On the Global Convergence of Stochastic Fictitious Play*

Josef Hofbauer
Institut für Mathematik
Universität Wien
Strudlhofgasse 4
A-1090 Vienna, Austria
josef.hofbauer@univie.ac.at
http://mailbox.univie.ac.at/Josef.Hofbauer

William H. Sandholm
Department of Economics
University of Wisconsin
1180 Observatory Drive
Madison, WI 53706, USA
whs@ssc.wisc.edu
http://www.ssc.wisc.edu/~whs

February 21, 2002

* This paper is an abridged and substantially revised version of our working paper, Hofbauer and Sandholm (2001). We thank Carlos Alós-Ferrer, Buz Brock, Glenn Ellison, Moe Hirsch, Ed Hopkins, Michael Kosfeld, Larry Samuelson, and three anonymous referees, as well as seminar audiences from Bonn, Cornell, UCL, Vienna, the 2001 Italian Game Theory Meeting, the Spring 2001 Midwest Economic Theory Meeting, and the 2001 Stony Brook Conference for helpful discussions and comments. The second author gratefully acknowledges financial support from the National Science Foundation (SES-0092145).

Abstract

We establish global convergence results for stochastic fictitious play for four classes of games: games with an interior ESS, zero sum games, potential games, and supermodular games. We do so by appealing to techniques from stochastic approximation theory, which relate the limit behavior of a stochastic process to the limit behavior of a differential equation defined by the expected motion of the process. The key result in our analysis of supermodular games is that the relevant differential equation defines a strongly monotone dynamical system. Our analyses of the other cases combine Lyapunov function arguments with a discrete choice theory result: that the choice probabilities generated by any additive random utility model can be derived from a deterministic model based on payoff perturbations that depend nonlinearly on the vector of choice probabilities.

Keywords: learning in games, stochastic fictitious play, supermodular games, discrete choice theory, chain recurrence, stochastic approximation theory

1. INTRODUCTION

Of the many existing models of learning and evolution in games, the oldest and best known is *fictitious play*, introduced by Brown (1951). In fictitious play, each player chooses best responses to his beliefs about his opponents, which are given by the time average of past play. Convergence of beliefs to Nash equilibrium has been established for two player zero sum games (Robinson (1951)), 2 x 2 games (Miyasawa (1961)), potential games (Monderer and Shapley (1996a)), games with an interior ESS (Hofbauer (1995b)), and certain classes of supermodular games (Milgrom and Roberts (1991), Krishna (1992), Hahn (1999)).

Since best responses are generically pure, a player's choices under fictitious play are quite sensitive to the exact value of his beliefs; small changes in beliefs can lead to discrete changes in behavior. Even when beliefs converge to Nash equilibrium, actual behavior may not; in particular, behavior can never converge to the mixed equilibrium of a game. For these reasons, the appropriateness of fictitious play as a model of learning has been called into question.

To contend with these issues, Fudenberg and Kreps (1993) introduced *stochastic fictitious play*. In this model, each player's payoffs are perturbed in each period by random shocks *a la* Harsanyi (1973a). As a consequence, each player's anticipated behavior in each period is a genuine mixed strategy. Fudenberg and Kreps (1993), Kaniovski and Young (1995), and Benaïm and Hirsch (1999a) are therefore able to extend Miyasawa's (1961) result for 2 x 2 games to stochastic fictitious play, proving not only convergence of beliefs to equilibrium, but also convergence of behavior. Benaïm and Hirsch (1999a) also establish convergence in certain p player, two strategy games. However, because of the complications created by the random payoff perturbations, results for other classes of games have proved difficult to obtain. In particular, nothing is known about convergence in games with more than two strategies per player.

In this paper, we establish convergence results for stochastic fictitious play for the remaining classes of games noted above in which standard fictitious play is known to converge: namely, games with an interior ESS, zero sum games, potential games, and supermodular games.

Our results should not be interpreted as suggesting that stochastic fictitious play converges in all games. Indeed, Benaïm and Hirsch (1999a) show that stochastic fictitious play fails to converge to equilibrium in Jordan's (1993) three player matching pennies game. Shapley (1964) and Gaunersdorfer and Hofbauer (1995) provide examples in which versions of standard fictitious play fail to converge, and it is clear that stochastic fictitious play can fail to converge in these examples as well. Nevertheless, the classes of games we consider are of economic interest, and for these games we can obtain global convergence results.

To establish these results, we rely on techniques from stochastic approximation theory (see, e.g., Benaïm (1999)). These techniques show that one can characterize the limit behavior of stochastic fictitious play in terms of the *perturbed best response dynamic*, a differential equation defined by the expected motion of the stochastic process. More precisely, all limit points of stochastic fictitious play must be contained in the *chain recurrent set* (Conley (1978)) of the perturbed best response dynamic. The chain recurrent set contains those states which can arise in the long run if the deterministic dynamic is subjected to small shocks occurring at isolated moments in time.

The perturbed best response dynamic is defined in terms of *perturbed best response functions*. These functions are perturbed versions of the underlying best response correspondences; the differences between the two are due to the random payoff disturbances. To understand behavior under the perturbed dynamic, we must characterize these functions.

In the case of supermodular games, we are able to show that the perturbed best response functions are monotone, where monotonicity is defined in terms of a stochastic dominance order on mixed strategies. This property enables us to show that the perturbed best response dynamic defines a strongly monotone dynamical system (Hirsch (1988)), which in turn allows us to describe the chain recurrent set. In particular, strong monotonicity implies that almost all solution trajectories of the perturbed dynamic converge to rest points, which themselves can be viewed as approximate Nash equilibria of the underlying game. Unlike the analyses mentioned above for settings without perturbations, our analysis provides general convergence results for supermodular games without appealing to assumptions beyond supermodularity.

The perturbed best response function can be expressed as the composition of a map from mixed strategy profiles to payoffs, and a map from payoffs to choice probabilities. The latter map is actually the standard choice probability function from the additive random utility model of discrete choice theory (see, e.g., McFadden (1981) or Anderson, de Palma, and Thisse (1992)). To analyze the remaining three classes of games, we prove a characterization theorem for this discrete choice model.

In the additive random utility model, an agent chooses from a set of n alternatives. The payoff to each alternative is the sum of a base utility and a random utility term; the probability with which an alternative is chosen is the probability that its overall utility is highest. We show that these choice probabilities can be derived from an alternative model in which the agent optimally chooses a probability distribution over the n alternatives; his overall payoff in this model is the sum of the expected base payoff and some nonlinear, deterministic function of the probability vector he selects.[1] By combining known results with an analysis based on Legendre transforms, we construct a deterministic representation of the random utility model which is valid regardless of the distribution of the random utility terms.

In the context of learning in games, this result is of interest because it allows us to express perturbed best response dynamics in terms of deterministic payoff perturbations. Hofbauer (2000) and Hofbauer and Hopkins (2000) have recently shown that these deterministically perturbed dynamics are susceptible to analysis via Lyapunov functions. By combining such an analysis with our discrete choice result, we are able to characterize the chain recurrent set of the stochastically perturbed best response dynamic in the remaining three classes of games.

Since the discrete choice theorem described above may be of interest outside game theory, we begin our analysis by presenting this result without reference to game theoretic concepts. Such concepts are introduced in Section 3, which defines stochastic fictitious play and derives the perturbed best response dynamic. Section 4 shows how the discrete choice result can be used to find Lyapunov functions for the perturbed best response dynamic in games with an interior ESS, zero sum games, and potential games, and uses these functions to characterize the chain recurrent set. Section 5 characterizes this set for supermodular games by showing that the perturbed dynamics form a strongly monotone dynamical system. Finally, Section 6 combines the analyses from the previous two sections with techniques from stochastic approximation theory to prove global convergence results for stochastic fictitious play. Proofs omitted from the text are provided in the Appendix.

---

[1] This dual description of choice probabilities is well known in the case of logit choice model – see Section 2.

2. A DISCRETE CHOICE THEOREM

We consider the standard additive random utility model as described, for example, by Anderson, de Palma, and Thisse (1992). In this model, an agent must choose from a set of alternatives $A = \{1, \ldots , n\}$ offering base payoffs of $\pi_1, \ldots, \pi_n$. But when choosing alternative i, the agent also obtains a stochastic payoff of ε_i. The random vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ takes values in $\mathbf{R}^n$ according to some strictly positive density f: $\mathbf{R}^n \to \mathbf{R}$; the distribution of ε does not depend on the base payoffs π. When the stochastic terms are realized, the agent chooses the alternative whose total payoff is highest. Hence, the probability that the agent chooses alternative i is given by the *choice probability function*

(1) $\qquad C_i(\pi) = P(\operatorname{argmax}_j \pi_j + \varepsilon_j = i).$

The best known example of a choice probability function which can be generated from a random utility model is the *logit choice function*,

$$L_i(\pi) = \frac{\exp(\eta^{-1}\pi_i)}{\sum_j \exp(\eta^{-1}\pi_j)}.$$

We call the parameter $\eta \in (0, \infty)$ the *noise level*. When η approaches zero, logit choice approaches unperturbed maximization; when η approaches infinity, it approaches uniform randomization. It is well known that the logit choice function is generated by a random utility model whose random utility terms ε_i are i.i.d. with the extreme value distribution $F(x) = \exp(-\exp(-\eta^{-1}x - \gamma))$, where γ is Euler's constant.

Interestingly, the logit choice function can also be derived using a quite different model of payoff perturbations. Consider an agent who directly chooses a probability distribution $y \in \Delta A = \{x \in \mathbf{R}^n_+ \colon \sum_j x_j = 1\}$ over the elements in A. If the agent chooses distribution y, he obtains an expected base payoff of $y \cdot \pi$, but must pay a cost of $V(y)$, where V is a nonlinear, deterministic function of the probability vector he chooses. Suppose that the agent always chooses the vector y which maximizes the combined payoff $y \cdot \pi - V(y)$. It is well known (Rockafellar (1970), Anderson, de Palma, and Thisse (1992), Fudenberg and Levine (1998)) and easily verified that if the function V is the entropy function $V(y) = \eta \sum_j y_j \ln y_j$, then this maximization yields the choice probabilities from the logit choice function L.

The main result of this section, Theorem 2.1, shows that a deterministic representation can be obtained for choice probabilities from the additive random utility model *regardless of the distribution of the random utility vector* ε*.* To state our result, we introduce one additional definition: following Fudenberg and Levine (1998), we call the deterministic perturbation V: $\text{int}(\Delta A) \to \mathbf{R}$ *admissible* if for all y, $D^2V(y)$ is positive definite on $\mathbf{R}_0^n = \{z \in \mathbf{R}^n: \sum_j z_j = 0\}$, the tangent space of ΔA, and if $\|\nabla V(y)\|$ approaches infinity as y approaches the boundary of ΔA.

Theorem 2.1: *Let* C: $\mathbf{R}^n \to \Delta A$ *be the choice probability function defined in equation* (1)*, where the random vector* ε *admits a strictly positive density on* $\mathbf{R}^n$ *and is such that the function* C *is continuously differentiable. Then there exists an admissible deterministic perturbation* V *such that*

$$C(\pi) = \operatorname*{arg\,max}_{y \in \text{int}(\Delta A)} \left(y \cdot \pi - V(y)\right). \tag{2}$$

The proof of this result proceeds as follows. First, we establish that the derivative matrix DC is symmetric and has negative off-diagonal terms. These properties of DC imply that the vector field C admits a convex potential function, which we call W.[2] We show that the required disturbance function V can be obtained as the Legendre transform of W. This choice of V ensures that the functions $(\nabla V)^{-1}$ and $\nabla W \equiv C$ are identical (in a sense to be made precise below), so that C satisfies the first order conditions for the maximization problem (2).

Proof: The probability that alternative i is chosen when the payoff vector is π is given by

$$\begin{aligned} C_i(\pi) &= P(\pi_i + \varepsilon_i \geq \max_j(\pi_j + \varepsilon_j)) \\ &= P(\varepsilon_j \leq \pi_i + \varepsilon_i - \pi_j \text{ for all } j) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\pi_i + x_i - \pi_1} \cdots \int_{-\infty}^{\pi_i + x_i - \pi_{i-1}} \int_{-\infty}^{\pi_i + x_i - \pi_{i+1}} \cdots \int_{-\infty}^{\pi_i + x_i - \pi_n} f(x)\, dx_n \ldots dx_{i+1} dx_{i-1} \ldots dx_1 dx_i \,. \end{aligned} \tag{3}$$

If we consider another alternative $j > i$ and perform the change of variable $\hat{x}_j = \pi_i +$

[2] These facts are well known. Indeed, the potential function W is known to describe the expected perturbed payoff resulting from an optimal choice among the n alternatives – see McFadden (1981) or Anderson, de Palma, and Thisse (1992). However, the remainder of our argument appears to be new.

