GOOD STRATEGIES FOR THE
ITERATED PRISONER’S DILEMMA:
SMALE VS. MARKOV

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

(Communicated by Josef Hofbauer)

ABSTRACT. In 1980 Steven Smale introduced a class of strategies for the Iterated Prisoner’s Dilemma which used as data the running average of the previous payoff pairs. This approach is quite different from the Markov chain approach, common before and since, which used as data the outcome of the just previous play, the memory-one strategies. Our purpose here is to compare these two approaches focusing upon good strategies which, when used by a player, assure that the only way an opponent can obtain at least the cooperative payoff is to behave so that both players receive the cooperative payoff. In addition, we prove a version for the Smale approach of the so-called Folk Theorem concerning the existence of Nash equilibria in repeated play. We also consider the dynamics when certain simple Smale strategies are played against one another.

1. Introduction. The Iterated Prisoner’s Dilemma has been an object of considerable study ever since Axelrod’s description of the results of computer tournaments [6] and Maynard Smith’s application of game theory to evolutionary competition [13]. Most of this work has focused upon what I call here Markov strategies, i.e. memory-one plans. Competition between two such strategies leads to a Markov process on the set of outcomes, e.g. [9], [15] and various surveys cited below. Considerable simulation work has been done to analyze numerically the competition between such strategies, e.g. [12] and [19]. In this context, I characterized in [3] and [5] certain so-called good strategies which ensured that an opponent could receive at least the cooperative payoff only by playing so that both players receive exactly the cooperative payoff. Recently, Mike Shub pointed out to me that Steve Smale had described in 1980 strategies which he called good [18]. These strategies use an entirely different way of aggregating the data of past outcomes. While there has been some work on Smale’s procedure, e.g. [1], [8] and [7], most of the game theory literature has ignored it. For example, Smale’s work is not referred to in [12], [14], [11], [16] or [17]. Our purpose here is to compare the Smale and Markov procedures and especially to use these to clarify the notion of a good strategy. In addition, we use the Taylor-Jonker equations for evolutionary dynamics [21] to analyze competition among certain simple Smale strategies.

2010 Mathematics Subject Classification. Primary: 91A20, 91A22; Secondary: 91A10.

Key words and phrases. Iterated Prisoner’s Dilemma, Smale, good strategies, simple Smale plan.
218 ETHAN AKIN

The author would like to express his appreciation to the referee for his helpful comments and corrections.

2. Plans for the Iterated Prisoner’s Dilemma. We will focus mostly on the symmetric version of the Prisoner’s Dilemma. Each of the two players, X and Y, has a choice between two strategies, c and d. Thus, there are four outcomes which we list in the order: \( cc, cd, dc, dd \), where, for example, \( cd \) is the outcome when X plays c and Y plays d. Either player can use a mixed strategy, randomizing by choosing c with probability \( p_c \) and d with the complementary probability \( 1 - p_c \).

Each then receives a payoff. The following \( 2 \times 2 \) chart describes the payoff to the X player. The transpose is the Y payoff.

| X \ Y | c | d |
|------|---|---|
| c    | R | S |
| d    | T | P |

Alternatively, we can define the payoff vectors for each player by

\[
S_X = (R, S, T, P) \quad \text{and} \quad S_Y = (R, T, S, P). \tag{2.2}
\]

The payoffs are assumed to satisfy

\[
T > R > P > S \quad \text{and} \quad 2R > T + S, \tag{2.3}
\]

but \( 2P \) might lie on either side of \( T + S \).

In the Prisoner’s Dilemma, the strategy c is cooperation. When both players cooperate they each receive the reward for cooperation (= R). The strategy d is defection. When both players defect they each receive the punishment for defection (= P). However, if one player cooperates and the other does not, then the defector receives the large temptation payoff (= T), while the hapless cooperator receives the very small sucker’s payoff (= S). The condition \( 2R > T + S \) says that the reward for cooperation is larger than the players would receive by dividing equally the total payoff of a \( cd \) or \( dc \) outcome. Thus, the maximum total payoff occurs uniquely at \( cc \) and that location is a strict Pareto optimum, which means that at every other outcome at least one player does worse. The cooperative outcome \( cc \) is clearly where the players “should” end up. If they could negotiate a binding agreement in advance of play, they would agree to play c and each receive R. However, the structure of the game is such that, at the time of play, each chooses a strategy in ignorance of the other’s choice. Furthermore, the strategy d strictly dominates strategy c. This means that, whatever Y’s choice is, X receives a larger payoff by playing d than by using c. In the array (2.1) each number in the d row is larger than the corresponding number in the c row above it. Hence, X chooses d, and for exactly the same reason, Y chooses d. So they are driven to the \( dd \) outcome with payoff \( P \) for each.

In the search for a way to avoid the mutually inferior payoff \( (P, P) \), attention has focused upon repeated play. X and Y play repeated rounds of the same game. For each round the players’ choices are made independently, but each is aware of all of the previous outcomes. The hope is that the threat of future retaliation will rein in the temptation to defect in the current round. It is this Iterated Prisoner’s Dilemma which we will consider here.

After the \( k^{th} \) round the players receive payoffs \( S^k = (S_X^k, S_Y^k) \) determined by the payoff matrix. For a single payoff after \( N \) rounds of play we use the time average:
\[ s^N = (s^N_X, s^N_Y) = \frac{1}{N} \Sigma_{k=1}^N S^k. \] (2.4)

Observe that
\[ s^{N+1} = \frac{N}{N+1} s^N + \frac{1}{N+1} S^{N+1}. \] (2.5)

and so
\[ s^{N+1} - s^N = \frac{1}{N+1} (S^{N+1} - s^N). \] (2.6)

The vector \( s^N = (s^N_X, s^N_Y) \) lies in \( S \) defined to be the convex hull of the four payoff pairs. If \( P < (T+S)/2 \) then \( S \) is a quadrilateral with vertices: \((S, T), (R, R), (P, P), (T, S)\). If \( P > (T+S)/2 \), then \( S \) is the triangle with vertices: \((S, T), (R, R), (T, S)\), which contains \((P, P)\).

The Euclidean diameter of the set \( S \) is \( \sqrt{2}(T-S) \). For \( \varepsilon > 0 \) we will call \( N^* \) an \( \varepsilon \) time-step when \( N^* > \sqrt{2}(T-S)/\varepsilon \). In that case, for \( N \geq N^* \) the distance \( ||s^{N+1} - s^N|| \) of the step from \( s^N \) to \( s^{N+1} \) is less than \( \varepsilon \).

**Proposition 2.1.** For any infinite sequence of outcomes, the set \( \Omega \) of limit points of the sequence \( \{s^N\} \) of payoff pairs is closed, connected subset of \( S \). If \( U \) is an open subset of \( S \) which contains \( \Omega \), then there exists \( N^* \) such that \( s^N \in U \) for all \( N \geq N^* \).

**Proof.** From (2.6) it is clear that \( ||s^{N+1} - s^N|| \to 0 \) as \( N \to \infty \). So the conclusion is immediate from the following well-known lemma. \( \square \)

For the lemma we introduce a bit of notation.

For \( A \) a nonempty, closed subset of a compact metric space \( X \) and \( x \in X \) we let \( d(x, A) = \inf\{d(x, a) : a \in A\} \). If \( B \) is another nonempty, closed subset we let \( d(A, B) = \inf\{d(a, b) : a \in A, b \in B\} \). By compactness, there exists \( a \in A \) such that \( d(x, A) = d(x, a) \), a point in \( A \) closest to \( x \), and there exist \( a \in A, b \in B \) such that \( d(A, B) = d(a, b) \). Given \( \varepsilon > 0 \) we let \( V_\varepsilon(A) \) denote the open set \( \{x : d(x, A) < \varepsilon\} \), i.e. the set of points less than \( \varepsilon \) away from a point of \( A \).

**Lemma 2.2.** Let \( \{x^N\} \) be a sequence in a compact metric space \( X \). If \( d(x^{N+1}, x^N) \to 0 \) as \( N \to \infty \) then the set of limit points \( \Omega \) is a nonempty, closed, connected subset of \( X \). If \( U \) is an open set containing \( \Omega \) then there exists \( N^* \) such that \( x^N \in U \) for all \( N \geq N^* \).

**Proof.** The set limit points \( \Omega \) is the intersection of the decreasing sequence \( \{X^k\} \) with \( X^k \) the closure of the tail \( \{x^N : N \geq k\} \) of the sequence. If \( U \) is an open set containing \( \Omega \) then \( \{U\} \) and \( \{X \setminus X^k : k = 1, \ldots\} \) is an open cover of \( X \) and so has a finite subcover. Since the \( X^k \)'s are decreasing, it follows that for some \( N^* \), \( \{U, X \setminus X^{N^*}\} \) is a cover of \( X \). Hence, \( \{x^N : N \geq N^*\} \subset U \). If \( \Omega \) were empty we could apply this to \( U = \emptyset \) and get a contradiction.

Now let \( A_0 \) and \( A_1 \) be disjoint nonempty, closed subsets of \( \Omega \). We will see that \( \Omega \setminus (A_0 \cup A_1) \) is nonempty and this implies that \( \Omega \) is connected. Let \( 3\varepsilon \) be the distance between the sets \( d(A_0, A_1) \). Choose \( N^* \) so that \( d(x^{N+1}, x^N) < \varepsilon \) for all \( N \geq n^* \).

The sequence repeatedly approaches arbitrarily closely to each point of \( \Omega \). Since \( A_0 \) and \( A_1 \) are nonempty we can define \( n_0 \) to be the minimum \( n \geq n^* \) such that \( x^n \) lies in \( V_\varepsilon(A_0) \). Inductively, for \( k \geq 0 \), define \( n_{2k+1} \) to be the minimum \( n \geq n_{2k} \) such that \( x_n \in V_\varepsilon(A_1) \), and for \( k \geq 1 \), define \( n_{2k} \) to be the minimum \( n \geq n_{2k-1} \) such that \( x_n \in V_\varepsilon(A_0) \). It is clear that \( x^{n_i-1} \in V_{2\varepsilon}(A_0) \setminus V_\varepsilon(A_0) \) if \( i \) is even and is
in \(V_2(A_1) \setminus V_1(A_1)\) if \(i\) is odd. Hence, the subsequence \(\{x^{n_1-1}\}\) lies in the closed set \(X \setminus (V(A_0) \cup V_c(A_1))\) and so it has limit points not in \(A \cup A_1\).

The choice of play for the first round is the initial play. A strategy is a choice of initial play together with what we call a plan. A plan is a choice of play, after the first round, to respond to any possible past history of outcomes in the previous rounds.

If \(X\) and \(Y\) use strategies with only pure strategy choices then the result is an infinite sequence of outcomes. However, if mixed strategy choices are used either on the initial plays or as part of the plan there results a probability measure on the space of all such sequences.

We will consider plans which use just a crucial portion of the past history data.

We will use the label Markov plan for a stationary, memory-one plan which bases its response entirely on the outcome of the previous round. For example, the Tit-for-Tat plan, hereafter TFT, due to Anatol Rapoport, plays the opponent’s response from the previous play.

With the outcomes listed in order as \(cc, cd, dc, dd\), a Markov plan for \(X\) is given by a vector \(p = (p_1, p_2, p_3, p_4) = (p_{cc}, p_{cd}, p_{dc}, p_{dd})\) where \(p_z\) is the probability of playing \(c\) when the outcome \(z\) occurred in the previous round. On the other hand, if \(Y\) also uses a Markov plan \(q = (q_1, q_2, q_3, q_4)\) then the response vector is \((q_{cc}, q_{cd}, q_{dc}, q_{dd}) = (q_1, q_3, q_2, q_4)\) and the successive outcomes follow a Markov chain with transition matrix given by:

\[
M = \begin{pmatrix}
    p_1q_1 & p_1(1-q_1) & (1-p_1)q_1 & (1-p_1)(1-q_1) \\
    p_2q_3 & p_2(1-q_3) & (1-p_2)q_3 & (1-p_2)(1-q_3) \\
    p_3q_2 & p_3(1-q_2) & (1-p_3)q_2 & (1-p_3)(1-q_2) \\
    p_4q_4 & p_4(1-q_4) & (1-p_4)q_4 & (1-p_4)(1-q_4)
\end{pmatrix}
\] (2.7)

Notice the switch in numbering from the \(Y\) strategy \(q\) to the \(Y\) response vector. This is done because switching the perspective of the players interchanges \(cd\) and \(dc\). This way the “same” plan for \(X\) and for \(Y\) is given by the same vector. For example, \(TFT\) for \(X\) and for \(Y\) is given by \(p = q = (1, 0, 1, 0)\), but the response vector for \(Y\) is \((1, 1, 0, 0)\). The plan Repeat is given by \(p = q = (1, 1, 0, 0)\) with the response vector for \(Y\) equal to \((1, 0, 0, 0)\). This plan just repeats the previous play, regardless of what the opponent did.

We can think of a Markov chain on a finite set \(I\) (in this case \(I = \{cc, cd, dc, dd\}\)) as representing motion on a directed graph with vertices \(I\) and an edge from \(i_1\) to \(i_2\) if there is a positive probability, according to \(M\), of moving from \(i_1\) to \(i_2\), i.e. \(M_{i_1i_2} > 0\). In particular, there is an edge from \(i\) to itself when \(M_{ii} > 0\). A path in the graph is a state sequence \(i^1, ..., i^n\) with \(n \geq 1\) such that there is an edge from \(i^k\) to \(i^{k+1}\) for \(k = 1, ..., n-1\). A set of states \(J \subseteq I\) is called a closed set when it is nonempty and no path that begins in \(J\) can exit \(J\). For example, the entire set of states is closed and for any \(i\) the set of states accessible via a path that begins at \(i\) is a closed set. The subset \(J\) is called a terminal set when it is closed and when for \(i_1, i_2 \in J\) there exists a path from \(i_1\) to \(i_2\). Equivalently, a terminal set is a minimal closed set. Since the set is closed, the path moves only on elements of \(J\).

A vector \(v\) is a stationary distribution for \(M\) when \(v_i \geq 0, \Sigma_i v_i = 1\) and \(vM = v\). For a terminal set \(J\) there is a unique stationary distribution vector \(v_J\) such that \((v_J)_i = 0\) for \(i \notin J\). Furthermore, \((v_J)_i > 0\) for \(i \in J\). Restricted to a terminal
set, the system is ergodic and so for any function \( f : I \rightarrow \mathbb{R} \) the sequence of time averages \( \frac{1}{N} \sum_{k=1}^{N} f(i^k) : N = 1, 2, \ldots \) converges with probability 1 to the space average \( \sum_{i \in J} f(i) \), the expected value with respect to \( v_J \). That is, such convergence occurs except on a set of outcome sequences which has probability 0. Think of the outcome sequence for a fair coin such that Heads comes up every time.

Distinct terminal sets are disjoint. If \( i \in I \) lies in some terminal set then it is called recurrent; otherwise, it is called transient. If \( i \) is transient then for each terminal set \( J \) there is a probability, determined by \( i \) and \( M \) so that the process beginning from \( i \) eventually enters \( J \). This probability might be zero but when summed over all of terminal sets the probabilities add up to 1. If the process enters \( J \) then the time average for any function \( f \) approaches the expected value with respect to \( v_J \) with probability 1.

The matrix \( M \) is called convergent when there is a unique terminal set \( J \). In that case, \( v_J \) is the unique stationary distribution for \( M \) and with probability 1, the time average for \( f : I \rightarrow \mathbb{R} \) converges to the \( v_J \) expected value regardless of the initial position.

Suppose \( X \) and \( Y \) play Markov plans leading to the \( 4 \times 4 \) matrix \( M \) of (2.7) and \( v_J \) is the stationary distribution for a terminal set \( J \subset \{cc, cd, dc, dd\} \). If the sequence of outcomes enters \( J \), then it remains in \( J \) and with probability 1 the average payoff \( (s^N_X, s^N_Y) \) converges with

\[
\lim_{N \rightarrow \infty} (s^N_X, s^N_Y) = v_1(R,R) + v_2(S,T) + v_3(T,S) + v_4(P,P), \tag{2.8}
\]

where \( v_J = (v_1, v_2, v_3, v_4) \).

