Working Paper

On the Nonconvergence of Fictitious Play in Coordination Games

Dean P. Foster and H. Peyton Young***

WP-96-33
May 1996 Revised

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

On the Nonconvergence of Fictitious Play in Coordination Games

Dean P. Foster and H. Peyton Young***

WP-96-33
May 1996 Revised

*Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104-6302

**Department of Economics, Johns Hopkins University
Baltimore, MD 21218-2685
and
International Institute for Applied Systems Analysis
Laxenburg, Austria

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

Abstract

It is shown by example that learning rules of the fictitious play type fail to converge in certain kinds of coordination games. By contrast, learning rules in which past actions are eventually forgotten and which incorporate small stochastic perturbations are better behaved: over the long run, players manage to coordinate with probability one.

On the Nonconvergence of Fictitious Play in Coordination Games

Dean Foster and Peyton Young

Although it is well-known that fictitious play does not converge to a Nash equilibrium in general (Shapley, 1964), fictitious play does converge for quite a few games having economic significance. These include zero-sum games (Robinson, 1951), two-person two-strategy games (Miyasawa, 1961), dominance solvable games (Milgrom and Roberts, 1991), two-person games with strategic complementarities and diminishing returns (Krishna, 1992), and games with identical interests, that is games that are best reply equivalent in mixed strategies to a game in which all players have identical payoff functions (Monderer and Shapley, 1993a).[1]

One might hope that fictitious play also works for coordination games. The reason is that coordination games have a natural positive reinforcement property: in a neighborhood of every coordination equilibrium every best-reply path gravitates toward the equilibrium, while in a neighborhood of any mixed equilibrium there is a best-reply path leading away from the equilibrium. We shall show by example, however, that this is not sufficient -- even in a coordination game, fictitious play can still get trapped in cyclic behavior that is far from equilibrium.

A finite, two-person game G is a *coordination game* if the players have the same number of strategies, which can be indexed so that is always a strict Nash equilibrium for both players to play strategies having the same index. Thus the payoff matrix has the form $\Pi = (a_{ij}, b_{ij})$, where $1 \leq i \leq n$, $1 \leq j \leq n$, and $a_{ii} > a_{ij}$, $b_{ii} > b_{ji}$ for all i and j.

Consider any infinite sequence of pure strategy pairs $S = \{(i_1, j_1), (i_2, j_2), \ldots (i_t, j_t), \ldots\}$. Let $p_t(S)$ denote the empirical frequency distribution of the choices $\{i_1, i_2, \ldots, i_t\}$ up through time t, and let $q_t(S)$ denote the empirical frequency distribution of the choices $\{j_1, j_2, \ldots, j_t\}$ up throught time t. The sequence S is a *fictitious play sequence* if there is some $t \geq 1$ such that for every $t' > t$, $i_{t'}$ is a best reply to $q_{t'-1}(S)$ and $j_{t'}$ is a best reply to

[1] For related work see Monderer and Sela (1992), Monderer and Shapley (1993b), Fudenberg and Kreps (1993), Kaniovski and Young (1995).

$p_{t'-1}(S)$. The game G has the *fictitious play property* if every limit point of every fictitious play sequence is a Nash equilibrium (pure or mixed) of G. In the next two sections we exhibit coordination games that do not have the fictitious play property.

2. The doctrines game.

Two groups of academics periodically announce a position on some matter of scholarly doctrine. There are three basic doctrines A, B, and C. Each doctrine has two variants: A' and A", B' and B", C' and C". Both groups would like to coordinate on the same version of the same doctrine. A *squabble* is a situation in which the groups choose different versions of the same doctrine. The payoffs are set up so that once a squabble begins, the two groups keep shifting position in a way that generates a new squabble. Specifically, although both players prefer to coordinate on B' or on B" in preference to an A-squabble, Row prefers B' to B" while Column prefers B" to B'. This generates a B-squabble. Both prefer C' and C" to a B-squabble, but Row prefers C' to C" while Column prefers C" to C', which generates a C-squabble. In short, once the academics disagree on the fine points of doctrine, they are unable to reach an agreement on the main points. This situation can be represented by the payoff matrix shown in Table 1.

	A'	A"	B'	B"	C'	C"	D'	D"
A'	24, 24	6, 6	0, 18	0, 18	18, 0	18, 0	5, 0	0, 0
A"	6, 6	24, 24	0, 18	0, 18	18, 0	18, 0	4, 0	0, 0
B'	18, 0	18, 0	24, 24	6, 6	0, 18	0, 18	3, 0	0, 0
B"	18, 0	18, 0	6, 6	24, 24	0, 18	0, 18	2, 0	0, 0
C'	0, 18	0, 18	18, 0	18, 0	24, 24	6, 6	1, 0	0, 0
C"	0, 18	0, 18	18, 0	18, 0	6, 6	24, 24	0, 0	0, 0
D'	0, 4	0, 5	0, 2	0, 3	0, 0	0, 1	25, 24	-25,-25
D"	0, 0	0, 0	0, 0	0, 0	0, 0	0, 0	-25, -25	24, 25

Table 1. Payoff matrix of the doctrines game.