$x_i - \pi_j$, we find that

$$(4) \qquad \frac{\partial C_i}{\partial \pi_j}(\pi) = -\int_{-\infty}^{\infty} \int_{-\infty}^{\pi_i + x_i - \pi_1} \cdots \int_{-\infty}^{\pi_i + x_i - \pi_{i-1}} \int_{-\infty}^{\pi_i + x_i - \pi_{i+1}} \cdots \int_{-\infty}^{\pi_i + x_i - \pi_{j-1}} \int_{-\infty}^{\pi_i + x_i - \pi_{j+1}} \cdots \int_{-\infty}^{\pi_i + x_i - \pi_n}$$

$$f(x_1, \ldots, x_{j-1}, \pi_i + x_i - \pi_j, x_{j+1}, \ldots, x_n)\, dx_n \ldots dx_{j+1} dx_{j-1} \ldots dx_{i+1} dx_{i-1} \ldots dx_1 dx_i$$

$$= -\int_{-\infty}^{\infty} \int_{-\infty}^{\pi_j + \hat{x}_j - \pi_1} \cdots \int_{-\infty}^{\pi_j + \hat{x}_j - \pi_{i-1}} \int_{-\infty}^{\pi_j + \hat{x}_j - \pi_{i+1}} \cdots \int_{-\infty}^{\pi_j + \hat{x}_j - \pi_{j-1}} \int_{-\infty}^{\pi_j + \hat{x}_j - \pi_{j+1}} \cdots \int_{-\infty}^{\pi_j + \hat{x}_j - \pi_n}$$

$$f(x_1, \ldots, x_{i-1}, \pi_j + \hat{x}_j - \pi_i, x_{i+1}, \ldots, x_n)\, dx_n \ldots dx_{j+1} dx_{j-1} \ldots dx_{i+1} dx_{i-1} \ldots dx_1 d\hat{x}_j$$

$$= \frac{\partial C_j}{\partial \pi_i}(\pi).$$

This equality shows that the derivative matrix $DC(\pi) \in \mathbf{R}^{n \times n}$ is symmetric. Moreover, $DC(\pi)$ is positive definite on $\mathbf{R}_0^n$. To see this, note that by equation (4), the off-diagonal terms of $DC(\pi)$ are strictly negative. Since $\sum_j C_j(\pi) = 1$ by definition, it follows that $\sum_j \frac{\partial C_j}{\partial \pi_i}(\pi) = 0$ for each i, and so that

$$(5) \qquad \frac{\partial C_i}{\partial \pi_i}(\pi) = -\sum_{j \neq i} \frac{\partial C_j}{\partial \pi_i}(\pi).$$

Hence, equations (4) and (5) imply that $DC(\pi)\, \mathbf{1} = 0$, where $\mathbf{1} \in \mathbf{R}^n$ denotes the vector of ones. Moreover, if z is not proportional to $\mathbf{1}$, then if we let $d_{ij} = \frac{\partial C_i}{\partial \pi_j}(\pi)$, equations (5) and (4) imply that

$$(6) \qquad z \cdot DC(\pi)\, z = \sum_i \sum_j d_{ij} z_i z_j = \sum_j \sum_{i \neq j} d_{ij} z_i z_j - \sum_j (\sum_{i \neq j} d_{ij}) z_j^2 = \sum_j \sum_{i \neq j} d_{ij}(z_i z_j - z_j^2)$$

$$= \sum_j \sum_{i < j} d_{ij}(2 z_i z_j - z_i^2 - z_j^2) = \sum_j \sum_{i < j} -d_{ij}(z_i - z_j)^2 > 0.$$

These observations imply that C is one-to-one on $\mathbf{R}_0^n$ and satisfies $C(\pi + c\mathbf{1}) = C(\pi)$ for all $c \in \mathbf{R}$: shifting payoffs by a constant vector does not affect choice probabilities.

Finally, we make an observation about the range of the function C: if components π_j, $j \in J \subset A$ stay bounded while the remaining components approach infinity, then $C_j(\pi) \to 0$ for all $j \in J$: that is, $C(\pi)$ converges to a subface of the simplex ΔA. It follows that there are points in the range of C arbitrarily close to each corner of the simplex.

Since the derivative matrix $DC(\pi)$ is symmetric, the vector field C admits a potential function W: $\mathbf{R}^n \to \mathbf{R}$ (that is, a function which satisfies $\nabla W \equiv C$). Equation (6) implies that W is strictly convex on $\mathbf{R}^n_0$.

Now consider the restrictions of W and $C \equiv \nabla W$ to $\mathbf{R}^n_0$, and let V: $\text{int}(\Delta A) \to \mathbf{R}$ denote the Legendre transform of W:

(7) $$V(y) = y \cdot C^{-1}(y) - W(C^{-1}(y)).$$

Since W: $\mathbf{R}^n_0 \to \mathbf{R}$ is strictly convex and C: $\mathbf{R}^n_0 \to \text{int}(\Delta A)$ takes values at points arbitrarily close to each corner of the simplex, Theorem 26.5 of Rockafellar (1970) implies that the following statements are true. First, the domain of V is convex and equals the range of C, which therefore must be all of $\text{int}(\Delta A)$. Second, V and W solve the dual optimization problems

(8) $$V(y) = \max_{\pi \in \mathbf{R}^n_0} \big(y \cdot \pi - W(\pi)\big) \text{ and}$$

(9) $$W(\pi) = \max_{y \in \text{int}(\Delta A)} \big(y \cdot \pi - V(y)\big).$$

Third, ∇V: $\text{int}(\Delta A) \to \mathbf{R}^n_0$ is invertible, with $(\nabla V)^{-1} \equiv \nabla W \equiv C$ on $\mathbf{R}^n_0$.[3]

We conclude by establishing the required properties of V. First, since $(\nabla V)^{-1} \equiv C$, the observation three paragraphs above shows that $\|\nabla V(y)\|$ approaches infinity as y approaches the boundary of ΔA. Furthermore, since $C(\nabla V(y)) = y$, differentiating yields $DC(\nabla V(y))\ D^2 V(y) = I$, where all expressions are interpreted as linear operators on $\mathbf{R}^n_0$. Since $DC(\nabla V(y))$ is symmetric and positive definite on $\mathbf{R}^n_0$ and inverts $D^2 V(y)$ on $\mathbf{R}^n_0$, it follows that $D^2 V(y)$ is also positive definite on $\mathbf{R}^n_0$.

Finally, solving for the maximizer

$$y^* = \operatorname*{arg\,max}_{y \in \text{int}(\Delta A)} \big(y \cdot \pi - V(y)\big),$$

we find that $\pi = \nabla V(y^*)$, and hence that $y^* = C(\pi)$. This completes the proof of the theorem. ■

---

[3] Since the domain of V is $\text{int}(\Delta A)$, the partial derivatives of V are not well defined. Consequently, $\nabla V(y)$ is defined to be the unique vector in $\mathbf{R}^n_0$ such that $V(y + hz) = V(y) + (\nabla V(y) \cdot z)h + o(h)$ for all unit length vectors z in $\mathbf{R}^n_0$.

The requirement that the function C be continuously differentiable is essentially a smoothness requirement on the distribution of the random vector ε. For example, if the components of ε are independent, standard results on convolutions imply that C is continuously differentiable as long as each component of ε admits a bounded density function.

It is natural to ask whether the converse of Theorem 2.1 also holds: that is, whether the choice function derived from any admissible deterministic perturbation can be derived from an appropriate stochastic perturbation. Proposition 2.2, which considers a logarithmic perturbation studied by Harsanyi (1973b), shows that such a reconstruction is not always possible.

Proposition 2.2: *When $n \geq 4$, there is no stochastic perturbation of payoffs which yields the same choice probabilities as the admissible deterministic perturbation $V(y) = -\sum_j \ln y_j$.*

More generally, we have the following characterizations of the two types of choice functions. The Legendre transform argument in the proof of Theorem 2.1 shows that a surjective choice function C: $\mathbf{R}^n \to \mathrm{int}(\Delta A)$ can be derived from an admissible deterministic payoff perturbation V if and only if $DC(\pi)$ is symmetric, positive definite on $\mathbf{R}^n_0$, and satisfies $DC(\pi)\,\mathbf{1} = 0$. On the other hand, the Williams-Daly-Zachary Theorem of McFadden (1981) implies that the choice functions C which can be derived from some stochastic payoff perturbation ε with a strictly positive density on $\mathbf{R}^n$ are characterized by these requirements, plus the additional requirement that the partial derivatives of C satisfy

$$(-1)^k \frac{\partial^k C_{i_0}}{\partial \pi_{i_1} \ldots \partial \pi_{i_k}} > 0$$

for each $k = 1, \ldots, n-1$ and each set of $k+1$ distinct indices $\{i_0, i_1, \ldots, i_k\}$.

In order to model boundedly rational choice in a simple fashion, Chen, Friedman, and Thisse (1997) consider choice functions of the form

$$C_i(\pi) = \frac{w(\pi_i)}{\sum_j w(\pi_j)}, \tag{10}$$

where the weighting function w: $\mathbf{R} \to (0, \infty)$ is some increasing and differentiable function of payoffs. We conclude this section by noting that the only choice function of this form which can be derived from either stochastic or deterministic perturbations of payoffs is the logit choice function L.

Proposition 2.3: *Suppose that the choice function C satisfies condition* (10) *and either condition* (1) *or condition* (2). *Then $C \equiv L$ for some noise level $\eta > 0$.*

3. STOCHASTIC FICTITIOUS PLAY AND PERTURBED BEST RESPONSE DYNAMICS

3.1 *Preliminaries*

In this section, we define the process of stochastic fictitious play that is our central interest in this paper. Before doing so, we introduce notation to describe normal form games. A p player normal form game G is defined by a collection of finite strategy sets $S^1, \ldots, S^p$ and a collection of utility functions $u^1, \ldots, u^p$. Player α's strategy set is $S^\alpha = \{1, \ldots, n^\alpha\}$, with typical elements s^α, i, and j. Player α's utility function u^α is a map from the set of strategy profiles $S = \prod_\beta S^\beta$ to the real line. Finally, $S^{-\alpha} = \prod_{\beta\neq\alpha} S^\beta$ denotes the set of strategy profiles of player α's opponents.

It will prove useful to define vector-valued functions which describe the payoffs to each of a player's pure strategies given the mixed strategies chosen by his opponents. Let $\Delta S^\alpha = \{x^\alpha \in \mathbf{R}^{n^\alpha}_+ : \sum_i x_i^\alpha = 1\}$ denote the set of player α's mixed strategies. Also, let $\Sigma = \prod_\beta \Delta S^\beta$, with typical element $x = (x^1, \ldots, x^p)$, denote the set of mixed strategy profiles, and let $\Sigma^{-\alpha} = \prod_{\beta\neq\alpha} \Delta S^\beta$. Then player α's payoff vector is denoted U^α: $\Sigma^{-\alpha} \to \mathbf{R}^{n^\alpha}$, and is defined by

$$U^\alpha_{s^\alpha}(x^{-\alpha}) = \sum_{s^{-\alpha}\in S^{-\alpha}} \left(u^\alpha(s^\alpha, s^{-\alpha}) \prod_{\beta\neq\alpha} x^\beta_{s^\beta} \right).$$

Using this definition, we can define player α's best response correspondence B^α: $\Sigma^{-\alpha} \Rightarrow \Delta S^\alpha$ by

$$B^\alpha(x^{-\alpha}) = \operatorname*{arg\,max}_{y\in\Delta S^\alpha} \; y \cdot U^\alpha(x^{-\alpha}).$$

If we let the random vector ε^α represent a random perturbation of player α's payoffs, and let C^α denote the corresponding choice probability function, we can define player α's *perturbed best response function* $\tilde{B}^\alpha\colon \Sigma^{-\alpha} \to \Delta S^\alpha$ by

$$\tilde{B}_i^\alpha(x^{-\alpha}) = P(\text{argmax}_j\ U_j^\alpha(x^{-\alpha}) + \varepsilon_j^\alpha = i) = C_i^\alpha(U^\alpha(x^{-\alpha})).$$

When the random utility terms are "small", the continuous function $\tilde{B}^\alpha$ is a perturbed version of the discontinuous correspondence B^α.

3.2 *Stochastic Fictitious Play*

In unperturbed fictitious play, each player chooses a best response to his beliefs about how his opponents will behave; these beliefs are determined by the time average of past play. In stochastic fictitious play, players make these choices after their payoffs are subjected to random shocks.

In *standard stochastic fictitious play*, a group of $p \geq 2$ players repeatedly plays a p player normal form game. The state variable is $Z_t \in \Sigma$, whose components Z_t^α describe the time averages of each player's past behavior.[4] Formally,

$$(11) \qquad Z_t^\alpha = \tfrac{1}{t}\sum_{u=1}^{t} \zeta_u^\alpha,$$

where $\zeta_t^\alpha \in \Delta S^\alpha$ represents the pure strategy played by player α at time t. The initial choices ζ_1^α are arbitrary pure strategies, while subsequent choices are best responses to beliefs Z_t. Best responses are determined after payoffs have been subjected to disturbances ε_t^α which are independent over time t and across players α. Player α's disturbance vectors are distributed according to a fixed density function $f^\alpha\colon \mathbf{R}^{n^\alpha} \to \mathbf{R}$ which satisfies the conditions of Theorem 2.1. Since the mixed strategy representation of player α's pure strategy $i \in S^\alpha$ is the standard basis vector $e_i \in \Delta S^\alpha \subset \mathbf{R}^{n^\alpha}$, player α's choice probabilities are described by

$$(12) \qquad P\left(\zeta_{t+1}^\alpha = e_i \,\middle|\, Z_t = z\right) = P(\text{argmax}_k\ U_k^\alpha(z^{-\alpha}) + (\varepsilon_t^\alpha)_k = i) = \tilde{B}_i^\alpha(z^{-\alpha}).$$

---

[4] This specification is based on the idea that players assess the behavior of each of their opponents separately. When there are three or more players, one could instead start with the alternative premise that each player tracks the time average of his opponents' past *joint* behavior, in which case the relevant state space would be the set of correlated strategies. For further discussion of this modeling issue, see Fudenberg and Levine (1998, Section 2.5).