If there is more than one terminal set \( J \), then with probability 1 the sequence will enter some terminal set \( J \) with the probabilities for different \( J \)’s depending upon the initial plays.

For example, suppose that \( p \) satisfies \( 0 < p_i < 1 \) for \( i = 1, \ldots, 4 \) and that \( q \) satisfies the analogous condition. Every entry of the associated Markov matrix \( M \) is positive and \( \{cc, cd, dc, dd\} \) is the unique terminal set. So there is a unique stationary distribution \( v \) with \( v_i > 0 \) for \( i = 1, \ldots, 4 \) and with probability one the outcome sequence passes repeatedly through each of the four outcomes with the average payoff sequence converging according to (2.8).

Smale in [18] aggregates the data in a different way. He suggest using as data the current average payoff given by (2.4). A Smale plan is a function \( \pi : S \rightarrow [0,1] \). If \( X \) uses the Smale plan \( \pi \) then in round \( N+1 \) he plays \( c \) with probability \( \pi(s^N_X, s^N_Y) \). Again we have the switch due to reverse in labeling for the other player. Let \( Switch : S \rightarrow S \) be defined by \( Switch(s_X, s_Y) = (s_Y, s_X) \). If \( Y \) uses the Smale plan \( \pi \) then she cooperates with probability \( \pi \circ Switch(s^N_X, s^N_Y) = \pi(s^N_Y, s^N_X) \). That is, \( Y \) responds to \( s \in S \) by using \( \pi \circ Switch \) applied to \( s \).

So if \( X \) uses \( \pi_X \) and \( Y \) uses \( \pi_Y \) then the outcomes \( cc, cd, dc, dd \) at time \( N+1 \) have probabilities given by

\[
\begin{align*}
p_{cc} &= \pi_X(s^N_X) \cdot \pi_Y(Switch(s^N_X)), \\
p_{cd} &= \pi_X(s^N_X) \cdot (1 - \pi_Y(Switch(s^N_X))), \\
p_{dc} &= (1 - \pi_X(s^N_X)) \cdot \pi_Y(Switch(s^N_X)), \\
p_{dd} &= (1 - \pi_X(s^N_X)) \cdot (1 - \pi_Y(Switch(s^N_X))). \tag{2.9}
\end{align*}
\]

After the randomization is applied, we obtain the time \( N+1 \) payoff \( S^{N+1} = (S^{N+1}_X, S^{N+1}_Y) \).
Finally, observe that

Proof. Since an affine map commutes with convex combinations, (2.5) implies

So

\[ \limsup_{N \to \infty} L(s^N) \leq 0. \]

Lemma 2.3. Let \( L : \mathbb{R}^2 \to \mathbb{R} \) be a nonconstant affine map, i.e. \( L(x, y) = ax + by + c \) with \( (a, b) \neq (0, 0) \). Let \( M \) be the maximum value of \(|L|\) on \( S \). X and Y use general strategies.

If there exists a time \( N^* \) so that for all \( N \geq N^* \), \( L(s^N) > 0 \) implies \( L(S^{N+1}) \leq 0 \), then for \( N \geq N^* \)

\[ L(s^N) \leq \frac{MN^*}{N}. \]  

(2.10)

So \( \limsup_{N \to \infty} L(s^N) \leq 0. \)

Proof. Since an affine map commutes with convex combinations, (2.5) implies

\[ L(s^{N+1}) = \frac{N}{N+1} L(s^N) + \frac{1}{N+1} L(S^{N+1}). \]  

(2.11)

So if \( N \geq N^* \)

\[ L(s^N) > 0 \implies L(s^{N+1}) \leq \frac{N}{N+1} L(s^N). \]  

(2.12)

On the other hand, since \( N^* \geq 1 \)

\[ L(s^N) \leq 0 \implies L(s^{N+1}) \leq \frac{1}{N+1} L(S^{N+1}) \leq \frac{MN^*}{N+1}. \]  

(2.13)

Finally, observe that \( L(s^{N^*}) \leq \frac{MN^*}{N}. \) So inequality (2.10) follows from (2.12) and (2.13) by mathematical induction. \( \square \)

Lemma 2.4. Let \( L : \mathbb{R}^2 \to \mathbb{R} \) be a nonconstant affine map. X and Y use general strategies.

(a) If \( L(P, P), L(T, S) < 0 \), then there is a positive integer \( k \) such that for all \( N \) there exists \( n \) with \( N \leq n \leq kN \) such that either \( L(s^n) < 0 \) or X plays \( c \) on round \( n \).

(b) If \( L(R, R), L(S, T) > 0 \), then there is a positive integer \( k \) such that for all \( N \) there exists \( n \) with \( N \leq n \leq kN \) such that either \( L(s^n) > 0 \) or X plays \( d \) on round \( n \).

Proof. (a) Let \( m = \min\{-L(P,P), -L(T,S)\} \) and let \( k \) be an integer with \( k \geq 2 \) such that \( \frac{M}{k-1} < m \). Assume X plays \( d \) in rounds \( N, \ldots, kN \) then in each round the outcome is either \( dc \) or \( dd \) and so \( L(S^n) \leq -m \) for \( n = N + 1, \ldots, kN \). On the other hand, \( L(S^n) \leq M \) for \( n = 1, \ldots, N \). Hence

\[ L(s^{kN}) = \frac{1}{kN} \sum_{t=1}^{kN} L(S^t) \leq \frac{1}{kN} (NM - (k-1)Nm) < 0. \]  

(2.14)

The proof of (b) is completely analogous. \( \square \)

Theorem 2.5. Let \( L : \mathbb{R}^2 \to \mathbb{R} \) be a nonconstant affine map with \( M \) be the maximum value of \(|L|\) on \( S \). Player X uses a Smale plan \( \pi \) from round \( N^* \) on and player Y uses an arbitrary plan and X and Y use arbitrary initial plays.

(a) Assume that \( L(s) > 0 \) implies \( \pi(s) = 0 \), i.e. \( \{L > 0\} \) is contained in the defection zone of \( \pi \).
(i) If \( L(P, P), L(T, S) \leq 0 \) then for all \( N \geq N^* \).

\[
L(s^N) \leq \frac{MN^*}{N}.
\]  

(2.15)

So \( \limsup_{N \to \infty} L(s^N) \leq 0 \).

(ii) If \( L(P, P), L(T, S) < 0 \) then there is a positive integer \( k \) so that for all \( N \geq N^* \) there exists \( n \) with \( N \leq n \leq kN \) such that \( L(s^n) \leq 0 \).

(iii) If \( L(R, R), L(S, T) > 0 \) then there is a positive integer \( k \) so that for all \( N \geq N^* \) there exists \( n \) with \( N \leq n \leq kN \) such that \( X \) plays \( d \) on round \( n \).

(b) Assume that \( L(s) < 0 \) implies \( \pi(s) = 1 \) i.e. \( \{L < 0\} \) is contained in the cooperation zone of \( \pi \).

(i) If \( L(R, R), L(S, T) \geq 0 \) then for all \( N \geq N^* \).

\[
L(s^N) \geq -\frac{MN^*}{N}.
\]  

(2.16)

Proof. (a)(i) If \( L(s^N) > 0 \) then \( X \) plays \( d \) and so the \( N + 1 \) outcome is either \( dc \) or \( dd \). Hence, \( S^{N+1} \) is either \( (T, S) \) or \( (P, P) \) which implies \( L(S^{N+1}) \leq 0 \). So we can apply Lemma 2.3 to get (2.15).

(a)(ii) Apply Lemma 2.4(b) to obtain \( k \) so that for some \( n \) between \( N \) and \( kN \) either \( L(s^n) < 0 \) or \( X \) plays \( c \) on round \( n \). By assumption, if \( X \) plays \( c \) on round \( n \) then \( L(s^n) \leq 0 \).

(a)(iii) Apply Lemma 2.4(a) to obtain \( k \) so that for some \( n \) between \( N \) and \( kN \) either \( L(s^n) > 0 \) or \( X \) plays \( d \) on round \( n \). If \( L(s^n) > 0 \) then \( X \) plays \( d \) on round \( n \).

The proof of (b) is completely analogous to that of (a) applying Lemma 2.3 to \(-L\) to get (2.16).

We will say that a player eventually uses a plan when there exists \( N^* \) so that the plan is used for all \( N \geq N^* \).

**Notation:** For distinct points \( A, B \in \mathbb{R}^2 \) we will use \([A, B]\) for the closed segment connecting the points, \([A, B]\) for the half-open segment \([A, B)\ \setminus B\), etc. We will denote by \( )A, B(\) the line through \( A \) and \( B \) and use \([A, B(\) for the ray from \( A \) through \( B \). In general, for a finite set of points \( \{A_1, \ldots, A_n\} \) we will use \([A_1, \ldots, A_n]\) for the convex hull. We will call the line \((P, P), (R, R)(\) the diagonal and the line \)((S, T), (T, S)(\) the co-diagonal. If \( \ell_1 \) and \( \ell_2 \) are non-parallel lines, then we will let \( \ell_1 \cap \ell_2 \) denote the point of intersection, abusively identifying the singleton set with the point contained therein.

Any line \( \ell \) in \( \mathbb{R}^2 \) is the intersection of two half-planes \( H^+ \) and \( H^- \). When the line is not vertical we will use \( H^+ \) for the upper half-plane. We will then refer to \( H^+ \setminus \ell \) as the set of points above \( \ell \), with \( H^- \setminus \ell \) the points below \( \ell \). Up to multiplication by a positive constant an affine map \( L : \mathbb{R}^2 \to \mathbb{R} \) is uniquely defined by the conditions that \( L \) is zero on \( \ell \) and is positive on \( H^+ \setminus \ell \). We will say that \( L \) is associated with \( \ell \) and vice-versa.

A line \( \ell \) is called a separation line for the game when \((S, T)\) and \((R, R)\) lie in one half-plane while \((P, P)\) and \((T, S)\) lie in the other.
A separation line intersects the segments \([S, T), (P, P]\) and \([R, R), (T, S]\) and so is determined by a choice of a point on each of these segments. Any point on \([S, T), (P, P]\) may be used. If \(P \leq \frac{1}{2}(T + S)\) then any point on \([R, R), (T, S]\) may be used. However, if \(P > \frac{1}{2}(T + S)\) then the line \((S, T), (P, P)\) intersects the open segment \((R, R), (T, S)\) at a point we will label \(W\). The point \(W\) is on or below any separation line. In general, a separation line has slope \(m\) with \(|m| \leq 1\), with \(m = 1\) only for the diagonal line and with \(m = -1\) only for the co-diagonal. The co-diagonal is a separation line only when \(P \leq \frac{1}{2}(T + S)\).

A separation line cannot be vertical, so \((S, T), (R, R) \in H^+\) and \((P, P), (T, S) \in H^−\). Hence, for an associated affine map \(L\),

\[
L(R, R), L(S, T) \geq 0 \geq L(P, P), L(T, S). \tag{2.17}
\]

**Corollary 2.6.** Assume that eventually \(X\) plays a Smale plan \(\pi\), \(Y\) uses an arbitrary plan and that the initial play is arbitrary. Let \(\Omega\) be the limit point set of an associated sequence of outcomes. Let \(C \subset S\) be a closed, convex set and \(\ell\) be a separation line.

(a) If \((P, P), (T, S) \in C\) and \(S \setminus C\) is contained in the defection zone \(\pi^{-1}(0)\), then \(\Omega \subset C\).

(b) If \((R, R), (S, T) \in C\) and \(S \setminus C\) is contained in the cooperation zone \(\pi^{-1}(1)\), then \(\Omega \subset C\).

(c) If \(S \cap \ell \subset C\) and \(\pi(s) = 0\) for \(s\) above \(C\) and \(\pi(s) = 1\) for \(s\) below \(C\), then \(\Omega \subset C\).

**Proof.** (a): For \(s \in S \setminus C\), let \(s'\) be the closest point in \(C\). The line \(\ell\) through \(s'\) which is perpendicular to \(s, s'\) is a line of support for \(C\). That is, if \(L\) is an affine function associated with \(\ell\) such that \(L(s) > 0\) then \(L \leq 0\) on \(C\). From Theorem 2.5 (a)(i) it follows that \(L \leq 0\) on \(\Omega\). In particular, \(s \not\in \Omega\).

(b): Proceed as above, using Theorem 2.5 (b)(i), instead.

(c): Let \(C^+\) consist of the points of \(S\) on or above \(C\) and \(C^-\) consist of the points of \(S\) on or below \(C\). To be precise, if \(H^±\) are the half-spaces associated with \(\ell\), then \(C^± = C \cup (H^± \cap S)\). These are each closed, convex sets. Because \(\ell\) is a separation line, \((P, P), (T, S) \in C^-\) and \((R, R), (S, T) \in C^+.\) From (a) it follows that \(\Omega \subset C^-\) and from (b) that \(\Omega \subset C^+.\) Thus, \(\Omega \subset C^- \cap C^+ = C\).

**Definition 2.7.** The map \(\pi : S \to [0, 1]\) is a simple Smale plan with separation line \(\ell\) if for an associated affine function \(L\) for \(\ell\)

\[
\pi(s) = \begin{cases} 
0 & \text{if } L(s) > 0, \\
1 & \text{if } L(s) < 0.
\end{cases} \tag{2.18}
\]

That is, the \(\pi\) player responds with \(d\) if \(s\) is above the line \(\ell\) and with \(c\) if \(s\) is below the line. We do not specify the value of \(\pi\) on the line.

**Corollary 2.8.** Assume that from some round \(N^*\) on, player \(X\) uses a simple Smale plan with separation line \(\ell\) and associated affine function \(L\). Let \(M\) be the maximum of \(|L|\) on \(S\). Assume that player \(Y\) uses an arbitrary plan and that the initial plays are arbitrary.

For all \(N \geq N^*\)

\[
|L(s^N)| \leq \frac{MN^*}{N}. \tag{2.19}
\]

So \(\lim_{N \to \infty} L(s^N) = 0\) and the limit point set \(\Omega\) is contained in \(\ell\).

**Proof.** Clearly, (2.19) follows from (2.15) and (2.16), given (2.17) and (2.18).
Thus, if eventually $X$ plays a simple Smale plan and player $Y$ uses an arbitrary plan, then after the randomization for mixed strategies has been applied, a sequence of outcomes is obtained and the limit point set $\Omega$ of the corresponding payoff sequence $\{s^N\}$ is a point or closed segment in the separation line $\ell$ by Proposition 2.1.

The plan All-C, which always cooperates, and so has $\pi = 1$ on $S$, is a simple Smale plan with separation line $\ell(R, R), (S, T)$. If $P \leq \frac{1}{2}(T + S)$ then All-D, which always defects, and so has $\pi = 0$ on $S$, is a simple Smale plan with separation line $\ell(P, P), (T, S)$. However, if $P > \frac{1}{2}(T + S)$ then every simple Smale plan cooperates below the line $\ell(P, P), (T, S)$ and so, in particular, has $\pi = 1$ on a neighborhood of $\frac{1}{2}(T + S, T + S)$.

If $P \leq E \leq R$ then the horizontal line $\{s_Y = E\}$ is a separation line. The associated simple Smale plan is the Smale version of an equalizer plan introduced in [9] and also described by Press and Dyson [15]. If $X$ uses this equalizer plan then the limiting payoff for $Y$ is $E$ regardless of $Y$’s play. On the other hand, the payoff to $X$ can be anything between $R$ and $P$, or even lower.

Figure 1. Competing Simple Smale Plans.

