Unified Convergence Proofs of Continuous-time Fictitious Play*

Jeff S. Shamma and Gurdal Arslan
Department of Mechanical and Aerospace Engineering
University of California Los Angeles
{shamma,garslan}@seas.ucla.edu

October 1, 2003
revised March 23, 2004

Abstract

We consider a continuous-time version of fictitious play, in which interacting players evolve their strategies in reaction to their opponents' actions without knowledge of their opponents' utilities. It is known that fictitious play need not converge, but that convergence is possible in certain special cases including zero-sum games, identical interest games, and two-player/two-move games. We provide a unified proof of convergence in all of these cases by showing that a Lyapunov function previously introduced for zero-sum games also can establish stability in the other special cases. We go on to consider a two-player game in which only one player has two-moves and use properties of planar dynamical systems to establish convergence.

1 Overview

The procedure of fictitious play was introduced in 1951 [5, 18] as a mechanism to compute Nash equilibria in matrix games. In fictitious play, game players repeatedly use strategies that are best responses to the historical averages, or *empirical frequencies*, of opponents. These empirical frequencies, and hence player strategies may or may not converge. The important implication of convergence of empirical frequencies is that it implies a convergence to a Nash equilibrium.

There is a substantial body of literature on fictitious play [9]. A selected timeline of results that establish convergence for special cases of games is as follows: 1951, two player zero-sum games [18]; 1961, two player two move games [16]; 1993, noisy two player two move games with a unique Nash equilibrium [8]; 1996, multiplayer games with identical player utilities [17]; 1999, noisy two-player/two-move games with countable Nash equilibria [2]; and two player games in which one player has only two moves [3]. A convergence *counterexample* due to Shapley in 1964 has two players with three moves each [20]. A 1993 counterexample due to Jordan has three players with two moves each [12]. Nonconvergence issues are also discussed in [6, 7, 10, 13, 21].

*Research supported by AFOSR/MURI grant #F49620-01-1-0361 and summer support by the UF Graduate Engineering Research Center. To appear in *IEEE Transactions on Automatic Control.*

In this paper, we consider a continuous-time form of fictitious play and provide a unified proof of convergence of empirical frequencies for the special cases of zero-sum games, identical interest games, and two-player/two-move games. The proofs are unified in the sense that they all employ an energy function that has the natural interpretation as an "opportunity for improvement". This energy function was used as a Lyapunov function in [11] for zero-sum games. We show that the same energy function can establish convergence for all of the above cases, in some cases by a Lyapunov argument and in other cases by an integrability argument.

We go on to consider games in which one of two players has only two moves. We provide an alternative proof that exploits some simple properties of planar dynamical systems.

The remainder of this paper is organized as follows. Section 2 sets up the problem of continuous time fictitious play. Section 3 contains convergence proofs for zero-sum, identical interest, and two-player/two-move games. Section 4 discusses games in which one of two players has only two moves. Finally, Section 5 has some concluding remarks.

Notation

— For $i \in \{1, 2, \ldots, n\}$, $-i$ denotes the complementary set $\{1, \ldots, i-1, i+1, \ldots, n\}$.

— Boldface $\mathbf{1}$ denotes the vector $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ of appropriate dimension.

— $\Delta(n)$ denotes the simplex in $\mathcal{R}^n$, i.e.,

$$\left\{ s \in \mathcal{R}^n | s \geq 0 \text{ componentwise, and } \mathbf{1}^T s = 1 \right\}$$

— $\text{Int}(\Delta(n))$ denotes the set of interior points of a simplex, i.e., $s > 0$ componentwise.

— $\mathbf{v}_i \in \Delta(n)$ denotes the i^{th} vertex of the simplex $\Delta(n)$, i.e., the vector whose i^{th} term equals 1 and remaining terms equal 0.

— $\mathcal{H} : \text{Int}(\Delta(n)) \to \mathcal{R}$ denotes the entropy function

$$\mathcal{H}(s) = -s^T \log(s)$$

— $\sigma : \mathcal{R}^n \to \text{Int}(\Delta(n))$ denotes the "logit" or "soft-max" function

$$(\sigma(x))_i = \frac{e^{x_i}}{e^{x_1} + \cdots + e^{x_n}}$$

This function is continuously differentiable. The Jacobian matrix of partial derivatives, denoted $\nabla\sigma(\cdot)$, is

$$\nabla\sigma(x) = \text{diag}(\sigma(x)) - \sigma(x)\sigma^T(x),$$

where $\text{diag}(\sigma(x))$ denotes the diagonal square matrix with elements taken from $\sigma(x)$.