The strategies D' and D'' serve solely as tie-breaking devices; the main action is on the remaining six strategies. Consider a fictitious play sequence that begins with Row choosing D' and Column choosing D''. In the next period the best replies are D'' for Row and D' for Column, and the process unfolds as shown below:

	t = 1	2	3	4	5	6	...	17	18	19	20	...	91	92	93	
Row	D'	D''	A'	A''	B'	B''	...	B'	B''	C'	C''	...	C'	C''	A'	...
Column	D''	D'	A''	A'	B''	B'	...	B''	B'	C''	C'	...	C''	C'	A''	...

Given the initial choices D' and D'', Row has a slight preference for prime over double-prime whereas Column prefers the reverse. This leads the players to coordinate on the same basic doctrine, but never on the same version of that doctrine. Once involved in a squabble they try to rectify the situation by imitating what the other did in the previous period, which leads to a new squabble. Thus the process cycles endlessly.

Theorem 1. *The doctrines game does not have the fictitious play property.*

Proof. Let t_k be the number of periods in squabble k. The first three squabbles are of length $t_1 = 2$, $t_2 = 14$, and $t_3 = 74$. In general, the t_k satisfy the following recursive equation

$$\text{for all } k \geq 1, \quad t_{k+2} = 6t_{k+1} - 5t_k. \qquad (1)$$

From this it follows that each squabble is about five times as long as the previous one. Hence the empirical frequency distribution of strategies (p_t, q_t) never converges. *Afortiori* it does not converge to an equilibrium. Indeed, we can show that no limit point of the process is close to a Nash equilibrium. Suppose that (p^*, q^*) is a limit point. Then there exist α, β, γ such that

$$p^*_{A'} = p^*_{A''} = q^*_{A'} = q^*_{A''} = \alpha,$$
$$p^*_{B'} = p^*_{B''} = q^*_{B'} = q^*_{B''} = \beta,$$
$$p^*_{C'} = p^*_{C''} = q^*_{C'} = q^*_{C''} = \gamma.$$

For this to be a Nash equilibrium, we must have $\alpha = \beta = \gamma$. But this is not the case, because at least one pair of these numbers must be in the ratio of about 5 to 1.

To prove (1) we proceed by induction on k. For $k = 1$ the result follows by plugging in the values $t_1 = 2$, $t_2 = 14$, and $t_3 = 74$. Suppose then that $k > 1$. Since the game is symmetric in A, B, C, and the squabbling proceeds in the cycle $A \to B \to C \to A$, there is no loss of generality in assuming that the $(k + 2)$nd is an A-squabble, the $(k + 1)$st is a C-squabble, and the kth is a B-squabble. To determine which strategy is a best response by Row at any given time t, it suffices to compute the hypothetical total payoff (to Row) of each strategy, assuming it were played against all previous choices by Column up through time t - 1. Call this the *score* of the strategy at time t. Fictitious play stipulates that in each period Row choose some strategy with highest score.

Consider the $(k + 2)$nd squabble, which by assumption is an A-squabble. Each time that Column plays A'A" in succession, both A-strategies for Row increase their score by $24 + 6 = 30$, both B-strategies increase their score by $18 + 18 = 36$, and both C-strategies increase their score by zero. Thus B' gains 6 points relative to A' in every two successive periods of an A-squabble.

Let $S_{A'}$ and $S_{B'}$ be the scores of A' and B' at the beginning of the current squabble. Let $[x]$ denote the least integer greater than or equal to x. Then it takes $t_{k+2} = 2[(S_{A'} - S_{B'})/6]$ periods for B' to overtake A' (i.e., for B' to become a better reply than A' by Row), which ends this squabble and starts the next one.

It remains to compute the difference $S_{A'} - S_{B'}$. Consider the first period of the kth squabble. At this point, B' has just overtaken A'. Moreover if their scores are $S^*_{A'}$ and $S^*_{B'}$, then we have $0 < S^*_{B'} - S^*_{A'} < 6$. (This is because they start period 3 with a difference that is less than 6, and all subsequent actions change the scores by multiples of 6.) During the ensuing B-squabble, which lasts for t_k periods, A' increases its score by 0, B' increases its score by $30t_k/2$, and C' increases its score by $36t_k/2$. After this a C-squabble commences. This increases the score of A' by $36t_{k+1}/2$, the score of B' by 0, and the score of C' by $30t_{k+1}/2$. Thus we have

$$S_{A'} = 36t_{k+1}/2 + S^*_{A'} \text{ and } S_{B'} = 30\, t_k/2 + S^*_{B'}. \qquad (2)$$

We may assume by induction that t_k and t_{k+1} are even. From (2) it follows that

$$(S_{A'} - S_{B'})/6 = 6t_{k+1}/2 - 5t_k/2 + (S^*_{A'} - S^*_{B'})/6.$$

We also know that $-5 \leq (S^*_{A'} - S^*_{B'}) \leq -1$. Hence

$$[(S_{A'} - S_{B'})/6] = 6t_{k+1}/2 - 5t_k/2,$$

and therefore

$$t_{k+2} = 2\,[(S_{A'} - S_{B'})/6] = 6t_{k+1} - 5t_k.$$

Hence t_{k+2} is even and formula (1) holds for k, from which it follows by induction that (1) holds for all k. This completes the proof of theorem 1.