If, eventually, $X$ and $Y$ play simple Smale plans $\pi_X$ and $\pi_Y$ with separation lines $\ell_X$ and $\ell_Y$, respectively, then $Y$ responds with $\pi_Y \circ Switch$ and so the set $\Omega$ of
limiting payoffs lies on the intersection $\ell_X \cap \text{Switch}(\ell_Y)$. Except for the extreme cases with $\ell_X = \ell_Y = \text{Switch}(\ell_Y)$ equal to the diagonal or $\ell_X = \ell_Y = \text{Switch}(\ell_Y)$ equal to the co-diagonal the intersection is a single point and so the payoff sequence $\{s^N\}$ converges to the intersection point $\ell_X \cap \text{Switch}(\ell_Y)$.

**Proposition 2.9.** (a) If, eventually, $X$ plays the plan All-C then for any strategy for $Y$ and any initial plays, the limit point set $\Omega$ is contained in the segment $[(R,R),(S,T)]$.

(b) If, eventually, $X$ plays the plan All-D then for any strategy for $Y$ and any initial plays, the limit point set $\Omega$ is contained in the segment $[(P,P),(T,S)]$.

**Proof.** (a) Since All-C is a simple Smale plan with separation line $\ell_1 = (R,R)$, the limit point set lies in $\ell_1 \cap S = [(R,R),(S,T)]$. The result also follows from Corollary 2.6 (b) with $C = [(R,R),(S,T)]$

(b) If $P \leq \frac{1}{2}(T + S)$ then All-D is a simple Smale plan with separation line $\ell_2 = (P,P),(T,S)$ and we can proceed as in (a). If $P > \frac{1}{2}(T + S)$, All-D is not a simple Smale plan. The result nonetheless follows from Corollary 2.6 (a) with $C = [(P,P),(T,S)]$.

**Corollary 2.10.** Assume that, eventually, $X$ uses a Smale plan $\pi$. Let $Y$ use an arbitrary strategy and let $X$ and $Y$ use arbitrary initial plays. If $\Omega$ is the limit point set, then $\Omega$ is not contained in the $S$ interior of $\pi^{-1}(1) \setminus [(R,R),(S,T)]$ and $\Omega$ is not contained in the $S$ interior of $\pi^{-1}(0) \setminus [(P,P),(T,S)]$.

**Proof.** Assume $X$ uses $\pi$ from some time $N_*$. Let $U$ denote the interior of $\pi^{-1}(1) \setminus [(R,R),(S,T)]$. If $\Omega \subset U$ then by Proposition 2.1 there exists $N^* \geq N_*$ such that $s^N \in U$ for all $N \geq N^*$. This implies that for every round beyond $N^*$, $X$ plays $c$. So the sequence of payoffs is the same as though $X$ plays All-C starting from round $N^*$. So by Proposition 2.9 (a) the limit point set would be contained in $[(R,R),(S,T)]$. This contradiction shows that $\Omega \subset U$ is impossible.

The second assertion similarly follows from Proposition 2.9 (b).

3. **Good plans.** We describe informally the conditions that we would like a **good plan** to satisfy.

- **(Cooperation Condition)** If the players $X$ and $Y$ use fixed good strategies, i.e. good plans together with initial cooperation, then for every $N$, $S^N = (R,R)$ and so, of course, the time averages $s^N = (R,R)$ for all $N$ as well.

- **(Protection Condition)** If $X$ eventually plays a fixed good plan, and $s^* = (s^*_X, s^*_Y)$ is a limit point for the sequence $\{s^N\}$ with arbitrary initial play and with $Y$ using any plan, then

$$s^*_Y \geq R \implies s^*_X = s^*_Y = R. \quad (3.1)$$

- **(Robustness Condition)** If eventually the players $X$ and $Y$ use fixed good plans, then regardless of the earlier play $\lim_{N \to \infty} s^N = (R,R)$ at least with probability one.

A Markov plan is called **agreeable** if the response to a $cc$ outcome is always $c$. That is, $p$ satisfies $p_1 = p_{cc} = 1$. A Markov plan is called **firm** if the response to a $dd$ outcome is always $d$. That is, $p$ satisfies $p_d = p_{dd} = 0$. For example, the $TFT$ plan with $p = (1,0,1,0)$ is both agreeable and firm.

If both $X$ and $Y$ use Markov plans then $\{cc\}$ is a terminal set if and only if both plans are agreeable.
We call a general plan is weakly agreeable if the response is \( c \) when every previous outcome is \( cc \). A general plan is called weakly firm if the response is \( d \) when every previous outcome is \( dd \). For example, a Smale plan \( \pi \) is weakly agreeable if and only if \( \pi(R, R) = 1 \) and is weakly firm if and only if \( \pi(P, P) = 0 \). Clearly, a Markov plan is weakly agreeable if and only if it is agreeable and is weakly firm if and only if it is firm.

If X and Y both use weakly agreeable plans and initially cooperate then every outcome is \( cc \) and \( s^N = (R, R) \) for all \( N \).

So to obtain the Cooperation Condition we demand that a good plan be weakly agreeable.

The agreeable Markov plans which satisfy the Protection Condition can be completely characterized.

**Theorem 3.1.** Let \( p = (p_1, p_2, p_3, p_4) \) be an agreeable Markov plan so that \( p_1 = 1 \).

The plan \( p \) satisfies the Protection Condition if and only if the following inequalities hold.

\[
\frac{T - R}{R - S} \cdot p_3 < (1 - p_2) \quad \text{and} \quad \frac{T - R}{R - P} \cdot p_4 < (1 - p_2). \tag{3.2}
\]

**Proof.** See [5] Theorem 1.5, where a plan is called good if it is agreeable and satisfies the Protection Condition. See also [3]. \( \square \)

**Remark:** Note that (3.2) implies \( p_3 < 1 \).

For a Markov plan \( p \) we will refer to (3.2), together with the assumption \( p_1 = 1 \), as the protection inequalities. Thus, a Markov plan is agreeable and satisfies the Protection Condition exactly when the the protection inequalities hold.

For Smale plans we have

**Theorem 3.2.** Let \( \ell \) be a separation line with associated affine function \( L \) such that

\[
L(R) = 0, \quad \text{and} \quad L(P) > 0. \tag{3.3}
\]

That is, \( \ell \) is a line through \((R, R)\) with slope \( m \) satisfying \( 0 < m \leq 1 \).

Let \( \pi : S \to [0, 1] \) be a Smale plan. If \( L(s) > 0 \) implies \( \pi(s) = 0 \), i.e. \( \{L > 0\} \) is contained in the defection zone, then \( \pi \) satisfies the Protection Condition.

In particular, if \( \pi \) is a simple Smale plan with separation line \( \ell \) then \( \pi \) satisfies the Protection Condition.

**Proof.** The line \( \ell \) contains \((R, R)\), and the point \((P, R)\) lies above \( \ell \). Since \( \ell \) is a separation line, it follows that \( s = (R, R) \) is the only point of \( S \cap H^- \) with \( s_Y \geq R \).

By Theorem 2.5 (a) \( L(s^*_X, s^*_Y) \leq 0 \) for every limit point \( s^* \), i.e. \( \Omega \subset S \cap H^- \).

Hence, \((R, R)\) is the only possible limit point \( s^* \) with \( s^*_Y \geq R \). \( \square \)

If \( \ell \) is a line through \((R, R)\) with slope \( m \) satisfying \( 0 < m \leq 1 \) then \( \ell \) is a separation line. and we call such a line \( \ell \) a protection line for \( \pi \) if \( \pi = 0 \) above the line. The above result says exactly that if a Smale plan admits a protection line then it satisfies the Protection Condition.

**Definition 3.3.** A Markov plan \( p = (p_1, p_2, p_3, p_4) \) is generous when

\[
p_1 = 1, \quad 1 > p_2 > 0, \quad p_4 > 0. \tag{3.4}
\]

That is, a generous plan is agreeable and with positive probability responds to an opponent’s defection with cooperation, but does not always cooperate from a cd outcome (which had payoff \((S, T)\)).
Theorem 3.4. Assume that $X$ and $Y$, eventually, play Markov plans $\mathbf{p}$ and $\mathbf{q}$ respectively. If both plans are generous then $\{cc\}$ is the only terminal set for the associated Markov chain. So from any initial play, with probability one, eventually the outcome sequence is constant at $cc$ and so $\lim_{N \to \infty} s^N = (R,R)$. \\

Proof. Since the two plans are agreeable, $\{cc\}$ is a terminal set. Since $p_1, q_1 > 0$, there is a positive probability of moving from $dd$ to $cc$ and so $dd$ is a transient state. From $cd$, $1 > p_2 > 0$ implies that $X$ plays either $c$ or $d$ with positive probability. If $Y$ plays $c$ (or $d$) with positive probability then from $cd$ there is a positive probability of moving to $cc$ (resp. to $dd$ and thence to $cc$). Hence, $cd$ is transient. Recall that $Y$ uses the response vector $(q_1, q_1, q_2, q_4)$ and so cooperates after $dc$ with probability $q_2$. Thus, a symmetric argument shows that $dc$ is transient. \\

If the players use $\mathbf{p}$ and $\mathbf{q}$ from time $N^*$ on then from that point, the play follows the Markov chain given by $\mathbf{M}$ and so with probability one the sequence of outcomes eventually arrives at the unique terminal set $\{cc\}$. \hfill $\Box$ \\

Definition 3.5. A Smale plan $\pi : S \to [0,1]$ is generous if $\pi(s) = 1$ when $s_X > s_X$, and there exists an open subset $U$ of $S$ which contains the half-open segment $[\frac{1}{2} (T + S, T + S), (R,R))$ such that $\pi(s) = 1$ for $s \in U$. That is, the cooperation zone contains $U$ and the points on and below the diagonal line.

The following is essentially a part of [18] Theorem 1.

Theorem 3.6. Assume that, eventually, $X$ and $Y$ play Smale plans $\pi_X$ and $\pi_Y$, respectively. If both plans are generous then from any initial state in $S$, we have $\lim_{N \to \infty} (s^N_X, s^N_Y) = (R,R)$. \\

Proof. Assume that both plans are adopted by time $N^*$. Let $L_0(s) = s_Y - s_X$. That is, $L_0$ is the affine function associated with the diagonal separation line. Let $L_1(s) = s_X + s_Y - T - S$. That is, $L_1$ is the affine function associated with the co-diagonal. Notice that the maximum value of $L_1$ on $\bar{S}$ is $L_1(R,R) = 2R - T - S$ and that $L_1 < L_1(R,R)$ on $S \setminus \{(R,R)\}$. \\

Since $L_0(s) < 0$ implies $\pi_X(s) = 1$ and $\pi_Y(s) = 1$, we can apply Theorem 2.5 (b) to $L_0$ and $\pi_X$ to get $L_0(s^N) \geq -M_0 N^*/N$, where $M_0 = T - S$ is the maximum value of $L_0$ on $S$. The $Y$ player uses $\pi_Y \circ \text{Switch}$ and so we apply the theorem to $L_0 \circ \text{Switch}$ and $\pi_Y \circ \text{Switch}$ to get $L_0(s^N) = L_0 \circ \text{Switch}(s^N) \geq -M_0 N^*/N$. Hence, \\

\[ |L_0(s^N)| \leq M_0 N^*/N. \] 

(3.5) \\

Thus, the limit point set $\Omega$ is contained in the diagonal.

Observe that everywhere in $\bar{S}$ either $X$ or $Y$ plays $c$ and so the only outcomes are $cd, dc$ and $cc$. Hence, we have $L_1(s^N) \geq 0$ for all $N$. We can apply Lemma 2.3 with $L = -L_1$ to get $L(s^N) \geq -\frac{M_1 N^*}{N}$ for all $N \geq N^*$, where $M_1$ is the maximum value of $|L_1|$ on $S$. Hence, on $\Omega$, $L_1 \geq 0$. Notice that in the case when $P \geq (T + S)/2$, $L_1 \geq 0$ on all of $S$.

Thus, $\Omega \subset \{ \frac{1}{2} (T + S, T + S), (R,R)\}$. \\

Let $U = U_X \cap \text{Switch}(U_Y)$ where $U_X$ and $U_Y$ are the open sets containing the half-open interval $[\frac{1}{2} (T + S, T + S), (R,R))$ on which $\pi_X = 1$ and $\pi_Y = 1$, respectively. By assumption, if $s^N \in U$ then both players play $c$ with outcome $cc$ at time $N + 1$. So $S^{N+1} = (R,R)$. \\

For $\epsilon > 0$, $\{ L_1 > L_1(R,R) - \epsilon \}$ is a neighborhood of $(R,R)$ and so $U \cup \{ L_1 > L_1(R,R) - \epsilon \}$ is a neighborhood of $[\frac{1}{2} (T + S, T + S), (R,R)]$ and so of $\Omega$. From Proposition 2.1 it follows that there exists $N_\epsilon \geq N^*$ so that $N \geq N_\epsilon$ implies
$s^N \in U\{L_1 > L_1(R,R)\}$. Hence, for $N \geq N_\epsilon$, $L_1(s^N) - L_1(R,R) + \epsilon \leq 0$ implies $s^N \in U$ and so $L_1(S^{N+1}) = L_1(R,R)$ and so $L_1(S^{N+1}) - L_1(R,R) + \epsilon = \epsilon > 0$. We can again apply Lemma 2.3 this time to $L = -(L_1 - L_1(R,R) + \epsilon)$ to get

$$N \geq N_\epsilon \implies L_1(s^N) \geq L_1(R,R) - \epsilon - \frac{(M + \epsilon)N_\epsilon}{N},$$

where $M$ is the maximum of $|L_1 - L_1(R,R)|$ on $S$. Hence, there exists $N'_\epsilon \geq N_\epsilon$ so that

$$N \geq N'_\epsilon \implies L_1(s^N) \geq L_1(R,R) - 2\epsilon.$$

Thus, $\lim_{N \to \infty} L_1(s^N) = L_1(R,R)$.

Since, $(R,R)$ is the unique maximum point for $L_1$, it follows that $\lim_{N \to \infty} s^N = (R,R)$. □

It would be nice to show that if $X$ plays a generous Smale plan and $Y$ plays a generous Markov plan, then with probability one $\lim_{N \to \infty} s^N = (R,R)$. I don’t know if the conjecture is true in full generality. However, we get the result we want if we strengthen the assumption.

**Definition 3.7.** A Smale plan $\pi$ is convex-generous if $\pi : S \to \{0,1\}$, i.e., no mixed strategy responses, and the cooperation zone $\pi^{-1}(1)$ is a closed convex set $C$ such that

(i) $(P,P), (R,R), (T,S) \in C$.

(ii) $(S,T) \notin C$.

(iii) $\frac{1}{2}(T + S, T + S) \in C^o$, where $C^o$ is the interior of $C$ with respect to $S$.

By (i) and (iii), there exists $t^* \in (\frac{1}{2}, 1)$ such that $(1 - t)(T,S) + t(S,T) \in C$ if and only if $0 \leq t \leq t^*$. Let $V$ denote the point $(1 - t^*)(T,S) + t^*(S,T)$.

Let $\hat{P} = \min(P, \frac{1}{2}(T + S))$. Thus, the diagonal line intersects $S$ in the segment $[(R,R), (P,\hat{P})]$ and $(P,\hat{P}) \in C$, by (i) and (iii). Hence, the convex hull of $[V, (P,\hat{P}), (R,R), (T,S)]$ is contained in $C$. This contains $(s_X, s_Y)$ if $s_X \geq s_Y$. Furthermore, its $S$ interior contains $[(\frac{1}{2}(T + S, T + S), (R,R))$, and so, as expected, $\pi$ is generous.

**Theorem 3.8.** Assume $X$ eventually plays a convex-generous Smale plan and $Y$ plays a generous Markov plan. With probability one there is a time after which both players play $c$ and so $\lim_{N \to \infty} s^N = (R,R)$.