2 Fictitious Play Setup

2.1 Static Game

We consider a two player game with players $\mathcal{P}_1$ and $\mathcal{P}_2$, each with positive integer dimensions m_1 and m_2, respectively. Each player, $\mathcal{P}_i$, selects a strategy, $p_i \in \Delta(m_i)$, and receives a real-valued reward according to the utility function $\mathcal{U}_i(p_i, p_{-i})$. These utility functions take the form

$$\mathcal{U}_1(p_1, p_2) = p_1^T M_1 p_2 + \tau \mathcal{H}(p_1)$$
$$\mathcal{U}_2(p_2, p_1) = p_2^T M_2 p_1 + \tau \mathcal{H}(p_2),$$

characterized by matrices M_i of appropriate dimension and $\tau > 0$.

The standard interpretation is that p_i represent probabilistic strategies. Each player selects an integer action $a_i \in \{1, \ldots, m_i\}$ according to the probability distribution p_i. The reward to player $\mathcal{P}_i$ is

$$\mathbf{v}_{a_i}^T M_i \mathbf{v}_{a_{-i}} + \tau \mathcal{H}(p_i),$$

i.e., the reward to player $\mathcal{P}_i$ is the element of M_i in the a_i^{th} row and a_{-i}^{th} column, plus the weighted entropy of its strategy. For a given strategy pair, (p_1, p_2), the utilities represent the expected rewards

$$\mathcal{U}_i(p_i, p_{-i}) = E\left[\mathbf{v}_{a_i}^T M_i \mathbf{v}_{a_{-i}}\right] + \tau \mathcal{H}(p_i).$$

Define the *best response* mappings,

$$\beta_i : \Delta(m_{-i}) \to \Delta(m_i),$$

by

$$\beta_i(p_{-i}) = \arg \max_{p_i \in \Delta(m_i)} \mathcal{U}_i(p_i, p_{-i}).$$

The best response turns out to be the logit or soft-max function (see Notation section)

$$\beta_i(p_{-i}) = \sigma(M_i p_{-i}/\tau).$$

A Nash equilibrium is a pair $(p_1^*, p_2^*) \in \Delta(m_1) \times \Delta(m_2)$ such that for all $p_i \in \Delta(m_i)$,

$$\mathcal{U}_i(p_i, p_{-i}^*) \leq \mathcal{U}_i(p_i^*, p_{-i}^*), \quad i \in \{1, 2\} \tag{1}$$

i.e., each player has no incentive to deviate from an equilibrium strategy provided that the other player maintains an equilibrium strategy. In terms of the best response mappings, a Nash equilibrium is pair (p_1^*, p_2^*) such that

$$p_i^* = \beta_i(p_{-i}^*), \quad i \in \{1, 2\}.$$

2.2 Discrete-time Fictitious Play

Now suppose that the game is repeated at every time $k \in \{0, 1, 2, \dots\}$. In particular, we are interested in an "evolutionary" version of the game in which the strategies at time k, denoted by $p_i(k)$, are selected in response to the entire prior history of an opponent's actions.

Towards this end, let $a_i(k)$ denote the action of player $\mathcal{P}_i$ at time k, chosen according to the probability distribution $p_i(k)$, and let $\mathbf{v}_{a_i(k)} \in \Delta(m_i)$ denote the corresponding simplex vertex. The *empirical frequency*, $q_i(k)$, of player $\mathcal{P}_i$ is defined as the running average of the actions of player $\mathcal{P}_i$, which can be computed by the recursion

$$q_i(k+1) = q_i(k) + \frac{1}{k+1}(\mathbf{v}_{a_i(k)} - q_i(k)).$$

In discrete-time fictitious play (FP), the strategy of player $\mathcal{P}_i$ at time k is the optimal response to the running average of the opponent's actions, i.e.,

$$p_i(k) = \beta_i(q_{-i}(k)).$$

The case where $\tau = 0$ corresponds to classical fictitious play. Setting τ positive rewards randomization, thereby imposing in so-called mixed strategies. As τ approaches zero, the best response mappings approximate selecting the maximal element since the probability of selecting a maximal element approaches one when the maximal element is unique. The game with τ positive then can be viewed as a smoothed version of the matrix game [8] in which rewards are subject to random perturbations.