We remark that this serves as a counterexample to various other conjectures that one might entertain about fictitious play. Consider any finite, two-person game G with strategy sets X and Y. A *(one-sided) best reply path* is a sequence of pairs $(x^1, y^1), (x^2, y^2)$, . . ., (x^k, y^k) such that x^{i+1} is a best reply to x^i or y^{i+1} is a best reply to y^i for $1 \leq i < k$. G is *acyclic* if no best reply path forms a cycle.[2] One might have thought that fictitious play converges for acyclic games. Since every coordination game is acyclic, however, the above example shows that this is not the case.

3. The merry-go-round game.

The doctrines game is somewhat delicate, because just a slight perturbation of the payoffs would cause the example to fail. Moreover, even if the payoffs are as given, it only takes one deviation by one player for the process to get onto a path leading to coordination. (If the players manage to coordinate once in the midst of a squabble, the squabble will be broken and they will coordinate in all subsequent periods.) The question therefore arises whether there are coordination games in which fictitious play fails to converge in a more robust sense.

Consider the following situation. Rick and Cathy are in love, but they are not allowed to communicate. Once a day, at an appointed time, they can take a ride on a merry-go-round. The merry-go-round has m pairs of horses, where m is odd. Before taking a

---

[2] Monderer and Shapley (1988) call this the *no-cycling condition for the one-sided best reply dynamic.* Young (1993a) considered the class of *weakly acyclic games,* which have the property that for any initial pair (x^1, y^1) there exists a one-sided best reply path ending in a sink (i.e., ending in a strict Nash equilibrium).

ride, each of them chooses one of the m pairs without communicating their choice to the other. There are no other riders. If they book the same pair they get to ride side-by-side, which is their preferred outcome. If they choose different pairs, their payoffs depend on how conveniently they can look at each other. The merry-go-round operates clockwise as shown in Figure 1.

If Rick chooses pair 1 and Cathy chooses 2, for example, then Rick can see Cathy but she cannot easily see him, because the horses all face in the clockwise direction. Say this outcome has payoff 4 for Rick and 0 for Cathy. If they sit side by side they can look at each other to their hearts' content (but not talk); this has a payoff of 6. If they are on opposite sides of the circle, they can both see each other easily, but the one who has to crane his neck less has a slightly higher payoff (5 compared to 4). (They are both stiff-necked.)

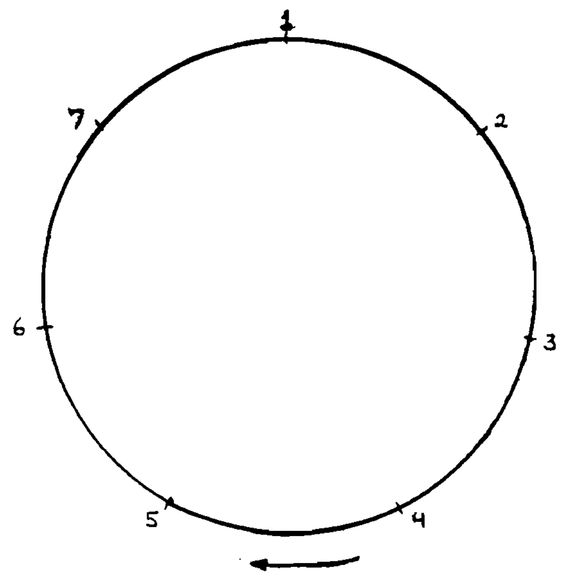

Figure 1. The merry-go-round game.

A similar story can be told for two players who want to coordinate on some *date* in the calendar. Suppose, for example, that two scientific societies hold their meetings once a year. Periodically they reconsider the date on which they hold their meetings, and both announce the new dates simultaneously without communicating their intentions. There are 365 possible dates. Ideally they would like to name the same date, because this generates a lot of media coverage about progress in science. (We shall assume their memberships are disjoint, so there is no competition for member participation). If one society holds its meeting "earlier" than the other, however, it casts a "publicity shadow" on the later one. Earlier and later are cyclical concepts: June 1 is earlier than December

1, which is earlier than January 1 (of the following year) and so forth. The new meeting dates are assumed to remain in place long enough so that there is no appreciable spill-over effect when a transition to new dates occurs. Over a period of years, the payoffs will have the same general form as in the merry-go-round game.