**Proof.** Let $X$ play $\pi$ with convex set $C = \pi^{-1}(1)$ and let $Y$ use $q = (q_1, q_2, q_3, q_4)$ with $q_1 = 1$ and $\epsilon < q_2, \epsilon < q_4$ for some $\epsilon > 0$. Recall that for $Y$, $q_2$ is the probability of cooperating in a round following the outcome $dc$.

**Claim 1.** With probability one, for every $N$ there exists $n \geq N$ such that $s^n \in C$.

**Proof.** It suffices to show that for every $N$ the event

$$E_N = \{s^n \in S \setminus C : \text{for all } n \geq N\}$$

has probability zero.

Let $\ell_1 = V(R,R)$, which is a separation line, and let $L_1$ be an affine function associated with $\ell_1$. Clearly, $(P,P)$ and $(T,S)$ lie below the line and so $L_1(P,P), L_1(T,S) < 0$. By Lemma 2.4 (a) there exists a positive constant $k_1$ (which depends only on $L_1$) such that for some $N_1$ between $N$ and $k_1N$ either $L_1(s^{N_1}) < 0$ or $X$ plays $c$ on round $N_1$. Note that the latter is equivalent to $s^{N_1} \in C$. 


Case 1. \([P \geq \frac{1}{2}(T+S)]\): The point \(V\) lies on the side \([S,T), (T,S)\) of the triangle \(S\). Thus, every point of \(S\) on or below \(\ell\) lies in \(C\). Hence, \(L_1(s) \leq 0\) implies \(s \in C\) and so \(s^{N_1} \in C\) in any case.

Case 2. \([P < \frac{1}{2}(T+S)]\): The line \(\ell_1\) intersects the open segment \((P,P), (S,T)\) in a point \(V'\). Let \(\Delta\) denote the triangle \([P,P), V', V]\) and \(\Delta' = \Delta \setminus C\). Observe that \(C \cup \Delta'\) contains the set \(\{L_1 \leq 0\} \cap S\). Assuming \(E_N\), \(s^{N_1}\) lies in \(\Delta'\) since it is not in \(C\). For \(n \geq N\) if \(s^n \in \Delta'\) then, since it is not in \(C\), the payoff \(S^{n+1}\) is either \((P,P)\) or \((T,S)\) and so by (2.5) \(L_1(s^{n+1}) \leq \frac{n}{n+1}L_1(s^n) \leq 0\). Since \(s^{n+1} \notin C\), it follows that \(s^{n+1} \in \Delta'\). Inductively, we have \(s^n \in \Delta'\) for all \(n \geq N_1\) and for all \(n \geq k_1N\).

Now if among the rounds \(k_1N, \ldots, M-1\) \(Y\) plays \(c\) exactly \(r\) times then

\[
s^{k_1N+M} = \frac{k_1Ns^{k_1N} + (M-r)(P,P) + r(T,S)}{k_1N + M},
\]

and so the vector from \((P,P)\) to \(s^{k_1N+M}\) is the nonzero vector

\[
s^{k_1N+M} - (P,P) = \frac{k_1N[s^{k_1N} - (P,P)] + r[(T,S) - (P,P)]}{k_1N + M}.
\]

Normalizing these vectors to obtain \([s^{k_1N+M} - (P,P)]\) of length 1.

Suppose that as \(M \to \infty\) the number \(r\) of \(c\) plays by \(Y\) is unbounded. Then, as \(M \to \infty\) these unit vectors have \([T,S) - (P,P)]\) as a limit point. On the other hand, the closed set of vectors \([s - (P,P)] : s \in \Delta \setminus \{(P,P)\}\) does not contain \([T,S) - (P,P)]\).

It follows that, assuming \(E_N\), there exists \(R < \infty\) such that \(Y\) plays \(c\) at most \(R\) times and so from some \(N_2\) onward \(Y\) always plays \(d\). For each \(N_2 \geq k_1N\) the event \(E_{N,N_2} = E_N\) and \(Y\) plays \(d\) on every round \(n\) with \(n \geq N_2\) has probability zero because at each such round \(Y\) is responding to \(dd\) by playing \(d\). These are independent events each with probability at most \(1-\epsilon\). So \(E_N = \bigcup_{N_2 \geq k_1N} E_{N,N_2}\) has probability zero.

This completes the proof of Claim 1.

If \(Y\) plays \(c\) at any time \(N\) when \(s^n \in C\) then the outcome of the \(N+1\) round is \(cc\) with payoff \((R,R)\). Since \(C\) is convex \(s^{N+1} \in C\) and so \(X\) next plays \(C\) and \(Y\) next plays \(C\) because \(q\) is agreeable. Inductively \(cc\) is the outcome and \(s^n \in C\) for every round \(n\) with \(n \geq N\). Let \(E_0\) denote the event \(Y\) plays \(d\) whenever \(s^n \in C\) and \(\bar{E} = E_0 \setminus \bigcup_N \{E_N\}\). From Claim 1, it suffices to show that \(\bar{E}\) has probability zero.

Claim 2. Assuming \(\bar{E}\), for every \(N\) there exists \(n \geq N\) such that \(s^n \in S \setminus C\).

Proof. If we assume \(E_0\), then whenever \(X\) plays \(c\), \(Y\) plays \(d\) leading to payoff \((S,T)\). If for some \(N\), \(s^n \in C\) for all \(n \geq N\) and \(E_0\) is true, then for all \(n > N\) we have \(S^n = (S,T)\) and the sequence \(\{s^n\}\) would converge to \((S,T)\), but the complement of \(C\) is a neighborhood of \((S,T)\) and so eventually \(s^n \notin C\). The contradiction proves Claim 2.

Assuming \(\bar{E}\), \(s^n \in C\) and \(s^n \in S \setminus C\) each occur infinitely often by Claim 1 and Claim 2.

Now let \(N_k\) be the \(k^{th}\) return time to \(C\) from \(X \setminus C\). This is an infinite sequence of Markov times. Since at time \(N_k - 1\) the payoff sequence was in \(S \setminus C\), \(X\) played \(d\) and so at time \(N_k\) \(Y\) plays \(d\) in response to either a \(dc\) or a \(dd\). Playing \(d\) in these cases has probability at most \(1 - \epsilon\). Furthermore, the \(Y\) play at time \(N_k\) is
independent of the previous plays. Thus, again \( E \) requires an infinite sequence of independent events, each with probability at most \( 1 - \epsilon \). Hence, \( E \) has probability zero.

**Remark:** Our assumptions on \( q \) allow the possibility \( q_3 = 0 \). So if \((S,T)\) were in \( C^o \) then from a neighborhood of \((S,T)\) it would be a limit point following an infinite sequence of \( cd \) outcomes. Notice that the assumption \((S,T) \notin C\) is analogous to the assumption that \( p_2 < 1 \) for a generous Markov plan. If \( p_2 = 1, p_3 = 0 \) and \( q \) satisfies the analogous condition then \{\( cd \)\} and \{\( dc \)\} are terminal sets when \( X \) plays \( p \) and \( Y \) plays \( q \).

**Definition 3.9.** We call a Markov plan *good* when it satisfies the protection inequalities (and so is agreeable) and is generous.

We call a Smale plan *good* (or *convex-good*) when it is weakly agreeable, admits a protection line and is generous (resp. and is convex-generous).

For example, if \( \pi \) is a simple Smale plan with separation line \( \ell \) then \( \pi \) is good if and only if it is weakly agreeable (i.e. \( \pi(R,R) = 1 \)) and \( \ell \) is a line through \((R,R)\) with slope strictly between 0 and 1, so that \( \ell \) is a protection line. It is convex-good if and only if, in addition, \( \pi = 1 \) on \( \ell \cap S \). Such a good simple Smale plan is the Smale version of what is called in [10] a *complier strategy* (also called a generous zero-determinant strategy, as in e.g. [20]). Any limit point \( s^* \), other than \((R,R)\) on the separation line \( \ell \) has \( s^*_Y \) larger than \( s^*_X \), albeit less than \( R \).

On the other hand, if \( \ell \) is a line through \((P,P)\) with slope between 0 and 1 then it is a separation line and the associated simple Smale plan is, when \( P \leq \frac{1}{2}(T+S) \), the Smale version of what Press and Dyson [15] call an *extortionate strategy*. Any limit point \( s^* \), other than \((P,P)\) on the separation line \( \ell \) satisfies \( s^*_Y < s^*_X \). Thus, if \( Y \) plays to avoid the \((P,P)\) payoff she always obtains less than \( X \) does from the change in policy. The best reply to such an extortionate strategy is All-C. The payoff point is then the intersection point \( B = \ell \cap ((R,R), (T,S)) \) with \( P < B_Y < R < B_X \).

In [5] a Markov plan, there called a memory-one plan, is called good when it satisfies the protection inequalities, or, equivalently, it is agreeable and satisfies the Protection Condition. The TFT plan with \( p = (1,0,1,0) \) satisfies the protection inequalities but is firm rather than generous. If both \( X \) and \( Y \) use the TFT plan then from initial outcome \( cc \) the sequence of outcomes is fixed at \( cc \), but from an initial \( dd \), the state \( dd \) is fixed. Following \( cd \) or \( dc \) the two states alternate leading to convergence of the payoff sequence to \( \frac{1}{2}(T+S,T+S) \). The phenomena illustrate the failure of robustness in the absence of generosity.

In [18] a Smale plan \( \pi \) is called good if it is generous, and so satisfies the Robustness Condition, and \( \pi(s) = 0 \) when \( s_Y > R \). The line \( \{ s_Y = R \} \) is an equalizer line and so Smale’s conditions allows the possibility of a limit outcome \((s_X,R)\) with \( s_X < R \).

In addition, Smale imposed the condition that \( \pi(s) = 0 \) when \( s_X < P \). If \( P > \frac{1}{2}(T+S) \), this would contradict the condition that \( \pi = 1 \) when \( s_X > s_Y \). Smale was only considering the case with \( P < \frac{1}{2}(T+S) \) and we will examine that situation first.

Assume \( P < \frac{1}{2}(T+S) \). Choose \( \ell_1 \) a line through \((R,R)\) with slope strictly between 0 and 1, so that the weakly agreeable simple Smale plan with separation line \( \ell_1 \) is good. Let \( A \) be the point of intersection of \( \ell_1 \) and the open segment \(((P,P), (S,T))\). Choose a point \( V \) on the open segment \(((R,R), A)\) with \( V_X \geq P \) or larger. Let \( \ell_2 \) be the line \( (P,P), V \). Let \( \pi \) be a Smale plan such that \( \pi(s) = 0 \)
if $s$ is above $\ell_1$ or above $\ell_2$ and $\pi(s) = 1$ at $(R, R)$, below the diagonal and on some open set containing $\frac{1}{2}(T + S, T + S), (R, R))$. Thus, $\pi$ is generous and since $\ell_1$ is a protection line it is a good Smale strategy. It follows from Corollary 2.6(c) that if $X$ eventually plays $\pi$ against any plan for $Y$ and any initial plays, then the limit point set $\Omega$ is contained in the triangle $[(P, P), (R, R), V]$.

The advantage of such a plan is that it excludes points above $\ell_2$ from the limit. In contrast, against the good simple Smale strategy with separation line $\ell_1$, any point of $[A, (R, R)]$ can occur as the limit if $Y$ plays a suitable simple Smale strategy. For example, recall that All-D is a simple Smale strategy with separation line $\ell_1$.

This sort of possible cost can always occur with a generous Markov plan $p$. If $Y$ plays All-D, which is a Markov plan with $q = (0, 0, 0, 0)$, then the unique terminal set is $\{cd, dd\}$ with stationary distribution $v = (0, p_4, 0, (1 - p_2))/[p_4 + (1 - p_2)]$. The limiting average payoff $(s_X^*, s_Y^*)$ given by (2.8) satisfies $(s_X^*, s_Y^*) - (P, P) = \epsilon(S - P, T - P)$ with $\epsilon = p_4/[p_4 + (1 - p_2)] > 0$. So $s_Y^* > P > s_X^*$.

On the other hand, for plans such as $\pi$ above we have seen that the limit point set $\Omega$ against any $Y$ play is a connected set contained in the triangle $[(P, P), V, (R, R)]$.

However:

**Example 3.10.** The limit set $\Omega$ need not be a point or interval and it may contain points in the interior of the triangle.

Assume $P < \frac{1}{2}(T + S)$. Choose $V \in S$ with $R > V_Y > V_X \geq P$. Let $\ell_1 = (R, R), V/\ell_2 = (P, P), V/\ell$ such that $\pi_X(s) = 0$ for above $\ell_1$ or above $\ell_2$ and $\pi_X(s) = 1$ otherwise. In that case, $\pi_X$ is convex-good with $\ell$ the quadrilateral $[(A, A), (B, B), (T, T), (S, S)]$. Let $C' \subset C$ denote the triangle $[(P, P), V, (R, R)]$. As we saw above, when $X$ plays $\pi_X$ against any strategy for $Y$, the limit point set $\Omega$ is contained in $C'$. Furthermore, by Corollary 2.10 $\Omega$ must intersect $[(P, P), V] \cup [V, (R, R)]$.

Choose a point $V'$ on the half-open segment $[V, (R, R)]$ and a point $W$ on $(S, T), V'$ between the lines $\ell_1$ and the diagonal. Let $\ell = (R, R), W$ so that the line $\ell$ lies between $\ell_1$ and the diagonal, with all three intersecting at $(R, R)$. Let $\ell' = (S, T), W = (S, T), V'$.

Label the following points:
- $\ell_1 \cap (P, P), (S, T) = A$.
- $\ell_2 \cap (R, R), (S, T) = B$.
- $\ell' \cap (P, P), (R, R) = (Q, Q)$ and $\ell' \cap \ell_1 = V'$.
- $\ell \cap \ell_2 = W'$ and $\ell \cap (P, P), (S, T) = W''$.

Let $C$ be the quadrilateral $[(P, P), (Q, Q), W, W'']$. Define $\pi_Y$ for $Y$ to be the Smale plan with $\pi_Y(s) = 1$ if $s$ lies in $\text{Switch}(C)$ and $= 0$ otherwise. Recall that $Y$ responds with $\pi_Y \circ \text{Switch}$ and so $Y$ plays $c$ if $s \in C$ and plays $d$ otherwise.

We prove that if, eventually, $X$ plays $\pi_X$ and $Y$ plays $\pi_Y$ then regardless of the initial play, the limit point set $\Omega$ is the boundary of the quadrilateral $\tilde{L} = [V, W', W, V']$ (If $V' = V$, $\Omega$ is the boundary of the triangle $[V, W', W]$).

**Proof.** Assume that $X$ and $Y$ play $\pi_X$ and $\pi_Y$, respectively, beyond time $N^*$.

The lines $\ell_1, \ell_2, \ell, \ell'$ and the diagonal subdivide $S$ into twelve polyhedral regions. For any $\epsilon > 0$, let $N_\epsilon > N^*$ be an $\epsilon$ step-time after which every move from $s^N$ to $s^{N+1}$ has distance less than $\epsilon$, i.e. let $N_\epsilon$ be greater than $\max(N^*, \sqrt{2}(T - S)/\epsilon)$. 

Let $\epsilon_0 > 0$ be smaller than the distance between any two non-intersecting regions and let $N_0 = N_{\epsilon_0}$. Thus, such a small move cannot jump between non-intersecting regions.

Let $\tilde{C} = C \cap C$, which is the quadrilateral $[(P, P), W', W, (Q, Q)]$.

**Claim.** For every $N \geq N^*$ there exists $n \geq N$ such that $s^n \in \tilde{C}$.

First we show that the sequence of payoffs must enter $\tilde{C}$. If not, then for every round beyond $N$, Y plays $d$. As in the proof of Proposition 2.10 the results after $N$ are the same as though Y plays All-D which is a simple Smale plan with separation line $(P, P), (T, S)$. Then $\Omega$ is contained in the intersection of the triangle $C'$ with $\text{Switch}(P, P), (T, S)$ of $(P, P), (S, T)$. This intersection contains only the point $(P, P)$. If $\Omega$ were just $(P, P)$ then for any small neighborhood $U$ of $(P, P)$ eventually $s^n \in U$. If $s^n$ is on or above the diagonal then $s^n \in \tilde{C}$. If $s^n$ is below the diagonal then the sequence of payoffs moves toward $(S, T)$ and eventually makes a small jump into $\tilde{C}$. Either way, this contradicts the assumption that the sequence never enters $\tilde{C}$.

