SMALE STRATEGIES
FOR THE n-PERSON ITERATED PRISONER'S DILEMMA

ETHAN AKIN — SŁAWOMIR PLASKACZ — JOANNA ZWIERZCHOWSKA

ABSTRACT. Adapting methods introduced by Steven Smale, we describe good strategies for a symmetric version of the Iterated Prisoner's Dilemma with n players.

1. Introduction

In [5] Smale introduced an approach to strategy for the Iterated Prisoner's Dilemma which was different from the popular Markov chain methods. He suggested using as data the current time average payoff to the players rather than using the results of the most recent round of play. Smale's results were extended in [1]. Here we apply these methods to a symmetric n player version of the Prisoner's Dilemma. The goal is to describe good plans which stabilize the cooperative outcome where each player receives (in long-term average) the payoff p_n obtained when all n players cooperate. A strategy is a choice of initial play together with a plan responding to previous play.

We describe good plans with the following properties:

- If all n players eventually use good plans, then from any initial position, convergence to cooperation is achieved.

- Suppose that for some k with $1 \leq k < n$, players j with $j = 1, \ldots, k$ eventually use good plans and suppose that $X = (x_1, \ldots, x_k, z_{k+1}, \ldots, z_n)$ is a limit point for the sequence of time averages of the payoffs. Let $\overline{x}$ be the mean of the good player payoffs and $\overline{z}$ the mean of the remaining, dissenting, player payoffs. The only way it can occur that $\overline{z} \geq p_n$ is if $X = (p_n, \ldots, p_n)$, i.e. the cooperative payoff. Otherwise, while $\overline{z} > \overline{x}$, i.e. on average the dissenters do better than the good players, their average payoff remains below p_n.

  If $k < n - 1$ it may happen that, while the mean $\overline{z} < p_n$, some of the dissenters do get a payoff larger than p_n. However, if $z_{k+1} \geq p_n$, then if $\widetilde{z}$ is the mean of the remaining dissenters, i.e. the average of $z_{k+2}, \ldots, z_n$ then $\widetilde{z} < \overline{x}$. That is, the remaining players on average do worse than the good players and thus it pays for some of them to switch to good plans.

Smale strategies for the Iterated Prisoner's Dilemma with n players were described by Behrstock, Benaim and Hirsch in [2]. There is overlap between our results and theirs, but because they were considering a variety of cases, their exposition is more complicated. Markov strategies for the Iterated Prisoner's Dilemma with n players were described in [3] and [4]. An extensive comparison of Smale and the usual Markov chain approach in the framework of the Iterated Prisoner's Dilemma is provided in [1].

We would like to thank the referee for helpful comments and corrections.

2. The game

We consider an n-player symmetric game. Each player has a choice of two strategies c or d (*cooperate* or *defect*). When no players cooperate, each obtains a payoff r_0 and, when all n players cooperate, each obtains p_n. For $k = 1, \ldots, n - 1$, when k players cooperate, each cooperator receives p_k and each defector receives r_k.

Our first assumption is monotonicity of the payoffs.

$$(2.1) \qquad r_0 < r_1 < \ldots < r_{n-1} \quad \text{and} \quad p_1 < p_2 < \ldots < p_n.$$

That is, the payoffs to both defectors and cooperators increase as the number of cooperators increase.

Our second assumption concerns the advantage of defection.

$$(2.2) \qquad \text{For } k = 1, \ldots, n, \quad p_k < r_{k-1}.$$

This says that any cooperator would do better by switching to defection, provided the choices of the other players remain fixed. Thus, the strategy of defection *dominates* cooperation regardless of the choices of the other players.

