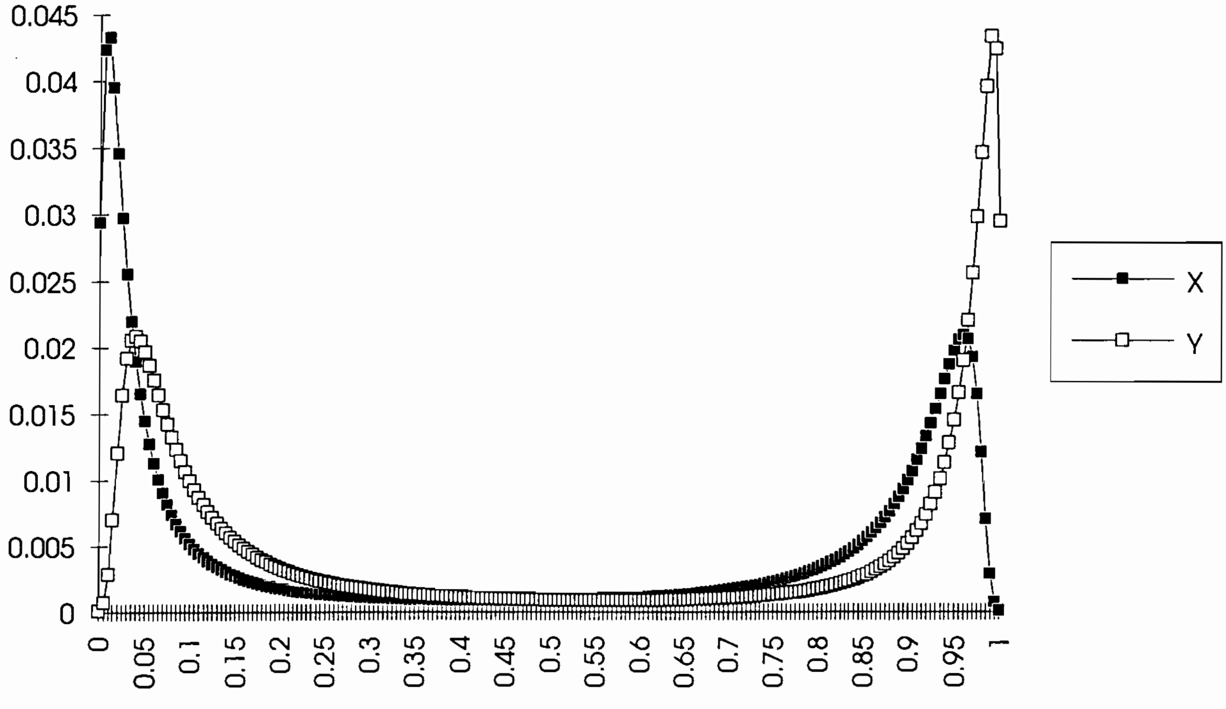

Figure 3: Distributions of X^t and Y^t for case 1, $t = 200$, $s = 7$ and $a_1^1 = a_2^1 = b_1^1 = b_2^1 = 4$.

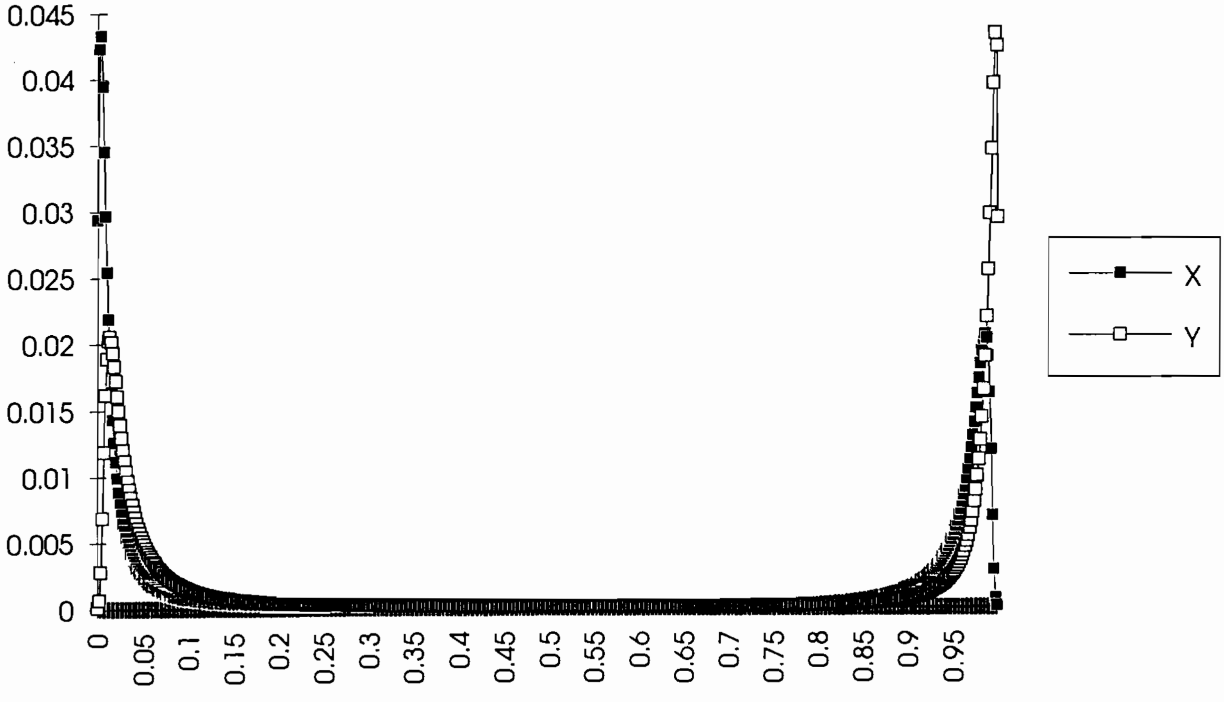

Figure 4: Distributions of X^t and Y^t for case 1, $t = 600$, $s = 7$ and $a_1^1 = a_2^1 = b_1^1 = b_2^1 = 4$.