Now for $Z$ on the open segment $((P, P), W')$, let $U_Z$ consist of the points of $S$ which are below both of the lines $(S, T)$, $Z$ (and $(T, S)$, $Z$. These are convex open
neighboring neighborhoods of \((P, P)\) which converge to \((P, P)\) as \(Z \to (P, P)\). Since \((P, P) \in \hat{C}\), if \(s^N \in C \setminus \hat{C}\) then there exists \(Z\) such that \(s^N \not\in U_Z\), i.e. \(s^N \in K_1 = C \setminus (C \cup U_Z)\). Assume \(n > N_0\). From such a point the sequence moves toward \((T, S)\). During the motion it remains above \(U_Z\). If the sequence does not enter \(\hat{C}\) from \(C \setminus \hat{C}\) then it jumps to below the diagonal to land in \(K_2\) the set of points outside \(U_Z\) which are on or below the diagonal and \(\ell'\). From such points the sequence moves back toward \((S, T)\). If it jumps over \(\hat{C}\) then it re-enters \(K_1\). This alternation cannot continue indefinitely. Notice that \(K_1\) and \(K_2\) are a positive distance \(\epsilon_Z\) apart. Once \(N \geq N_{\epsilon_Z}\) the move from \(K_1\) or \(K_2\) must land in \(\hat{C}\).

This completes the proof of the Claim.

For any \(\epsilon > 0\), let \(s^n \in \hat{C}\) with \(n > N_\epsilon\). From this point the sequence moves toward \((R, R)\) exiting \(\hat{C}\) at a point above, and \(\epsilon\) close to, the line \(\ell'\). The sequence now moves toward \((S, T)\), \(\epsilon\) close to and above the line \(\ell'\). It exits \(C\) at a point above \(\ell_1\) and \(\epsilon\) close to \(V'\). Now the sequence moves toward \((P, P)\) entering \(L\) on or below, and \(\epsilon\) close to \(\ell_1\) and then moving back toward \((S, T)\). These \(PP\) and then \(ST\) alternate motions may continue for a long time but it must eventually cease since the sequence must eventually return to \(\hat{C}\). The exit occurs when the sequence lands above \(\ell_2\), \(\epsilon\) close to \(V\). The sequence then moves toward \((P, P)\) above and \(\epsilon\) close to the line \(\ell_2\) until it enters \(C \setminus \hat{C}\), \(\epsilon\) close to \(V'\). The sequence then moves toward \((T, S)\) crossing \(\ell_2\) close to \(W'\) to re-enter \(\hat{C}\) now below and \(\epsilon\) close to the line \(\ell\).

As \(N \to \infty, \epsilon \to 0\) and the motion gets close to motion from \(W'\) to \(W\), from \(W\) to \(V'\), from \(V'\) to \(V\), and then from \(V\) back to \(W\). □

Turning now to the case when \(P \geq 1/2(T + S)\) we see that the additional condition imposed by Smale does not work so well.

Suppose you demand that \(\pi\) satisfy \(\pi(s) = 0\) when \(s_X \leq 1/2(T + S)\) or even just \(\pi = 0\) on the half-open segment \((1/2(T + S, T + S), (S, T)]\) with \(\pi = 1\) on \((1/2(T + S, T + S), (T, S)]\). If \(\pi_X\) and \(\pi_Y\) both satisfy this condition then for \(s^N \in (1/2(T + S, T + S), (S, T)]\) the payoff \(S^{N+1} = (T, S)\) and for \(s^N \in (1/2(T + S, T + S), (T, S)]\) the payoff \(S^{N+1} = (S, T)\). Thus, the sequence remains on \([(S, T), (T, S)]\) moving back and forth as the point \(1/2(T + S, T + S)\) is passed with limit point \(1/2(T + S, T + S)\), unless the sequence lands exactly on the point \(1/2(T + S, T + S)\). If that happens then, the result depends on the choices at the point \(1/2(T + S, T + S)\).

For most initial points, this cannot happen. For example, if the initial point is an irrational mixture of \((S, T)\) and \((T, S)\) then hitting \(1/2(T + S, T + S)\) does not occur. However, for actual play all outcomes are rational mixture of the four points \((S, T), (T, S), (P, P)\) and \((R, R)\). Focusing upon actual play leads to odd results.

**Proposition 3.11.** Suppose that \(\pi(S, T) = 0, \pi(T, S) = 1\) and \(\pi(s) = 1\) when \(s\) lies on the diagonal. If both \(X\) and \(Y\) play Smale plans which satisfy these conditions then for any initial plays, \(\lim s^N = (R, R)\).

**Proof.** If the initial outcome is \(cc\) or \(dd\) then the initial payoff lies on the diagonal and so every successive outcome is \(cc\). If the initial outcome is \(cd\), with payoff \((S, T)\) then the next outcome is \(dc\) with payoff \((T, S)\) so that \(s^2 = 1/2(T + S, T + S)\) which lies on the diagonal. Hence, in any case, the outcome on the \(n\)th round is \(cc\) for \(n \geq 3\) and the limit result follows. □

This version of “robustness” is very unsatisfying. For real robustness one wants approach to \((R, R)\) even if errors occur in the computations and if the plans are
Example 3.12. Returning to the case with $P > \frac{1}{2}(T + S)$ we describe what seems to me to be the best version of robustness we can hope for if we demand protection against payoffs below $P$.

Let $\ell$ be a protection line with slope less than 1 so that $(P, P)$ lies below $\ell$. Let $V \in \ell$ with $P \leq V_X < R$, $\ell_1 = (P, P), V$, and $\ell_2$ the vertical line $\{s_X = P\}$. Let $W$ be the point of intersection $\ell_2 \cap (S, T), (T, S)$. Assume that $\pi_X(s) = 0$ whenever $s$ is above $\ell$ or $\ell_1$, or if it is on or to the left of $\ell_2$. Otherwise, $\pi_X(s) = 1$.

It follows from Corollary 2.6(a) and (b) that if $X$ eventually plays $\pi_X$ then against any $Y$ play, the limit point set $\Omega$ is contained in the convex hull $[W, (P, P), V, (R, R)] = [W, (P, P), V, (R, R), (T, S)] \cap [W, (R, R), (S, T)]$.

Now assume that $Y$ also eventually plays such a plan $\pi_Y$ with lines $\ell', \ell'_1, \ell'_2 = \ell_2$ and with points $V'$ and $W$.

Let $D$ consist of the set of points of $S$ which are either on or below $(P, P), (S, T)$ (or on or below $(P, P), (T, S)$). Let $D_1 = [W, Switch(W), (P, P)]$, the set of $s \in S$ with $s_X, s_Y \leq P$. It is easy to check that if $s^{N*} \in D$ then $s^n \in D$ for all $n \geq N^*$.
and that for some $N^{**} \geq N^{*}$, $s^{N^{**}} \in D_{1}$. All subsequent outcomes are $dd$ and so $\lim s^{N} = (P, P)$. On the other hand, if $\{s^{N}\}$ does not converge to $(P, P)$ then it is easy to check that some $s^{N^{**}}$ lies in $[(P, P), V, (R, R), Switch(V')] \setminus \{(P, P)\}$. All subsequent outcomes are $cc$ and so $\lim s^{N} = (R, R)$. Thus, we always have either convergence to $(P, P)$ or to $(R, R)$.

Return now to consider a simple Smale plan with separation line $\ell$. For such a plan the choices on the line $\ell$ were, in general, left unspecified. Recall that if $\pi_{X}$ and $\pi_{Y}$ are simple Smale plans with separation lines $\ell_{X}$ and $\ell_{Y}$ then, except for the extreme cases when both $\ell_{X}$ and $\ell_{Y}$ are the diagonal or both the co-diagonal (which requires $P \leq \frac{1}{2}(T + S)$), then if, eventually, $X$ uses $\pi_{X}$ and $Y$ uses $\pi_{Y}$ the sequence $\{s^{N}\}$ converges to the point of intersection $\ell_{X} \cap Switch(\ell_{Y})$ regardless of earlier play and regardless of the choices on the separation lines.

To illustrate where the choices on the lines become important, let us consider the extreme cases.

Suppose $P < \frac{1}{2}(T + S)$ and $\pi$ is a simple Smale plan with separation line the co-diagonal, $\ell = (S, T), (T, S)$. If both players use $\pi$ from time $N^{*}$ on then if $s^{N^{*}}$ is above $\ell$ then the sequence remains above $\ell$ converging to $(T, S)$. Similarly, if $s^{N^{*}}$ is below $\ell$ then the sequence remains below $\ell$ and converges to $(S, T)$. If $s^{N^{*}} \in \ell$ then the result depends on the choice of $\pi$ on $\ell$. If $\pi(s^{N^{*}}) = \pi(Switch(s^{N^{*}})) = 1$ then $s^{N^{*}+1}$ is above $\ell$ with convergence to $(T, S)$. If $\pi(s^{N^{*}}) = \pi(Switch(s^{N^{*}})) = 0$, then $s^{N^{*}+1}$ is below $\ell$ with convergence to $(S, T)$. Suppose that $\pi(s) = 0$ if $s \in \ell$ with $s$ above the diagonal and $\pi(s) = 1$ if $s \in \ell$ on or below the diagonal, then we again have alternating $(S, T)$ and $(T, S)$ motion with limit $\frac{1}{2}(T + S, T + S)$ unless the sequence lands on the point $\frac{1}{2}(T + S, T + S)$ in which case we have convergence to $(T, S)$.

Now let $\pi$ be the simple Smale plan with separation line the diagonal. Suppose that for $s = (Q, Q) \in \ell$, $\pi(s) = 1$ if $Q \geq \frac{1}{2}(T + S)$ and $= 0$ if $Q < \frac{1}{2}(T + S)$. I think of this as the Smale version of Tit-for Tat. Suppose both players use $\pi$ for $N \geq N^{*}$ and $s^{N^{*}} \not\in \ell$. We obtain alternating motion towards $(T, S)$ and $(S, T)$ with limit point $\frac{1}{2}(T + S, T + S)$ unless at some time $N \geq N^{*}$, $s^{N} = (Q, Q) \in \ell$. If $Q \geq \frac{1}{2}(T + S)$ then we obtain outcomes $cc$ for all rounds after $N$ with convergence to $(R, R)$. If $Q < \frac{1}{2}(T + S)$ then we obtain outcomes $dd$ for all rounds after $N$ with convergence to $(P, P)$.

Thus, in both extreme cases, the limit results depend upon the $\pi$ choices on the separation line.

4. Separation paths and the Folk Theorem. This section is the result of some suggestions and questions raised by Christian Hilbe in response to an earlier version.

Definition 4.1. We call a pair of Smale plans $\pi_{X}, \pi_{Y}$ a Nash Equilibrium when the following hold:

(a) There is a point $s^{*} = (s_{X}^{*}, s_{Y}^{*})$ such that if $X$ eventually plays $\pi_{X}$ and $Y$ eventually plays $\pi_{Y}$ then the outcome sequence $\{s^{N}\}$ converges to $s^{*}$. We then call $s^{*}$ the payoff to the pair.

(b) If $Y$ eventually plays $\pi_{Y}$ then any limit point $V$ of an outcome sequence satisfies $V_{X} \leq s_{X}^{*}$, regardless of the play of $X$.

If $X$ eventually plays $\pi_{X}$ then any limit point $V$ of an outcome sequence satisfies $V_{Y} \leq s_{Y}^{*}$, regardless of the play of $Y$.

That is, neither player can obtain an improved payoff by unilaterally changing plans.
The so-called Folk Theorem of Iterated Play in this context should say that for any \( s^* \in \mathcal{S} \) with \( s^*_X, s^*_Y \geq P \), there exists a Nash Equilibrium with payoff \( s^* \).

With \( R \geq s^*_X, s^*_Y \geq P \) this is easy to obtain. If \( \pi_X \) and \( \pi_Y \) are the equalizer simple Smale plans with horizontal separation lines \( \ell_X \) given by \( s_Y = s^*_Y \) and \( \ell_Y \) given by \( s_Y = s^*_X \) then the pair \( \pi_X, \pi_Y \) is a Nash equilibrium. In fact, as long as \( Y \) plays \( \pi_Y \), the limit set \( \Omega \) lies in the vertical line \( \text{Switch}(\ell_Y) \) and \( X \) obtains \( s^*_X \), regardless of his play. Consequently, \( X \) has no incentive to use \( \pi_X \). Similarly, for \( Y \). So we would like to strengthen condition (b) to the analogue of the Protection Condition in this context.

**Definition 4.2.** We call a pair of Smale plans \( \pi_X, \pi_Y \) a Strong Nash Equilibrium when the following hold:

(a) There is a point \( s^* = (s^*_X, s^*_Y) \) such that if \( X \) eventually plays \( \pi_X \) and \( Y \) eventually plays \( \pi_Y \), then the outcome sequence \( \{s^N\} \) converges to \( s^* \).

(b) If \( Y \) eventually plays \( \pi_Y \), then the point \( V = s^* \) is the only limit point \( V \) of an outcome sequence with \( V_X \geq s^*_X \), regardless of the play of \( X \).

If \( X \) eventually plays \( \pi_X \), then the point \( V = s^* \) is the only limit point \( V \) of an outcome sequence with \( V_Y \geq s^*_Y \), regardless of the play of \( Y \).

Also, we would like to obtain Nash equilibrium results when \( s^*_X \) or \( s^*_Y \) is greater than \( R \).

To deal with the cases when \( P > \frac{1}{2}(T + S) \), we again let \( \bar{P} = \min(P, \frac{1}{2}(T + S)) \).

**Definition 4.3.** If \( s \in \mathcal{S} \) then the upper triangle with vertex \( s \), denoted \( T(s) \), is the triangle \( [s, (S, T), (R, R)] \). The lower quadrangle with vertex \( s \), denoted \( Q(s) \), is the convex set \( [s, (T, S), (P, P), (P, \bar{P})] \).

A non-empty subset \( C \subseteq \mathcal{S} \) is upper full if \( s \in C \) implies \( T(s) \subseteq C \), and lower full if \( s \in C \) implies \( Q(s) \subseteq C \).

For \( C \subseteq \mathcal{S} \) we let \( C^\circ \) denote the \( \bar{s} \) interior. Observe that if \( s \) is on the segment \( (\mathcal{S}, T), (R, R)] \) then \( T(s) \) is the segment with empty interior. If \( P \leq \frac{1}{2}(T + S) \) then \( \bar{P} = P \), then \( Q(s) \) is the triangle \( [s, (T, S), (P, P)] \).

If \( C \subseteq \mathcal{S} \) is nonempty, then the distance from \( s \) to \( C \) is \( d(s, C) = \inf \{||s - \bar{s}|| : \bar{s} \in C\} \). For \( \epsilon \geq 0 \), let \( C_\epsilon = \{s \in \mathcal{S} : d(s, C) \leq \epsilon\} \). Observe that \( C_0 \) is the closure of \( C \).

**Lemma 4.4.** Let \( C \) be a nonempty subset of \( \mathcal{S} \) and let \( \epsilon \geq 0 \).

(a) If \( C \) is upper full, then it is connected and contains \( [(S, T), (R, R)] \). If \( C \) is lower full, then it is connected and contains \( [(T, S), (P, P), (P, \bar{P})] \).

(b) If \( C \) is closed and is upper (or lower) full then \( C_\epsilon \) is upper full (resp. lower full). In particular, the closure of \( C \) is upper full (resp. lower full).