Theorem 2. *The merry-go-round game does not have the fictitious play property whenever the number of positions is odd and at least seven.*

Proof. We shall construct an initial sequence, such that fictitious play from that point on generates empirical distributions (p_t, q_t) whose limit points are not close to any Nash equilibrium of the game.

Each pair of horses will be called a *position*. Proceeding clockwise, number the positions 1, 2, . . . , m, where m = 2k + 1 is odd. Rick (R) is the row player and Cathy (C) is the column player. Let the process begin in period 1 with R in position 1 and C in position k. We are going to create a sequence of moves of the following form: after a certain number of periods t_1, C moves to position k + 1. After t_2 more periods, R moves to position 2. After t_3 more periods, C moves to position k + 2, and so forth. In other words, R and C alternate in moving clockwise around the circle, with C always k or k + 1 steps "ahead" of R. Such a sequence is called a *chase*. The payoffs from a chase are fairly good, because both parties can see each other without craning their necks too much. But both would be better off if they sat side-by-side. We shall demonstrate the existence of an initial chase (not necessarily generated by fictitious play), such that in every subsequent period, fictitious play continues the chase*ad infinitum*.

Let the initial chase be represented by the sequence $t_1, t_2, \ldots, t_j, \ldots$, where t_j is the number of consecutive periods that the process spent in a situation where neither R nor C moved. We shall call this *stage* j of the chase. Assume that when j is even, R will be the next to move, whereas when j is odd, C will be the next to move. A *cycle* is a sequence of 2m stages that takes the players once around the circle.

Let us focus on some stage j, which without loss of generality we can suppose is even. Assume that we are in the last period of the current stage. In the next period, R will move clockwise one step and stay there for t_{j+1} periods. For this move to be generated by fictitious play, it must be R's best reply to the history. To compute the best reply, refer to table 2, which shows the last 14 stages of the most recent cycle within the payoff matrix for the case m = 7.

6, 6	4, 0	4, 0	5, 4 t_{j-1}	4, 5 t_j	0, 4	0, 4
0, 4	6, 6	4, 0	4, 0	5, 4 t_{j-13}	4, 5 t_{j-12}	0, 4
0, 4	0, 4	6, 6	4, 0	4, 0	5, 4 t_{j-11}	4, 5 t_{j-10}
4, 5 t_{j-8}	0, 4	0, 4	6, 6	4, 0	4, 0	5, 4 t_{j-9}
5, 4 t_{j-7}	4, 5 t_{j-6}	0, 4	0, 4	6, 6	4, 0	4, 0
4, 0	5, 4 t_{j-5}	4, 5 t_{j-4}	0, 4	0, 4	6, 6	4, 0
4, 0	4, 0	5, 4 t_{j-3}	4, 5 t_{j-2}	0, 4	0, 4	6, 6

Table 2.

Let us adopt the tie-breaking convention that a player moves one position to the right as soon as that position has an equal or higher payoff to his current position. Thus, in the above cycle, R moves as soon as the second row has at least as high an expected payoff as the first row, that is, as soon as

$$6(t_{j-8} + t_{j-7}) + 4(t_{j-6} + t_{j-5}) + 4(t_{j-4} + t_{j-3}) + 5(t_{j-1} + t_{j-2}) + 4(t_j + t_{j-13}) + \xi$$
$$\leq 6(t_{j-5} + t_{j-6}) + 4(t_{j-3} + t_{j-4}) + 4(t_{j-2} + t_{j-1}) + 5(t_{j-13} + t_j) + 4(t_{j-12} + t_{j-11}),$$

where ξ involves all terms t_{j-14} and earlier. This is equivalent to

$$t_j \geq t_{j-1} + t_{j-2} + 0t_{j-3} + 0t_{j-4} - 2t_{j-5} - 2t_{j-6} + 6t_{j-7} + 6t_{j-8} + 0t_{j-9} + 0t_{j-10} - 4t_{j-11} - 4t_{j-12} - t_{j-13} + \xi.$$

In the period just before this one, R did *not* move to the right, so the above expression with t_{j-1} substituted for t_j must be less than zero. It follows that it exactly equals zero now, that is,

$$t_j = t_{j-1} + t_{j-2} + 0t_{j-3} + 0t_{j-4} - 2t_{j-5} - 2t_{j-6} + 6t_{j-7} + 6t_{j-8} + 0t_{j-9} + 0t_{j-10} - 4t_{j-11} - 4t_{j-12} - t_{j-13} + \xi.$$

Observe that T_j periods ago, R was faced with the same choice that he is now, and he moved right. Therefore, T_j periods ago, the payoff difference between staying put and moving right must have been exactly equal to zero, which implies that $\xi = 0$. We conclude that, *if fictitious play has generated a chase for the last 2m stages, and if it continues to generate a chase from stage j on, then the following recursive equation holds*:

$$t_j = t_{j-1} + t_{j-2} + 0t_{j-3} + 0t_{j-4} - 2t_{j-5} - 2t_{j-6} + 6t_{j-7} + 6t_{j-8} + 0t_{j-9} + 0t_{j-10} - 4t_{j-11} - 4t_{j-12} - t_{j-13}.$$