We also consider a formulation of stochastic fictitious play where in each period, two players are chosen from a larger group to play a symmetric two player game. This formulation corresponds to the single population framework commonly studied in evolutionary game theory. A two player game is *symmetric* if the two players' strategy sets are the same and if a player's payoffs do not depend on whether he is called player 1 or player 2: $S^1 = S^2$ and $u^1(s^1, s^2) = u^2(s^2, s^1)$. If each player's payoffs are subject to random disturbances ε^1 and ε^2 drawn from the same distribution, the resulting perturbed best response functions $\tilde{B}^1$ and $\tilde{B}^2$ are identical.

In *symmetric stochastic fictitious play*, the state variable is the time average of all past plays of the game: if $\hat{\zeta}_t^\alpha$ represents the strategy chosen by the player assigned to role α at time t, then the state variable $\hat{Z}_t \in \Delta S^1$ is given by

$$\hat{Z}_t = \tfrac{1}{2t}\sum_{u=1}^{t}\left(\hat{\zeta}_u^1 + \hat{\zeta}_u^2\right).$$

Players choose best responses after their payoffs have been subjected to the random disturbances ε_t^1 and ε_t^2, which are independently and identically distributed over time and across players. Choice probabilities are therefore described by

$$P\left(\hat{\zeta}_{t+1}^\alpha = e_i \,\middle|\, \hat{Z}_t = z\right) = P(\operatorname{argmax}_k\ U_k^\alpha(z) + (\varepsilon_t^\alpha)_k = i) = \tilde{B}_i^1(z).$$

3.3 *Perturbed Best Response Dynamics*

The first step in analyzing stochastic fictitious play is to determine its expected motion. In the case of standard stochastic fictitious play, we do so by first rearranging equation (11) in order to obtain a recursive definition of Z_t^α:

$$Z_{t+1}^\alpha = \tfrac{1}{t+1}(t\,Z_t^\alpha + \zeta_{t+1}^\alpha).$$

We can then use equation (12) to compute the expected increments of Z_t^α:

$$E\left(Z_{t+1}^\alpha - Z_t^\alpha \,\middle|\, Z_t = z\right) = \tfrac{1}{t+1}\left[E\left(\zeta_{t+1}^\alpha \,\middle|\, Z_t = z\right) - z\right] = \tfrac{1}{t+1}(\tilde{B}^\alpha(z^{-\alpha}) - z^\alpha).$$

Thus, we see that after a reparameterization of time, expected changes in the time average Z_t are governed by the *perturbed best response dynamic*

(P) $$\dot{x}^\alpha = \tilde{B}^\alpha(x^{-\alpha}) - x^\alpha.$$

This dynamic is defined on the space of mixed strategy profiles Σ. Similarly, the expected motion of symmetric fictitious play is governed by a symmetric version of the perturbed best response dynamic, defined on the set of mixed strategies ΔS^1:

$$\text{(SP)} \qquad \dot{x} = \tilde{B}^1(x) - x.$$

The dynamics (P) and (SP) can be viewed as perturbed versions of the multipopulation and single population *best response dynamics* (Gilboa and Matsui (1991), Matsui (1992)):

$$\text{(BR)} \qquad \dot{x}^\alpha \in B^\alpha(x^{-\alpha}) - x^\alpha;$$

$$\text{(SBR)} \qquad \dot{x} \in B^1(x) - x.$$

The latter dynamics can be used to approximate the time average of behavior under unperturbed fictitious play. However, since the best response correspondences B^α are set valued, the dynamics (BR) and (SBR) can exhibit complicated solution trajectories, and in fact may admit multiple solution trajectories from a single initial condition – see Matsui (1992) and Hofbauer (1995b). In contrast, since the perturbed best response functions $\tilde{B}^\alpha$ are single-valued and smooth, the perturbed dynamics (P) and (SP) are well behaved, and in particular possess unique solution trajectories.

We begin our analysis of stochastic fictitious play by studying its expected motion in four classes of games (Sections 4 and 5). Once this is accomplished, we can use techniques from stochastic approximation theory to characterize the original stochastic processes (Section 6). Before beginning our analysis of equations (P) and (SP), we take care of a few additional preliminaries.

3.4 *ω-Limit Sets and Chain Recurrence*

Most analyses of evolutionary dynamics use the notion of an ω-limit set to describe limit behavior. However, understanding stochastic fictitious play requires the more general notion of chain recurrence. We now introduce and contrast these two concepts. Consider a dynamic

$$\text{(D)} \qquad \dot{x} = F(x)$$

which generates a *semiflow* ϕ: $\mathbf{R}_+ \times X \to X$ on the compact set $X \subset \mathbf{R}^n$. The point $\phi(t, x) \in X$ is the position at time t of the solution to (D) which begins at $x \in X$. The set of rest points of (D) can be defined as $RP(\mathrm{D}) = \{x \in X\colon \phi(t, x) = x \text{ for all } t \geq 0\} = \{x \in X\colon F(x) = 0\}$.

The *ω-limit set* of the state x, $\omega(x) = \{z \in X\colon \lim_{k\to\infty}\phi(t_k, x) = z \text{ for some } t_k \to \infty\}$, is the set of limit points of the solution trajectory starting at x. We let $\Omega(\mathrm{D}) = \bigcup_{x\in\Delta}\omega(x)$ denote the union of the ω-limit sets. Clearly, $RP(\mathrm{D}) \subset \Omega(\mathrm{D})$.

Knowledge of $\Omega(\mathrm{D})$ is generally not sufficient to characterize limit behavior under stochastic fictitious play. To accomplish this, we require the more general notion of chain recurrence (Conley (1978)), which allows for states which can arise in the long run if the flow is subject to small shocks occurring at isolated moments in time. Call a sequence $\{x = x_0, x_1, \dots , x_k = y\}$ an *ε-chain* from x to y if for each $i \in \{1, \dots, k\}$, there is a $t_i \geq 1$ such that $\left|\phi(t_i, x_{i-1}) - x_i\right| < \varepsilon$. The ε-chain specifies $k + 1$ segments of solution trajectories to (D). The first begins at x, and the last is simply the point y; the jumps between the ends and beginnings of consecutive segments are never longer than ε. We call the state *x chain recurrent* if there is an ε-chain from x to itself for all $\varepsilon > 0$, and we let $CR(\mathrm{D})$ denote the set of chain recurrent points. The set $CR(\mathrm{D})$ contains all rest points, periodic orbits, quasiperiodic motions, and chaotic orbits of the flow. It can be shown that $\Omega(\mathrm{D}) \subset CR(\mathrm{D})$, and that in general this inclusion is strict. For further discussion, see Conley (1978), Akin (1993), Robinson (1995), or Benaïm (1999).

To see why chain recurrence is needed here, consider a flow on a circle which moves clockwise everywhere except at a single rest point. This rest point is the unique ω-limit point of the flow. Now suppose that the flow represents the expected motion of some underlying stochastic process. If the stochastic process reaches the rest point, its expected motion is nil. Nevertheless, actual motions may occur with positive probability, and in particular the process can jump past the rest point and begin another circuit. Therefore, in the long run all regions of the circle are visited infinitely often. Since the only ω-limit point of the flow is the rest point, this notion of recurrence does not capture the long run behavior of the underlying stochastic process. However, this long run behavior is captured by chain recurrence, as one can easily verify that all points on the circle are chain recurrent under the flow.

3.5 *Rest Points and Nash Equilibria*

The rest points of the perturbed best response dynamics will ultimately constitute our predictions of long run behavior under stochastic fictitious play. For these predictions to accord with standard game theoretic analyses, these rest points should approximate Nash equilibria of the underlying game. The following result ensures that this is true whenever the perturbations generating the dynamics are sufficiently small.

Proposition 3.1: *Fix a game G. For each $k \in \mathbf{Z}_+$, let $\varepsilon^k = (\varepsilon^{1,k}, \ldots, \varepsilon^{p,k})$ be a collection of disturbance vectors, and let $x^k \in \Sigma$ be a rest point of* (P) *under ε^k. If the sequence $\{\varepsilon^k\}$ converges weakly to a mass point at the origin, and if the sequence $\{x^k\}$ converges to x^*, then x^* is a Nash equilibrium of G. If instead each $x^k \in \Delta S^1$ is a rest point of* (SP), *then x^* is a symmetric Nash equilibrium of G.*

4. DISCRETE CHOICE THEORY AND LYAPUNOV FUNCTIONS

The perturbed best response dynamics (P) and (SP) are defined in terms of stochastic payoff perturbations. In this section, we use Theorem 2.1 to express these dynamics in terms of deterministic payoff perturbations. This transformation is useful because in certain cases, one can characterize behavior under the deterministically perturbed dynamics by introducing suitable Lyapunov functions. Doing so enables us to characterize the chain recurrent sets of the original dynamics (P) and (SP) in cases where the underlying game admits an interior ESS, is a zero sum game, or is a potential game.

If we write out the perturbed best response dynamic (P) explicitly in terms of the stochastic perturbations ε_t^α, we obtain

$$\dot{x}_i^\alpha = P(\text{argmax}_k\ U_k^\alpha(x^{-\alpha}) + (\varepsilon_t^\alpha)_k = i) - x_i^\alpha,$$

Theorem 2.1 tells us that the choice probabilities induced by the random vectors ε_t^α can always be represented in terms of some deterministic perturbations V^α. Consequently, the dynamic (P) is equivalent to the dynamic

$$\text{(PV)} \qquad \dot{x}^\alpha = \underset{y \in \text{int}(\Delta S^\alpha)}{\arg\max} \left(y \cdot U^\alpha(x^{-\alpha}) - V^\alpha(y) \right) - x.$$

for the appropriate choices of V^α. We can also use this transformation in the symmetric case, obtaining the dynamic

(SPV) $$\dot{x} = \underset{y \in \mathrm{int}(\Delta S^1)}{\arg\max} \left(y \cdot U^1(x) - V^1(y) \right) - x.$$

We call (PV) and (SPV) *deterministically perturbed best response dynamics.*

We call a real valued function Λ a *strict Lyapunov function* for a dynamic if its value increases strictly along every non-constant solution trajectory. The existence of a Lyapunov function for a dynamic ensures that limit behavior is very simple. In particular, if its state space is compact, all solution trajectories of the dynamic must converge to connected sets of rest points. While the original perturbed dynamics (P) and (SP) do not seem especially conducive to admitting Lyapunov functions, Hofbauer (2000) and Hofbauer and Hopkins (2000) have constructed Lyapunov functions for the deterministically perturbed dynamics (PV) and (SPV) for certain classes of games. In the two subsections which follow we review and extend their results.

In addition to the games we explicitly consider in the coming sections, we can also establish results for related games obtained through certain transformations of payoffs. As usual, any result which holds for the game G also holds for the game $\hat{G}$, where $\hat{u}^\alpha(s) = u^\alpha(s) + \phi^\alpha(s^{-\alpha})$ for some functions ϕ^α: $S^{-\alpha} \to \mathbf{R}$. This invariance holds because shifting player α's payoffs by a term which he cannot influence does not alter his incentives. More notably, our convergence results which hold for G also hold for $\overline{G}$, where $\overline{u}^\alpha(s) = \kappa^\alpha\, u^\alpha(s) + \psi^\alpha(s^\alpha)$. The constant $\kappa^\alpha > 0$ rescales the payoffs of the original game, while the function ψ^α: $S^\alpha \to \mathbf{R}$ captures a shift in player α's payoffs which depends only on his own behavior. To see why our convergence results are not affected by these changes, note that the perturbed best response function generated by the game $\overline{G}$ and the disturbance vector ε^α is identical to the perturbed best response function generated by the game G and the disturbance vector $\frac{1}{\kappa^\alpha}(\varepsilon^\alpha + \psi^\alpha)$. In other words, the effect of the affine transformation $\kappa^\alpha u^\alpha$ + ψ^α on the payoffs of the underlying game can always be mimicked by a corresponding transformation of the payoff disturbances. Of course, such transformations can alter Nash equilibria and rest points of the perturbed dynamics, while shifts by ϕ^α cannot.

4.1 *Games with an Interior ESS and Zero Sum Games*

The notion of an *evolutionarily stable strategy* (Maynard Smith and Price (1973)) is the original solution concept of evolutionary game theory. Consider a symmetric two player game G. A mixed strategy $x^* \in \Delta S^1$ is an ESS if $x^* \cdot U^1(x) > x \cdot U^1(x)$ for all mixed strategies x in a neighborhood of x^*. In other words, an ESS is a mixed strategy with the property that after any invasion by a mutant mixed strategy, the ESS performs better than the mutant in the post-entry population.

We focus on games with an ESS in the interior of the state space ΔS^1.[5] Hofbauer (2000) shows that if G admits an interior ESS, then the dynamic (SPV) admits a strict Lyapunov function which is strictly concave. In our notation, this function can be expressed as

$$\hat{\Lambda}(x) = x \cdot U^1(x) - V^1(x) - W^1(U^1(x)).$$

Because $\hat{\Lambda}$ is strictly concave, the maximizer of $\hat{\Lambda}$ is globally asymptotically stable under (SPV), and indeed is the unique chain recurrent point of (SPV). Consequently, Theorem 2.1 implies that this maximizer is also the unique chain recurrent point of the original dynamic (SP).