SMALE STRATEGIES FOR THE n-PERSON IPD 353

For our third assumption, we define the population mean payoff, or just *mean payoff*, m_k by

(2.3)
$$m_0 = r_0, \qquad m_n = p_n,$$
$$m_k = \frac{1}{n}[k \cdot p_k + (n - k) \cdot r_k] \quad \text{for } k = 1, \ldots, n - 1.$$

From (2.1) and (2.2) it is clear that

(2.4)
$$p_k < m_k < r_k \quad \text{for } k = 1, \ldots, n - 1.$$

Our third assumption is monotonicity of the mean payoffs.

(2.5)
$$m_0 < m_1 < \ldots < m_{n-1} < m_n.$$

If there are currently $k - 1$ cooperators and one defector switches to cooperation, then he bears a cost of $r_{k-1} - p_k$. This assumption says that the aggregate increase to the other players exceeds this cost, i.e.

(2.6)
$$r_0 - p_1 < (n - 1) \cdot (r_1 - r_0), \quad r_{n-1} - p_n < (n - 1) \cdot (p_n - p_{n-1}),$$
$$r_{k-1} - p_k < (k - 1) \cdot (p_k - p_{k-1}) + (n - k) \cdot (r_k - r_{k-1}),$$
$$\text{for} \quad k = 2, \ldots, n - 1.$$

In the case when $n = 2$ the game is the classic Prisoner's Dilemma. In the notation of [1]

(2.7)
$$r_0 = P, \qquad r_1 = T, \qquad p_1 = S, \qquad p_2 = R,$$

with the assumptions

(2.8)
$$S < P < \frac{1}{2}(T + S) < R < T.$$

EXAMPLE 2.1. The values $(p_1, r_0, p_2, r_1, \ldots) = (0, 1, 2, 3, \ldots)$ for $n \geq 2$ is an example.

PROOF. $p_k = 2k - 2, r_k = 2k + 1$ and so

$$n \cdot m_k = kp_k + (n - k)r_k = k(2k - 2) + (n - k)(2k + 1) = (2n - 3)k + n.$$

Thus,

(2.9)
$$m_k = \left(2 - \frac{3}{n}\right)k + 1.$$

which is positive and increasing in k for $n \geq 2$. $\qquad \square$

3. The outcome and strategy regions

We consider repeated play. After the Tth round the payoff to the n players is a vector S^T in $\mathbb{R}^n$. There are 2^n possible payoff vectors according to the choices made by the players. For example, if in round T a single player cooperated, then the vector S^T has a single entry p_1 and $n-1$ entries r_1. There are n such payoff vectors according to which a single player cooperated.

For a single payoff after T rounds we use the time-average of the payoff vectors:

$$(3.1) \qquad s^T = \frac{1}{T} \sum_{t=1}^{T} S^t.$$

Observe that

$$(3.2) \qquad s^{T+1} = \frac{T}{T+1} s^T + \frac{1}{T+1} S^{T+1}$$

and so

$$(3.3) \qquad s^{T+1} - s^T = \frac{1}{T+1}(S^{T+1} - s^T).$$

The sequence $\{s^T\}$ lies in the *outcome set* $\varnothing \subset \mathbb{R}^n$ which is a convex hull of the 2^n payoff vectors. The sequence need not converge in $\varnothing$ but the set Ω of limit points is always a nonempty, compact, connected subset of $\varnothing$, see e.g. [1, Proposition 2.1].

The approach introduced by Smale in [5] is for each player to use the current time average s^T as data in order to determine his choice in round $T+1$. In this case, the symmetry of the game suggests that it would be sufficient for each player to keep track of how he is doing and to compare it merely with the time average of the mean payoff. That is, define $\overline{\pi} \colon \mathbb{R}^n \to \mathbb{R}$ by

$$\overline{\pi}(x_1, \ldots, x_n) = \frac{1}{n} \sum_{j=1}^{n} x_j.$$

For player j the linear map $\Pi_j = \pi_j \times \overline{\pi} \colon \mathbb{R}^n \to \mathbb{R}^2$ maps the outcome set $\varnothing$ onto the set $\S \subset \mathbb{R}^2$ which is the convex hull of the set of $2n$ points

$$(3.4) \qquad \{(r_0, r_0), (p_n, p_n)\} \cup \{(p_k, m_k), (r_k, m_k) : k = 1, \ldots, n-1\}.$$

Thus, Π_j maps s^T to a pair $(x, y) \in \mathbb{R}^2$ with x the j player's current average payoff and with y the current average population mean.