(c) The union \( C^+ \) (or \( C^- \)) of the upper triangles (resp. the lower quadrangles) with vertices in \( C \) is an upper full (resp. lower full) subset which is closed if \( C \) is.

(d) If \( C \) is convex then it is upper full if \( (S, T), (R, R) \in C \) and it is lower full if \( (T, S), (P, P), (P, \bar{P}) \in C \).

**Proof.** We prove the results for the upper case as the lower is similar.

(a) If \( C \) is upper full then it is a union of upper triangles, all of which are connected and all of which contain \( [(S, T), (R, R)] \). Such a union is connected.

(b) Let \( \epsilon_1 > \epsilon \). If \( r \in \mathcal{S} \) then the segment \( [s, r] \) consists of the points \( ts + (1 - t)r \) for all \( t \in [0, 1] \). The upper triangle \( T(s) \) is the union of the segments \( [s, r] \) with \( r \in [(S, T), (R, R)] \). If \( s_1 \in \mathcal{S} \) then \( ||(ts + (1 - t)r) - (ts_1 + (1 - t)r)|| = t||s - s_1|| \leq \)
Assume that eventually $X$ plays a Smale plan $\pi$, $Y$ uses an arbitrary plan and that the initial play is arbitrary. Let $\Omega$ be the limit point set of an associated sequence of outcomes. Let $C \subset S$ be a nonempty closed set.

(a) If $C$ is lower full and $S \setminus C$ is contained in the defection zone $\pi^{-1}(0)$, then $\Omega \subset C$.

(b) If $C$ is upper full and $S \setminus C$ is contained in the cooperation zone $\pi^{-1}(1)$, then $\Omega \subset C$.

Proof. (a) Assume $X$ plays $\pi_X$ after $N^*$. For $\epsilon > 0$ let $N_\epsilon > N^*$ be an $\epsilon$ step-time. If for all $N \geq N_\epsilon$, $s^N \in S \setminus C$, then after $N_\epsilon$ the outcomes are just as though $X$ played All D. Hence, by Proposition 2.9 (b), $\Omega \subset [(T, S), (P, P)] \subset C$. So we may assume that at some time $N_1 > N_\epsilon$, $s^{N_1} \in C$. We show by induction that for all $N \geq N_1$, $s^N \in C$.

If $s^N \in C$, then since $N > N_\epsilon$, $|s^{N+1} - s^N| < \epsilon$, and so $s^{N+1} \in C$.

If $s^N \in C \setminus C \subset \pi^{-1}(0)$, then $s^{N+1} \in [(T, S), (P, P)]$ and so $s^{N+1}$ is in $Q(s^N)$ and so is contained in the lower full set $C$.

Since, eventually, $s^N$ is in the closed set $C$, it follows that $\Omega \subset C$. As $\epsilon$ was arbitrary, $\Omega \subset \bigcap_{\epsilon > 0} C_\epsilon = C$.

The proof of (b) is completely analogous.

Definition 4.6. A separation path is a closed, connected subset $C \subset S$ which meets the segments $[(P, P), (S, T)]$ and $[(R, R), (T, S)]$ and which satisfies

(*) If $s \in C$ then $T(s) \cap Q(s) \cap C$ is disjoint from $C$.

Note that if $s \in [(S, T), (R, R)]$ then $T(s)$ has empty interior and so is disjoint from $C$.

Recall that when $P > \frac{1}{2}(T + S)$ the line $\{(S, T), (P, P)\}$ intersects $((R, R), (T, S))$ at a point we label $\bar{W}$. We let $W$ denote the point $(T, S)$ when $P \leq \frac{1}{2}(T + S)$.

Theorem 4.7. Assume that $C \subset S$ is a separation path with $s \in C$. Let $\text{proj} : C \to [S, T]$ be the restriction to $C$ of the first coordinate projection, i.e. $\text{proj}(s) = s_X$.

$$C^+ \cup C^- = S \quad \text{and} \quad C^+ \cap C^- = C. \quad (4.1)$$

$$S \setminus C^- = C^+ \setminus C = (C^+) \quad \text{and} \quad S \setminus C^+ = C^- \setminus C = (C^-). \quad (4.2)$$

$$T(s) \cap C^- = \emptyset \quad \text{and} \quad Q(s) \cap C^+ = \emptyset. \quad (4.3)$$

$$\text{proj}(\bar{W}) \cap C = \emptyset. \quad (4.4)$$

The map $\text{proj}$ is injective mapping $C$ onto an interval $[a, b]$ with $S \leq a \leq \bar{P}$ and $R \leq b \leq T$. 

Proof. Since \( C \) meets \([(S, T), (\bar{P}, \bar{P})]\) and \([(R, R), (T, S)]\), \( C \) meets \([V_-, V_+]\) if \( V_-, V_+ = (S, T), (\bar{P}, \bar{P}) \) or if \( V_-, V_+ = (R, R), (T, S) \). If \( V_- \in ((\bar{P}, \bar{P}), (T, S)) \) and \( V_+ \in ((S, T), (R, R)) \) then the segment \([V_-, V_+]\) separates the edges \([(S, T), (\bar{P}, \bar{P})]\) and \([(R, R), (T, S)]\) and so meets the connected set \( C \).

Let \( s_0 \in C \cap [V_-, V_+] \). The half-open segment \([(s_0, V_+)]\) is contained in \( T(s_0) \) and so \((s_0, V_+) \subset C^+ \setminus C\) by Condition (*). Similarly, \([V_-, s_0) \subset C^- \setminus C\). On the other hand, if \( s_1 \in Q(s) \cap [V_-, V_+] \), then \([V_-, s_1) \subset Q(s_1) \subset Q(s) \). So Condition (*) implies \( s_0 \notin [V_-, s_1) \). Thus, \( Q(s) \cap [V_-, V_+] \subset [V_-, s_0) \). Similarly, \( T(s) \cap [V_-, V_+] \subset [s_0, V_+] \). Since \( s \) was an arbitrary point of \( C \), including the possibility \( s = s_0 \), it follows that \( C^- \cap [V_-, V_+] = [V_-, s_0) \) and \( C^+ \cap [V_-, V_+] = [s_0, V_+] \). Thus, we have

\[
[V_-, V_+] \cap (\{ s \setminus C^- \} = [V_-, V_+] \cap (C^- \setminus C) = [V_-, V_+] \cap (C^+) = (s_0, V_+].
\]

with analogous equations for \([V_-, s_0)\). Because every point of \( S \) lies on some such interval \([V_-, V_+]\), equations (4.2) follow and clearly imply (4.1). Since \( T(s) \subset (C^+)^o \) and \( Q(s) \subset (C^-) \), (4.2) implies (4.3), as well.

If \( P \leq \frac{1}{7}(T+S) \) then \( W = (T, S) \) and \( Q(W) = ((\bar{P}, \bar{P}), (T, S)] \) has empty interior. \( P > \frac{1}{2}(T+S) \) then for all \( s \in Q(W)^o \), \( (P, P) \subset T(s)^o \). Since \( (P, P) \subset C^- \), (4.3) implies that \( s \notin C \), proving (4.4).

The image \( \text{proj}(C) \) is a compact connected subset of \([S, T] \subset \mathbb{R}\). Hence, it is a closed interval \([a, b]\). Since \( C \) meets \([(S, T), (\bar{P}, \bar{P})]\) and \([(R, R), (T, S)]\), \( \text{proj}(C) \) meets the closed intervals \([S, \bar{P}]\) and \([R, T]\) in \( \mathbb{R} \).

For \( t \in \mathbb{R} \) with \( S \leq t \leq T \) let \( \ell_t \) be the vertical line \( \{ s_X = t \} \). For \( t \in \text{proj}(C) \) with \( S < t < R \), the point \( V'_+ = \ell_t \cap \{ S, T \}, (R, R) \) lies in \((S, T), (R, R)\). Let \( s_t \) be the point of \( C \cap \ell_t \) with the smallest \( s_Y \) coordinate. Since \( (s_t, V'_+) \subset T(s)^o \) it follows from Condition (*) that \( C \cap \ell_t \) is the singleton \( \{ s_t \} \). Similarly, for \( t \in \text{proj}(C) \) with \( P < t < T \), \( \ell_t \) meets \((\bar{P}, \bar{P}), (T, S)\) at a point \( V'_- \). Let \( s'_t \) be the point of \( C \cap \ell_t \) with the largest \( s_Y \) coordinate. \([V'_-, s'_t] \subset Q(s^t)^o \) and so \( C \cap \ell_t \) is the singleton \( \{ s'_t \} \). Finally, \((S, T)\) is the only point of \( S \) with \( s_X = S \) and \((T, S)\) is the only point of \( S \) with \( s_X = T \). It follows that each vertical line meets \( C \) in at most one point. Thus, \( \text{proj} \) is injective.

If \( \ell \) is a separation line, then \((S, T)\) and \((R, R)\) are on or above \( \ell \) and so for \( s \in \ell \cap S \), \( T(s) \subset H^+ \) and \( T(s)^o \subset H^+ \setminus \ell \). Also, \((P, P)\) and \((T, S)\) are on or below \( \ell \). Since \((R, R)\) is on or above \( \ell \) it follows that \((\bar{P}, \bar{P})\) is also on or below \( \ell \). Thus, \( Q(s) \subset H^- \) and \( Q(s)^o \subset H^- \setminus \ell \). It follows that \( \ell \cap S \) is a separation path.

Call \( \ell \) a strict separation line when it is a separation line which does not contain \((S, T), (R, R), (T, S)\) or \((P, P)\). That is, for a strict separation line \( \ell \) the points \((S, T)\) and \((R, R)\) are strictly above \( \ell \) and the points \((T, S)\) and \((P, P)\) are strictly below \( \ell \). Furthermore, \((P, P)\) is strictly below \( \ell \) as well, because if \((P, P) \neq (P, P)\) then \( P > \frac{1}{2}(T+S) \) and the only separation line with \((P, P)\) on or above it is the diagonal which contains \((P, P)\). Thus, with \( L \) an affine map associated with \( \ell \), we have \( L(S, T), L(R, R) > 0 \) and \( L(T, S), L(P, P) \), \( L(\bar{P}, \bar{P}) < 0 \). It follows that:

\[
s \in \ell \cap S \implies (T(s) \cup Q(s)) \cap \ell = \{ s \}. \quad (4.5)
\]

By analogy, if \( C \) meets the segments \([(S, T), (\bar{P}, \bar{P})]\) and \([(R, R), (T, S)]\) and satisfies the strengthening of Condition (*)

\[
s \in C \implies (T(s) \cup Q(s)) \cap C = \{ s \}, \quad (4.6)
\]

then we will call \( C \) a strict separation path.
Recall that a function $\gamma : [a, b] \to \mathbb{R}$ is called piecewise $C^1$ when $\gamma$ is continuous and there is a finite sequence $a = a_0 < a_1 < \cdots < a_n = b$ so that $\gamma$ is continuously differentiable on each subinterval $[a_{i-1}, a_i]$ for $i = 1, \ldots, n$. Then at each point $(t, \gamma(t))$ with $t \in (a, b)$ there is a tangent line from the right and at each point with $t \in (a, b]$ there is a tangent line from the left. Except at the points with $t = a_i$ for $i = 0, \ldots, n$ these two agree are both are the true tangent line at the point.

**Theorem 4.8.** Assume that $C$ is a subset of $S$.

(a) If $C$ is a separation path then there exists a continuous function $\gamma : [a, b] \to [S, T]$ with $S \leq a \leq P, R \leq b \leq T$ such that $C$ is the graph of $\gamma$, i.e. $C = \{(t, \gamma(t)) : t \in [a, b]\}$.

(b) If $C$ is a separation path which is the graph of a piece-wise $C^1$ function $\gamma$, then every tangent line (i.e. every line tangent to a point from the left or the right) is a separation line.

(c) Assume that $C$ is the graph of a piece-wise $C^1$ function with $(a, \gamma(a)) \in [(P, P), (S, T)]$ and $(b, \gamma(b)) \in [(R, R), (T, S)]$. If every tangent line is a strict separation line, then $C$ is a strict separation path.

(d) If $C_1$ and $C_2$ are separation paths, then $C_1 \cap \text{Switch}(C_2)$ is non-empty. If, in addition, one of them is a strict separation path, then $C_1 \cap \text{Switch}(C_2)$ is a singleton set.

**Proof.** (a) As it is a continuous bijection on a compact set, $\text{proj} : C \to [a, b]$ is a homeomorphism. If $\gamma$ is the composition of $\text{proj}^{-1}$ with the projection to the $s_Y$ coordinate then $\gamma$ is a continuous map and $\text{proj}^{-1}$ is given by $t \mapsto (t, \gamma(t))$.

For (b) and (c) we use $s(t)$ for $(t, \gamma(t))$ with $t \in [a, b]$.

(b) If $t \in [a, b]$ then Condition (*) implies that for every $t_1 \in (t, b]$ the secant line $s(t), s(t_1)$ passes through $[(R, R), (T, S)]$. Furthermore, if $t < P$, then $(P, P)$ is on or below the secant line. Passing to the limit, the same is true for every tangent line from the right. Similarly, for $t \in (a, b]$ every tangent line from the left passes through $[(S, T), (P, P)]$ and if $t \geq P$ then $(P, P)$ is on or below the tangent line from the left.

Thus, at the points of $[a, b] \setminus \{a_0, \ldots, a_n\}$ the tangent line is a separation line. The remaining tangent lines from the left and right are limits of true tangent lines and so are separation lines as well.

(c) The graph is a closed, connected set which meets $[(S, T), (P, P)]$ and $[(R, R), (T, S)]$ by hypothesis. If $s \in Q(W)^\circ$ then no separation line passes through $s$ and so the graph is disjoint from $Q(W)^\circ$.

Let $a \leq t_0 < b$. Then (4.5) for the tangent line from the right implies that, for sufficiently small $h > 0$, $s(t_0+h) \notin T(s(t_0)) \cup Q((s_0))$. If $s(t) \in T(s(t_0)) \cup Q((s_0))$ for some $t > t_0$ then we can let $t^*$ be the first entrance time, i.e. $t^* = \inf\{t > t_0 : s(t) \in T(s(t_0)) \cup Q((s_0))\}$ and let $\ell$ be the tangent line from the left at $s(t^*)$. Thus, $s(t^*)$ is in the $S$ topological boundary of $T(s(t_0)) \cup Q((s_0))$. The portion of the boundary with $x = s_X(t_0)$ consists of $(s(t_0), (R, R)]$ and either $(s(t_0), (T, S)]$ or if $(P, P)$ lies above this segment then $(s(t_0), (P, P)) \cup ((P, P), (T, S)]$ (since $s(t_0) \notin Q(W)^\circ$). If $s(t^*) \in (s(t_0), (R, R)]$ then because $s(t_1)$ is below $(s(t_0), (R, R)]$ for $t_0 < t_1 < t^*$ the point $(R, R)$ lies below the secant line $s(t_1), s(t^*)$. It follows that, in the limit, $(R, R)$ is on or below $\ell$. Similarly, if $s(t^*) \in (s(t_0), (T, S)]$ then $(T, S)$ is on or above $\ell$. Finally, if $s(t^*) \in ((P, P), (T, S)]$ then $(P, P)$ is on or above $\ell$. 
Let \( a < t_0 \leq b \). Then, for sufficiently small \( h > 0 \), \( s(t_0 - h) \notin T(s(t_0)) \cup Q((s_0)) \). If \( s(t) \in T(s(t_0)) \cup Q((s_0)) \) for some \( t < t_0 \) then we can let \( t^* \) be the first entrance time moving left, i.e. \( t^* = \sup \{ t < t_0 : s(t) \in T(s(t_0)) \cup Q((s_0)) \} \) and let \( \ell \) be the tangent line from the right at \( s(t^*) \). The portion of the boundary with \( s_X < s_X(t_0) \) consists of \( (s(t_0), (S, T)) \) and either \( (s(t_0), (P, P)) \) or if \( (P, P) \) lies above this segment then \( (s(t_0), (P, P)) \cup ((P, P), (P, P)) \) (since \( s(t_0) \notin Q(W)^s \)). As before \( s(t^*) \in (s(t_0), (S, T)) \) implies \( (S, T) \) is on or below \( \ell \), and \( s(t^*) \in ((P, P), (P, P)) \) implies \( (P, P) \) is on or above \( \ell \). Finally, if \( s(t^*) \in (s(t_0), (P, P)) \) then \( (P, P) \) is on or above \( \ell \).