One can also find Lyapunov functions for two player zero sum games. A two player game is *zero sum* if $u^1(s) = -u^2(s)$ for every strategy profile $s \in S$. In the symmetric case, Hofbauer (2000) shows that the Lyapunov function $\hat{\Lambda}$ defined above is again a strict Lyapunov function for the dynamic (SPV).[6] For cases where the game is not necessarily symmetric, Hofbauer and Hopkins (2000) show that the strictly concave function

$$\Lambda(x^1, x^2) = -V^1(x^1) - W^1(U^1(x^2)) - V^2(x^2) - W^2(U^2(x^1))$$

is a strict Lyapunov function for the dynamic (PV). Theorem 2.1 again implies that the maximizer of Λ is the unique chain recurrent point of (P).

4.2 *Potential Games*

Potential games include pure coordination games and congestion games, and also arise in applications of evolutionary techniques to implementation problems (see Sandholm (2000)). We call the game G a *potential game* if $u^\alpha(s) = u^\beta(s)$ for all players α and β and all strategy profiles $s \in S$. That is, G is a potential game if all

---

5 An interior ESS cannot exist outside of the current symmetric framework – see Selten (1980).

6 Note that in the zero-sum case, the first term of $\hat{\Lambda}$ is identically zero.

players always receive the same payoff.[7] Hofbauer (1995a) and Sandholm (2001) show that in games of this form, the aggregate payoff function serves as a Lyapunov function for a broad class of evolutionary dynamics. By adding an appropriate perturbation, one can also obtain Lyapunov functions for the deterministically perturbed best response dynamics. In the case where G is a two player symmetric potential game, Hofbauer (2000) shows that the function

$$\hat{\Pi}(x) = \tfrac{1}{2} \sum_{(s^1,s^2)\in S} u^1(s^1,s^2)\, x_{s^1} x_{s^2} - V^1(x)$$

is a strict Lyapunov function for the dynamic (SPV). The first term in $\hat{\Pi}(x)$ equals one half of the payoff to mixed strategy x when played against itself, while the second term is the perturbation.

By building on the previous result and on a result of Hofbauer and Hopkins (2000) for two player games, we can construct Lyapunov functions for the dynamic (PV) for any p player potential game.

Proposition 4.1: *If G is a potential game, then the function*

$$\Pi(x^1,\dots,x^p) = \sum_{s\in S}\left(u^1(s)\prod_{\alpha} x^{\alpha}_{s^{\alpha}} \right) - \sum_{\alpha} V^{\alpha}(x^{\alpha})$$

is a strict Lyapunov function for the dynamic (PV).

The Lyapunov functions above guarantee that the sets of ω-limit points for (PV) and (SPV) are equal to the sets of rest points. But remarkably, the existence of a strict Lyapunov function is not enough to ensure that all chain recurrent points are rest points.[8] However, this equivalence can be established under slightly stronger assumptions. We now present results which establish this equivalence for the case of the dynamic (PV). Given the proofs of these results, analogues for the dynamic (SPV) are easily obtained.

---

[7] Typically, the definition of potential games also includes games in which payoffs are common up to shifts which do not affect the players' incentives – see Monderer and Shapley (1996b). However, since these shifts have no bearing on perturbed best responses (see the discussion preceding Section 4.1), it is enough to consider games in which payoffs are identical. More generally, the discussion preceding Section 4.1 implies that our results for games with identical payoffs extend immediately to the weighted potential games of Monderer and Shapley (1996b) and to the rescaled partnership games of Hofbauer and Sigmund (1988).

[8] For counterexamples, see Akin (1993, p. 25-26 and 55-56) and Benaïm (1999, p. 27).

One way to obtain the conclusion that all chain recurrent points are rest points is to require that the potential function Π be sufficiently smooth. This will be true so long as the perturbations V^α are sufficiently smooth.

Proposition 4.2: *Suppose that G is a potential game and that each function V^α is C^N, where $N = \sum_\alpha (n^\alpha - 1)$ is the dimension of the state space Σ. Then $CR(\mathrm{PV}) = RP(\mathrm{PV})$.*

If we begin with the dynamic (P) based on the stochastic perturbations ε^α, the deterministic perturbations V^α which correspond to the random vectors ε^α will be C^N if the distributions of the ε^α are sufficiently smooth.

Alternatively, we can reach stronger conclusions by imposing a generic regularity condition.[9]

Proposition 4.3: *If G is a potential game and all rest points of* (PV) *are hyperbolic, then $RP(\mathrm{PV})$ is finite and $CR(\mathrm{PV}) = RP(\mathrm{PV})$.*

Again, Theorem 2.1 shows that these results also apply to the stochastically perturbed dynamics (P) and (SP).

5. SUPERMODULAR GAMES AND STRONGLY MONOTONE DYNAMICS

Supermodular games describe situations in which different players' actions are strategic complements. In these games, an order relation is placed on the strategy sets; strategic complementarity means that the advantage of switching to a higher strategy increases when opponents choose higher strategies. Supermodular games arise in many economic applications; for examples, see Topkis (1979), Milgrom and Roberts (1990), Vives (1990), and Fudenberg and Tirole (1992).

In this section, we show that when the underlying game is supermodular, the perturbed dynamics (P) and (SP) form a strongly monotone dynamical system (Hirsch (1988)). This fact allows us to establish a number of important properties of the dynamics, and in particular to describe the sets of chain recurrent states.

We say that the game G is (*strictly*) *supermodular* if for all distinct players α and β and all strategy profiles s and $\hat{s}$ such that $s^\alpha > \hat{s}^\alpha$ and $s^{-\alpha} = \hat{s}^{-\alpha}$, the difference $u^\alpha(s)$

---

[9] A rest point x^* of the dynamic (D) is *hyperbolic* if all eigenvalues of the derivative matrix $DF(x^*)$ corresponding to eigenvectors in the relevant tangent space have nonzero real part.

– $u^\alpha(\hat{s})$ is strictly increasing in $s^\beta = \hat{s}^\beta$. In other words, a game is supermodular if the advantage a player obtains from choosing a higher strategy is increasing in the strategy choices of each of his opponents.[10]

A fundamental property of supermodular games is that they possess increasing best response correspondences. This property can be used to show that every supermodular game admits a minimal and a maximal Nash equilibrium, and it is also important for studying learning processes.[11] A related monotonicity property is fundamental for studying the perturbed best response dynamics. For each player α, define the invertible linear operator T^α: $\Delta S^\alpha \to \mathbf{R}^{n^\alpha - 1}$ by

$$(T^\alpha x^\alpha)_i = \sum_{j=i+1}^{n^\alpha} x_j^\alpha .$$

If $x^\alpha \in \Delta S^\alpha$ is a mixed strategy for player α, then the ith component of $T^\alpha x^\alpha$ equals the mass in x^α placed on pure strategies larger than i. If we view points in the simplex ΔS^α as probability distributions on the strategy set $S^\alpha = \{1, 2, \ldots, n^\alpha\}$, then $T^\alpha y^\alpha \geq T^\alpha x^\alpha$ if and only if y^α stochastically dominates x^α. To compare full mixed strategy profiles, we define T: $\Sigma \to \prod_\alpha \mathbf{R}^{n^\alpha - 1}$ by $T(x^1, \ldots, x^p) = (T^1 x^1, \ldots, T^p x^p)$; $T^{-\alpha}$: $\Sigma^{-\alpha} \to \prod_{\beta \neq \alpha} \mathbf{R}^{n^\beta - 1}$ is defined analogously.

Theorem 5.1 shows that in supermodular games, each player's perturbed best response function is monotone with respect to this stochastic dominance order.[12]

Theorem 5.1: *Suppose that G is supermodular. If $T^{-\alpha} y^{-\alpha} \geq T^{-\alpha} x^{-\alpha}$, then $T^\alpha \tilde{B}^\alpha(y^{-\alpha}) \geq T^\alpha \tilde{B}^\alpha(x^{-\alpha})$.*

The monotonicity of the perturbed best response functions allows us to characterize the behavior of the perturbed dynamics (P) and (SP). To avoid

[10] Of course, it is enough for this property to hold after the names of the strategies have been permuted in an appropriate way.

[11] For the properties of pure strategy equilibria, see the aforementioned references. Milgrom and Roberts (1991) show that fictitious play converges in supermodular games with a unique Nash equilibrium. Krishna (1992) proves convergence in supermodular games satisfying a diminishing returns condition; Hahn (1999) proves convergence in 3 x 3 and 3 x 2 supermodular games. Finally, Kandori and Rob (1995) use the monotonicity of best responses to characterize the stochastically stable states of the Kandori, Mailath, and Rob (1993) model in certain supermodular games.

[12] It is worth noting that while monotonicity holds for any perturbed best response function defined in terms of stochastic perturbations, it does not extend to all perturbed best response functions defined in terms of deterministic perturbations. For an explanation of this point, see the Appendix.

repetition, we state our next four results for the asymmetric dynamic (P); nearly identical results can be established for the symmetric dynamic (SP).

With the monotonicity of $\tilde{B}^{\alpha}$ in hand, one can establish that the dynamic (P) possesses a minimal and a maximal rest point.

Theorem 5.2: *If G is a supermodular game, there exist rest points* $\underline{x}$, $\bar{x} \in RP$(P) *such that* $RP(\mathrm{P}) \subset [\underline{x}, \bar{x}]$, *where* $[\underline{x}, \bar{x}] = \{x \in \Sigma\colon T\underline{x} \leq Tx \leq T\bar{x}\}$.

To relate this result to our earlier remarks, recall (from Proposition 3.1) that the rest points of (P) represent approximate Nash equilibria of G.

To obtain more precise information about behavior under (P), it is helpful to study this dynamic after applying the change of coordinates T. The next result shows that this yields the dynamic

$$\text{(T)} \qquad \dot{v}^{\alpha} = T^{\alpha}\tilde{B}^{\alpha}((T^{-\alpha})^{-1}v^{-\alpha}) - v^{\alpha}$$

on the set $T(\Sigma) = \{(v^1, \dots, v^p) \in \prod_{\alpha} \mathbf{R}^{n^{\alpha}-1}\colon\ 1 \geq v_1^{\alpha} \geq \dots \geq v_{n^{\alpha}-1}^{\alpha} \geq 0 \text{ for all } \alpha\}$.

Proposition 5.3: *The dynamic* (P) *and the dynamic* (T) *are linearly conjugate*: $\{x_t\}_{t\geq 0}$ *solves* (P) *if and only if* $\{Tx_t\}_{t\geq 0}$ *solves* (T).

A differential equation $\dot{v} = g(v)$ on $T(\Sigma)$ is called *cooperative* if $\frac{\partial g_i^{\alpha}}{\partial v_j^{\beta}}(v) \geq 0$ for all $v \in T(\Sigma)$ and all distinct pairs (α, i) and (β, j). That is, an increase in any component of the state increases the rates of change of all other components. The equation is *irreducible* if for each $v \in T(\Sigma)$ and each nonempty proper subset I of the components of v, there is an $(\alpha, i) \in I$ and a $(\beta, j) \in I^C$ such that $\frac{\partial g_i^{\alpha}}{\partial v_j^{\beta}}(v) \neq 0$ for all $v \in T(\Sigma)$. The transformed dynamics (T) are of interest because they possess both of these properties.

Theorem 5.4: *If G is supermodular, the dynamic* (T) *is cooperative and irreducible.*

Observe that if ι_j is a standard basis vector in $\mathbf{R}^{n^{\alpha}-1}$, then $\iota_j = T^{\alpha}(e_{j+1} - e_j)$, where the latter vectors are standard basis vectors in $\mathbf{R}^{n^{\alpha}}$. In light of this observation and Proposition 5.3, the fact that (T) is cooperative has the following interpretation for the perturbed best response dynamics (P): if we shift mass in the strategy distribution $x^\beta \in \Delta S^\beta$ from strategy j to strategy $j + 1$, then for each player $\alpha \neq \beta$, the growth rate of every strategy $i + 1 \in S^\alpha$ increases relative to that of strategy i.

Theorem 5.4 is important because dynamics which are cooperative and irreducible are strongly monotone, and so have desirable monotonicity and convergence properties. In the next result, we list a number of useful implications of Theorem 5.4 for the perturbed best response dynamic (P).

Corollary 5.5: *If G is supermodular, then*

(i) *The dynamic* (P) *is* strongly monotone *with respect to the stochastic dominance order*: *if* $\{x_t\}_{t\geq 0}$ *and* $\{y_t\}_{t\geq 0}$ *are two solutions to* (P) *with* $Ty_0 \geq Tx_0$ *and* $y_0 \neq x_0$, *then* $Ty_t > Tx_t$ *for all* $t > 0$.

(ii) *The chain recurrent set lies between the minimal and maximal rest points of* (P): $CR(\mathrm{P}) \subset [\underline{x}, \overline{x}]$. *In particular, if* $RP(\mathrm{P}) = \{x^*\}$, *then* $CR(\mathrm{P}) = \{x^*\}$ *as well.*

(iii) *There is an open dense set of initial conditions from which solutions to* (P) *converge to unique limit points in* $RP(\mathrm{P})$.