Notice that from (2.5) and convexity it follows that (r_0, r_0) is the unique point of $\S$ with y coordinate less than or equal to r_0 and (p_n, p_n) is the unique point of $\S$ with y coordinate greater than or equal to p_n. Hence, these two points are extreme points of $\S$.

We will call $\{(p_k, m_k) : k = 1, \ldots, n\}$ the *cooperation points* and $\{(r_k, m_k) : k = 0, \ldots, n-1\}$ the *defection points* of $\S$.
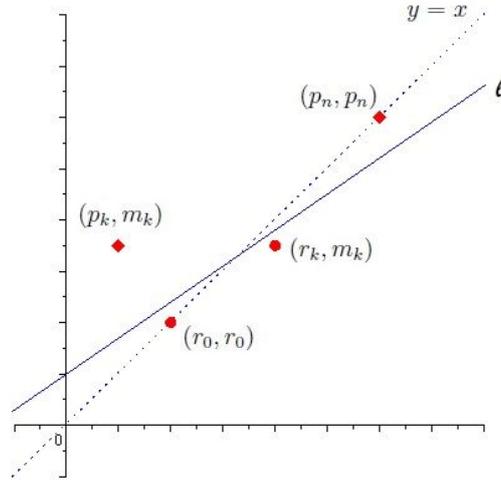
FIGURE 1. A separation line

DEFINITION 3.1. For a line ℓ in $\mathbb{R}^2$ with equation $y = lx + b$ we call the map $L\colon \mathbb{R}^2 \to \mathbb{R}$ given by $(x, y) \mapsto y - lx - b$ the *associated affine map*. We call ℓ a *separation line* for the game if each cooperation point is on or above ℓ and each defection point is on or below ℓ, i.e. L is non-negative on the cooperation points and non-positive on the defection points.

The concept of a separation line is depicted on Figure 1.

PROPOSITION 3.2. *If a line ℓ with equation $y = lx + b$ is a separation line, then $0 \le l \le 1$ and if $n > 2$ then $0 < l$.*

PROOF. By the Intermediate Value Theorem, a separation line intersects the intervals $[(p_1, m_1), (r_0, r_0)]$ and $[(p_n, p_n), (r_{n-1}, m_{n-1})]$. Hence, the steepest possible separation line connects (r_0, r_0) and (p_n, p_n) with slope 1. On the other hand, the slope must be at least that of the line connecting (p_1, m_1) with (r_{n-1}, m_{n-1}). This line has positive slope if $n > 2$ and slope 0 if $n = 2$. □

We will call the line through (r_0, r_0) and (p_n, p_n) the *diagonal line* or just the diagonal. It is the line with equation $y = x$ and is the unique separation line with slope equal to 1.

If a line ℓ with equation $y = lx + b$ passes through the point (p_n, p_n) then it is a separation line if and only if $l \in [l_0, 1]$, when $l_0 = \max\limits_{1 \le k < n} (p_n - m_k)/(p_n - p_k)$. By (2.4), we have $l_0 < 1$.

All our results are based on the following.

Proposition 3.3. *Assume that X is a limit point of the sequence $\{s^T\}$, i.e.*
$X \in \Omega \subset \emptyset$. *Let ℓ be a line in $\mathbb{R}^2$ with finite slope.*

(a) *Assume that every defection point lies on or below ℓ. If from some time onward, player j defects whenever $\Pi_j(s^T)$ lies above ℓ, then $\Pi_j(X)$ lies on or below ℓ.*

(b) *Assume that every cooperation point lies on or above ℓ. If from some time onward, player j cooperates whenever $\Pi_j(s^T)$ lies below ℓ, then $\Pi_j(X)$ lies on or above ℓ.*

Proof. Assume that $y = lx + b$ is the equation of ℓ and that L is the associated affine map. Let M be a positive number greater than the maximum of L on §.