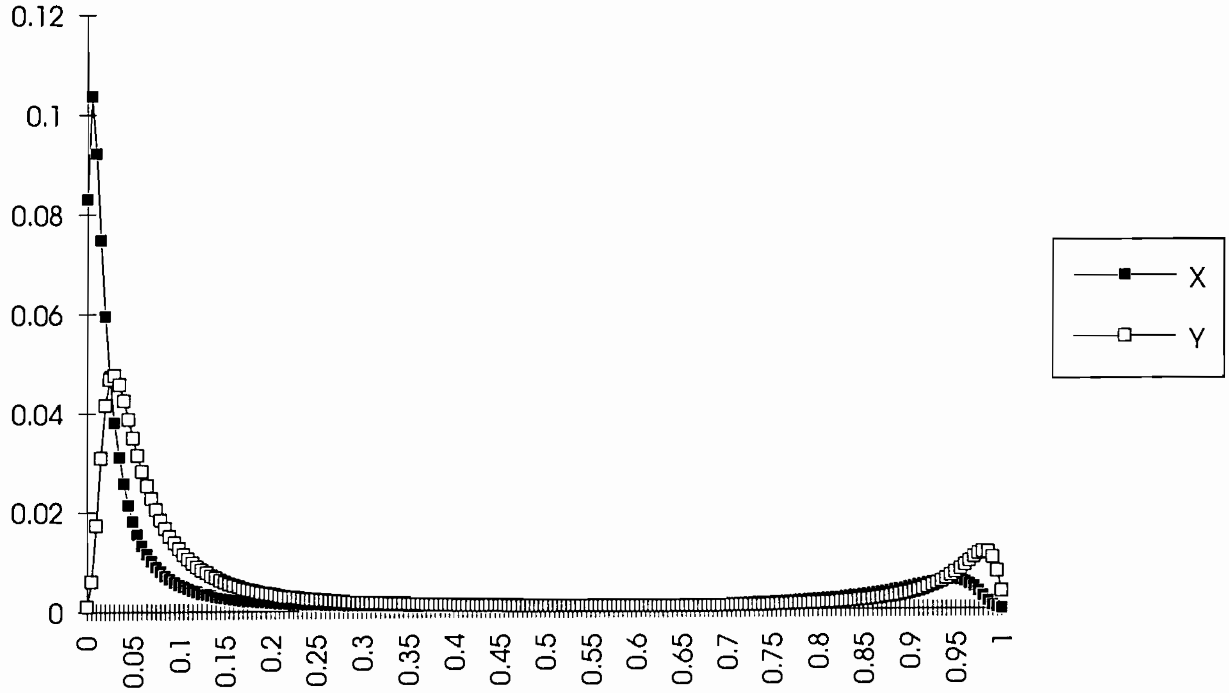

Figure 5: Distributions of X^t and Y^t for case 1, $t = 200$, $s = 7$ and $a_1^1 = b_1^1 = b_2^1 = 4, a_2^1 = 6$.

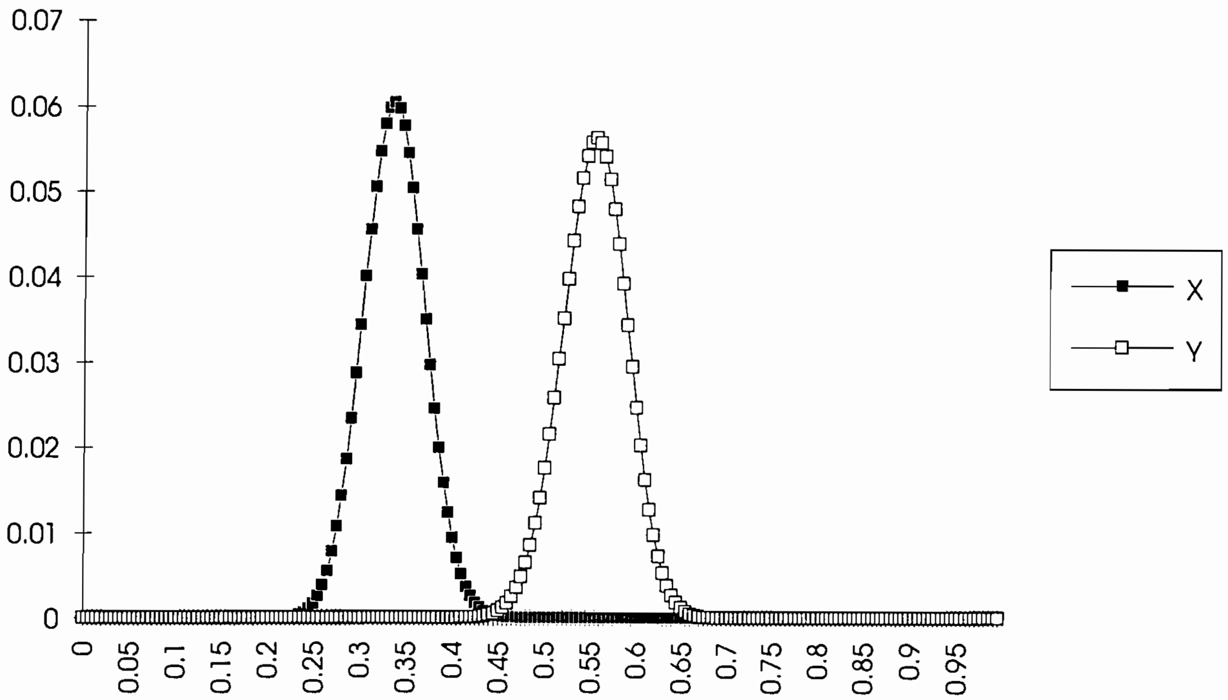

Figure 6: Distributions of X^t and Y^t for case 2, $t = 200$, $s = 7$ and $a_1^1 = a_2^1 = b_1^1 = b_2^1 = 4$.

The unique Nash equilibrium is $(1/3, 2/3)$, so this is a case 2 game. Figure 6 shows the distributions of X^t and Y^t for $t = 200$, where $s = 7$ and the process starts at $(4,4,4,4)$. One can discern some bias (due to small sample size) from the theoretical value $(1/3, 2/3)$.

5 Mistakes and other source of noise

In this section we relax the assumption that agents always choose best replies by supposing that they sometimes make "mistakes", that is, they do things for unexplained reasons. To make this idea concrete, assume that Row samples from the previous actions by Column and chooses a best reply with probability $1 - 2\epsilon$, but with probability 2ϵ she chooses 1 or 2 at random. For notational convenience we shall assume that Row selects 1 and 2 with equal probability, though all that really matters is that she chooses both 1 and 2 with fixed positive probabilities whenever she randomizes. The column player acts similarly.