(iv) *The remaining initial conditions are contained in a finite or countable union* $\bigcup_i M_i$ *of invariant manifolds of codimension* 1, *and hence have measure zero.*

(v) *Chain recurrent points are either rest points or are contained in these invariant manifolds*: $CR(\mathrm{P}) \subset RP(\mathrm{P}) \cup \bigcup_i M_i$.

Proof: In light of Proposition 5.3 and Theorem 5.4, part (i) follows from Theorem 4.1.1 of Smith (1995), part (iii) from Theorem 2.4.7 of Smith (1995), part (iv) (after a reversal of time) from Theorem 1.1 of Hirsch (1988), and part (v) from Theorems 1.6 and 1.7 of Hirsch (1999) (also see Theorem 3.3 and Corollary 3.4 of Benaïm and Hirsch (1999b)). The proof of part (ii) is provided in the Appendix. ■

Suppose we restrict attention to supermodular games with exactly two strategies per player. In this case, supermodularity requires only that the payoff advantage of every player's second strategy is increasing in the mass that each of his opponents places on her second strategy. Benaïm and Hirsch (1999a) observe that games with this property yield strongly monotone perturbed best response dynamics. Because there are only two strategies per player, there is no need to introduce the stochastic dominance order to obtain this conclusion. Our analysis shows that by introducing this order, one establish strong monotonicity of the dynamics for all supermodular games.[13]

Corollary 5.5 shows that solution trajectories from almost all initial conditions converge to rest points of (P), but allows the possibility that convergence does not occur from the remaining initial conditions. We conclude this section with an example in which convergence fails on a measure zero set.

Example 5.6: Consider a p player game, $p \geq 5$, with two strategies per player. Each player α wishes to coordinate his behavior with player $\alpha + 1$ (with the convention that $p + 1 = 1$). More specifically, $u^\alpha(s)$ equals 1 if $s^\alpha = s^{\alpha+1}$ and equals 0 otherwise. This game is (weakly) supermodular.[14] It has three Nash equilibria: two strict equilibria in which all players coordinate on the same strategy, and the mixed equilibrium $x^* = ((\frac{1}{2}, \frac{1}{2}), \ldots, (\frac{1}{2}, \frac{1}{2}))$.

Suppose that player α's payoffs are augmented by the random perturbation $\varepsilon^\alpha = (\varepsilon_1^\alpha, \varepsilon_2^\alpha)$. Let g denote the common density function for the differences $\varepsilon_1^\alpha - \varepsilon_2^\alpha$, and suppose that g is symmetric about zero, is decreasing on $\mathbf{R}_+$, and satisfies $g(0) > \frac{1}{2}$. We show in the Appendix that the resulting perturbed dynamic (P) possesses exactly three rest points: the mixed equilibrium x^*, and two stable symmetric rest points which approximate the two pure Nash equilibria. We establish below that the rest point x^* is unstable. Since Corollary 5.5 tells us that an open dense set of initial conditions must converge to a rest point, it follows that the two stable rest points attract almost all initial conditions in Σ, and that the basins of attraction for these rest points are separated by a $p - 1$ dimensional invariant manifold $M \subset \Sigma$ which contains x^*.

Since there are two strategies per player, the transformed dynamic (T) tracks the mass each player places on his second strategy. It is easily verified that the derivative matrix for (T) evaluated at Tx^* is the $p \times p$ circulant matrix whose rows are permutations of the vector $(-1, 2g(0), 0, \ldots, 0)$. This matrix has an eigenvalue of $2g(0) - 1 > 0$ corresponding to the unstable direction $\mathbf{1}$, and eigenvalues of $2g(0) \exp(2\pi ik/p) - 1$, $k = 1, \ldots, p - 1$, corresponding to directions tangent to M (see

---

[13] Furthermore, while in the two strategy case every reasonable evolutionary dynamic generates a strongly monotone flow, with more strategies this property is particular to the perturbed best response dynamics (P) and (SP). Indeed, this property fails for the replicator dynamics, and it does not even hold for all specifications of the deterministically perturbed best response dynamics (PV) and (SPV).

[14] In our analysis above, *strict* supermodularity is only used to establish irreducibility; the full strength of the strictness assumption is not needed for this conclusion. In the present example, one can easily versify that the dynamics (T) are irreducible, and hence strongly monotone.

Hofbauer and Sigmund (1988, p. 66)). Since $p \geq 5$, at least two of the latter eigenvalues will have positive real parts if $g(0)$ is sufficiently large. In this case, the rest point x^* is also unstable with respect to the restriction of (P) to the manifold M. It follows that solutions to (P) from almost all initial conditions in M do not converge to a rest point. In fact, the theory of positive feedback loops (Mallet-Paret and Smith (1990)) can be used to show that all solutions from these initial conditions converge to a periodic orbit in M.

6. CONVERGENCE OF STOCHASTIC FICTITIOUS PLAY

Using techniques from stochastic approximation theory, Fudenberg and Kreps (1993), Kaniovski and Young (1995), and Benaïm and Hirsch (1999a) show how the limit behavior of stochastic fictitious play can be characterized in terms of a perturbed best response dynamic. However, as the perturbations make this dynamic difficult to analyze, Fudenberg and Kreps (1993) and Kaniovski and Young (1995) only establish convergence in 2 x 2 games, while Benaïm and Hirsch (1999a) also prove convergence in certain p player, two strategy games.

By combining our analysis of the perturbed best response dynamics with results of Pemantle (1990), Benaïm and Hirsch (1999a), and Benaïm (2000), we can establish convergence of beliefs and choice probabilities in four important classes of games. Our results are stated for beliefs Z_t and $\hat{Z}_t$; however, since the perturbed best response functions $\tilde{B}^\alpha$ are continuous, corresponding results hold for choice probabilities as well. In particular, if beliefs Z_t converge to some rest point x of (P), then player α's choice probabilities converge to $C^\alpha(U^\alpha(x^{-\alpha})) = \tilde{B}^\alpha(x^{-\alpha}) = x^\alpha$ as well.

To state our results, we let $LS(\text{D}) \subset RP(\text{D})$ denote the set of linearly stable rest points of the dynamic (D), and let $LU(\text{D}) \subset RP(\text{D})$ denote the set of linearly unstable rest points of (D).

Theorem 6.1: *Consider standard stochastic fictitious play Z_t and symmetric stochastic fictitious play $\hat{Z}_t$ starting from arbitrary initial conditions.*

(*i*) *Suppose that G is a two player symmetric game with an interior ESS. Then $P(\lim_{t\to\infty} \hat{Z}_t = x^*) = 1$, where x^* is the unique rest point of* (SP).

(*ii*) *Suppose that G is a two player zero sum game. Then $P(\lim_{t\to\infty} Z_t = x^*) = 1$, where x^* is the unique rest point of* (P).

(iii) *Suppose that G is a p player potential game. If the distributions of the vectors ε^{α} are sufficiently smooth, then* $P(\omega(Z_t)$ is a connected subset of $RP(\mathrm{P})) = 1$. *If all rest points of* (P) *are hyperbolic and* (P) *is* C^2, *then* $P(\lim_{t\to\infty} Z_t \in LS(\mathrm{P})) = 1$.

(iv) *Suppose that G is a p player supermodular game. Then* $P(\omega\{Z_t\} \subset RP(\mathrm{P})$ or $\omega\{Z_t\} \subset M_i \cap [\underline{x}, \bar{x}]$ for some $i) = 1$. *In particular, if* $RP(\mathrm{P}) = \{x^*\}$, *then* $P(\lim_{t\to\infty} Z_t = x^*) = 1$. *If the state space of* (P) *is one- or two-dimensional and* (P) *is* C^2, *then* $P(\lim_{t\to\infty} Z_t \in RP(\mathrm{P}) - LU(\mathrm{P})) = 1$.

Finally, if G is a two player symmetric game, then results (ii), (iii), *and* (iv) *also hold for symmetric stochastic fictitious play if Z_t is replaced by $\hat{Z}_t$ and* (P) *by* (SP).

Proof: Theorem 3.3 of Benaïm and Hirsch (1999a) and Proposition 5.3 of Benaïm (1999) imply that with probability one, stochastic fictitious play must converge to a connected component of the chain recurrent set of the appropriate perturbed best response dynamic, (P) or (SP). Theorem 2.1 shows that these dynamics are equivalent to the dynamics (PV) and (SPV) for appropriate choices of the deterministic perturbations V^{α}. In the games considered in parts (i) and (ii), Theorem 4.2 of Hofbauer (2000), Theorem 3.2 of Hofbauer and Hopkins (2000) and Proposition 6.4 of Benaïm (1999) imply that latter dynamics admit unique rest points which are also the unique chain recurrent points. This proves parts (i) and (ii). The proofs of the parts (iii) and (iv), which rely on the analyses from Sections 2, 4.2, and 5 and on results of Pemantle (1990), Benaïm and Hirsch (1999a), and Benaïm (2000), can be found in the Appendix. ■

Parts (i) and (ii) of the theorem guarantee convergence of stochastic fictitious play to the unique rest point of the perturbed best response dynamics in games which are zero sum or which admit an interior ESS. Part (iii) shows that in potential games, convergence to the set of rest points is ensured if the disturbance distribution is sufficiently smooth; if all rest points are hyperbolic, convergence is always to a unique limit point which is Lyapunov stable, and hence a local maximizer of the relevant Lyapunov function, Π or $\hat{\Pi}$.

Part (iv) of the theorem offers a global convergence result for supermodular games, but it only guarantees convergence to rest points of the perturbed dynamics if the rest point is unique or if the dimension of the state variable is 1 or 2. Even in the symmetric case, the latter condition requires that there are at most three strategies in the underlying game. When there are more strategies, we cannot rule out convergence to one of the unstable invariant manifolds M_i. However, Benaïm

(2000) conjectures (and proves under additional assumptions) that such manifolds cannot be limits of stochastic approximation processes. If this conjecture is correct, convergence of stochastic fictitious play to rest points of (P) and (SP) can be established in all supermodular games.

In this paper, we combined an analysis of the perturbed best response dynamics with results from stochastic approximation theory to prove global convergence results for stochastic fictitious play. Interestingly, these perturbed dynamics also arise in other disturbance-based models of evolution and learning in games. In the stochastic evolutionary model of Blume (1993, 1997) and Young (1998), these dynamics describe expected changes in the behavior of large populations of myopic agents. Similarly, in Ellison and Fudenberg's (2000) model of population fictitious play, and in Ely and Sandholm's (2000) model of evolution in a diverse population, the perturbed best response dynamics arise as descriptions of changes in aggregate behavior. By combining the analysis of the dynamics (P) and (SP) presented here with other mathematical techniques, one can establish global convergence results for these three models of evolution and learning in games. For a presentation of these results, we refer the reader to Hofbauer and Sandholm (2001).

APPENDIX

The Proof of Proposition 2.2

Substituting $V(y) = -\sum_j \ln y_j$ into equation (2), we find that this selection of V yields the choice probability function $C_i(\pi) = (c(\pi) - \pi_i)^{-1}$, where $c(\pi)$ is the unique number satisfying $c(\pi) > \max_j \pi_j$ and $\sum_j (c(\pi) - \pi_i)^{-1} = 1$. Now suppose that $n \geq 4$, and let $i, j,$ and k be distinct strategies. A computation reveals that

$$\frac{\partial^2 C_i}{\partial \pi_j \, \partial \pi_k} = \frac{2\, C_i^2\, C_j^2\, C_k^2}{\left(\sum_l C_l^2\right)^3} \left(\left(C_i + C_j + C_k \right) \sum_l C_l^2 - \sum_l C_l^3 \right).$$

This expression is negative whenever C_i, C_j, and C_k are all close enough to zero.

Now suppose that the choice function $\hat{C}$ is derived from a stochastic perturbation of payoffs. Then differentiating expression (4) with respect to π_k reveals that $\frac{\partial^2 \hat{C}_i}{\partial \pi_j \, \partial \pi_k}$ must always be strictly positive. Hence, the choice function $C_i(\pi)$ = $(c(\pi) - \pi_i)^{-1}$ cannot be derived from a stochastic perturbation of payoffs. ■

The Proof of Proposition 2.3

Suppose that the choice function C satisfies equation (10) and equation (2) for some admissible V. (Theorem 2.1 implies that if C satisfies (1), it satisfies (2) as well.) Then if we define W: $\mathbf{R}_0^n \to \mathbf{R}$ via equation (9), Theorem 26.5 of Rockafellar (1970) implies that $\nabla W \equiv C$ on $\mathbf{R}_0^n$. Moreover, it is clear from equation (2) that $C(\pi + c\mathbf{1}) = C(\pi)$ for all $c \in \mathbf{R}$. Thus, if we define W on the remainder of $\mathbf{R}^n$ by $W(\pi + c\mathbf{1}) = W(\pi) + c$ for any $\pi \in \mathbf{R}_0^n$ and $c \in \mathbf{R}$, it can be verified that $\nabla W \equiv C$ on all of $\mathbf{R}^n$. It follows that $DC(\pi)$ is symmetric for all $\pi \in \mathbf{R}^n$.