(a) Assume that after some time $T_0 \geq 1$ player j defects when $L(\Pi_j(s^T)) > 0$. Let $M_0 = T_0 \cdot M$. We show that for all $T \geq T_0$, $L(\Pi_j(s^T)) \leq M_0/T$. We proceed by induction. Observe that the result holds for $T = T_0$.

If $L(\Pi_j(s^T)) > 0$ then player j defects on the next round and so $L(\Pi_j(S^{T+1})) \leq 0$. Hence, because L is an affine map, (3.2) implies that

$$(3.5) \qquad L(\Pi_j(s^{T+1})) = \frac{T}{T+1} L(\Pi_j(s^T)) + \frac{1}{T+1} L(\Pi_j(S^{T+1}))$$
$$\leq \frac{T}{T+1} L(\Pi_j(s^T)) \leq \frac{T}{T+1} M_0/T = M_0/(T+1),$$

by induction hypothesis.

On the other hand, if $L(\Pi_j(s^T)) \leq 0$

$$(3.6) \qquad L(\Pi_j(s^{T+1})) = \frac{T}{T+1} L(\Pi_j(s^T)) + \frac{1}{T+1} L(\Pi_j(S^{T+1}))$$
$$\leq \frac{1}{T+1} L(\Pi_j(S^{T+1})) \leq \frac{1}{T+1} M \leq M_0/(T+1).$$

It follows that $\limsup\{L(\Pi_j(s^T))\} \leq 0$ and so any limit point lies on or below ℓ.

(b) The proof is completely analogous and left to the reader. It is convenient to use $-L$ instead of L.   □

4. Simple Smale plan

Definition 4.1. Let ℓ be a separation line for the game with positive slope. The *simple Smale plan* for player j defects on round $T+1$ when $\Pi_j(s^T)$ is above ℓ and cooperates when $\Pi_j(s^T)$ is on or below ℓ.

We use the term "plan" because a strategy includes a choice of initial play as well as a plan responding to previous outcomes. Observe that the demand that the slope be positive is automatically satisfied if $n > 2$ by Proposition 3.2.

We will say that a player *eventually uses a simple Smale plan* when there is a fixed simple Smale plan which is used by the player from some time onward.

Theorem 4.2. *Let Ω be the set of limit points of the sequence $\{s^T\}$. If player j eventually uses the simple Smale plan with separation line ℓ, then $\Pi_j(\Omega) \subset \ell$.*

Proof. This immediately follows from Proposition 3.3.   □

Corollary 4.3. *For $j = 1, \ldots, n$ let ℓ_j be a separation line with positive slope and assume that at least one has slope less than 1. There exists a unique point $X \in \emptyset$ such that if each player j eventually uses the simple Smale plan associated with ℓ_j, then the sequence $\{s^T\}$ converges in $\emptyset$ to X. If each ℓ_j passes through (p_n, p_n), then $X = (p_n, \ldots, p_n)$ and so we obtain convergence to the position of complete cooperation.*

Proof. Since line ℓ_j has positive slope we can write its equation as $y = l_j(x - a_j)$. Now let $(x_1, \ldots, x_n)$ be an arbitrary limit point of the sequence $\{s^T\}$. If player j uses the ℓ_j strategy, then with $\overline{x} = \overline{\pi}(x_1, \ldots, x_n)$ Theorem 4.2 implies that $(x_j, \overline{x}) = \Pi_j(x_1, \ldots, x_n)$ lies on ℓ_j and so $\overline{x} = l_j(x_j - a_j)$.