In place of the expressions in Lemma 1 we obtain the following

$$\text{case 1}: \quad \begin{array}{l} P\{\xi^t(y) = 1\} = (1-2\epsilon)[f_\alpha^s(y) + \delta_\alpha^s(b^t, y)] + \epsilon, \\ P\{\psi^t(x) = 1\} = (1-2\epsilon)[f_\beta^s(x) + \delta_\beta^s(a^t, x)] + \epsilon; \end{array}$$

$$\text{case 2}: \quad \begin{array}{l} P\{\xi^t(y) = 1\} = (1-2\epsilon)[f_\alpha^s(y) + \delta_\alpha^s(b^t, y)] + \epsilon, \\ P\{\psi^t(x) = 1\} = (1-2\epsilon)[1 - f_\beta^s(x) - \delta_\beta^s(a^t, x)] + \epsilon. \end{array}$$

Consider case 1. Let $f_\alpha^{s,\epsilon}(y) = (1-2\epsilon)f_\alpha^s(y) + \epsilon$, and $f_\beta^{s,\epsilon}(x) = (1-2\epsilon)f_\beta^s(x) + \epsilon$. An argument similar to the preceding shows that if the stochastic process defined by (4) converges, then it converges to a solution of the system

$$x = f_\alpha^{s,\epsilon}(y), \quad y = f_\beta^{s,\epsilon}(f_\alpha^{s,\epsilon}(y)).$$

The situation is similar to the earlier one, since for sufficiently small ϵ the curves $f_\alpha^{s,\epsilon}(\cdot)$ and $f_\beta^{s,\epsilon}(\cdot)$ become more and more S-shaped as s increases. Hence they cross at exactly three points. The crossing points are the only possible limits of the stochastic process, but the interior crossing point is unstable, and the process converges to it with probability zero. In case 1, the stable crossing points are close to the two pure Nash equilibria (when ϵ is small and s is large), hence the process converges almost surely to a neighborhood of the pure Nash equilibria. In case 2, there is a unique crossing point, and for all sufficiently small ϵ the process converges to a neighborhood of the mixed Nash equilibrium of the game when s is sufficiently large.

6 Perturbed best reply dynamics

We now formulate a general model that captures the preceding examples as special cases. Let G be a non-degenerate 2×2 game with mixed Nash equilibrium (β, α). We say that the co-evolving process given by equations (4) is a *perturbed best reply dynamic* of G if it can be written as a slightly perturbed step function plus a term that goes to zero at least as fast as $t^{-\gamma}$ for some $\gamma > 0$. More precisely, suppose that for small $\epsilon, \delta > 0$, and positive integers k, s there are functions $q_\alpha^{k,\epsilon}(\cdot), g_\beta^{s,\delta}(\cdot), \nu_\alpha^{k,\epsilon}(t,\cdot), \mu_\beta^{s,\delta}(t,\cdot)$ such that for some choice of the parameters the Bernoulli random variables $\xi^t(\cdot)$ and $\psi^t(\cdot)$ can be represented in the following form

$$\text{case 1}: \quad \begin{array}{l} P\{\xi^t(y)=1\} = q_\alpha^{k,\epsilon}(y) + \nu_\alpha^{k,\epsilon}(t,y), \\ P\{\psi^t(x)=1\} = g_\beta^{s,\delta}(x) + \mu_\beta^{s,\delta}(t,x); \end{array}$$

$$\text{case 2}: \quad \begin{array}{l} P\{\xi^t(y)=1\} = q_\alpha^{k,\epsilon}(y) + \nu_\alpha^{k,\epsilon}(t,y), \\ P\{\psi^t(x)=1\} = 1 - g_\beta^{s,\delta}(x) - \mu_\beta^{s,\delta}(t,x). \end{array}$$

We assume that:

i) the functions $q_\alpha^{k,\epsilon}(\cdot)$ and $g_\beta^{s,\delta}(\cdot)$ are strictly increasing, continuously differentiable and convex-concave;

ii)

$$\lim_{\min(k,1/\epsilon)\to\infty} q_\alpha^{k,\epsilon}(y) = \begin{cases} 0 & \text{if} \quad y \in [0,\alpha), \\ 1 & \text{if} \quad y \in (\alpha,1], \end{cases}$$

$$\lim_{\min(s,1/\delta)\to\infty} g_\beta^{s,\delta}(x) = \begin{cases} 0 & \text{if} \quad x \in [0,\beta), \\ 1 & \text{if} \quad x \in (\beta,1]; \end{cases}$$

iii) $\sup_{y\in[0,1]} \nu_\alpha^{k,\epsilon}(t,y) \le c_{k,\epsilon} t^{-\gamma}$ and $\sup_{x\in[0,1]} \mu_\beta^{s,\delta}(t,x) \le c_{s,\delta} t^{-\gamma}$ for some $\gamma > 0$.

Theorem 2 *Let G be a non-degenerate 2×2 game. Every perturbed best reply dynamic converges with probability one to a neighborhood of the stable Nash equilibria of G. That is, for every $\sigma > 0$ there exist $\epsilon', \delta', k', s'$ such that for all $\epsilon \le \epsilon', \delta \le \delta'$ and $k \ge k', s \ge s'$ the process (4) converges with probability one to a pair of strategies that lies within a σ-neighborhood of the set of stable Nash equilibria of G.*

The proof is given in Appendix. We remark that, in case 1, the curves $q_\alpha^{k,\epsilon}(\cdot)$ and $g_\beta^{s,\delta}(\cdot)$ cross at exactly three points: (x_0, y_0), (x_β, y_α) and (x_1, y_1). One of them, (x_β, y_α), lies close to the mixed equilibrium and is unstable, so the process converges to it with zero probability. The other two lie close to (or coincide with) $(0,0)$ and $(1,1)$. It can be shown that the process (4) *converges to each of them with positive probability* as long as $P\{\xi^t(y) = 1\} \in (0,1)$ and $P\{\psi^t(x) = 1\} \in (0,1)$ for all $y \in (0,1)$, $x \in (0,1)$ and $t \ge 1$. Similarly, in case 2 there is a unique crossing point, which lies close to the unique equilibrium of the game, and the process converges to it with probability one.

7 Perturbed payoffs

In this section we shall show how slight perturbations in the players' payoffs can also be treated within this framework. (This is the case treated by Fudenberg and Kreps (1993).) Assume that in each period t the payoffs α_{ij}^t, β_{ij}^t are random variables. Specifically let us assume that

$$\alpha_{ij}^t = \alpha_{ij} + e_{ij}^t(\epsilon), \quad \beta_{ij}^t = \beta_{ij} + e_{ij}^t(\delta),$$

where α_{ij}, β_{ij} are the mean payoffs and $e_{ij}^t(\epsilon)$, $e_{ij}^t(\delta)$ are independent random variables with mean zero and distribution functions $F_\epsilon(\cdot)$ and $R_\delta(\cdot)$. We want the errors to be "small" as $\epsilon \to 0$ (or $\delta \to 0$); the most straightforward way to ensure this is to assume that ϵ and δ are the standard deviations of the corresponding distributions. For simplicity of exposition we shall assume that $F_\epsilon(\cdot)$ and $R_\delta(\cdot)$ are normal, though it will be apparent that the analysis holds much more generally. We shall also assume for convenience that players have full information about past actions by the other side. Sampling, mistakes, and other perturbations can be included without substantially changing the conclusions.