2.3 Continuous-time Fictitous Play

Now consider the continuous-time dynamics,

$$\begin{aligned} \dot{q}_1(t) &= \beta_1(q_2(t)) - q_1(t) \\ \dot{q}_2(t) &= \beta_2(q_1(t)) - q_2(t). \end{aligned} \tag{2}$$

We will call these equations *continuous-time FP*. These are the dynamics obtained by viewing discrete-time FP as stochastic approximation iterations and applying associated ordinary differential equation (ODE) analysis methods [1, 14].

3 Convergence Proofs for zero-sum, Identical Interest, and Two-Move Games

We will derive a unified framework which establishes convergence of (2) to a Nash equilibrium of the static game (1) in the aforementioned special cases of zero-sum, identical interest, and two-move games.

Zero-sum and identical interest here refer to the portion of the utility other than the weighted entropy. In other words, zero-sum means that for all p_1 and p_2,

$$p_1^T M_1 p_2 = -p_2^T M_2 p_1,$$

and identical interest means that

$$p_1^T M_1 p_2 = p_2^T M_2 p_1.$$

Strictly speaking, the inclusion of the entropy term does not result in a zero-sum or identical interest game, but we will use these terms nonetheless.

Define the function $V_1 : \Delta(m_1) \times \Delta(m_2) \to [0, \infty)$ as

$$\begin{aligned} V_1(q_1, q_2) &= \max_{s \in \Delta(m_1)} \mathcal{U}_1(s, q_2) - \mathcal{U}_1(q_1, q_2) \\ &= (\beta_1(q_2) - q_1)^T M_1 q_2 + \tau(\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)). \end{aligned}$$

Similarly define

$$V_2(q_2, q_1) = \max_{s \in \Delta(m_2)} \mathcal{U}_2(s, q_1) - \mathcal{U}_2(q_2, q_1).$$

Each V_i has the natural interpretation as the maximum possible reward improvement to player $\mathcal{P}_i$ by using the best response to q_{-i} rather than the specified q_i. Note that by definition,

$$V_i(q_i, q_{-i}) \geq 0,$$

with equality if and only if

$$q_i = \beta_i(q_{-i}).$$

The above functions were used in [11] for zero-sum games, i.e., $M_1 = -M_2^T$, through a Lyapunov argument using $V_1 + V_2$ to show that the continuous-time empirical frequencies converge to a Nash equilibrium.

We will show that the *same* functions can be used to establish convergence to a Nash equilibrium in the case of identical interest games and in the case of two-move games. The identical interest case will not be a Lyapunov argument. Rather, we will show that the sum, $V_1 + V_2$, is integrable. For two-move games, we will show that an appropriately scaled sum, $\alpha_1 V_1 + \alpha_2 V_2$, is either a Lyapunov function or is integrable.

The following lemma reveals a special structure for the derivatives of the V_i along trajectories of continuous-time FP (2).

Lemma 3.1 *Define*

$$\tilde{V}_i(t) = V_i(q_i(t), q_{-i}(t))$$

along solutions of continuous-time FP (2). Then

$$\begin{aligned} \dot{\tilde{V}}_1 &\leq -\tilde{V}_1 + \dot{q}_1^T M_1 \dot{q}_2, \\ \dot{\tilde{V}}_2 &\leq -\tilde{V}_2 + \dot{q}_2^T M_2 \dot{q}_1. \end{aligned}$$

The proof uses the following lemma.

Lemma 3.2 ([4], Lemma 3.3.1) *Let $F(x,u)$ be a continuously differentiable function of $x \in \mathcal{R}^n$ and $u \in \mathcal{R}^m$. Let U be a convex subset of $\mathcal{R}^m$. Assume that $\mu^*(x)$ is a continuously differentiable function such that for all x,*

$$\mu^*(x) = \arg\max_{u \in U} F(x,u).$$

Then

$$\nabla_x \left(\max_{u \in U} F(x,u)\right) = \nabla_x F(x, \mu^*(x)).$$

Proof (Lemma 3.1) By definition

$$\begin{aligned}\dot{\tilde{V}}_1 &= -(M_1 q_2)^T(\beta_1(q_2) - q_1) - \tau \frac{d}{dt}\mathcal{H}(q_1(t)) + (\beta_1(q_2) - q_1)^T M_1 \dot{q}_2 \\ &= -\tilde{V}_1 + \tau(\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) - \frac{d}{dt}\mathcal{H}(q_1(t)) + \dot{q}_1^T M_1 \dot{q}_2,\end{aligned}$$

where we used Lemma 3.2 to show that

$$\nabla_{q_2} \max_{s \in \Delta(m_2)} \left(s^T M_1 q_2 + \tau \mathcal{H}(s)\right) = \beta_1^T(q_2) M_1.$$