Now, applying equation (10), we find that if $i \neq j$, then

$$\frac{\partial C_i}{\partial \pi_j}(\pi) = -\frac{w(\pi_i)\,w'(\pi_j)}{\left(\sum_k w(\pi_k)\right)^2}.$$

Thus, the symmetry of DC implies that $w(\pi_i)w'(\pi_j) = w(\pi_j)w'(\pi_i)$ for all π. Since w is strictly positive, it follows that $w'(\pi_i) = \eta\, w(\pi_i)$ for some constant η, and hence that $w(\pi_i) = K\exp(\eta\,\pi_i)$ for some constants η and K. Since w is strictly positive and increasing, it must be that η and K are strictly positive, and hence that $C \equiv L$. ■

The Proof of Proposition 3.1

We only consider the case of the dynamic (P); the proof of the result for the dynamic (SP) is similar. Recall that the perturbed best response $\tilde{B}^\alpha(x^{-\alpha})$ can be written as $C^\alpha(U^\alpha(x^{-\alpha}))$, where C^α is a perturbed version of the maximizer correspondence $M^\alpha(\pi) = \arg\max_{y \in \Delta S^\alpha} y \cdot \pi$. For each disturbance vector $\varepsilon^{\alpha,k}$, let $C^{\alpha,k}$ denote the corresponding choice probability function: $C_i^{\alpha,k}(\pi) = P(\operatorname{argmax}_j \pi_j + \varepsilon_j^{\alpha,k} = i)$.

We first prove a lemma.

Lemma A.1: *Suppose that* $\varepsilon^{\alpha,k} \Rightarrow \delta_{\{\mathbf{0}\}}$ *and that* $\pi^k \to \pi^*$. *If* $i \notin \operatorname{argmax}_j \pi_j^*$, *then* $C_i^{\alpha,k}(\pi^k) \to 0$.

Proof: Let $l \in \operatorname{argmax}_j \pi_j^*$, and let $\delta = \pi_l^* - \pi_i^* > 0$. Then for all large enough k,

$$\begin{aligned} C_i^{\alpha,k}(\pi^k) &= P(\operatorname{argmax}_j \pi_j^k + \varepsilon_j^{\alpha,k} = i) \\ &\leq P(\pi_i^k + \varepsilon_i^{\alpha,k} \geq \pi_l^k + \varepsilon_l^{\alpha,k}) \end{aligned}$$

$$\leq P(\varepsilon_i^{\alpha,k} - \varepsilon_l^{\alpha,k} \geq \tfrac{\delta}{2}).$$

Since the random variables $(\varepsilon_i^{\alpha,k} - \varepsilon_l^{\alpha,k})$ converge in distribution to the constant 0 as k approaches infinity, we conclude that $C_i^{\alpha,k}(\pi^k) \to 0$. ❑

Now suppose that $x^{\alpha,k} = \tilde{B}^\alpha(x^{-\alpha,k}) = C^{\alpha,k}(U^\alpha(x^{-\alpha,k}))$ for all α and that $x^k \to \hat{x}$. To prove the result, it is enough to show that $\hat{x}^\alpha \in M^\alpha(U^\alpha(\hat{x}^{-\alpha}))$. Clearly, $U^\alpha(x^{-\alpha,k}) \to U^\alpha(\hat{x}^{-\alpha})$. Hence, Lemma A.1 tells us that if $i \notin \text{argmax}_j\ U_j^\alpha(\hat{x}^{-\alpha})$, then $\hat{x}_i^\alpha = \lim_{k\to\infty} x_i^{\alpha,k} = \lim_{k\to\infty} C_i^{\alpha,k}(U^\alpha(x^{-\alpha,k})) = 0$. Consequently, $\hat{x}^\alpha \in M^\alpha(U^\alpha(\hat{x}^{-\alpha}))$. This completes the proof of the proposition. ■

The Proof of Proposition 4.1

The lemma preceding Theorem 4.2 of Hofbauer (2000) shows that $(U^\alpha(x^{-\alpha}) - \nabla V^\alpha(x^\alpha)) \cdot (\tilde{B}^\alpha(x^{-\alpha}) - x^\alpha) \geq 0$, with equality only when $\tilde{B}^\alpha(x^{-\alpha}) - x^\alpha = 0$. Furthermore, a calculation reveals that $\nabla\hat{\Pi}(x) = (U^1(x^{-1}) - \nabla V^1(x^1), \ldots, U^p(x^{-p}) - \nabla V^p(x^p))$. Since $\dot{x}^\alpha = \tilde{B}^\alpha(x^{-\alpha}) - x^\alpha$, we can conclude that

$$\tfrac{d}{dt}\hat{\Pi}(x_t) = \sum_\alpha (U^\alpha(x_t^{-\alpha}) - \nabla V^\alpha(x_t^\alpha)) \cdot \dot{x}_t^\alpha \geq 0,$$

with equality only if $\dot{x} = 0$. ■

The Proof of Proposition 4.2

Let $\hat{C}^\alpha(\pi) = \arg\max_{y\in\Delta S^\alpha}(y \cdot \pi - V^\alpha(y))$. Then $\hat{C}^\alpha(\pi) = (\nabla V^\alpha)^{-1}(\pi)$ for all $\pi \in \mathbf{R}_0^n$, and $\hat{C}^\alpha(\pi + c^\alpha\mathbf{1}) = \hat{C}^\alpha(\pi)$. Since Π is defined on Σ, $x \in \Sigma$ is a critical point of Π if and only if $U^\alpha(x^{-\alpha}) - \nabla V^\alpha(x^\alpha) = c^\alpha\mathbf{1}$ for all players α, which is true if and only if $\hat{C}^\alpha(U^\alpha(x^{-\alpha})) = \hat{C}^\alpha(U^\alpha(x^{-\alpha}) - c^\alpha\mathbf{1}) = \hat{C}^\alpha(\nabla V^\alpha(x^\alpha)) = x^\alpha$ for all α. Thus, the critical points of Π are precisely the rest points of (PV).

Our assumption about the functions V^α implies that the Lyapunov function Π: $\Sigma \to \Delta$ is C^N. Since $N > \max\{0, N - 1\}$, Sard's Theorem tells us that the set of critical values of Π has measure zero. The previous paragraph shows that this set is equal to $\{\Pi(x)\text{: } x \in RP\}$. We therefore conclude from Propositions 5.3 and 6.4 of Benaïm (1999) (also see Exercises 3.16 and 6.11 of Akin (1993)) that CR(PV) = RP(PV). ■

The Proof of Theorem 4.3

The rest points of (PV) are hyperbolic by assumption, and therefore are isolated. If $RP(\text{PV}) \subset \Sigma$ were an infinite set, then as Σ is compact, $RP(\text{PV})$ would have an accumulation point $\bar{x}$. But since (PV) is continuous, this would imply that $\bar{x} \in RP(\text{PV})$, contradicting that all rest points are isolated. We therefore conclude that $RP(\text{PV})$, and hence $\{\Pi(x)\colon\ x \in RP(\text{PV})\}$, are finite. Thus, Propositions 5.3 and 6.4 of Benaïm (1999) again imply that $CR(\text{PV}) = RP(\text{PV})$. ■

The Proof of Theorem 5.1

It is enough to show that if $T^\beta y^\beta \geq T^\beta x^\beta$ and $y^{-\beta} = x^{-\beta}$, then $T^\alpha \tilde{B}^\alpha(y^{-\alpha}) \geq T^\alpha \tilde{B}^\alpha(x^{-\alpha})$. Since the behavior of all players besides α and β is fixed, it is enough to consider a two player game and to suppose that $\alpha = 1$ and that $\beta = 2$. It will be convenient to represent player 1's payoffs as a matrix, and so we define $A \in \mathbf{R}^{n^\alpha \times n^\beta}$ by $A_{ij} = u^1(i, j)$. Thus, if player 2 chooses mixed strategy x^2, player 1's vector of payoffs is $A\,x^2 \in \mathbf{R}^{n^\alpha}$.

We begin with a lemma.

Lemma A.2: *Suppose that* $\{b_k\}_{k=1}^n$ *is strictly increasing and that* $\{c_k\}_{k=1}^n$ *satisfies*

(13) $$\sum_{k=1}^{j} c_k \leq 0 \text{ for all } j \leq n\text{, with a strict inequality for some } j \text{ and equality at } j = n.$$

Then $\sum_{i=1}^n b_k\, c_k > 0$.

Proof: For all $j < n$, let $d_j = b_{j+1} - b_j > 0$. Then

$$\sum_{i=1}^{n} b_k\, c_k = b_n \sum_{k=1}^{n} c_k - \sum_{j=1}^{n-1} d_j \left(\sum_{k=1}^{j} c_k \right) = -\sum_{j=1}^{n-1} d_j \left(\sum_{k=1}^{j} c_k \right) > 0. \ ❑$$

The next lemma shows that the increasing differences property of supermodular games still holds when we consider ordered pairs of opponent's *mixed* strategies. Its proof makes use of the following observation:

(14) $$T^\alpha y^\alpha \geq T^\alpha x^\alpha \text{ if and only if } \sum_{i=1}^{m} (y_i^\alpha - x_i^\alpha) \leq 0 \text{ for all } m < n^\alpha.$$

Lemma A.3: *If* $T^2 y^2 \geq T^2 x^2$ *and* $y^2 \neq x^2$, *then* $(A\,y^2)_i - (A\,x^2)_i$ *is strictly increasing in* i.

Proof: Fix $i < j$. We want to show that $(A\,y^2)_j - (A\,x^2)_j > (A\,y^2)_i - (A\,x^2)_i$, or equivalently, that

$$(A\,y^2)_j - (A\,y^2)_i = \sum_{k=1}^{n^2}(A_{jk} - A_{ik})y_k^2 > \sum_{k=1}^{n^2}(A_{jk} - A_{ik})x_k^2 = (A\,x^2)_j - (A\,x^2)_i.$$

Since G is strictly supermodular, $b_k = A_{jk} - A_{ik} = u^1(j,k) - u^1(i,k)$ is strictly increasing in k, while since $T^2y^2 \geq T^2x^2$ and $y^2 \neq x^2$, observation (14) implies that $c_k = y_k^2 - x_k^2$ satisfies condition (13). Thus, Lemma A.2 yields the result. ❑

Now suppose that $T^2y^2 \geq T^2x^2$; we want to show that $T^1\tilde{B}^1(y^2) \geq T^1\tilde{B}^1(x^2)$. If $y^2 = x^2$ this is obviously true, so we suppose instead that $y^2 \neq x^2$. By observation (14), it is enough to show that for all $m < n^1$,

$$0 > \sum_{i=1}^{m}(\tilde{B}_i^1(y^2) - \tilde{B}_i^1(x^2)) = \sum_{i=1}^{m}\int_0^1 \nabla\tilde{B}_i^1(\lambda y^2 + (1-\lambda)x^2)\cdot(y^2 - x^2)\,d\lambda.$$

If we let $z^2 = \lambda y^2 + (1-\lambda)\,x^2$, then it is enough to show that

(15) $$\sum_{i=1}^{m}\nabla\tilde{B}_i^1(z^2)\cdot(y^2 - x^2) < 0 \text{ for all } m < n^1.$$

Since $\tilde{B}^1(z^2) = C^1(Az^2)$, we see that

$$\sum_{i=1}^{m}\nabla\tilde{B}_i^1(z^2)\cdot(y^2 - x^2) = \sum_{i=1}^{m}\sum_{j=1}^{n^2}\left(\sum_{k=1}^{n^1}\frac{\partial C_i^1}{\partial \pi_k}(Az^2)A_{kj}\right)(y_j^2 - x_j^2).$$

This expression is negative if

(16) $$\sum_{k=1}^{n^1}(A(y^2 - x^2))_k\left(-\sum_{i=1}^{m}\frac{\partial C_i^1}{\partial \pi_k}(Az^2)\right) > 0.$$

Now $b_k = (A(y^2 - x^2))_k$ is strictly increasing by Lemma A.3, while equations (4) and (5) imply that $c_k = -\sum_{i=1}^{m}\frac{\partial C_i^1}{\partial \pi_k}$ satisfies condition (13). Therefore, Lemma A.2 implies that inequality (16) holds for all $m < n^1$, and hence that $T^1\tilde{B}^1(y^2) \geq T^1\tilde{B}^1(x^2)$. This completes the proof of the proposition. ■

It is worth noting that only two properties of the choice probability function C^α were used to prove Theorem 5.1. To establish condition (13), we used these two facts:

$$\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{\partial C_i^\alpha}{\partial \pi_j} > 0 \text{ for all } k, l < n^\alpha; \tag{17}$$

$$\sum_{j=1}^{n^\alpha}\frac{\partial C_i^\alpha}{\partial \pi_j} = 0 \text{ for all } i \leq n^\alpha. \tag{18}$$

Notably, the symmetry of DC^α, which was essential for establishing our results for other classes of games, was not needed here. In fact, all of our results for supermodular games extend immediately to dynamics based on any choice probability function satisfying (17) and (18). On the other hand, our results for supermodular games *cannot* be extended to all perturbed best response dynamics based on deterministic perturbations of payoffs: the characterization of the corresponding choice functions given after the proof of Theorem 2.1 shows that the inequalities in expression (17) are not even weakly satisfied by all such functions.

The Proof of Theorem 5.2

Define $\tilde{B}: \Sigma \to \Sigma$ by $\tilde{B}(x) = (\tilde{B}^1(x^{-1}), \ldots, \tilde{B}^p(x^{-p}))$. Theorem 5.1 implies that $T\tilde{B}(x) \leq T\tilde{B}(y)$ whenever $Tx \leq Ty$.