Our task is to show that this process satisfies the conditions of a perturbed best reply dynamic. Assume that we have a case 1 game, and let us focus on the row player. Suppose that, at some period t, the proportion of column players choosing strategy 1 is y. Then the row player chooses 1 with probability

$$q_\alpha^\epsilon(y) = P\{\alpha_{22}^t - \alpha_{21}^t < (\alpha_{11}^t - \alpha_{12}^t - \alpha_{21}^t + \alpha_{22}^t)y\} =$$

$$P\{(1-y)e_{22}^t(\epsilon) - (1-y)e_{21}^t(\epsilon) + ye_{12}^t(\epsilon) - ye_{11}^t(\epsilon) < (\alpha_{11} - \alpha_{12} - \alpha_{21} + \alpha_{22})y - (\alpha_{22} - \alpha_{21})\}.$$

Let $a = \alpha_{11} - \alpha_{12} - \alpha_{21} + \alpha_{22}$, so that $\alpha = (\alpha_{22} - \alpha_{21})a^{-1}$. Let us also recall that $a > 0$. We have

$$q_\alpha^\epsilon(y) = P\{N(0,1) < \frac{a}{\epsilon}\frac{y-\alpha}{\sqrt{2[(1-y)^2+y^2]}}\},$$

where $N(0,1)$ is a normal random variable with zero mean and variance 1. In particular,

$$N(0,1) = \frac{(1-y)e_{22}^t(\epsilon) - (1-y)e_{21}^t(\epsilon) + ye_{12}^t(\epsilon) - ye_{11}^t(\epsilon)}{\epsilon\sqrt{2[(1-y)^2+y^2]}}.$$

Clearly,

$$\lim_{\epsilon\to 0} q_\alpha^\epsilon(y) = 0 \text{ if } y < \alpha \text{ and } \lim_{\epsilon\to 0} q_\alpha^\epsilon(y) = 1 \text{ if } y > \alpha.$$

It remains to be shown that, for all sufficiently small ϵ, $q_\alpha^\epsilon(\cdot)$ is increasing and convex-concave.

Let

$$h(y) = a\frac{y-\alpha}{\sqrt{2[(1-y)^2+y^2]}}.$$

Then

$$\frac{d}{dy}q_\alpha^\epsilon(\cdot) = \frac{1}{\epsilon\sqrt{2\pi}}\exp[-\frac{h^2(\cdot)}{2\epsilon^2}]h'(\cdot).$$

A straightforward calculation shows that $h'(y) > 0$ for all $y \in [0,1]$, hence $q_\alpha^\epsilon(\cdot)$ is increasing. To demonstrate that $q_\alpha^\epsilon(\cdot)$ is convex-concave we need to show that $\frac{d^2}{dy^2}q_\alpha^\epsilon(\cdot)$ is first positive then negative. Now

$$\frac{d^2}{dy^2}q_\alpha^\epsilon(\cdot) = \frac{1}{\epsilon\sqrt{2\pi}}\exp[-\frac{h^2(\cdot)}{2\epsilon^2}]\{h''(\cdot) - \frac{1}{\epsilon^2}h(\cdot)[h'(\cdot)]^2\}.$$

Let us examine the behavior of

$$H_\epsilon(\cdot) = h''(\cdot) - \frac{1}{\epsilon^2}h(\cdot)[h'(\cdot)]^2.$$

Since $h(y)(y-\alpha) > 0$ for $y \neq \alpha$ and $h''(\cdot)$ is bounded on $[0,1]$, for all small enough ϵ the function $H_\epsilon(\cdot)$ changes sign over $[0,1]$ from plus to minus. Hence there is at least one point y_ϵ where $H_\epsilon(y_\epsilon) = 0$. Moreover, this point must belong to a neighborhood of α, which shrinks to zero as $\epsilon \to 0$. Let us show that the root is unique.

If there were more than one such root, there would exist y'_ϵ (in the same neighborhood) such that $H'_\epsilon(y'_\epsilon) = 0$. But

$$H'_\epsilon(\cdot) = h'''(\cdot) - \frac{1}{\epsilon^2}[h'(\cdot)]^3 - \frac{2}{\epsilon^2}h(\cdot)h'(\cdot)h''(\cdot).$$

Since $h(y'_\epsilon) \to 0$ as $\epsilon \to 0$, the sign of this derivative coincides with the sign of $-[h'(\cdot)]^3$, which is negative. Consequently $H'_\epsilon(\cdot)$ cannot have zeros in a sufficiently small neighborhood of α, hence y_ϵ is unique.

Thus we have established that there is a unique point y_ϵ such that $H_\epsilon(y) > 0$ for $y \in [0, y_\epsilon)$ and $H_\epsilon(y) < 0$ for $y \in (y_\epsilon, 1]$. This implies that $q_\alpha^\epsilon(\cdot)$ is convex-concave.

We have proved that $q_\alpha^\epsilon(\cdot)$ has all the properties of a perturbed best reply dynamic for the row player in a case 1 game. The other cases are verified similarly.

Note that

$$q_\alpha^\epsilon(0) = P\{N(0,1) < \frac{a\alpha}{\epsilon\sqrt{2}}\} > 0 \text{ and } q_\alpha^\epsilon(1) = P\{N(0,1) < \frac{a(1-\alpha)}{\epsilon\sqrt{2}}\} < 1.$$

Similarly $g_\beta^\delta(0) > 0$ and $g_\beta^\delta(1) < 1$, where $g_\beta^\delta(x)$ stands for the probability that Column chooses strategy 1 given that the proportion of row players choosing 1 is x. The inequalities imply that in case 1 there are two limits (x_0, y_0) and (x_1, y_1) for the random process generated by the perturbed best reply dynamic, which *do not coincide* with $(0,0)$ and $(1,1)$, though they approach these values as $\max(\epsilon, \delta) \to 0$.

From this and Theorem 2 it follows that, *when the variance (standard deviation) of perturbations is sufficiently small, fictitious play with normally perturbed payoffs converges almost surely* *to a neighborhood of the stable Nash equilibria of the game, and the size of the neighborhood goes to zero with the variance.*

A similar argument holds for a wide class of perturbations. For example, we could consider bounded perturbations that are uniformly distributed over some finite interval. In this case the corresponding functions $q_\alpha^\epsilon(\cdot)$ and $g_\beta^\delta(\cdot)$ attain the values 0 and 1 for *all* sufficiently small ϵ and δ, hence (for case 1 games) the dynamic converges to the pure equilibria *exactly.*

8 Rate of convergence

We conclude by studying the rate of convergence of perturbed best reply dynamics. As we have already mentioned, (4) can be thought as a two-dimensional stochastic approximation procedure. The central limit theorem for stochastic approximation (see Fabian (1968) or Nevelson and Hasminskii (1976)) shows that that when the process has a unique limit, then the deviation from the limit is approximately $\vec{N}(\vec{0}, t^{-1}K)$ as $t \to \infty$, where $\vec{N}(\vec{0}, t^{-1}K)$ stands for a normal distribution with mean zero and variance matrix $t^{-1}K$. Consequently for case 2 games, where the limit is a singleton, we can apply this result directly to our processes and conclude that that they converge at the rate $1/\sqrt{t}$, which is the convergence rate for most estimation processes in statistics. In case 1 there are two points that are attainable with positive probability in the limit. In this situation we need to use conditional limit theorems (see Arthur, Ermoliev and Kaniovski (1987)). These results imply the following.