If this holds for all j then $(l_j)^{-1}\overline{x} = x_j - a_j$. Let $\overline{a} = \overline{\pi}(a_1, \ldots, a_n)$ and let $\widehat{l}$ be the harmonic mean of the l_j's, i.e. $\widehat{l} = [\overline{\pi}((l_1)^{-1}, \ldots, (l_n)^{-1})]^{-1}$. It is clear that $\overline{x} = \widehat{l}(\overline{x} - \overline{a})$. Since not all the $l_j = 1$, Proposition 3.2 implies that $0 < \widehat{l} < 1$ and so the line $y = \widehat{l}(x - \overline{a})$ intersects the line $y = x$ at a unique point $(\overline{x}, \overline{x})$. That is, the mean, $\overline{x}$, is determined by the lines ℓ_j. Since $\overline{x} = l_j(x_j - a_j)$, all of the x_j's are determined as well.

Thus, the limit point $X = (x_1, \ldots, x_n)$ is uniquely determined by the equations $\overline{x} = \widehat{l}(\overline{x} - \overline{a})$ and $\overline{x} = l_j(x_j - a_j)$ for $j = 1, \ldots, n$.

Since the sequence $\{s^T\}$ has a unique limit point, compactness implies convergence.

Now assume that $(p_n, p_n) \in \ell_j$ for all j. Hence, $(l_j)^{-1}p_n = p_n - a_j$ for all j and so, averaging, we have $p_n = \widehat{l}(p_n - \overline{a})$. Uniqueness in the above argument implies $\overline{x} = p_n$. The point $X = (p_n, \ldots, p_n)$ is the unique point of $\emptyset$ with $\overline{\pi}(X) = p_n$.   □

Remark 4.4. We can characterize the limit point $(x_1, \ldots, x_n)$ as follows. It is the unique point such that there exists $y \in \mathbb{R}$ so that $X_k = (x_k, y) \in \ell_k$ and $\overline{X} = (1/n) \sum_{k=1}^{n} X_k$ is on the diagonal line with $y = x$.

Definition 4.5. Let ℓ be a separation line with positive slope. We call simple Smale plan a *good simple Smale plan* when it is associated with a separation line ℓ which passes through (p_n, p_n) and which has slope l satisfying $(n - 1)/n < l < 1$.

That is, the equation of a separation line for a good simple Smale plan can be written $p_n - y = l(p_n - x)$ with $(n - 1)/n < l < 1$.

It follows from Corollary 4.3 that, if each player eventually uses a good simple Smale plan, then we achieve convergence to complete cooperation. Smale calls

this *stability*. However, we are also interested in what happens when only some of the players are good.

Suppose that, for $j = 1, \ldots, k$, player j uses a good simple Smale plan associated with separation line ℓ_j having slope l_j. Suppose further that $X = (x_1, \ldots, x_k, z_{k+1}, \ldots, z_n)$ is a limit point for $\{s^T\}$. We assume that $1 \leq k \leq n-1$. We will call the players j with $j = k+1, \ldots, n$ the *dissenting players*. Define

$$\overline{x} = \frac{1}{k} \sum_{j=1}^{k} x_j, \qquad \overline{z} = \frac{1}{n-k} \sum_{j=k+1}^{n} z_j,$$

(4.1)
$$y = \frac{1}{n} \left[\sum_{j=1}^{k} x_j + \sum_{j=k+1}^{n} z_j \right] = \frac{1}{n} [k\overline{x} + (n-k)\overline{z}],$$

$$\widehat{l} = \left[\frac{1}{k} \sum_{j=1}^{k} (l_j)^{-1} \right]^{-1}.$$

Thus, $\overline{x}$ is the mean payoff among the good players, $\overline{z}$ is the mean payoff among the dissenters and y is the mean for the entire population.