Let $\underline{y}$ and $\bar{y}$ be the minimal and maximal elements of Σ. That is, $\underline{y}$ is the mixed strategy profile at which each player chooses his lowest pure strategy with probability 1, and $\bar{y}$ is the profile at which each player chooses his highest pure strategy with probability 1. Let $\tilde{B}^{[k]}$ denote the k-fold iteration of the function $\tilde{B}$. Since $T\tilde{B}(\underline{y}) \geq T\underline{y}$, iteration of $\tilde{B}$ starting from $\underline{y}$ yields an increasing sequence $\{\tilde{B}^{[k]}\}_{k=0}^{\infty}$. Because this sequence is contained in the compact set Σ, its limit $\underline{x}$ exists. Moreover, the continuity of $\tilde{B}$ implies that

$$\underline{x} = \lim_{k\to\infty}\tilde{B}^{[k]}(\underline{y}) = \tilde{B}(\lim_{k\to\infty}\tilde{B}^{[k-1]}(\underline{y})) = \tilde{B}(\underline{x}),$$

and so $\underline{x}$ is a rest point of (P). Similarly, if we iterate $\tilde{B}$ starting from $\bar{y}$, then in the limit we obtain a rest point $\bar{x}$.

Now let x^* be any rest point of (P), so that $\tilde{B}(x^*) = x^*$. Since $x^* \in \Sigma$, $T\underline{y} \leq Tx^* \leq T\bar{y}$. Therefore, if we iteratively apply $\tilde{B}$ to $\underline{y}$, x^*, and $\bar{y}$, then in the limit we obtain $T\underline{x} \leq Tx^* \leq T\bar{x}$, proving the theorem. ■

The Proof of Proposition 5.3

If $\{x_t\}$ solves (P), then

$$\tfrac{d}{dt}T^\alpha x_t^\alpha = T^\alpha(\tfrac{d}{dt}x_t^\alpha) = T^\alpha(\tilde{B}^\alpha(x_t^{-\alpha}) - x_t^\alpha) = T^\alpha\tilde{B}^\alpha((T^{-\alpha})^{-1}(T^{-\alpha}x_t^{-\alpha})) - T^\alpha x_t^\alpha,$$

so $\{Tx_t\}$ solves (T). Conversely, if $\{Tx_t\}$ solves (T), then

$$\tfrac{d}{dt}x_t^\alpha = (T^\alpha)^{-1}(\tfrac{d}{dt}T^\alpha x_t^\alpha) = (T^\alpha)^{-1}(T^\alpha\tilde{B}^\alpha((T^{-\alpha})^{-1}(T^{-\alpha}x_t^{-\alpha})) - T^\alpha x_t^\alpha) = \tilde{B}^\alpha(x_t^{-\alpha}) - x_t^\alpha,$$

so $\{x_t\}$ solves (P). ■

The Proof of Theorem 5.4

Given any subset I of the components of v, we can find a pair of components $(\alpha, i) \in I$ and $(\beta, j) \notin I$ such that $\alpha \neq \beta$. Define the function $\hat{B}^\alpha$: $T^{-\alpha}(\Sigma^{-\alpha}) \to T^\alpha(\Delta S^\alpha)$ by $\hat{B}^\alpha(v^{-\alpha}) = T^\alpha\tilde{B}^\alpha((T^{-\alpha})^{-1}v^{-\alpha})$. To prove the theorem, it is enough to show that $\frac{\partial \hat{B}_i^\alpha}{\partial v_j^\beta}(v^{-\alpha}) > 0$. As in the proof of Theorem 5.1, it is enough to consider a two player game, and to let $\alpha = 1$ and $\beta = 2$.

Let $x^2 = (T^2)^{-1}v^2$. Observe that if e_{j+1} and e_j are standard basis vectors in $\mathbf{R}^{n^2}$, then $T^2(e_{j+1} - e_j) = \iota_j$, a standard basis vector in $\mathbf{R}^{n^2-1}$. It follows that

$$\begin{aligned}
\frac{\partial \hat{B}_i^1}{\partial v_j^2}(v^2) &= \left[\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left(\hat{B}^1(v^2 + \varepsilon\iota_j) - \hat{B}^1(v^2)\right)\right]_i \\
&= \left[\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left(T^1\tilde{B}^1((T^2)^{-1}(v^2 + \varepsilon\iota_j)) - T^1\tilde{B}^1((T^2)^{-1}(v^2))\right)\right]_i \\
&= \left[T^1\left(\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left(\tilde{B}^1(x^2 + \varepsilon(e_{j+1} - e_j)) - \tilde{B}^1(x^2)\right)\right)\right]_i \\
&= \left[T^1(D\tilde{B}^1(x^2)(e_{j+1} - e_j))\right]_i \\
&= \sum_{k=i+1}^{n^1}\nabla\tilde{B}_k^1(x^2)\cdot(e_{j+1} - e_j)
\end{aligned}$$

Since $\sum_{k=1}^{n^1}\tilde{B}_k^1(y^2) = 1$ for all $y^2 \in \Delta S^2$, $\sum_{k=1}^{n^1}\nabla\tilde{B}_k^1(y^2)\cdot z^2 = 0$ for all $y^2 \in \Delta S^2$ and all $z^2 \in \mathbf{R}_0^{n^2}$. We can therefore conclude from equation (15) that

$$\frac{\partial \hat{B}_i^1}{\partial v_j^2}(v^2) = -\sum_{k=1}^{i}\nabla\tilde{B}_k^1(x^2)\cdot(e_{j+1} - e_j) > 0. \;■$$

The Proof of Corollary 5.5 (ii)

Let ϕ: $\mathbf{R}_+ \times \Sigma \to \Sigma$ denote the semiflow of (P), and let $\underline{y}$ and $\bar{y}$ denote the minimal and maximal points in Σ. By part (*i*) of the corollary, we have that $T\phi(t, \underline{y}) \le T\phi(t, x) \le T\phi(t, \bar{y})$ for all $t \ge 0$ and $x \in \Sigma$. The forward invariance of Σ and Theorem 1.2.1 of Smith (1995) imply that $\phi(t, \underline{y})$ and $\phi(t, \bar{y})$ both converge to rest points of (P) as t grows large. Consequently, the previous inequalities and Theorem 5.2 imply that

$$T\underline{x} \le T\hat{x} \le T\bar{x} \quad \text{for all } \hat{x} \in GA(\mathrm{P}) \equiv \bigcap_{t \ge 0} \phi(t, \Sigma).$$

(Moreover, these facts also imply that $\phi(t, \underline{y}) \to \underline{x}$ and that $\phi(t, \bar{y}) \to \bar{x}$.) The set GA(P) is called the *global attractor* of the dynamic (P). It is well known that GA(P) is the maximal invariant subset of Σ and that it is compact and asymptotically stable. Most importantly, Theorem 9.1.3 of Robinson (1995) shows that $CR(\mathrm{P}) \subset GA(\mathrm{P})$, establishing the result. ■

Characterization of Equilibria in Example 5.6

Let G denote the common distribution function for $\varepsilon_1^\alpha - \varepsilon_2^\alpha$. Our assumptions about this difference imply that $G(0) = \frac{1}{2}$, that $G'(0) > \frac{1}{2}$, that G is concave on $(0, \infty)$, and that G is convex on $(-\infty, 0)$. It follows that the equality $G(2p - 1) = p$ holds for exactly three values of p: one above, one below, and one equal to $\frac{1}{2}$.

Since player α's payoffs only depend on the behavior of player $\alpha + 1$ and since each player has exactly two strategies, we can write

$$B_2^\alpha(x_2^{\alpha+1}) = P(x_2^{\alpha+1} + \varepsilon_2^\alpha \ge (1 - x_2^{\alpha+1}) + \varepsilon_1^\alpha) = P(\varepsilon_1^\alpha - \varepsilon_2^\alpha \le 2x_2^{\alpha+1} - 1) = G(2x_2^{\alpha+1} - 1).$$

It follows that there is a rest point with $x_2^\alpha = p$ for all players p if and only if $p = G(2p - 1)$, and so there are exactly three symmetric rest points of (P). Moreover, if $p^+ > \frac{1}{2}$ and $p^- < \frac{1}{2}$ are the weights placed on strategy 2 in the rest points corresponding to the pure equilibria of the underlying game, then our assumptions about the distribution of $\varepsilon_1^\alpha - \varepsilon_2^\alpha$ imply that $g(p^+) = g(p^-) < \frac{1}{2}$; a variation on the stability analysis provided in the text then reveals that these two rest points are stable.

We now show that (P) does not admit any asymmetric rest points. By symmetry, we can write $B(\cdot)$ for $B_2^\alpha(\cdot)$. Now suppose that x is an asymmetric rest point of (P). Then $x_2^\alpha \ne x_2^{\alpha+1}$ for some α. Suppose that $x_2^\alpha < x_2^{\alpha+1}$. Then since each player is playing a best response and since B is strictly increasing, we see that $x_2^{\alpha-1} = B(x_2^\alpha) < B(x_2^{\alpha+1}) = x_2^\alpha$. Iterating (and using the fact that player p's payoffs depend on player 1's behavior), we conclude that $x_2^{\alpha+1} < x_2^{\alpha+1}$, which is a contradiction. If $x_2^{\alpha} > x_2^{\alpha+1}$, we can establish that $x_2^{\alpha+1} > x_2^{\alpha+1}$, which is again a contradiction. ■

To prove parts (*iii*) and (*iv*) of Theorem 6.1, we need to establish that standard stochastic fictitious play Z_t satisfies a global version of Pemantle's (1990) nondegeneracy condition. To state this condition, we let $U = \{\theta \in \prod_\alpha \mathbf{R}_0^{n^\alpha}: \sum_\alpha \sum_i (\theta_i^\alpha)^2 = 1\}$ denote the set of unit vectors in $\prod_\alpha \mathbf{R}_0^{n^\alpha}$, the tangent space of Σ.

Lemma A.4: $\min\limits_{z\in\Sigma} \min\limits_{\theta\in U} E\Big(\max\Big\{\sum_\alpha (\zeta_{t+1}^\alpha - \tilde{B}^\alpha(z^{-\alpha}))\cdot\theta^\alpha, 0\Big\} \Big| Z_t = z\Big) > 0.$

To interpret this condition, recall that if $Z_t = z$, then the increment in beliefs at time $t + 1$ is described by $Z_{t+1}^\alpha - Z_t^\alpha = \frac{1}{t+1}(\zeta_{t+1}^\alpha - z^\alpha)$, while the expected increment is described by $E\big(Z_{t+1}^\alpha - Z_t^\alpha \,\big|\, Z_t = z\big) = \frac{1}{t+1}(\tilde{B}^\alpha(z^{-\alpha}) - z^\alpha)$. Thus, the condition in the lemma requires that there are significant random deviations of the process Z_t from its expected motion; these deviations must be possible from any current state z and in any direction θ. The proof of Lemma A.4 requires this preliminary result.

Lemma A.5: *If $\theta \in U$, then there exists a player β such that* $\max_{i\in S^\beta} \theta_i^\beta > \frac{1}{n^\beta\sqrt{p}}$ *and* $\min_{i\in S^\beta} \theta_i^\beta < -\frac{1}{n^\beta\sqrt{p}}$.

Proof of Lemma A.5: Since θ is of unit length, there must be a player β such that $\tau^\beta \equiv \sum_{i\in S^\beta}(\theta_i^\beta)^2 \geq \frac{1}{p}$. Since $\theta^\beta \in \mathbf{R}_0^{n^\beta}$, we know that k, the number of strictly positive components of θ^β, is between 1 and $n^\beta - 1$; we may suppose the first k components are strictly positive. Now suppose that $\max_{i\in S^\beta}\theta_i^\beta < \sqrt{\tau^\beta/(n^\beta(n^\beta-1))}$. In this case, $\sum_{i=1}^k \theta_i^\beta < k\sqrt{\tau^\beta/(n^\beta(n^\beta-1))}$ and $\sum_{i=1}^k(\theta_i^\beta)^2 < (k\tau^\beta)/(n^\beta(n^\beta-1))$. Since $\theta^\beta \in \mathbf{R}_0^{n^\beta}$, $\sum_{i=k+1}^{n^\beta}\theta_i^\beta = -\sum_{i=1}^k\theta_i^\beta$; moreover, $\sum_{i=k+1}^{n^\beta}(\theta_i^\beta)^2 < (k^2\tau^\beta)/(n^\beta(n^\beta-1))$, as this sum is maximized if exactly one term is non-zero. But then $\sum_{i=1}^{n^\beta}(\theta_i^\beta)^2 < ((k+1)k\tau^\beta)/(n^\beta(n^\beta-1)) \leq \tau^\beta$, which is a contradiction. Therefore, $\max_{i\in S^\beta}\theta_i^\beta \geq \sqrt{\tau^\beta/(n^\beta(n^\beta-1))} > \frac{1}{n^\beta\sqrt{p}}$. The proof of the other claim is similar. ❑

Proof of Lemma A.4: Since the density of each disturbance vector ε^α has full support on $\mathbf{R}^{n^\alpha}$, we can place a uniform lower bound on the probability of an arbitrary strategy being the best response after the payoff disturbances are realized:

$$m \equiv \min_{\alpha} \min_{x^{-\alpha}\in\Sigma^{-\alpha}} \min_{i\in S^{\alpha}} \tilde{B}_i^{\alpha}(x^{-\alpha}) > 0.$$