In none of these cases can \( \ell \) be a strict separation line. This proves (4.6) for any \( t = t_1 \). This in turn implies implies \( C \) is a separation path.

(d) The connected set \( \text{Switch}(C_2) \) meets \( ([S, T], (R, R)] \subset C_1^+ \) and \( [(P, P), (T, S)] \subset C_1^- \). It follows from (4.1) for \( C_1 \) that \( \text{Switch}(C_2) \) meets \( C_1 \).

For any \( s \in S \), \( (T(s) \cup Q(s)) \cup \text{Switch}((T(s)^o \cup Q(s)^o) = S \). Assume that \( C_1 \) is a strict separation path and that \( s_1 \in S \setminus \{s\} \). If \( s_1 \in C_1 \) then \( s_1 \notin (T(s) \cup Q(s)) \). If \( s_1 \in \text{Switch}(C_2) \) then \( s_1 \notin \text{Switch}((T(s)^o \cup Q(s)^o) \). It follows that \( s_1 \notin C_1 \cap \text{Switch}(C_2) \).

**Remark.** It is possible to extend (c) somewhat. It is easy to check that the set of separation paths is itself closed in the space of closed subsets of \( S \) equipped with the Hausdorff topology. So if \( \{\gamma_n\} \) is a sequence of piecewise \( C^1 \) functions which satisfy the conditions of (c) and the sequence of graphs converges in the appropriate sense to the graph of \( \gamma \) then the graph of \( \gamma \) is a separation path.

Using simple geometric arguments similar to those in (c) above, we can describe what occurs in a non-strict separation path. Since we will not need the results, we leave the proof to the interested reader.

**Proposition 4.9.** Assume that \( s_1 \) and \( s_2 \) are distinct points of a separation path \( C \).

(i) Let \( V_-, V_+ \) be one of the pairs \((S, T), (R, R) \) or \((P, P), (T, S) \) or \((P, P), (P, P) \).

If \( C \) meets the open segment \( [V_-, V_+] \), then it contains the closed segment \( [V_-, V_+] \).

(ii) Let \( V \) be one of the points \((S, T), (R, R), (T, S) \), \( (P, P) \) or \((P, P) \). If \( s_2 \) lies in

the open segment \( (s_1, V) \) then \( C \) contains the closed segment \( [s_1, s_2] \).

**Example 4.10.** A differential equations construction for strict separation paths.

For the example, we will restrict to the case \( P < \frac{1}{2}(T+S) \) and we will just sketch the argument, leaving the details to the reader.

Define for \( s \in S \)

\[
\begin{align*}
m^+(s) &= \begin{cases} 
(R - sy)/ (R - sx) & \text{if } sy \geq sx \\
(sy - P)/(sx - P) & \text{if } sy \leq sx.
\end{cases} \\
m^-(s) &= \begin{cases} 
(T - sy)/(sx - s) & \text{if } sy + sx \geq T + S \\
(sy - s)/(T - sx) & \text{if } sy + sx \leq T + S.
\end{cases}
\end{align*}
\]

(4.7)

So \( m^+(s) = 1 \) if \( sy = sx \) and \( m^-(s) = 1 \) if \( sy + sx = T + S \). If \( s \in [(S, T), (R, R)] \) or \( s \in [(P, P), (T, S)] \) then \( m^+(s) = -m^-(s) \). Otherwise, \( m^+(s) > -m^-(s) \). A line \( \ell \) through \( s \) with slope \( m \) is a separation line iff \( m^+(s) \geq m \geq -m^-(s) \). So \( m^+(s) \geq m \geq -m^-(s) \) if \( m \) is a strict separation line iff both inequalities are strict.
Now let \( m \) be a real-valued, differentiable function on \( S \) with \( m(s) = m^+(s) = -m^-(s) \) on \( [(S,T), (R,R)] \cup [(P,P), (T,S)] \) and with \( m^+(s) > m(s) > -m^-(s) \) otherwise.

Consider the differential equation \( \frac{dy}{dx} = m(x,y) \) defined for \( s = (x,y) \in S \). Observe that \( [(S,T), (R,R)] \) and \( [(P,P), (T,S)] \) are both solution curves for the differential equation. So by the uniqueness theorem for ode’s every other solution curve remains in \( S \) \( [(S,T), (R,R)] \cup [(P,P), (T,S)] \) and so by Theorem 4.8 (c) each of the remaining solution curves is a strict separation path.

Since the \( S \) intersection with a separation line is a separation path, the following extends Corollary 2.8.

**Corollary 4.11.** Assume that eventually \( X \) plays a Smale plan \( \pi \), \( Y \) uses an arbitrary plan and that the initial play is arbitrary. Let \( \Omega \) be the limit point set of an associated sequence of outcomes. If \( C \) is a separation path such that \( C^+ \backslash C \subset \pi^{-1}(0) \) and \( C^- \backslash C \subset \pi^{-1}(1) \), then \( \Omega \subset C \).

**Proof.** \( C^+ \) is upper full with \( S \backslash C^+ = C^- \backslash C \) contained in the cooperation zone and \( C^- \) is lower full with \( S \backslash C^- = C^+ \backslash C \) contained in the defection zone. It follows from Proposition 4.5 that \( \Omega \subset C^+ \cap C^- = C \).

**Proposition 4.12.** If \( s^* = (s^*_X, s^*_Y) \in S \) with \( P < s^*_Y < R \) then there is a strict separation path \( C \) such that \( s^* \) is the unique point \( s \in C \) with \( s_Y \geq s^*_Y \).

**Proof.** First, assume \( s^* \not\in [(S,T), (P,\bar{P})] \cup [(R,R), (T,S)] \).

Choose \( V \) a point on the open segment \( ((S,T), (P,\bar{P})) \) with \( P < V_Y < s^*_Y \) and so that \( )s^*( \) intersects \( ((R,R), (T,S)) \). Choose \( W \) a point on the open segment \( ((R,R), (T,S)) \) with \( P < W_Y < s^*_Y \) and so that \( )s^*( \) intersects \( ((S,T), (P,\bar{P})) \). The lines \( )s^*( \) and \( )s^*, W( \) are strict separation lines.

Let \( C = [V, s^*] \cup [s^*, W] \). While it is easy to check directly that \( C \) satisfies (4.6), it follows from Theorem 4.8(c) that \( C \) is a strict separation path. Clearly, \( s^* \) is the unique point of \( C \) with maximum height.

If \( s^* \in [(S,T), (P,\bar{P})] \) then choose \( W \) as above so that \( \ell = )s^*, W( \) is a strict separation line. If \( s^* \in [(R,R), (T,S)] \) then choose \( V \) as above so that \( \ell = )s^*, V( \) is a strict separation line. In either case, \( s^* \) is the point of maximum height on the segment \( \ell \cap S \).

**Theorem 4.13.** If \( s^* = (s^*_X, s^*_Y) \in S \) with \( P < s^*_X, s^*_Y < R \) then there is a pair of Smale plans \( \pi_X, \pi_Y \) which is a strong Nash equilibrium with \( s^* \) the payoff to the pair.

**Proof.** \textbf{Case 1.} \( (P < s^*_X, s^*_Y < R) \): Apply Proposition 4.12 to choose \( C_X \) a strict separation path with maximum point \( s^* \) and \( C_Y \) be a strict separation path with maximum point \( \text{Switch}(s^*) = (s^*_Y, s^*_X) \). Let \( \pi_X \) and \( \pi_Y \) be Smale plans with

\[
\begin{align*}
C^+_X \backslash C_X &\subset (\pi_X)^{-1}(0), & C^-_X \backslash C_X &\subset (\pi_X)^{-1}(1), \\
C^+_Y \backslash C_Y &\subset (\pi_Y)^{-1}(0), & C^-_Y \backslash C_Y &\subset (\pi_Y)^{-1}(1).
\end{align*}
\]

(4.8)

Corollary 4.11 and Theorem 4.8 (d) imply that when \( \pi_X \) plays \( \pi_Y \) the payoff is \( s^* = C_X \cap \text{Switch}(C_Y) \). If \( X \) uses an alternative plan, then the limit set is contained in \( \text{Switch}(C_Y) \), and similarly if \( Y \) varies against \( \pi_X \).

\textbf{Case 2.} \( (s^*_X = s^*_Y = R) \): This is the good plan case studied in Section 3. Use \( \pi_X = \pi_Y \) a simple Smale plan with the separation line through \( (R,R) \) with slope \( m \) satisfying \( 0 < m < 1 \).
Figure 4. Theorem 4.10

Case 3. \((P < s^*_X < R \leq s^*_Y, \text{ or } P < s^*_Y < R \leq s^*_X)\): Notice that if \(s \in S\) with \(s_Y > R\), then \(s_X < R\). Furthermore, for the case \(P < s^*_Y < R \leq s^*_X\) it suffices to apply \(\text{Switch}\) to the other case. So we will assume that \(P < s^*_X < R \leq s^*_Y\). To begin with we will also assume that \(s^* \notin [(R, R), (S, T)]\).

First we use Proposition 4.12, or more precisely its proof, to choose a separation path with maximum point \(\text{Switch}(s^*) = (s^*_Y, s^*_X)\). That is, \(C_Y = \text{Switch}(V), \text{Switch}(s^*)]\) or \(\text{Switch}(s^*), \text{Switch}(W)\) with \(V \in ((P, P), (T, S)), W \in ((R, R), (S, T))\), with \(P < V_X, W_X < s^*_X\) and with the other conditions as described above. Let \(\ell = V, s^*\). Let \(\pi_Y = 0\) on \(C_Y \setminus C_Y \) and \(1\) on \(C_Y \setminus C_Y \). So regardless of the \(X\) plan, the Switched limit set \(\text{Switch}(\Omega)\) is contained in \(C_Y\) which has \(\text{Switch}(s^*)\) as it unique point of maximum height.

Next choose \(W' \in ((R, R), (T, S))\) and \(V' \in ((S, T), (P, P))\) with \(P < W'_X = V'_X < R\). So the horizontal line \(\ell' = V', W'(\) is a separation line. Let \(C_X = [V', s^*] \cup [s^*, W']\). Let \(\pi_X = 0\) above \(C_X\) and \(1\) below \(C_X\). Now despite the labeling, \(C_X\) is not a separation path. The lower full set \(\overline{C_X}\) is the set of points on or below each of the lines \(V', s^*\) and \(s^*, W'(\) but the set of points on or above \(C_X\) is not upper full. For every point \(r \in [V', s^*], (r, s^*) \subset T(r)^*\). On the other hand, Corollary 2.8 (c) applies to \(\pi_X\) with \(C \subset C_X\) equal to the triangle \([V', s^*, W']\). So regardless of the \(Y\) plan, the limit set \(\Omega\) is contained in \(C\) which has \(s^*\) as its unique point of maximum height.
To complete the proof we must show that if, after some $N^*$, $X$ uses $\pi_X$ and $Y$ uses $\pi_Y$ then the solution sequence converges to $s^*$, i.e. $\Omega = \{s^*\}$. We know that $\Omega$ is contained in $C \cap \text{Switch}(C_Y)$. If $s = \ell \cap \ell'$, then $C \cap \text{Switch}(C_Y)$ is the segment $K = [s^*, \bar{s}]$. Thus, given an arbitrary $\delta > 0$ there is an $N^\delta > N^*$ after which that the sequence remains $\delta$ close to $K$.

Let $\delta > 0$ be arbitrary. Within the $\delta$ ball $V_\delta(s^*)$ we will build a box which contains $s^*$ in its interior. Choose $A \in (V', s^*) \cap V_\delta(s^*)$ close enough to $s^*$ that the line $\ell_1 = (S, T), A(\ell)$ crosses $\ell$ within $V_\delta(s^*)$, i.e. $s' = \ell_1 \cap \ell \in V_\delta(s^*)$. Choose $B \in \ell_1 \cap V_\delta(s^*)$ so that $s'' = (A, B) \subset V_\delta(s^*)$. We may choose $A$ close enough that the horizontal segment from $B$ to $(s^*, V')$ is also in $V_\delta(s^*)$. Now let $\ell_2 = (P, P), B(\ell$ and $\ell_3 = (P, P), A(\ell$. Because $B(P, P)$ lies below $)V', s'(, and above $\ell = V, s'(,$ it follows that $s^*$ is below $\ell_3$ and above $\ell_2$. We can choose $A$ close enough to $s^*$ that $B' = \ell_2 \cap (s^*, W')$ and $A' = \ell_1 \cap (s^*, W)$ lie in $V_\delta(s^*)$. This is where we use $s^* \notin [(R, R), (S, T)], \ell_3 = \ell_1 \cap (s^*, W)$, $A' = (A, A'')$ lies to the right of $\ell_3$ and $A'' = (A, A'')$. The region of $\Omega$ above both $s^*, W'(\ell, s^*), W((), is convex, as is $V_\delta(s^*)$. It follows that the quadrilateral $\text{Box} = [A, B, B''(, A''],$ lies in $V_\delta(s^*)$ and contains $s^*$ in its interior.

The various lines cut $\delta$ into a -finite- number of regions and we choose $\epsilon < \delta$ to be smaller than the distances between any two disjoint such regions and small enough that the open neighborhood $V_\epsilon(V'(K))$ is contained in the union of $\text{Box}$ and $C^*$, the convex set of points of $S$ in $C$ or below $\ell$, and finally, small enough that $V_\epsilon(\text{Box})$ is contained in $V_\delta(s^*)$.

Let $N_\epsilon > N^*$ be an $\epsilon$ step-time so that for every $N \geq N_\epsilon$ the length $||s^{N+1} - s^*||$ is less than $\epsilon$. Hence, no such small move crosses between disjoint regions. Since $\Omega \subset K$ we can also choose $N_\epsilon$ so that for every $N \geq N_\epsilon, s^* \in V_\epsilon(K)$.

Suppose that for some $N \geq N_\epsilon, s^N \notin \text{Box}$. If $s^N$ lies to the left of $\ell$ then $s^{N+1} = (R, R)$ and if $s^N$ lies to the right of $\ell$ then $s^{N+1} = (S, T)$. For $s^N \in \ell$ either $s^{N+1} = (R, R)$ or $s^{N+1} = (S, T)$. For any such $s$ the (negative) slope of $\ell(S, T), s(\ell$ is less than the slope of $\ell(R, R), s(\ell$. Thus from the left we move toward $(R, R)$ and after crossing we move toward $(S, T)$. The successive crossings are higher on $K$ and thus our net motion is upward until we enter $\text{Box}$. From $\text{Box}$ it is easy to see that exit could only occur from the triangle $[s', B, s^*]$ landing above the horizontal line through $B$ and $\epsilon$ close to $\text{Box}$ and so is still in $V_\delta(s^*)$. If it lands to the left of $\ell$ then the motion toward $(R, R)$ is closer to $\text{Box}$ and so $s^N, s^{N+1} \subset V_\delta(s^*)$. If it lands to the right of $\ell$ then the motion toward $(S, T)$ is upward, remaining above the $B$ horizontal line and so $[s^N, s^{N+1}] \subset V_\delta(s^*)$. Subsequent alternating moves are above these initial ones and so remain in $V_\delta(s^*)$ until the sequence re-enters $\text{Box}$. It follows that eventually the sequence lies in $V_\delta(s^*)$ and so the limit point set $\Omega$ is contained in the closed ball of radius $\delta$ about $s^*$. As $\delta > 0$ was arbitrary, it follows that $\Omega = \{s^*\}$ as required.