Theorem 3 *Let G be a non-degenerate 2×2 game. Every perturbed best reply dynamic converges to its limit at rate at least $1/\sqrt{t}$ as $t \to \infty$. More precisely, there exist $\epsilon', \delta', k', s'$ such that for all $\epsilon \leq \epsilon', \delta \leq \delta'$, $k \geq k', s \geq s'$ and $\vec{z} \in R^2$*

$$\lim_{t\to\infty} P\{\sqrt{t}[(X^t, Y^t) - (x, y)] < \vec{z}, \lim_{m\to\infty}(X^m, Y^m) = (x, y)\} =$$

$$P\{\vec{N}(\vec{0}, K(x, y)) < \vec{z}\} P\{\lim_{m\to\infty}(X^m, Y^m) = (x, y)\}.$$

Here (x, y) can be any of the points to which our process converges with positive probability, and $\vec{N}(\vec{0}, K(x, y))$ is a two-dimensional normal random vector with mean zero and variance matrix

$$K(x, y) = \int_0^\infty \exp\{[J(x, y) + \frac{1}{2}I]t\} B(x, y) \exp\{[J(x, y)^T + \frac{1}{2}I]t\} dt,$$

where the sign ^{T} designates transposition, I stands for the identity matrix in R^2, and

$$B(x, y) = \begin{pmatrix} x(1-x) & 0 \\ 0 & y(1-y) \end{pmatrix},$$

$$\text{case 1}: \quad J(x, y) = \begin{pmatrix} -1 & \frac{d}{dy} q_\alpha^{\epsilon,k}(y) \\ \frac{d}{dx} g_\beta^{\delta,s}(x) & -1 \end{pmatrix},$$

$$\text{case 2}: \quad J(x,y) = \begin{pmatrix} -1 & \frac{d}{dy} q_\alpha^{\epsilon,k}(y) \\ -\frac{d}{dx} g_\beta^{\delta,s}(x) & -1 \end{pmatrix}.$$

This theorem is proved in the Appendix.

Let us consider some implications of this result on the convergence of best reply dynamics. For case 1 games, if a limit coincides with a *pure* equilibrium, then Theorem 3 gives a zero variance matrix for the limiting normal distribution. This indicates that the actual convergence rate is faster than $1/\sqrt{t}$. In fact, if the payoff perturbations are *bounded*, and these are the only perturbations, then it can be shown (we shall not do so here) that convergence occurs faster than $1/t^{\tau-1}$ for *every* small $\tau > 0$. A similar result holds if the perturbations arise solely from sampling.

For case 2 games, by contrast, under both types of perturbations convergence occurs only at rate $1/\sqrt{t}$, and the limit lies only in a neighborhood of the mixed Nash equilibrium. (The same holds for case 1 games if agents make *systematic* random errors.) In other words, sampling variability and payoff perturbations lead to a much faster rate of learning when the stable equilibria are pure as opposed to mixed. When the players make small systematic errors, the rates are relatively slow in both cases. We suspect that similar results hold for a wide class of games, but for the moment this remains a matter of conjecture.

Acknowledgments. We thank Dr. Luigi Marengo for providing the numerical simulations given in the paper.

International Institute for Applied Systems Analysis, A-2361, Laxenburg, Austria

Department of Economics, Johns Hopkins University, Baltimore, MD 21218, U.S.A.

and

International Institute for Applied Systems Analysis, A-2361, Laxenburg, Austria

Appendix

This appendix gives proofs of Theorems 2 and 3.

Theorem 2 *Let G be a non-degenerate 2×2 game. Every perturbed best reply dynamic converges with probability one to a neighborhood of the stable Nash equilibria of G. That is, for every $\sigma > 0$ there exist $\epsilon', \delta', k', s'$ such that for all $\epsilon \leq \epsilon', \delta \leq \delta'$ and $k \geq k', s \geq s'$ the process (4) converges with probability one to a pair of strategies that lies within a σ-neighborhood of the set of stable Nash equilibria of G.*

Proof. Consider the system of ordinary differential equations

$$\begin{aligned} &\text{case } 1: \quad \dot{x} = q_\alpha^{\epsilon,k}(y) - x, \quad \dot{y} = g_\beta^{\delta,k}(x) - y, \\ &\text{case } 2: \quad \dot{x} = g_\alpha^{\epsilon,k}(y) - x, \quad \dot{y} = 1 - g_\beta^{\delta,s}(x) - y. \end{aligned} \tag{a1}$$

Since

$$\frac{\partial}{\partial x}[q_\alpha^{\epsilon,k}(y) - x] + \frac{\partial}{\partial y}[g_\beta^{\delta,s}(x) - y] = -2$$

and

$$\frac{\partial}{\partial x}[q_\alpha^{\epsilon,k}(y) - x] + \frac{\partial}{\partial y}[1 - g_\beta^{\delta,s}(x) - y] = -2,$$

Dulac's criterion (see Hahn (1967), p. 66) says that neither of equations $(a1)$ has cycles. From the theorem of Bendixon (Hahn (1967), p. 66), we conclude that every trajectory of $(a1)$ is either identically equal to some stationary point of these equations, or approaches one of them.

Let us first consider case 1. Arguing as in the proof of Lemma 3, we conclude that there are three stationary points: (x_0, y_0), (x_β, y_α) and (x_1, y_1) such that

$$(x_0, y_0) \to (0,0), \ (x_\beta, y_\alpha) \to (\beta, \alpha), \ \text{ and } \ (x_1, y_1) \to (1,1), \tag{a2}$$

as $\max(\epsilon, \delta) \to 0$ and $\min(k, s) \to \infty$. The argument also shows that

$$\frac{d}{dy} q_\alpha^{\epsilon,k}(y_\alpha) \frac{d}{dx} g_\beta^{\delta,s}(x_\beta) > 1 \tag{a3}$$

for all sufficiently small ϵ, δ and all sufficiently large k, s.