Now fix $z \in \Sigma$ and $\theta \in U$, and let $S^+ = \{s \in S\colon \sum_{\alpha}(e^{\alpha}_{s^{\alpha}} - \tilde{B}^{\alpha}(z^{-\alpha}))\cdot\theta^{\alpha} > 0\}$, where $e^{\alpha}_{s^{\alpha}}$ is a standard basis vector in $\mathbf{R}^{n^{\alpha}}$. Then

$$\begin{aligned}
&E\Big(\max\Big\{\textstyle\sum_{\alpha}(\zeta^{\alpha}_{t+1} - \tilde{B}^{\alpha}(z^{-\alpha}))\cdot\theta^{\alpha}, 0\Big\}\Big| Z_t = z\Big) \\
&\quad = \sum_{s\in S} P\Big(\zeta^1_{t+1} = e^1_{s^1}, \ldots, \zeta^p_{t+1} = e^p_{s^p} \,\Big|\, Z_t = z\Big) \max\Big\{\textstyle\sum_{\alpha}(e^{\alpha}_{s^{\alpha}} - \tilde{B}^{\alpha}(z^{-\alpha}))\cdot\theta^{\alpha}, 0\Big\} \\
&\quad = \sum_{s\in S^+}\left(\prod_{\gamma}\tilde{B}^{\gamma}_{s^{\gamma}}(z^{-\gamma})\right)\left(\sum_{\alpha}(e^{\alpha}_{s^{\alpha}} - \tilde{B}^{\alpha}(z^{-\alpha}))\cdot\theta^{\alpha}\right) \\
&\quad \geq m^p \max_{s\in S} \sum_{\alpha}(e^{\alpha}_{s^{\alpha}} - \tilde{B}^{\alpha}(z^{-\alpha}))\cdot\theta^{\alpha} \\
&\quad = m^p\left(\max_{s\in S}\sum_{\alpha}\theta^{\alpha}_{s^{\alpha}} - \sum_{\gamma}\tilde{B}^{\gamma}(z^{-\gamma})\cdot\theta^{\gamma}\right) \\
&\quad \geq m^p\left(\sum_{\alpha}\max_{i\in S^{\alpha}}\theta^{\alpha}_i - \sum_{\gamma}\left((1-m)\max_{i\in S^{\gamma}}\theta^{\gamma}_i + m\min_{i\in S^{\gamma}}\theta^{\gamma}_i\right)\right) \\
&\quad = m^{p+1}\sum_{\alpha}\left(\max_{i\in S^{\alpha}}\theta^{\alpha}_i - \min_{i\in S^{\alpha}}\theta^{\alpha}_i\right) \\
&\quad \geq \tfrac{2m^{p+1}}{n^{\beta}\sqrt{p}},
\end{aligned}$$

where β is the player specified in Lemma A.5. ■

The Proof of Theorem 6.1 (iii)

We only consider the case of standard stochastic fictitious play; the proof for the symmetric case is similar. Suppose that G is a potential game and that the distributions of the vectors ε^{α} are smooth enough that the corresponding deterministic perturbations V^{α} are C^N. Then Proposition 4.2 and Theorem 2.1 imply that CR(P) = RP(P). Therefore, Theorem 3.3 of Benaïm and Hirsch (1999a) and Proposition 5.3 of Benaïm (1999) imply that $P(\omega(Z_t)$ is a connected subset of RP(P)) = 1. (In addition, Proposition 6.4 of Benaïm (1999) implies that the Lyapunov function Π is almost surely constant on $\omega(Z_t)$.)

Next, suppose that G is a potential game, that all rest points of (P) are hyperbolic, and that (P) is C^2. Then Proposition 4.3 and Theorem 2.1 show that CR(P) = RP(P) and that this set is finite. Thus, Theorem 3.3 of Benaïm and Hirsch (1999a) and Proposition 5.3 of Benaïm (1999) imply that $P(\lim_{t\to\infty} Z_t \in RP(\text{P})) = 1$. Furthermore, since the rest points of (P) are hyperbolic, each is either linearly stable or linearly unstable. Given this observation, the fact that (P) is C^2, and Lemma A.4, Theorem 1 of Pemantle (1990) implies that $P(\lim_{t\to\infty} Z_t \in LS) = 1$. ■

The Proof of Theorem 6.1 (iv)

Again, we only consider the case of standard stochastic fictitious play; the proof for the symmetric case is similar. Suppose that G is a supermodular game. By Theorem 5.2 and Corollary 5.5, $CR(\text{P}) \subset RP(\text{P}) \cup (\bigcup_i M_i \cap [\underline{x}, \bar{x}])$, where each set M_i is an unstable invariant manifold of (P). Hence, Theorem 3.3 of Benaïm and Hirsch (1999a) and Proposition 5.3 of Benaïm (1999) establish the first statement in the result. If Σ is two dimensional, we can appeal to Proposition 5.3, Theorem 5.4, and Theorem 4.3 of Benaïm (2000), which establishes that when (P) is a C^2, two dimensional, cooperative, and irreducible dynamic, Z_t converges with probability one to a rest point of (P) which is not linearly unstable. Once again, Lemma A.4 provides the nondegeneracy condition which is needed to apply this result. (Under symmetric stochastic fictitious play, the state space ΔS^1 is one dimensional if $n^1 = 2$. In this case, each unstable invariant manifold is simply a rest point, and so our result for this case follows from the C^2 smoothness of (P), Lemma A.4, and Theorem 1 of Pemantle (1990).) ■

REFERENCES

Akin, E. (1993). *The General Topology of Dynamical Systems*. AMS Graduate Studies in Mathematics 1. Providence, RI: American Mathematical Society.

Anderson, S. P., A. de Palma and J.-F. Thisse (1992). *Discrete Choice Theory of Product Differentiation*. Cambridge: MIT Press.

Benaïm, M. (1999). "Dynamics of Stochastic Approximation Algorithms," In *Séminaire de Probabilités XXXIII*, J. Azéma *et. al.*, Eds. Lecture Notes in Mathematics 1709. Berlin: Springer.

— (2000). "Convergence with Probability One of Stochastic Approximation Algorithms Whose Average is Cooperative," *Nonlinearity* **13**, 601-616.

Benaïm, M., and M. W. Hirsch (1999a). "Mixed Equilibria and Dynamical Systems Arising from Repeated Games," *Games Econ. Behav.* **29**, 36-72.

— (1999b). "On Stochastic Approximation Algorithms with Constant Step Size Whose Average is Cooperative," *Ann. Appl. Prob.* **30**, 850-869.

Blume, L. E. (1993). "The Statistical Mechanics of Strategic Interaction," *Games Econ. Behav.* **5**, 387-424.

— (1997). "Population games." In *The Economy as an Evolving Complex System II*, W. B. Arthur, S. N. Durlauf, and D. A. Lane, Eds. Reading, MA: Addison-Wesley.

Brown, G. W. (1951). "Iterative Solutions of Games by Fictitious Play," in *Activity Analysis of Production and Allocation*, T. C. Koopmans, Ed., New York: Wiley.

Chen, H.-C., J. W. Friedman, and J.-F. Thisse (1997). "Boundedly Rational Nash Equilibrium: A Probabilistic Choice Approach," *Games Econ. Behav.* **18**, 32-54.

Conley, C. C. (1978). *Isolated Invariant Sets and the Morse Index*. Regional Conference Series in Mathematics 38. Providence, RI: American Mathematical Society.

Ellison, G., and D. Fudenberg (2000). "Learning Purified Mixed Equilibria," *J. Econ. Theory* **90**, 84-115.

Ely, J. C., and W. H. Sandholm (2000). "Evolution with Diverse Preferences," mimeo, Northwestern University and University of Wisconsin.

Fudenberg, D., and D. M. Kreps (1993). "Learning Mixed Equilibria," *Games Econ. Behav.* **5** (1993), 320-367.

Fudenberg, D., and D. K. Levine (1998). *Theory of Learning in Games*. Cambridge: MIT Press.

Fudenberg, D., and J. Tirole (1992). *Game Theory*. Cambridge: MIT Press.

Gaunersdorfer, A., and J. Hofbauer (1995). "Fictitious Play, Shapley Polygons, and the Replicator Equation," *Games Econ. Behav.* **11**, 279-303.

Gilboa, I., and A. Matsui (1991). "Social Stability and Equilibrium," *Econometrica* **59**, 859-867.

Hahn, S. (1999). "The Convergence of Fictitious Play in 3 x 3 Games with Strategic Complementarities," *Econ. Lett.* **64**, 57-60.

Harsanyi, J. C. (1973a). "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points." *Int. J. Game Theory* **2**, 1-23.

— (1973b). "Oddness of the Number of Equilibrium Points: A New Proof." *Int. J. Game Theory* **2**, 235-250.

Hirsch, M. W. (1988). "Systems of Differential Equations That Are Competitive or Cooperative III: Competing Species," *Nonlinearity* **1**, 51-71.

— (1999). "Chain Transitive Sets for Smooth Strongly Monotone Dynamics," *Dynamics of Continuous, Discrete, and Impulsive Systems* **5**, 529-543.

Hofbauer, J. (1995a). "Imitation Dynamics for Games," mimeo, Universität Wien.

— (1995b). "Stability for the Best Response Dynamics," mimeo, Universität Wien.

— (2000). "From Nash and Brown to Maynard Smith: Equilibria, Dynamics, and ESS," *Selection* **1**, 81-88.

Hofbauer, J., and E. Hopkins (2000). "Learning in Perturbed Asymmetric Games," mimeo, Universität Wien and University of Edinburgh.

Hofbauer, J., and W. H. Sandholm (2001). "Evolution and Learning in Games with Randomly Disturbed Payoffs," mimeo, Universität Wien and University of Wisconsin.

Hofbauer, J., and K. Sigmund (1988). *Theory of Evolution and Dynamical Systems*. Cambridge: Cambridge University Press.

Jordan, J. S. (1993). "Three Problems in Learning Mixed Strategy Nash Equilibria," *Games Econ. Behav.* **5**, 368-386.

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

Kandori, M., and R. Rob (1995). "Evolution of Equilibria in the Long Run: A General Theory and Applications," *J. Econ. Theory* **65**, 383-414.

Kaniovski, Y. M., and H. P. Young (1995). "Learning Dynamics in Games with Stochastic Perturbations," *Games Econ. Behav.* **11**, 330-363.

Krishna, V. (1992). "Learning in Games with Strategic Complementarities," mimeo, Harvard University.

Mallet-Paret, J., and H. L. Smith (1990). "The Poincaré-Bendixson Theorem for Monotone Cyclic Feedback Systems." *J. Dynam. Differential Equations* **2**, 367-421.

Matsui, A. (1992). "Best Response Dynamics and Socially Stable Strategies," *J. Econ. Theory* **57**, 343-362.

Maynard Smith, J., and G. Price (1973). "The Logic of Animal Conflicts," *Nature* **246**, 15-18.

McFadden, D. (1981). "Econometric Models of Probabilistic Choice." In *Structural Analysis of Discrete Data with Econometric Applications*, C. F. Manski and D. McFadden, Eds. Cambridge: MIT Press.

Milgrom, P., and J. Roberts (1990). "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica* **58**, 1255-1278.

— (1991). "Adaptive and Sophisticated Learning in Normal Form Games," *Games Econ. Behav.* **3**, 82-100.

Miyasawa, K. (1961). "On the Convergence of Learning Processes in a 2 x 2 Non-Zero Sum Game," Research Memorandum No. 33, Princeton University.

Monderer, D., and L. Shapley (1996a). "Fictitious Play Property for Games with Identical Interests," *J. Econ. Theory* **68**, 258-265.

— (1996b). "Potential Games," *Games Econ. Behav.* **14**, 124-143.

Pemantle, R. (1990). "Nonconvergence to Unstable Points in Urn Models and Stochastic Approximations," *Ann. Prob.* **18**, 698-712.

Robinson, C. (1995). *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*. Boca Raton, FL: CRC Press.

Robinson, J. (1951). "An Iterative Method of Solving a Game," *Ann. Math.* **24**, 296-301.

Rockafellar, R. T. (1970). *Convex Analysis.* Princeton: Princeton University Press.

Sandholm, W. H. (2000). "Evolutionary Implementation and Congestion Pricing," forthcoming, *Rev. Econ. Stud.*

— (2001). "Potential Games with Continuous Player Sets," *J. Econ. Theory* **97**, 81-108.

Selten, R. (1980). "A Note on Evolutionarily Stable Strategies in Asymmetric Animal Conflicts," *J. Theor. Biol.* **84**, 93-101.

Shapley, L. S. (1964). "Some Topics in Two-Person Games," in *Advances in Game Theory*, M. Drescher, L. S. Shapley, and A. W. Tucker, Eds. Annals of Mathematics Study 52. Princeton: Princeton University Press.

Smith, H. L. (1995). *Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems*. AMS Mathematical Monographs and Surveys 41. Providence, RI: American Mathematical Society.

Topkis, D. (1979). "Equilibrium Points in Nonzero-Sum *n*-Person Submodular Games," *SIAM J. Control Opt.* **17**, 773-787.

Vives, X. (1990). "Nash Equilibrium with Strategic Complementarities," *J. Math. Econ.* **19**, 305-321.

Young, H. P. (1998). *Individual Strategy and Social Structure*. Princeton: Princeton University Press.