The lemma follows by noting that concavity of $\mathcal{H}(\cdot)$ implies that [19, Theorem 25.1]

$$\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1) \leq \nabla \mathcal{H}(q_1)(\beta_1(q_2) - q_1) = \frac{d}{dt}\mathcal{H}(q_1(t)).$$

Similar statements apply for V_2. ∎

3.1 Zero-sum and Identical Interest Games

Theorem 3.1 *Assume that either*

$$M_1 = -M_2^T$$

or

$$M_1 = M_2^T$$

Then solutions of continuous-time FP (2) satisfy

$$\begin{aligned}\lim_{t\to\infty} \left(q_1(t) - \beta_1(q_2(t))\right) &= 0 \\ \lim_{t\to\infty} \left(q_2(t) - \beta_2(q_1(t))\right) &= 0.\end{aligned}$$

Proof Define

$$V_{12}(t) = \tilde{V}_1(t) + \tilde{V}_2(t).$$

Zero-sum (see also [11]): $M_1 = -M_2^T$

In case $M_1 = -M_2^T$, summing the expressions for $\dot{\tilde{V}}_i$ in Lemma 3.1 results in a cancellation of terms, thereby producing

$$\dot{V}_{12} + V_{12} \leq 0.$$

Since,

$$V_{12} \geq 0$$

with equality only at an equilibrium point of (2), the theorem follows from standard Lyapunov arguments.
Identical interest: $M_1 = M_2^T$

By definition,

$$\begin{aligned}\tilde{V}_1 - \tau(\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) &= \dot{q}_1^T M_1 q_2 \\ \tilde{V}_2 - \tau(\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)) &= \dot{q}_2^T M_2 q_1.\end{aligned}$$

Therefore

$$V_{12} - \tau(\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) - \tau(\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)) = \dot{q}_1^T M_1 q_2 + \dot{q}_2^T M_2 q_1.$$

Under the assumption $M_1 = M_2^T = M$,

$$\frac{d}{dt}\left(q_1^T(t) M q_2(t)\right) = \dot{q}_1^T M_1 q_2 + \dot{q}_2^T M_2 q_1.$$

Therefore

$$V_{12} = \tau(\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) + \tau(\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)) + \frac{d}{dt}\left(q_1(t)^T M q_2(t)\right).$$

By concavity of $\mathcal{H}(\cdot)$ and [19, Theorem 25.1],

$$\begin{aligned}\tau(\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) &\leq \tau \frac{d}{dt}\mathcal{H}(q_1(t)) \\ \tau(\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)) &\leq \tau \frac{d}{dt}\mathcal{H}(q_2(t)),\end{aligned}$$

which implies that for any $T > 0$

$$\int_0^T V_{12} \leq \left(q_1^T(t) M q_2(t) + \tau\mathcal{H}(q_1(t)) + \tau\mathcal{H}(q_2(t))\right)\Big|_{t=0}^{T}.$$

The integrand is positive, and $T > 0$ is arbitrary. Furthermore, one can show that $\dot{V}_{12}$ is bounded. Therefore $V_{12}(t)$ asymptotically approaches zero as desired. ∎

We comment that the integrability argument above can be viewed a version of the discrete-time argument in [17], but applied to a smoothed game (i.e., $\tau > 0$) in continuous-time.

3.2 Two-Move Games

We now consider the case in which each player in the original static game has two moves, i.e., $m_1 = m_2 = 2$.