Introduce the Jacobian of $(a1)$

$$J(x,y) = \begin{pmatrix} -1 & \frac{d}{dy} q_\alpha^{\epsilon,k}(y) \\ \frac{d}{dx} g_\beta^{\delta,s}(x) & -1 \end{pmatrix}.$$

Its eigenvalues are

$$-1 \pm \sqrt{\frac{d}{dy} q_\alpha^{\epsilon,k}(y) \frac{d}{dx} g_\beta^{\delta,s}(x)}.$$

Due to $(a3)$ one of the eigenvalues of $J(x_\beta, y_\alpha)$ is positive, so the matrix is unstable. Using standard results in stochastic approximation on the non-attainability of unstable points (see, for example, Nevelson and Hasminskii (1976), p. 113), we conclude that

$$P\{(X^t, Y^t) \to (x_\beta, y_\alpha)\} = 0 \tag{a4}$$

for every initial state of the system.

Next we show that the states (x_0, y_0) and (x_1, y_1) are stable. To prove this, note that $q_\alpha^{\epsilon,k}(\cdot)$ is increasing over $[0,1]$ and converges to the step function at α. Hence its derivative converges uniformly to 0 on any closed set that does not contain α. The same holds for the derivative of $g_\beta^{\delta,s}(\cdot)$ on any closed set which does not contain β. Consequently $J(x_i, y_i) \to -I$ as $\max(\epsilon, \delta) \to 0$ and $\min(k, s) \to \infty$ for $i = 0, 1$. Hence the matrices are stable for all sufficiently small ϵ, δ and all sufficiently large k, s.

Now we can prove convergence with probability one of the vector (X^t, Y^t) as $t \to \infty$. Rewrite (4) in the vector form

$$\vec{X}^{t+1} = \vec{X}^t + (1/a^{t+1})[\vec{R}(\vec{X}^t) + \vec{R}(t, \vec{X}^t) + \vec{\zeta}(t, \vec{X}^t)],\ t \geq 1,\ \vec{X}^1 = (X^1, Y^1), \tag{a5}$$

where for every $\vec{x} = (x, y)$ we have

$$\vec{R}(\vec{x}) = (q_\alpha^{\epsilon,k}(y) - x, g_\beta^{\delta,s}(x) - y), \quad \vec{\zeta}(t, \vec{x}) = (\Xi^t(y), \frac{a^1 - b^1}{b^{t+1}}\Psi^t(x)),$$

$$\Psi^t(x) = \psi^t(x) - E\psi^t(x) \quad \Xi^t(y) = \xi^t(y) - E\xi^t(y),$$

$$\vec{R}(t, \vec{x}) = (\nu_\alpha^{\epsilon,k}(t, y), \frac{a^{t+1}}{b^{t+1}}\mu_\beta^{\delta,s}(t, x) + \frac{a^1 - b^1}{b^{t+1}}[g_\beta^{\delta,s}(x) - y]).$$

The martingale convergence theorem implies that

$$\sum_{i=1}^{\infty}(1/a^{i+1})\vec{\zeta}(i, \vec{X}^i)$$

exists with probability one. Designate by Ω_0 the joint event that it exists and $\vec{X}^t$ does not converge to (x_β, y_α) as $t \to \infty$. Then owing to $(a4)$, $P\{\Omega_0\} = 1$. Fix an elementary outcome ω from Ω_0. Then the stochastic sequence $(a5)$ converts to the following deterministic sequence

$$\vec{x}^{t+1} = \vec{x}^t + (1/a^{t+1})\vec{R}(\vec{x}^t) + \vec{\sigma}^t,\ t \geq 1,\ \vec{x}^1 = \vec{X}^1, \tag{a6}$$

where $\vec{x}^t$ and $\vec{\sigma}^t$ stand for the realizations of $\vec{X}^t$ and $(1/a^{t+1})[\vec{R}(t, \vec{X}^t) + \vec{\zeta}(t, \vec{X}^t)]$ and

$$\lim_{t\to\infty} ||\sum_{i=t}^{\infty}\vec{\sigma}^i|| = 0. \tag{a7}$$

What remains to be shown that $\{\vec{x}^t\}$ converges either to (x_0, y_0) or (x_1, y_1). Assume to the contrary that there exists a subsequence of $\{\vec{x}^t\}$ converging to a point different from (x_0, y_0) and (x_1, y_1). We shall show show that this assumption leads to a contradiction.

Suppose there is a subsequence $\{n_p\}$ such that $\vec{x}^{n_p} \to (\bar{x}, \bar{y})$ as $p \to \infty$ and $(\bar{x}, \bar{y}) \neq (x_0, y_0)$, $(\bar{x}, \bar{y}) \neq (x_1, y_1)$.

For all positive integers n and all real $t \geq 0$ set

$$\vec{x}^n(t) = \vec{x}^i \quad \text{where} \quad \sum_{j=n}^{i}(1/a^{j+1}) \leq t < \sum_{j=n}^{i+1}(1/a^{j+1}).$$

Let $\vec{y}((\bar{x},\bar{y}),\cdot)$ stand for the solution (in vector form) of the system of ordinary differential equations $(a1)$ satisfying the initial condition $\vec{y}((\bar{x},\bar{y}),0) = (\bar{x},\bar{y})$. One can show (see, for example, Benveniste, Métivier and Priouret (1990), p.p. 230-231) that for every finite $T > 0$

$$\lim_{p\to\infty} \sup_{t\in[0,T]} ||\vec{x}^{n_p}(t) - \vec{y}((\bar{x},\bar{y}),t)|| = 0. \tag{a8}$$

Since the trajectories of the system $(a1)$ approach its stationary points, $\vec{y}((\bar{x},\bar{y}),t)$ converges as $t \to \infty$ either to (x_0,y_0) or (x_1,y_1). Without loss of generality let us consider the case when $\lim_{t\to\infty} \vec{y}((\bar{x},\bar{y}),t) = (x_0,y_0)$.

We conclude that there is a subsequence $\{m_p\}$ such that $n_p < m_p < n_{p+1}$ and $\vec{x}^{m_p} \to (x_0,y_0)$ as $p \to \infty$.

Since $J(x_0,y_0) \to -I$, there exists τ^0 such that $||\vec{x} - (x_0,y_0)|| \leq \tau^0$ implies

$$\langle \vec{R}(\vec{x}), \vec{x} - (x_0,y_0)\rangle \leq -\frac{1}{2}||\vec{x} - (x_0,y_0)||^2 \tag{a9}$$

whenever $\max(\epsilon,\delta)$ is sufficiently small and $\min(k,s)$ is sufficiently large.