In the general case $m = 2k + 1$, the analogous equation is

$$t_j = t_{j-1} + t_{j-2} \, . \, . \, 0 \, . \, . \, - 2t_{j-2k+3} - 2t_{j-2k+2} + 6t_{j-2k+1} + 6t_{j-2k} \, . \, . \, 0 \, . \, . \, - 4t_{j-2m+3} - 4t_{j-2m+2} - t_{j-2m+1}. \quad (3)$$

Lemma 1. Let $m = 2k + 1$, where $k \geq 3$, and let $j^* \geq 2m$. Suppose that the following two relations hold for all j, $3 \leq j < j^*$:

$$t_j > 2t_{j-2}, \quad (4)$$

$$t_j > t_{j-1} > t_{j-2} \quad (5)$$

Suppose further that (3) holds for $j = j^*, j^*-1, \dots, j^* - 6$. Then (4) and (5) hold for $j = j^*$.

Proof. The recursive equation (3) is easier to work with if we write it down as a table of coefficients:

	j	j-1	j-2	...	j-2k+1	j-2k	j-2k-1	j-2k-2	...	j-2m+3	j-2m+2	j-2m+1
0 =	-1	1	1	0	-2	-2	6	6	0	-4	-4	-1 .

Let $j = j^*$ and assume that (4) and (5) hold for $3 \leq j < j^*$. Assume further that $k \geq 4$. (We shall deal separately with the case $k = 3$.) In this case, t_{j-2} is at least four stages ahead of t_{j-2k+1}, so $t_{j-2} > 4t_{j-2k+1}$. Furthermore, $t_{j-2k+1} > t_{j-2k}$, so

$$t_{j-2} - 2t_{j-2k+1} - 2t_{j-2k} > t_{j-2} - 4t_{j-2k+1} > 0.$$

By condition (2), the terms $6t_{j-2k-1} + 6t_{j-2k-2}$ outweigh the last three negative terms. Putting all of this together we conclude that $t_j > t_{j-1}$, so that (5) holds for j.

Next we shall show that $t_j > 2t_{j-2}$. By assumption the recursive equation holds for j - 1 as well as j. Adding the resulting two equations we obtain

	j	j-1	j-2	j-3	. . .	j-2k+1	j-2k	j-2k-1	j-2k-2	j-2k-3	. . .	j-2m+3	j-2m+2	j-2m+1	j-2m	
0 =	-1	1	1	0	. . .	-2	-2	6	6	0	. . .	-4	-4	-1	0	
0 =	0	-1	1	1	. . .	0	-2	-2	6	6	. . .	0	-4	-4	-1	
0 =	-1	0	2	1	. . .	-2	-4	4	12	6	. . .	-4	-8	-5	-1	(6)

By assumption, the last equation applies for j - 3 as well as j, so

	j							. . .															j-2m-3
0 =	-1	0	2	1	0	0	0	. . .	-2	-4	4	12	6	0	0	0	. . .	-4	-8	-5	-1	0	0	0

Adding this to the bottom line of the preceding table we obtain

	j							. . .															j-2m-3
0 =	-1	0	2	1	0	0	0	. . .	-2	-4	4	12	6	0	0	0	. . .	-4	-8	-5	-1	0	0	0
0 =	0	0	0	-1	0	2	1	. . .	0	0	0	-2	-4	4	12	6	. . .	0	0	0	4	8	5	1
0 =	-1	0	2	0	0	2	1	. . .	-2	-4	4	10	2	4	12	6	. . .	-4	-8	-5	3	8	5	1
								A									B							

From the assumption that (4) and (5) hold for j < j*, it follows that the sum of the terms in A is strictly positive. Similarly, the sum of the terms in B are strictly positive. Since all other terms are positive, we conclude that $t_j > 2t_{j-2}$ as claimed.

Next we need to dispose of the case k = 3. Using the same substitution procedure as above, we obtain the following analog of (6):

$$0 = -1 \quad 0 \quad 2 \quad 0 \quad 0 \quad 0 \quad -3 \quad 4 \quad 10 \quad 2 \quad 4 \quad 16 \quad -2 \quad -5 \quad -5 \quad -8 \quad -5 \quad -1 \qquad (7)$$

Take three times the analogous equation for j - 6 and subtract it from the above to obtain

	j	j-1	j-2	j-3	. . .							j-11	j-12	j-13	j-14	j-15	j-16	j-17						j-23
0 =	-1	0	2	0	0	0	-3	4	10	2	4	16	-2	-5	-5	-8	-5	-1	0	0	0	0	0	0
0 =	0	0	0	0	0	0	3	0	-6	0	0	0	9	-12	-30	-6	-12	-48	6	15	15	24	15	3
0 =	-1	0	2	0	0	0	0	4	4	2	4	16	7	-17	-35	-14	-17	-49	6	15	15	24	15	3