Continuous-time FP dynamics (2) involve differences of probability distributions. Since these distributions live on the simplex, their elements sum to one. Therefore, the sum of the elements of the *difference* of two distributions must equal zero, i.e., differences of distributions must lie in the subspace spanned by the vector

$$N = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Using this fact, we see that necessarily

$$\dot{q}_1(t) = \beta_1(q_2(t)) - q_1(t) = N w_1(t)$$

for an appropriately defined scalar variable $w_1(t)$. Similarly

$$\dot{q}_2(t) = N w_2(t).$$

This observation will be the key to proving the desired results in the two-move game. Two separate cases will emerge,

$$(N^T M_1 N)(N^T M_2 N) < 0$$

or

$$(N^T M_1 N)(N^T M_2 N) > 0.$$

In the first case, the proof will follow the same Lyapunov argument of the zero-sum proof. In the second case, the proof will follow the integrability argument of the identical interest proof. Reference [17] suggests a link between the proofs for zero-sum, identical interest, and two-player two-move games. Namely, it states that non-degenerate two-player two-move games are best-response equivalent in mixed-strategies to an appropriate zero-sum or identical interest game, and since fictitious play relies on best responses, this equivalence establishes convergence. The present approach does not utilize this equivalence, but does exploit the present zero-sum and identical interest proofs by establishing a direct link in terms of the constructed storage functions, V_i.

Theorem 3.2 *Assume that* $m_1 = m_2 = 2$ *and*

$$(N^T M_1 N)(N^T M_2 N) \neq 0.$$

Then solutions of continuous-time FP (2) satisfy

$$\lim_{t \to \infty} \big(q_1(t) - \beta_1(q_2(t))\big) = 0$$
$$\lim_{t \to \infty} \big(q_2(t) - \beta_2(q_1(t))\big) = 0.$$

Proof First suppose that $N^T M_1 N$ and $N^T M_2 N$ have opposite signs. Then there exist positive scalars, α_1 and α_2, such that

$$(N^T M_1 N)\alpha_1 + (N^T M_2 N)\alpha_2 = 0. \tag{3}$$

From Lemma 3.1,

$$\dot{\tilde{V}}_1 + \tilde{V}_1 \leq \dot{q}_1^T M_1 \dot{q}_2$$
$$\dot{\tilde{V}}_2 + \tilde{V}_2 \leq \dot{q}_2^T M_2 \dot{q}_1.$$

Since $m_1 = m_2 = 2$,

$$\dot{\tilde{V}}_1 + \tilde{V}_1 \leq N^T M_1 N w_1 w_2$$
$$\dot{\tilde{V}}_2 + \tilde{V}_2 \leq N^T M_2 N w_1 w_2.$$

Scaling the above equations by α_1 and α_2, respectively, and summing them results in a cancellation of the $N^T M_i N$ terms and leads to

$$\dot{V}_{12} + V_{12} \leq 0,$$

where V_{12} is now defined as

$$V_{12} = \alpha_1 \tilde{V}_1 + \alpha_2 \tilde{V}_2,$$

with α_i calculated in (3). As in the zero-sum case,

$$V_{12} \geq 0,$$

with equality only at an equilibrium point of continuous-time FP (2). Standard Lyapunov arguments imply the desired result.

Now suppose that $N^T M_1 N$ and $N^T M_2 N$ have the same sign. Then there exist positive scalars, α_1 and α_2, such that

$$(N^T M_1 N)\alpha_1 = (N^T M_2 N)\alpha_2. \tag{4}$$

By definition,

$$\tilde{V}_1 - \tau(\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) = \dot{q}_1^T M_1 q_2$$
$$\tilde{V}_2 - \tau(\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)) = \dot{q}_2^T M_2 q_1.$$

Since $\frac{1}{2} N N^T$ is a projection matrix and $\dot{q}_i = N w_i$,

$$\dot{q}_1^T M_1 q_2 = (\dot{q}_1^T N N^T M_1 q_2)/2$$
$$\dot{q}_2^T M_2 q_1 = (\dot{q}_2^T N N^T M_2 q_1)/2.$$

Define

$$f(t) = (q_1^T(t) N N^T M_1 q_2(t)\alpha_1 + q_2^T(t) N N^T M_2 q_1(t)\alpha_2)/2.$$

Then

$$
\begin{aligned}
2\dot{f} &= \dot{q}_1^T N N^T M_1 q_2 \alpha_1 + \dot{q}_2^T N N^T M_2 q_1 \alpha_2 \\
&\quad + q_1^T N N^T M_1 \dot{q}_2 \alpha_1 + q_2^T N N^T M_2 \dot{q}_1 \alpha_2 \\
&= (\tilde{V}_1 - \tau(\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)))\alpha_1 + (\tilde{V}_2 - \tau(\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)))\alpha_2 \\
&\quad + q_1^T N N^T M_1 N w_2 \alpha_1 + q_2^T N N^T M_2 N w_1 \alpha_2.
\end{aligned}
$$