Fix $\tau > 0$ such that $\tau < \min(\tau^0, \sqrt{\bar{x}^2+\bar{y}^2})$. There is a subsequence $\{l_p\}$ such that $l_p = \max\ n > m_p :\ n < n_{p+1}$ and $||\vec{x}^n - (x_0,y_0)|| \leq \tau$. Selecting a subsequence of $\{l_p\}$ if necessary, we can assume that $\vec{x}^{l_p} \to \vec{x}(\tau)$ as $p \to \infty$, where $||\vec{x}(\tau) - (x_0,y_0)|| = \tau$. Fix small enough $\tau' > 0$ and define a subsequence $\{j_p\}$ such that $j_p = \max\ n > l_p :\ n < n_{p+1}$ and $||\vec{x}^n - \vec{x}^{l_p}|| \leq \tau'$. Then from $(a6)$

$$\vec{x}^{j_p} - \vec{x}^{l_p} = \sum_{i=l_p}^{j_p-1} (1/a^{i+1})\vec{R}(\vec{x}^i) + \vec{\sum}^p, \tag{a10}$$

where

$$\vec{\sum}^p = \sum_{i=l_p}^{j_p-1} \vec{\sigma}^i.$$

Using the Lipschitz property of $\vec{R}(\cdot)$, one obtains from $(a10)$

$$||\vec{x}^{j_p} - \vec{x}^{l_p}|| \leq [||\vec{R}(\vec{x}^{l_p})|| + L\tau']T^p + ||\vec{\sum}^p||$$

and

$$||\vec{x}^{j_p} - \vec{x}^{l_p}|| \geq [||\vec{R}(\vec{x}^{l_p})|| - L\tau']T^p + ||\vec{\sum}^p||,$$

where L stands for the Lipschitz constant and

$$T^p = \sum_{i=l_p}^{j_p-1} (1/a^{i+1}).$$

From $(a7)$ it follows that, for small enough τ' and all sufficiently large p, there are positive constants c and C such that

$$c\tau' \leq T^p \leq C\tau'. \tag{a11}$$

Using $(a7)$, $(a9)$, $(a10)$ and $(a11)$, we obtain, for large enough p,

$$||\vec{x}^{j_p} - (x_0, y_0)||^2 = ||\vec{x}^{l_p} - (x_0, y_0)||^2 + 2T^p\langle \vec{x}^{l_p} - (x_0, y_0), \vec{R}(\vec{x}^{l_p})\rangle +$$

$$2\langle \vec{x}^{l_p} - (x_0, y_0), \sum_{i=l_p}^{j_p-1}(1/a^{i+1})[\vec{R}(\vec{x}^i) - \vec{R}(\vec{x}^{l_p})]\rangle + 2\langle \vec{x}^{l_p} - (x_0, y_0), \vec{\sum}^p\rangle +$$

$$2\langle \sum_{i=l_p}^{j_p-1}(1/a^{i+1})\vec{R}(\vec{x}^i), \vec{\sum}^p\rangle + ||\sum_{i=l_p}^{j_p-1}(1/a^{i+1})\vec{R}(\vec{x}^i)||^2 + ||\vec{\sum}^p||^2 \leq$$

$$||\vec{x}^{l_p} - (x_0, y_0)||^2[1 - c\tau' + L^2C^2(\tau')^2] + 2||\vec{x}^{l_p} - (x_0, y_0)||LC(\tau')^2 + o_p(1),$$

where $o_p(1) \to 0$ as $p \to \infty$. Passing to the limit as $p \to \infty$ we conclude that

$$\limsup_{p\to\infty} ||\vec{x}^{j_p} - (x_0, y_0)||^2 \leq \tau^2[1 - c\tau' + L^2C^2(\tau')^2] + 2\tau LC(\tau')^2.$$

Consequently, if τ' is so small that $\tau^2[1 - c\tau' + L^2C^2(\tau')^2] + 2\tau LC(\tau')^2 < \tau^2$, then

$$\limsup_{p\to\infty} ||\vec{x}^{j_p} - (x_0, y_0)|| < \tau.$$

However, $j_p > l_p$ and l_p is the last time instant before n_{p+1} when the sequence is inside the τ-neighborhood of (x_0, y_0). Hence $\vec{x}^{j_p}$ must lie *outside* the τ-neighborhood of (x_0, y_0), that is,

$$\liminf_{p\to\infty} ||\vec{x}^{j_p}|| \geq \tau.$$

This contradiction shows that there is no subsequence of $\{\vec{x}^t\}$ converging to a limit different from (x_0, y_0) or (x_1, y_1).

This completes the proof for case 1.

In case 2, we find the stationary points of the system

$$q_\alpha^{\epsilon,k}(y) = x \quad \text{and} \quad 1 - g_\beta^{\delta,s}(x) = y.$$

This reduces to the solutions of

$$1 - g_\beta^{\delta,s}(q_\alpha^{\epsilon,k}(y)) = y.$$

Since the left hand side of this equation is monotonically decreasing, it has a unique solution (x'_β, y'_α). An argument similar to proof of Lemma 3 shows that, as $\max(\epsilon,\delta) \to 0$ and $\min(k,s) \to \infty$,

$$(x'_\beta, y'_\alpha) \to (\beta, \alpha), \quad \text{and} \quad \begin{array}{l} \frac{d}{dy}q_\alpha^{\epsilon,k}(y'_\alpha) \to \infty, \\ \frac{d}{dx}g_\beta^{\delta,k}(x'_\beta) \to \infty. \end{array} \tag{a12}$$

The Jacobian of this system is

$$J(x,y) = \begin{pmatrix} -1 & \frac{d}{dy}q_\alpha^{\epsilon,k}(y) \\ -\frac{d}{dx}g_\beta^{\delta,s}(x) & -1 \end{pmatrix}$$

and its eigenvalues are

$$-1 \pm \sqrt{-\frac{d}{dy} q_\alpha^{\epsilon,k}(y) \frac{d}{dx} g_\beta^{\delta,s}(x)}.$$

From ($a12$) we see that for sufficiently small ϵ, δ and sufficiently large k, s, the eigenvalues of $J(x'_\alpha, y'_\beta)$ are negative and arbitrarily large in absolute value. Consequently the matrix is stable. Using an argument similar to the one given above[3], it follows that, with probability one, (X^t, Y^t) converges to (x'_β, y'_α). This concludes the proof of Theorem 2.