By the doubling property, $t_{j-11} > 8t_{j-17}$, and $t_{j-11} > 4t_{j-15}$. Hence

$$16t_{j-11} - 14t_{j-15} - 49t_{j-17} > 0.$$

Similarly,

$$7t_{j-12} - 17t_{j-16} > 0,$$

By the doubling property, $t_{j-8} > 8t_{j-14}$, so $4t_{j-8} - 35t_{j-14} > -3t_{j-14}$. Similarly, $4t_{j-7} - 17t_{j-13} \geq > 15t_{j-13}$. By monotonicity, $t_{j-13} > t_{j-14}$, from which it follows that

$$4t_{j-7} + 4t_{j-8} - 17_{j-13} - 35t_{j-14} > -15t_{j-13} - 3t_{j-14} > 12t_{j-14} > 0.$$

Since all other terms are positive, we conclude that $t_j > 2t_{j-2}$, which completes the proof of Lemma 1.

Lemma 1 shows in particular that if a sequence $t_1, t_2, \ldots, t_{j^*}, \ldots$ is generated from some point $j^* \geq 2m$ on by equation (3), and if the initial sequence is strictly monotone increasing and satisfies the doubling property, then the whole sequence has these two properties.

To complete the proof of the theorem, we need to establish that given a suitable initial sequence, fictitious play generates all subsequent stages t_j according to the recursive equation (3). This will follow if, whenever a player moves under fictitious play, he always moves one position clockwise. In other words, we need to show that if the clockwise position has the same or higher expected payoff than the current position, then no other other position can have a still higher expected payoff. For this it suffices to show that coordinating with the other player's most recent move never offers a higher expected payoff than staying put or moving one position clockwise. (It is straightforward to check that if coordinating is inferior to moving one position clockwise, then all other moves are inferior to moving one position clockwise.)

Coordination has the highest payoff relative to other strategies when some player has occupied the same position for a long time, since this creates the strongest incentive for the other player to imitate him (or her). Thus the crucial point to check is the attractiveness of coordination *when some player has just completed two stages in the same position*. Without loss of generality, consider the situation at the end of an even stage j: R has occupied the same position for $t_j + t_{j-1}$ periods, and is about to move. (This situation is depicted in table 2.) C does not have an incentive to coordinate with R if and only if the following inequality holds:

	j	j-1	j-2	...	j-2k+2	j-2k+1	j-2k	j-2k-1	j-2k-2	j-2k-3	j-2k-4	...	j-2m+3	j-2m+2	j-2m+1	
0 <	-1	-1	4	4	4	4	2	2	-5	-5	-4	-4	-4	0	0	(8)

Note that we are omitting all stages older than the last complete cycle. This is sufficient, because if the inequality holds now, and it held 2m stages ago, 2m stages before that, and so forth, then it must hold now with all previous stages included. In particular, if a sequence satisfies the recursive equation (3) and inequality (8) for all $j \geq j^* \geq 2m$, then from stage j on fictitious play will generate precisely this sequence.

Lemma 2. Suppose that doubling (4) and monotonicity (5) hold for all j', $3 \leq j' \leq j$ and (3) holds for j and j - 1. Then (8) holds for j.

Proof. By assumption, the recursive equation holds for j and j - 1. Adding these equations to inequality (8) results in the equivalent inequality shown below:

	j	j-1	j-2	j-3		j-2k+1											j-2m
0 <	-1	-1	4	4 ..	4 ..	4	4	2	2	-5	-5	-4 ..	-4 ..	-4	0	0	0
0 =	1	-1	-1	0 ..	0 ..	0	2	2	-6	-6	0	0 ..	0 ..	4	4	1	0
0 =	0	2	-2	-2 ..	0 ..	0	0	4	4	-12	-12	0 ..	0 ..	0	8	8	2
0 <	0	0	1	2 ..	4 ..	4	6	8	0	-23	-17	-4 ..	-4 ..	0	12	9	2

We shall show that the last row is positive, from which it will follow that the first row is positive, that is, inequality (8) holds for j. To prove that the last row is positive, we need to show that each of the negative terms in the bottom row is dominated by some combination of positive terms to the left. We shall refer to these terms by the values of their coefficients. By the doubling property, the -23 term is dominated by the 8 term, which is two periods ahead, plus the 4 term, which is four periods ahead. The -17 is dominated by the 6, which is four periods ahead. The first -4 is dominated by the first 2. Each subsequent -4 in the second ellipsis is dominated by the corresponding 4 in the first ellipsis. (Note that this argument applies for k = 3 as well as k > 3.) This completes the proof of Lemma 2.