Define κ as

$$
\kappa = (N^T M_1 N)\alpha_1 = (N^T M_2 N)\alpha_2.
$$

Then

$$
\begin{aligned}
q_1^T N N^T M_1 N w_2 \alpha_1 + q_2^T N N^T M_2 N w_1 \alpha_2 &= (q_1^T N w_2 + q_2^T N w_1)\kappa \\
&= (q_1^T \dot{q}_2 + q_2^T \dot{q}_1)\kappa \\
&= \kappa \frac{d}{dt}\left(q_1^T q_2\right).
\end{aligned}
$$

Finally, define

$$
V_{12} = \alpha_1 \tilde{V}_1 + \alpha_2 \tilde{V}_2,
$$

with α_i calculated in (4). Then similarly to the identical interest case

$$
V_{12} - \alpha_1 \tau(\mathcal{H}(\beta_1(q_2)) - \mathcal{H}(q_1)) - \alpha_2 \tau(\mathcal{H}(\beta_2(q_1)) - \mathcal{H}(q_2)) = 2\dot{f} - \kappa \frac{d}{dt}\left(q_1^T q_2\right),
$$

which, using concavity of $\mathcal{H}(\cdot)$ and [19, Theorem 25.1], implies that

$$
V_{12} \le 2\dot{f} - \kappa \frac{d}{dt}\left(q_1^T q_2\right) + \alpha_1 \tau \frac{d}{dt}\mathcal{H}(q_1(t)) + \alpha_2 \tau \frac{d}{dt}\mathcal{H}(q_2(t)).
$$

Therefore

$$
\int_0^\infty V_{12}(t) < \infty,
$$

which again leads to the desired result. ∎

4 Two Player Games with One Player Restricted to Two Moves

The previous sections used energy arguments, either Lyapunov or integrability, with the same energy functions that represent the "opportunity for improvement". Reference [2] uses properties of planar dynamical systems to establish convergence for two-player/two-move games. Reference [3] considers games in which only *one* of the players has two moves, and uses a relatively extended argument to establish convergence by eliminating the possibility of so-called Shapley polygons. In this section, we also consider games in which only one player has two moves, but we will apply properties of planar dynamical systems to provide an alternative proof.

Suppose that player $\mathcal{P}_1$ has only two moves, i.e., $m_1 = 2$. Following [3], we will introduce a change of variables that will lead to planar dynamics that describe the evolution of the player P_1's strategy.

Define

$$\phi : \mathcal{R} \rightarrow \text{Int}(\Delta(2))$$

as

$$\phi(\delta) = \begin{pmatrix} \frac{e^\delta}{e^\delta+1} \\ \frac{1}{e^\delta+1} \end{pmatrix}.$$

Then it is straightforward to show that the two-dimensional softmax function, $\sigma : \mathcal{R}^2 \rightarrow \text{Int}\Delta(2)$ can be written as

$$\sigma(\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}) = \phi(v_1 - v_2),$$

i.e., $\sigma(\cdot)$ only depends on the *difference* of v_1 and v_2.

We will exploit this equivalence as follows. Define

$$N = \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

and define the scalar

$$w_2 = \frac{1}{\tau} N^T M_1 q_2.$$

Then a subset of the continuous fictitious play dynamics can be written as

$$\begin{aligned} \dot{q}_1 &= \beta_1(q_2) - q_1 = \phi(w_2) - q_1 \\ \dot{w}_2 &= \frac{1}{\tau} N^T M_1 \left(\beta_2(q_1) - q_2\right) = \frac{1}{\tau} N^T M_1 \beta_2(q_1) - w_2. \end{aligned}$$

Since q_1 evolves in the simplex interior, the scalar $w_1(t)$ is uniquely defined by

$$q_1(t) = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} + N w_1(t).$$

Furthermore, $\dot{w}_1$ satisfies

$$\dot{w}_1(t) = \frac{1}{2} N^T \dot{q}_1(t).$$

The result of introducing w_1 and w_2 is that a subset of the continuous fictitious play dynamics can be expressed completely in terms of w_1 and w_2, namely,

$$\begin{aligned} \dot{w}_1(t) &= \frac{1}{2} N^T \phi(w_2(t)) - w_1(t) \\ \dot{w}_2(t) &= \frac{1}{\tau} N^T M_1 \beta_2 \left(\begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} + N w_1(t) \right) - w_2(t). \end{aligned} \quad (5)$$