From Theorem 4.2 it follows that $(x_j, y) = \Pi_j(X) \in \ell_j$ and so $p_n - y = l_j(p_n - x_j)$ for $j = 1, \ldots, k$. Dividing by l_j and averaging as before we obtain $p_n - y = \widehat{l}(p_n - \overline{x})$. Hence,

(4.2)
$$\frac{1}{n} [k(p_n - \overline{x}) + (n-k)(p_n - \overline{z})] = \widehat{l}(p_n - \overline{x}).$$

We can rewrite this as

$$(n-k)(p_n - \overline{z}) = (n\widehat{l} - k)(p_n - \overline{x}),$$

(4.3)
$$\overline{z} = \overline{x} + \frac{n(1 - \widehat{l})}{n-k} (p_n - \overline{x}).$$

Now assume that $k \leq n-2$ and suppose that $z_{k+1} \geq p_n$ and so $p_n - z_{k+1} \leq 0$. Define

(4.4)
$$\widetilde{z} = \frac{1}{n-k-1} \sum_{j=k+2}^{n} z_j.$$

So, we have

(4.5)
$$\frac{1}{n} [k(p_n - \overline{x}) + (n-k-1)(p_n - \widetilde{z})] \geq \widehat{l}(p_n - \overline{x}).$$

We can rewrite this as

(4.6)
$$\widetilde{z} \leq \overline{x} - \frac{1 - n(1 - \widehat{l})}{n-k-1} (p_n - \overline{x}).$$

Now we apply all this.

THEOREM 4.6. *With $1 \leq k \leq n-1$, assume that players $j = 1, \ldots, k$ eventually play good simple Smale plans. For a limit point $X = (x_1, \ldots, x_k, z_{k+1}, \ldots, z_n)$ let $\overline{x}$ be the mean payoff among the good players and let $\overline{z}$ be the mean payoff*

among the dissenting players. If $\overline{z} \geq p_n$, then $X = (p_n, \ldots, p_n)$ the complete cooperation point. Otherwise, $\overline{x} < \overline{z} < p_n$. In particular, if $k = n - 1$ then $z_n \geq p_n$ only when $X = (p_n, \ldots, p_n)$. Now assume that $k \leq n - 2$. If $z_{k+1} \geq p_n$ and $\widetilde{z}$ is the mean payoff among the remaining dissenting players, then either $X = (p_n, \ldots, p_n)$ or $\overline{x} > \widetilde{z}$.

PROOF. Because the strategies are assumed to be good simple Smale plans, it follows that $1 > \widehat{l} > (n-1)/n$. Hence, $n\widehat{l} - k > 0$. It follows from (4.3) that if $\overline{z} \geq p_n$, then $\overline{x} \geq p_n$. So $y \geq p_n$ and this implies X is the cooperation point. From the second equation in (4.3) we see that the dissenters on average do better than the good players but not as well as if they switched to cooperation and achieved the p_n payoff.

If there is a single dissenter, player n then $\overline{z} = z_n$ and so $z_n \geq p_n$ only when $X = (p_n, \ldots, p_n)$.

The second result follows from (4.6) because $1 - n(1 - \widehat{l}) > 0$.   □

Thus, if even one player uses a good simple Smale plan and the dissenting players do not allow the group to reach the cooperation point, then on average they do worse than the p_n payoff although they do better than the good players.

With $\overline{z} < p_n$ it may still happen that some among the dissenters can reach a payoff of p_n or greater. However, if this happens then the remaining dissenters on average do worse than the good players and so it pays for some of them to switch to a good simple Smale plan. That is, at least one of the remaining dissenters would do better by switching to a good simple Smale plan.

The case of a dissenter obtaining a payoff greater than p_n can happen. It seems to require exploiting naive behavior by the other dissenters. Here is an extreme case.

EXAMPLE 4.7. With $n \geq 3$, if, from some time on, player k for $k = 1, \ldots, n-2$ always cooperates, and player n always defects, then player n always receives a payoff between r_{n-2} and r_{n-1} regardless of the behavior of player $n - 1$, e.g. even if player $n-1$ uses a good simple Smale strategy. In particular, if $r_{n-2} > p_n$, then player n always receives a payoff greater than p_n.