To complete the proof of the theorem, we shall demonstrate an initial sequence $t_1, t_2, \ldots, t_{j^*}$, $j^* = 2m+2$, such that: i) t_j satisfies the recursive equation for $3 \leq j \leq j^*$ (truncated if necessary to the number of preceding terms); ii) monotonicity and doubling hold for $3 \leq j \leq j^*$.

Let fictitious play be allowed to run beginning in period $T = t_1 + t_2 + \ldots + t_{j^*} + 1$. By Lemma 2 we know that inequality (8) holds for j^*. Hence C has no incentive to move. Assume without loss of generality that j^* is even (so that R is about to move). Since the recursive equation holds for t_{j^*-2m}, R was indifferent between staying put and moving clockwise one step at the end of this stage. Since the recursive equation holds for j^*, he is also indifferent now. Thus the tie-breaking rule implies that R does in fact move one step clockwise in period T. Both players stay in their positions for t_{j^*+1} periods, that is, until C is indifferent between staying put and moving one step clockwise. This happens when the value of t_{j^*+1} is determined by the recursive equation. From Lemma 1 it follows that monotonicity and doubling hold for $3 \leq j \leq j^* + 1$. Thus properties i) and ii) hold for the sequence $t_1, t_2, \ldots, t_{j^*+1}$. Inductively, we conclude that fictitious play generates an infinite chase $t_1, t_2, \ldots, t_{j^*}, \ldots$ where all terms beginning with t_3 are determined by the recursive equation.

Now consider the empirical distributions (p_t, q_t) generated by such a chase. Because of monotonicity and the doubling property, any limit point of the process also is monotonic and has the doubling property. But such a point cannot be close to any Nash equilibrium of the game.

It remains to actually demonstrate the required sequence. For this purpose let us fix $k = 3$ and $m = 7$. (The construction is quite general however.) Fix any pair of positions on the circle in which R is 3 steps ahead of C. This will be the starting point for the chase. Let $t_1 = 1$ and $t_2 = 2$. Let each subsequent t_i be generated by the recursive equation for $k = 3$, truncated to the existing number of preceding terms. This guarantees indifference on the part of the player about to move next. Recall that this equation is

$$t_j = t_{j-1} + t_{j-2} + 0t_{j-3} + 0t_{j-4} - 2t_{j-5} - 2t_{j-6} + 6t_{j-7} + 6t_{j-8} + 0t_{j-9} + 0t_{j-10} - 4t_{j-11} - 4t_{j-12} - t_{j-13}.$$

Note that this is a coordination game, and that strategies 1-7 for the two players are the same as in the merry-go-round game. Let the process start in period one with the strategy choices (α, β). It can be verified that, beginning in period 3, fictitious play generates a chase with the following stage-lengths:

stage	strategy pair	number of periods
-	(α, β)	1
-	(β, α)	1
1	(1, 4)	8
2	(1, 5)	14
3	(2, 5)	23
4	(2, 6)	54
5	(3, 6)	78
6	(3, 7)	133
7	(4, 7)	216
8	(4, 1)	297
9	(5, 1)	453
10	(5, 2)	709
11	(6, 2)	1196
12	(6, 3)	1968
13	(7, 3)	3307
14	(7, 4)	5703
15	(1, 4)	9442
16	(1, 5)	15284
17	(2, 5)	24472
18	(2, 6)	39162
19	(3, 6)	62413
20	(3, 7)	99719
	.	.
	.	.

The recursive equation is satisfied beginning at stage 15, that is, t_j is given by the recursive equation for all $j \geq 15$. Unlike the situations analyzed earlier, there is not perfect payoff indifference at the end of each stage between staying put and moving clockwise one step (because of the noninteger payoffs attached to α and β). Nevertheless, there is near-perfect indifference, and the number of periods in each stage is determined by the same equation as before. It can be verified that the numbers in the initial sequence are strictly increasing and more than double every two periods. It follows as in Lemmas 1 and 2 that fictitious play generates an infinite chase beginning at stage 15.

4. Modifications of fictitious play

The above examples do not show that coordination games *in general* are hard to learn, but they do show that some are harder to learn than others. More importantly, fictitious play is not the best way to learn them. Fictitious play has two defects as a learning rule: i) it does not contain enough randomization to escape from unproductive learning paths, and ii) it has too much inertia, that is, it puts too much weight on moves in the distant past. To deal with the latter problem, we could truncate fictitious play by having the players only react to the frequence distributions of the last m periods, where m is a (large) positive integer. It can be shown, however, that this is not enough to achieve convergence: indeed, there are 2 x 2 coordination games in which, for any m, the players miscoordinate forever if they miscoordinate in the first period (Fudenberg and Kreps, 1993; Young, 1993).