Theorem 4.1 *Assume a finite number of Nash equilbria satisfying (1). Assume further that* $m_1 = 2$. *Then solutions of continuous-time FP (2) satisfy*

$$\begin{aligned} \lim_{t\to\infty} \left(q_1(t) - \beta_1(q_2(t))\right) &= 0 \\ \lim_{t\to\infty} \left(q_2(t) - \beta_2(q_1(t))\right) &= 0. \end{aligned}$$

Proof Equations (5)a–b are planar dynamics that describe the evolution of $q_1(t)$. These dynamics form area contracting flow, due to the negative divergence of the right-hand-side. Furthermore, solutions evolve over a bounded rectangular set. A suitable modification of Bendixson's criterion [15] leads to the conclusion that the only ω-limit points are equilibria. In the original coordinates, this implies that $q_1(t)$ converges, and hence so does $q_2(t)$. ∎

5 Concluding Remarks

This paper has provided unified energy based convergence proofs for several special cases of games under fictitious play. These proofs of convergence of continuous-time fictitious play, in themselves, do not immediately guarantee the almost sure convergence of discrete-time fictitious play. Additional arguments are needed to establish that the deterministic continuous-time limits completely capture the stochastic discrete-time limits. Such issues are discussed in general in [1] and specifically for fictitious play in [2].

References

[1] M. Benaim and M.W. Hirsch. A dynamical systems approach to stochastic approximation. *SIAM Journal on Control and Optimization*, **34**:437–472, 1996.

[2] M. Benaim and M.W. Hirsch. Mixed equilibria and dynamical systems arising from fictitious play in perturbed games. *Games and Economic Behavior*, **29**:36–72, 1999.

[3] U. Berger. Fictitious play in $2 \times n$ games. Economics Working Paper Archive at WUSTL, mimeo, 2003.

[4] D.P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont, MA, 1995.

[5] G.W. Brown. Iterative solutions of games by fictitious play. In T.C. Koopmans, editor, *Activity Analysis of Production and Allocation*, pages 374–376. Wiley, New York, 1951.

[6] G. Ellison and D. Fudenberg. Learning purified mixed equilibria. *Journal of Economic Theory*, **90**:83–115, 2000.

[7] D.P. Foster and H.P. Young. On the nonconvergence of fictitious play in coordination games. *Games and Economic Behavior*, **25**, 79–96.

[8] D. Fudenberg and D. Kreps. Learning mixed equilibria. *Games and Economic Behavior*, **5**:320–367, 1993.

[9] D. Fudenberg and D.K. Levine. *The Theory of Learning in Games*. MIT Press, Cambridge, MA, 1998.

[10] S. Hart and A. Mas-Colell. Uncoupled dynamics do not lead to Nash equilibrium. *American Economic Review*, **93**(5):1830–1836, 2003.

[11] J. Hofbauer and W. Sandholm. On the global convergence of stochastic fictitious play. *Econometrica*, **70**:2265–2294, 2002.

[12] J. Jordan. Three problems in game theory. *Games and Economic Behavior*, **5**:368–386, 1993.

[13] V. Krishna and T. Sjöström. On the convergence of fictitious play. *Mathematics of Operations Research*, **23**(2):479–511, 1998.

[14] H.J. Kushner and G.G. Yin. *Stochastic Approximation Algorithms and Applications*. Springer-Verlag, 1997.

[15] C.C. McCluskey and J.S. Muldowney. Stability implications of Bendixson's criterion. *SIAM Review*, **40**:931–934, 1998.

[16] K. Miyasawa. On the convergence of learning processes in a 2×2 non-zero-sum two person game. Technical Report 33, Economic Research Program, Princeton University, 1961.

[17] D. Monderer and L.S. Shapley. Fictitious play property for games with identical interests. *Journal of Economic Theory*, **68**:258–265, 1996.

[18] J. Robinson. An iterative method of solving a game. *Ann. Math.*, **54**:296–301, 1951.

[19] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, 1996.

[20] L.S. Shapley. Some topics in two-person games. In L.S. Shapley M. Dresher and A.W. Tucker, editors, *Advances in Game Theory*, pages 1–29. University Press, Princeton, NJ, 1964.

[21] A. Vasin. On stability of mixed equilibria. *Nonlinear Analysis*, **38**:793–802, 1999.