If, in addition, player $n - 1$ eventually uses a simple Smale strategy then we obtain convergence to a unique point X regardless of the early plays.

PROOF. The strategies All-C and All-D, always cooperating and always defecting, respectively, trivially include Smale plans. If players $1, \ldots, n - 2$ always use All-C and player n always uses All-D, then every outcome $S^T \in \varnothing$ is either $V_0 = (p_{n-2}, \ldots, p_{n-2}, r_{n-2}, r_{n-2})$ or $V_1 = (p_{n-1}, \ldots, p_{n-1}, p_{n-1}, r_{n-1})$, depending whether player $n - 1$ defects or cooperates. It follows that every average point s^T lies on the segment between them, which we parameterize by $V_a = V_0 + a(V_1 - V_0)$ for $0 \leq a \leq 1$. If the strategies All-C and All-D are only

adopted after some time T_0, the effect of the initial terms on the average sequence $\{s^T\}$ tends to 0 as $T \to \infty$ and so in any case, the limit set Ω is contained in the segment $[V_0, V_1]$. So for any $X \in \Omega$, $\pi_n(X) \geq r_{n-2}$.

The line δ through $\Pi_{n-1}([V_0, V_1]) = [(r_{n-2}, m_{n-2}), (p_{n-1}, m_{n-1})]$ has negative slope. As a separation line ℓ has positive slope it follows that $\delta \cap \ell$ is a single point. Thus, $\Pi_{n-1}^{-1}(\ell)$ is a hyperplane in $\mathbb{R}^n$ which intersects $[V_0, V_1]$ in a unique point X. By Theorem 4.2 $\Omega = \{X\}$ and so we have convergence to X.   □

In the above example, suppose ℓ_l is a separation line through (p_n, p_n) with equation $y - p_n = l(x - p_n)$ so that $l \leq 1$. We then let $X_l = V_{a_l}$ denote the intersection point with the line δ. The value a_l is obtained by solving the equation $\overline{\pi}(V_a) - p_n = l(\pi_{n-1}(V_a) - p_n)$.

In the special case $l = 1$, ℓ_1 is the diagonal with equation $y = x$ and so a_1 satisfies

$$m_{n-2} + a_1(m_{n-1} - m_{n-2}) = r_{n-2} + a_1(p_{n-1} - r_{n-2}).$$

That is, a_1 equals

$$\frac{r_{n-2} - m_{n-2}}{(m_{n-1} - m_{n-2}) + (r_{n-2} - p_{n-1})} = \frac{r_{n-2} - m_{n-2}}{(r_{n-2} - m_{n-2}) + (m_{n-1} - p_{n-1})}.$$

Thus, $\pi_n(X_1) = r_{n-2} + a_1(r_{n-1} - r_{n-2}) \geq p_n$ when $a_1 \cdot r_{n-1} + (1 - a_1) \cdot r_{n-2} \geq p_n$.

If $l < 1$, then ℓ_l intersects δ between V_{a_1} and V_1. So, if $l < 1$, then $a_l > a_1$. Because $\pi_n(V_a)$ is increasing in a, we see that

(4.7) $l < 1 \Rightarrow \pi_n(X_l) > \pi_n(X_1)$.

With $n = 3$, $(p_1, r_0, p_2, p_3, r_1, r_2) = (0, 2, 4, 6, 7, 8)$ provides an example with $r_1 > p_3$, since $3(m_0, m_1, m_2, m_3) = (6, 14, 16, 18)$. Thus, if player 1 always cooperates and player 3 always defects, then player 3 receives a payoff of at least $7 > 6$ regardless of the behavior of player 2.

Even without the $r_{n-2} > p_n$ assumption, it can happen that a dissenter obtains greater than p_n even when one of the other players uses a good, simple, Smale plan.