Better results are obtained if there is some "noise" in the learning process. One way to introduce stochastic variation is to suppose that the players have incomplete information about what the others have done in the past. Assume that in each period $t > 1$, each agent draws a random sample of size k without replacement from the last m plays. (If $t < k$, we can assume that all previous plays are sampled.) The draws are independent for the two agents. Each agent then chooses a best reply to the empirical frequency distribution in his sample. Once a given coordination equilibrium has been played m times in succession, it is played forever because the sampling in this case yields no deviation from the equilibrium. In other words, m successive repetitions of a coordination equilibrium is an absorbing state of the learning process. It can be shown further that these are the only absorbing states of the process. Moreover, if the ratio of information k/m is sufficiently *small* (in particular if $k/m \leq 1/2$), the process converges with probability one to an absorbing state, that is, a coordination equilibrium will eventually be played with probability one (Young, 1993, Theorem 1). The reason this works is that the stochastic variability created by incomplete sampling eventually jostles the process out of uncoordinated cycles. Once the process hits an absorbing state, however, the sampling variability vanishes and the process stays there forever.

Similar results obtain under other kinds of stochastic perturbation. Suppose, for example, that there is some systematic "error" in the players' responses. Let δ be a small positive number. Suppose that with probability $1 - \delta$ a given agent chooses a best reply to the frequency distribution of the other side's actions in a random sample drawn from the truncated history, but with probability δ she chooses a strategy at random. The probabilities of these events are independent for the two agents. We then obtain a Markov process P^δ on the finite state space H consisting of all truncated histories. The process is ergodic because there is a positive probability of moving from any state to any other in m periods or less. It can be shown that, for all sufficiently small δ, the players play a coordination equilibrium with near certainty over the long run. More precisely, given the process P^δ, let π_j^δ be the long-run probability that the jth coordination equilibrium (x_j, x_j) is played in any given period t as $t \rightarrow \infty$. This probability exists because the process is aperiodic and ergodic. It can be shown that, given any $\varepsilon > 0$, $\sum_{j=1,n} \pi_j^\delta \geq 1 - \varepsilon$ for all sufficiently small δ (Young, 1993).[3] In other words, the probability is at least $1 - \varepsilon$ that over the long run the players coordinate at any given time. Furthermore, in the absence of ties (i.e., in a generic coordination game), it can be shown that the players coordinate almost all of the time on *exactly one* of the coordination equilibria when the noise δ is small.[4]

Kaniovski and Young (1995) analyze a learning model in which there is sampling and systematic error, but the history is not truncated and all past plays are given equal weight. They show that, for 2 x 2 games, such a model converges with probability one to a neighborhood of a Nash equilibrium. They show further that for 2 x 2 coordination games, the process converges with probability one to a coordination equilibrium.[5] We conjecture that this type of learning model does not converge for the merry-go-round game, though we shall not attempt to prove this here. More generally, we conjecture that both random perturbations and finite memory (or sufficiently rapid discounting) are necessary conditions for a learning rule to converge with probablity one to a coordination equilibrium in a general coordination game.

---

[3]Kandori Mailath and Rob (1993) prove a similar result for symmetric 2 x 2 coordination games.

[4]Similar results hold for games that are *weakly acyclic* in the sense that, from any initial pair (x_1, x_2) there is a sequence of best replies (one player at a time) that ends in a strict Nash equilibrium (Young, 1993).

[5]This generalizes results of Fudenberg and Kreps (1993), who show that similar types of learning processes converge with probability one to a neighborhood of the mixed equilibrium *provided* the game has a unique equilibrium, which is mixed (in particular, the game is not a coordination game).

References

Fudenberg, Drew, and David Kreps (1993): "Learning Mixed Equilibria," *Games and Economic Behavior*, 5, 320-367.

Kandori, Michihiro, George Mailath, and Rafael Rob (1993): " Learning, Mutation, and Long-Run Equilibria in Games," *Econometrica*, 61, 29-56.

Kaniovski, Yuri, and H. Peyton Young (1995): "Learning Dynamics in Games with Stochastic Perturbations," *Games and Economic Behavior*, 11, 330-363.

Krishna, Vijay (1991): "Learning in Games with Strategic Comple-mentarities." Preprint, Pennsylvania State University.

Milgrom, Paul, and John Roberts (1991): "Adaptive and Sophisticated Learning in Normal Form Games," *Games and Economic Behavior*, 3, 82-100.

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

Monderer, Dov, and Aner Sela (1993): "Fictitious Play and No-Cycling Conditions." Preprint, The Technion, Israel.

Monderer, Dov, and Lloyd S. Shapley, (1993a): "Fictitious Play Property for Games with Identical Interests." Preprint, The Technion, Israel.

Monderer, Dov, and Lloyd S. Shapley (1993b): "Potential Games." Preprint, The Technion, Israel.

Robinson, Julia (1951): "An Iterative Method of Solving a Game," *Annals of Mathematics*, 54, 296-301.

Shapley, Lloyd S. (1964): "Some Topics in Two-Person Games, " in M. Dresher, L. S. Shapley, and A.W. Tucker, eds., *Advances in Game Theory*, pp. 1-28, Princeton: Princeton University Press.

Young, H. P., 1993: "The Evolution of Conventions," *Econometrica*, 61, 57-84.