Theorem 3 *Let G be a non-degenerate 2×2 game. Every perturbed best reply dynamic converges to its limit at rate at least $1/\sqrt{t}$ as $t \to \infty$. More precisely, there exist $\epsilon', \delta', k', s'$ such that for all $\epsilon \leq \epsilon', \delta \leq \delta'$, $k \geq k', s \geq s'$ and $\vec{z} \in R^2$*

$$\lim_{t\to\infty} P\{\sqrt{t}[(X^t, Y^t) - (x, y)] < \vec{z}, \lim_{m\to\infty}(X^m, Y^m) = (x, y)\} =$$

$$P\{\vec{N}(\vec{0}, K(x, y)) < \vec{z}\} P\{\lim_{m\to\infty}(X^m, Y^m) = (x, y)\}.$$

Here (x, y) can be any of the points to which our process converges with positive probability, and $\vec{N}(\vec{0}, K(x, y))$ is a two-dimensional normal random vector with mean zero and variance matrix

$$K(x, y) = \int_0^\infty \exp\{[J(x, y) + \frac{1}{2}I]t\} B(x, y) \exp\{[J(x, y)^T + \frac{1}{2}I]t\} dt,$$

where the sigh ^{T} designates transposition, I stands for the identity matrix in R^2,

$$B(x, y) = \begin{pmatrix} x(1-x) & 0 \\ 0 & y(1-y) \end{pmatrix},$$

$$\text{case 1}: \quad J(x, y) = \begin{pmatrix} -1 & \frac{d}{dy} q_\alpha^{\epsilon,k}(y) \\ \frac{d}{dx} g_\beta^{\delta,s}(x) & -1 \end{pmatrix},$$

$$\text{case 2}: \quad J(x, y) = \begin{pmatrix} -1 & \frac{d}{dy} q_\alpha^{\epsilon,k}(y) \\ -\frac{d}{dx} g_\beta^{\delta,s}(x) & -1 \end{pmatrix}.$$

Proof. Consider our game dynamic given as a stochastic approximation procedure ($a5$). Since the noise term $\vec{\zeta}(t, \cdot)$ is bounded, the Lindeberg condition holds (see, for example, Loève (1955), p. 280).

Let us check that $E\vec{\zeta}(t, (x, y))\vec{\zeta}(t, (x, y))^T$ is close to $B(x, y)$. We have

$$E\vec{\zeta}(t, (x, y))\vec{\zeta}(t, (x, y))^T = \begin{pmatrix} q_\alpha^{\epsilon,k}(y)[1 - q_\alpha^{\epsilon,k}(y)] & 0 \\ 0 & g_\beta^{\delta,s}(x)[1 - g_\beta^{\delta,s}(x)] \end{pmatrix} + o_t(1),$$

[3]Alternatively, if we set $X = x - x'_\beta$, $Y = y - y'_\alpha$, $a(X) = -1$, $b(Y) = [1 - q_\alpha^{\epsilon,k}(Y + y'_\alpha) - x'_\beta]Y^{-1}$, $d(Y) = -1$, $c(X) = [1 - g_\beta^{\delta,s}(X + x'_\beta) - y'_\alpha]X^{-1}$, then, taking into account monotonicity of $q_\alpha^{\epsilon,k}(\cdot)$ and $g_\beta^{\delta,s}(\cdot)$, the global stability of the point (x'_β, y'_α) follows from a result of Krasovskii (see Hahn (1967), p. 127). Hence we can apply in this case the standard stochastic approximation technique (see, Benveniste, Métivier and Priouret (1990)).

where $o_t(1)$ stands for a matrix whose elements converge uniformly to 0 as $t \to \infty$. But if (x, y) is a limit (stationary point) of our dynamic, then

$$\text{case 1}: \quad q_\alpha^{\epsilon,k}(y) = x, \quad g_\beta^{\delta,k}(x) = y,$$

$$\text{case 2}: \quad g_\alpha^{\epsilon,k}(y) = x, \quad 1 - g_\beta^{\delta,s}(x) = y.$$

In both cases this leads to

$$q_\alpha^{\epsilon,k}(y)[1 - q_\alpha^{\epsilon,k}(y)] = y(1-y) \quad \text{and} \quad g_\beta^{\delta,s}(x)[1 - g_\beta^{\delta,s}(x)] = x(1-x).$$

Hence

$$E\vec{\zeta}(t,(x,y))\vec{\zeta}(t,(x,y))^T = B(x,y) + o_t(1).$$

To apply the limit theorem for stochastic approximation, we need to check the stability of the matrix $J(\cdot,\cdot) + 1/2I$ at the stable stationary points. But we already showed in the proof of Theorem 2 that, for small ϵ, δ and large k, s, $J(\cdot,\cdot)$ is close to $-I$ for case 1 games, hence it is stable. In case 2, $J(\cdot,\cdot)$ has two arbitrarily large negative eigenvalues, so again it is stable. Thus in both cases $J(\cdot,\cdot) + 1/2I$ is stable at the stationary points of the process, from which Theorem 3 follows.

References

Arthur, W. B., Y. M. Ermoliev, and Y. M. Kaniovski (1987): "Adaptive Process of Growth Being Modeled by Urn Schemes," *Kibernetika,* No. 6, 49–57 (in Russian). (Translated into English in *Cybernetics*, **23**, 779–789.)

Benveniste, A., M. Métivier, and P. Priouret (1990): *Adaptive Algorithms and Stochastic Approximations.* New York: Springer–Verlag.

David, H. A. (1981): *Order Statistics, 2-nd edition.* New York: John Wiley.

Dosi, G., and Y. Kaniovski (1994): "On "Badly Behaved" Dynamics. Some Applications of Generalized Urn Schemes to Technological and Economic Change," *Journal of Evolutionary Economics*, **4**, 1–31.

Ellison, G., and D. Fudenberg (1994): *World of Mouth Communication and Social Learning*, Economic Theory Discussion Paper No. 12, Harward Institute of Economic Research, Harward University.

Fabian, V. (1968): "On Asymptotic Normality in Stochastic Approximation", *Ann. Math. Statist.*, **39**, 1327-1332.

Foster, D., and H. P. Young (1990): "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology,* **38**, 219–232.

Fudenberg, D., and D. M. Kreps (1993): "Learning Mixed Equilibria," *Games and Economic Behavior,* **5**, 320–367.

Hahn, W. (1967): *Stability of Motion.* New York: Springer–Verlag.

Kandori, M., G. J. Mailath, and R. Rob (1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica,* **61**, 29–56.

Loève, M. (1955): *Probability Theory.* New York: D. Van Nostrand Company, Inc.

Nevelson, M. B., and R. Z. Hasminskii (1976): *Stochastic Approximation and Recursive Estimation.* **47** American Math. Soc., Providence, RI.

Posch, M. (1994): *Cycling with a Generalized Urn Scheme and a Learning Algorithm for* 2×2 *Games*, Working Paper WP-94-76, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Vorob'ev, N. N. (1977): *Game Theory.* New York: Springer–Verlag.

Young, H. P. (1993a): "An Evolutionary Model of Bargaining," *Journal of Economic Theory,* **59**, 145–168.

Young, H. P. (1993b): "The Evolution of Conventions," *Econometrica,* **61**, 57–83.