EXAMPLE 4.8. Assume that for $n \geq 3$, $(p_1, r_0, p_2, r_1, \ldots) = (0, 1, 2, 3, \ldots)$ as in Example 2.1. If, eventually, player k for $k = 1, \ldots, n-2$ always cooperates, player $n-1$ uses a good, simple Smale plan and player n always defects, then there is a unique limit point $X = (x_1, \ldots, x_n)$ with $x_n > p_n = 2n - 2$.

PROOF. In this case, we have

$$\pi_j(V_a) = 2n - 6 + 2a = 2n - 3 + (2a - 3) \quad \text{for } j = 1, \ldots, n-2,$$
$$\pi_{n-1}(V_a) = 2n - 3 - a, \qquad \pi_n(V_a) = 2n - 3 + 2a,$$
$$\overline{\pi}(V_a) = 2n - 3 + \frac{1}{n}[a(2n - 3) - 3(n - 2)].$$

Equating $\pi_{n-1}(V_a) = \overline{\pi}(V_a)$ to solve for a_1 we obtain $a_1 = (n-2)/(n-1)$. Hence,

$$(4.8) \qquad \pi_n(X_1) = 2n - 3 + 2a_1 = 2n - 2 + \frac{n-3}{n-1}.$$

Since $p_n = 2n - 2$, (4.8) and (4.7) imply $\pi_n(X_l) > \pi_n(X_1) \geq p_n$ with the latter inequality strict if $n > 3$. Recall that for a good, simple Smale plan, the separation line passes through (p_n, p_n) and has slope less than 1. $\qquad\square$

REMARK 4.9. In the above example with $n = 3$ we easily see that if player 2 uses the simple Smale plan with separation line the diagonal, i.e. $l = 1$, then $a_1 = 1/2$ and the payoff vector is $X_1 = (1, 2.5, 4)$ when player 1 uses All-C and player 3 uses All-D. In general, if player 2 uses the diagonal simple Smale plan, then no player can obtain a payoff greater than $4 = p_3$, the cooperative payoff. For, if $X = (x_1, x_2, x_3)$ is a limit point, then $\Pi_2(X) = (x_2, \overline{x})$ lies on the diagonal by Theorem 4.2 and so $x_2 = \overline{x}$. This implies that $\overline{x} = (x_1 + x_3)/2$. Observe that $\Pi_1(X) = (x_1, \overline{x}) \in \Pi_1(\emptyset) = \S'$ which is the convex hull of $\{(1,1),(0,2),(3,2),(2,3),(5,3),(4,4)\}$. It follows that $x_1 \geq 2\overline{x} - 4$. Thus, $x_3 = 2\overline{x} - x_1 \leq 4$.

References

[1] E. Akin, *Good strategies for the Iterated Prisoner's Dilemma*: *Smale vs Markov*, J. Dyn. Games, **4** (2017), no. 3, 217–253.

[2] K. Behrstock, M. Benaim and M. Hirsch, *Smale strategies for network Prisoner's Dilemma Games*, J. Dyn Games **2** (2015), no. 2, 141–155.

[3] C. Hilbe, B. Wu, A. Traulsen and M. Nowak, *Cooperation and control in multiplayer social dilemmas*, PNAS **111** (2014), no. 46, 16425–16430.

[4] C. Hilbe, B. Wu, A. Traulsen and M. Nowak, *Evolutionary performance of zero-determinant strategies in multiplayer games*, J. Theoret. Biol. **374** (2014), 115–124.

[5] S. Smale, *The Prisoner's Dilemma and dynamical systems associated to non-cooperative games*, Econometrica **48** (1980), no. 7, 1617–1634.

Ethan Akin
Mathematics Department
The City College
137 Street and Convent Avenue
New York City, NY 10031, USA

E-mail address: ethanakin@earthlink.net

Sławomir Plaskacz and Joanna Zwierzchowska
Faculty of Mathematics and Computer Science
Nicolaus Copernicus University in Torun

E-mail address: plaskacz@mat.umk.pl, janna.zwierzchowska@mat.umk.pl