Finally, we adjust the proof to deal with the case $s^* \in [(S, T), (R, R)]$. If $s^* = (R, R)$ we are in Case 2. Since $s^*_X > P$, we have $s^* \notin (S, T)$. We use $C_X$ as before. Now $C_Y$ is the separation line $\ell = s^*, V$ with $W$ undefined. The points $A, B, B'$ and the lines $\ell_1, \ell_2, \ell_3$ are chosen as before. This time, $B'' = \ell_2 \cap (S, T), (R, R)(A'' = \ell_1 \cap (S, T), (R, R). Now $s^*$ is in the upper edge $(A'', B'')$ which is nonetheless in the $\delta$ interior of $\text{Box} = [A, B, B'', A'']$. With an easy adjustment of the previous argument, one can again show that $\Omega$ is contained in the closed ball of radius $\delta$ about $s^*$ and so conclude that $\Omega = \{s^*\}$. □
5. **Competition among simple Smale plans.** In this section we move beyond the classical question which motivated our original interest in good strategies. We consider now the evolutionary dynamics among simple Smale plans. We follow Hofbauer and Sigmund [11] Chapter 9 and Akin [2].

The dynamics that we consider takes place in the context of a symmetric two-person game, but generalizing our initial description, we merely assume that there is a set of strategies indexed by a finite set $I$. When players $X$ and $Y$ use strategies with index $i,j \in I$, respectively, then the payoff to player $X$ is given by $A_{ij}$ and the payoff to $Y$ is $A_{ji}$. Thus, the game is described by the payoff matrix $\{A_{ij}\}$. We imagine a population of players each using a particular strategy for each encounter and let $\xi_i$ denote the ratio of the number of $i$ players to the total population.

The frequency vector $\{\xi_i\}$ lives in the unit simplex $\Delta \subset \mathbb{R}^I$, i.e. the entries are nonnegative and sum to 1. The vertex $v(i)$ associated with $i \in I$ corresponds to a population consisting entirely of $i$ players. Thus, $\xi = v(i)$ exactly when $\xi_i = 1$. We assume the population is large so that we can regard $\xi$ as changing continuously in time.

Now we regard the payoff in units of fitness. That is, when an $i$ player meets a $j$ player in an interval of time $dt$, the payoff $A_{ij}$ is an addition to the background reproductive rate $\rho$ of the members of the population. So the $i$ player is replaced by $1 + (\rho + A_{ij}) dt$ players. Averaging over the current population distribution, the expected relative reproductive rate for the subpopulation of $i$ players is $\rho + A_i \xi$, where

\begin{align}
A_{i\xi} &= \Sigma_{j \in I} \xi_j A_{ij} \quad \text{and} \\
A_{\xi\xi} &= \Sigma_{i \in I} \xi_i A_{i\xi} = \Sigma_{i,j \in I} \xi_i \xi_j A_{ij}.
\end{align}

(5.1)

The resulting dynamical system on $\Delta$ is given by the *Taylor-Jonker Game Dynamics Equations* introduced in Taylor and Jonker [21].

\begin{align}
\frac{d\xi_i}{dt} &= \xi_i (A_{i\xi} - A_{\xi\xi}).
\end{align}

(5.2)

This system is an example of the *replicator equation* systems studied in great detail in Hofbauer and Sigmund [11].

We will need some general game dynamic results for later application. Fix the game matrix $\{A_{ij}\}$.

A subset $A$ of $\Delta$ is called *invariant* if $\xi(0) \in A$ implies that the entire solution path lies in $A$. That is, $\xi(t) \in A$ for all $t \in \mathbb{R}$. An invariant point is an *equilibrium*.

Each nonempty subset $J$ of $I$ determines the *face* $\Delta_J$ of the simplex consisting of those $\xi \in \Delta$ such that $\xi_i = 0$ for all $i \notin J$. Each face of the simplex is invariant because $\xi_i = 0$ implies that $\frac{d\xi_i}{dt} = 0$. In particular, for each $i \in J$ the vertex $v(i)$, which represents fixation at the $i$ strategy, is an equilibrium.

In general, $\xi$ is an equilibrium when, for all $i,j \in J$, $\xi_i,\xi_j > 0$ imply $A_{i\xi} = A_{j\xi}$, or, equivalently, $A_{i\xi} = A_{\xi\xi}$ for all $i$ such that $\xi_i > 0$, i.e. for all $i$ in the *support* of $\xi$.

An important example of an invariant set is the *omega limit point set of an orbit*. Given an initial point $\xi \in \Delta$ with associated solution path $\xi(t)$, it is defined by intersecting the closures of the tail values.

\begin{align}
\omega(\xi) &= \bigcap_{t > 0} \{\xi(s) : s \geq t\}.
\end{align}

(5.3)
By compactness this set is nonempty. A point is in \( \omega(\xi) \) iff it is the limit of some sequence \( \{\xi(t_n)\} \) with \( \{t_n\} \) tending to infinity. The set \( \omega(\xi) \) consists of a single point \( \xi^* \) iff \( \lim_{t \to \infty} \xi(t) = \xi^* \). In that case, \( \{\xi^*\} \) is an invariant point, i.e. an equilibrium.

Notice that this is the analogue for the solution path of the limit point set \( \Omega \) of a payoff sequence, considered in the previous sections.

**Definition 5.1.** We call a strategy \( i^* \) a *evolutionarily stable strategy* (hereafter, an ESS) when

\[
A_{ji^*} < A_{i^*i^*} \quad \text{for all } j \neq i^* \text{ in } I. \tag{5.4}
\]

**Remark.** We follow [5] in labeling this condition ESS although it is stronger than the condition originally introduced in [13]. In [11] precisely this condition is called a *strict Nash equilibrium* and so this language requires a bit of justification.

In the pure game theory context we regard a distribution \( \xi \) over \( I \) as a mixed-strategy rather than a population distribution of pure strategists. Then a pair \( \xi_1, \xi_2 \) is a *Nash equilibrium* when each is a best reply against the other. That is, for all distributions \( \eta \),

\[
A_{\eta \xi_2} \leq A_{\eta \xi_1} \quad \text{and} \quad A_{\eta \xi_1} \leq A_{\eta \xi_2}. \tag{5.5}
\]

Notice that from this we see that

\[
(\xi_1)_j > 0 \quad \Rightarrow \quad A_{j\xi_2} = A_{\xi_1\xi_2} \quad \text{and} \quad A_{j\xi_1} = A_{\xi_2\xi_1}. \tag{5.6}
\]

That is, all of the pure strategies active in \( \xi_1 \) are best replies to \( \xi_2 \) and vice-versa. In the context of Smale strategies this is the concept used in the previous section.

Following [2] we call the pair a *regular Nash equilibrium* when, in addition to (5.5),

\[
(\xi_1)_j = 0 \quad \Rightarrow \quad A_{j\xi_2} < A_{\xi_1\xi_2} \quad \text{and} \quad A_{j\xi_1} < A_{\xi_2\xi_1}. \tag{5.7}
\]

That is, the pure strategies active in \( \xi_1 \) are all of the pure strategies which give the best reply to \( \xi_2 \) and vice-versa.

Returning to the dynamic context, a distribution \( \xi \) is called a (regular) Nash equilibrium when the pair \( \xi, \xi \) is a (regular) Nash equilibrium in the above sense. From (5.6) we see that a Nash equilibrium is an equilibrium for the Taylor-Jonker equations as defined above. When \( \xi \) is the vertex \( v(i^*) \) then it is a regular Nash equilibrium exactly when it is a strict Nash equilibrium as defined on page 62 of [11], or, equivalently, (5.4) holds.

**Proposition 5.2.** If \( i^* \) is an ESS then the vertex \( v(i^*) \) is an attractor, i.e. a locally stable equilibrium, for the system (5.2). In fact, there exists \( \epsilon > 0 \) such that

\[
1 > \xi_{i^*} \geq 1 - \epsilon \quad \Rightarrow \quad \frac{d\xi_{i^*}}{dt} > 0. \tag{5.8}
\]

Thus, near the equilibrium \( v(i^*) \), which is characterized by \( \xi_{i^*} = 1 \), \( \xi_{i^*}(t) \) increases monotonically, converging to 1 and the alternative strategies are eliminated from the population in the limit.

**Proof.** When \( i^* \) is an ESS, \( A_{ji^*} < A_{i^*i^*} \) for all \( j \neq i^* \). It then follows for \( \epsilon > 0 \) sufficiently small that \( \xi_{i^*} \geq 1 - \epsilon \) implies \( A_{i^*i^*} > A_{j\xi} \) for all \( j \neq i^* \). If also \( 1 > \xi_{i^*} \), then \( A_{i^*i^*} > A_{i^*\xi} \). So (5.2) implies (5.8). \( \square \)
Definition 5.3. For $\mathcal{J}$ a nonempty subset of $\mathcal{I}$ we say a strategy $i$ weakly dominates a strategy $j$ in $\mathcal{J}$ when $i, j \in \mathcal{J}$ and
\[ A_{jk} \leq A_{ik} \quad \text{for all} \quad k \in \mathcal{J}, \quad (5.9) \]
with strict inequality for $k = i$ or $k = j$. If the inequalities are strict for all $k$ then we say that $i$ dominates $j$ in $\mathcal{J}$.

We say that $i \in \mathcal{J}$ weakly dominates a sequence $\{j_1, ..., j_n\}$ in $\mathcal{J}$ when there exists $1 \leq m \leq n$ such that $i$ weakly dominates $j_p$ in $\mathcal{J}$ for $p = 1, \ldots, m$ and for $p = m + 1, \ldots, n$, $i$ dominates $j_p$ in $\mathcal{J} \setminus \{j_1, ..., j_{p-1}\}$. When $\mathcal{J}$ equals all of $\mathcal{I}$ we will omit the phrase “in $\mathcal{J}$”.

Proposition 5.4. For $i \in \mathcal{I}$, let $\xi(t)$ be a solution path with $\xi_i(0) > 0$
(a) If $i$ weakly dominates $j$ then $\lim_{t \to \infty} \xi_j(t) = 0$.
(b) If $i$ weakly dominates the sequence $\{j_1, ..., j_n\}$ then for $j = j_1, ..., j_n$, $\lim_{t \to \infty} \xi_j(t) = 0$.

Proof. (a): The face $\{\xi : \xi_j = 0\}$ is invariant. So if $\xi_j(0) = 0$ then $\xi_j(t) = 0$ for all $t$ and so it is 0 in the limit. Thus, we may assume $\xi_j(0) > 0$.

For $i, j \in \mathcal{J}$, define the open set $Q_{ij}$ and on it the real valued function $H_{ij}$ by
\[
Q_{ij} = \{\xi \in \Delta : \xi_i, \xi_j > 0\} \quad H_{ij}(\xi) = \ln(\xi_i) - \ln(\xi_j) . \quad (5.10)
\]
Let $h_0 = H_{ij}(\xi(0))$.

Observe that on $Q_{ij}$
\[ dH_{ij}/dt = A_{ik} - A_{jk} = \sum_{k \in J} \xi_k (A_{ik} - A_{jk}) > 0 \quad (5.11) \]
Hence, $H_{ij}(\xi(t))$ is a strictly increasing function of $t$ on the open invariant set $Q_{ij}$. Thus, as $t$ tends to infinity $H_{ij}(\xi(t))$ approaches $h_\infty = \sup\{H_{ij}(\xi(t)) : t \geq 0\}$ with $h_0 < h_\infty \leq +\infty$.

We must prove that $\xi_j = 0$ on the omega limit set. Assume instead that $\xi^* \in \omega(\xi(0))$ with $\xi_j^* > 0$. If $\xi_j^*$ were 0 then $H_{ij}(\xi(t))$ would not be bounded below on $\{\xi(t) : t \geq 0\}$. Hence, $\xi^*$ lies in $Q_{ij}$ with $h_\infty = H_{ij}(\xi^*) < \infty$. So on the invariant set $\omega(\xi(0)) \cap Q_{ij}$, which contains $\xi^*$, and so is nonempty, $H_{ij}$ would be constantly $h_\infty < \infty$. Since this set is invariant, $dH_{ij}/dt$ would equal zero. This contradicts (5.11) which implies that the derivative is positive on $\omega(\xi(0)) \cap Q_{ij}$.

The proof of (b) is a variation of the proof of (a). We refer to [5] Proposition 4.6.

An obvious adjustment of the initial step in the inductive proof of (b) there yields the proof here. \[ \square \]

Corollary 5.5. Assume $I = \{i^*, j_1, ..., j_n\}$ and $i^* \in I$ weakly dominates the sequence $\{j_1, ..., j_n\}$. If $\xi_\infty(0) > 0$ then $\lim_{t \to \infty} \xi_{i^*}(t) = 1$.

Proof. By Proposition 5.4 $\xi_{j_p}(t) \to 0$ for all $p = 1, \ldots, n$ and so $\xi_{i^*}(t) = 1 - \sum_{p=1}^n \xi_{j_p}(t) \to 1$.

\[ \square \]

In [5], see also [3], we examined competition among certain special Markov plans called Zero-Determinant Plans. It was proved that good Markov plans among them are attractors when competing against plans which are not agreeable. In addition, global stability was proved when the class of competitors was further restricted.

Here we will similarly consider competition among simple Smale plans. Let $I = \{i^*, j_1, ..., j_n\}$ index a list of simple Smale plans with $\pi_i$ associated with separation line $\ell_i$ for $i \in I$. Except for the extreme cases the intersection $\ell_i \cap \text{Switch}(\ell_j)$ is
a single point. For \( \pi \) the plan with \( \ell \) the diagonal we will assume that the plan is weakly agreeable and that the initial play is \( c \). So if \( X \) and \( \tilde{Y} \) both play \( \pi \) the payoff is \((R,R)\). If \( P \leq \frac{1}{2}(T+S) \) then the co-diagonal is a separation line and we will adopt the convention that if both players use co-diagonal plans then the payoff is \( \frac{1}{2}(T+S,T+S) \). If \( X \) plays \( \pi_i \) and \( \tilde{Y} \) plays \( \pi_j \) we will let \((A_{ij}, A_{ji})\) be the coordinates of the payoff point with the above conventions in the extreme cases. We will then use (5.2) to represent the dynamics of the competition with \( \xi_i \) the fraction of the \( \pi_i \) players in the population.

Notice that if \( \tilde{Y} \) plays an equalizer plan \( \pi_j \) with \( \ell_j \) horizontal then \( A_{ij} = A_{i,j} \) for any plans \( \pi_i \) and \( \pi_{i'} \) for \( X \). In particular, if all of the plans are equalizer plans then \( A_{i\xi} = A_{i\xi} \) for all \( i, i' \in I \) and so \( A_{i\xi} = A_{\xi\xi} \) for all \( i \in I \) and for any population state \( \xi \). Thus, the dynamics is trivial and every state \( \xi \) is an equilibrium.

Now we consider the case when \( \ell_\ell \) is a protection line. That is, \( \ell_\ell \) is a line through \((R,R)\) with slope \( m \) satisfying \( 0 < m \leq 1 \). Notice that \( m = 1 \) is the diagonal line case. If \( m < 1 \) and \( \pi_\ell (R,R) = 1 \) then \( \pi_\ell \) is a good simple Smale plan.

**Theorem 5.6.** If \( \ell_\ell \) is a protection line and \((R,R) \not\in \ell_j \) for any \( j \in I \setminus \{i^*\} \) then \( i^* \) is an ESS and so fixation at \( i^* \) is an attractor.

**Proof.** \( A_{i^*,\ell} = R \). If \( \tilde{Y} \) plays \( \pi_{i^*} \) and \( X \) plays \( \pi_j \) for \( j \in I \setminus \{i^*\} \) then the payoff point is not \((R,R)\) and so \( A_{ji^*} \) and \( A_{i^*j} \) are both less than \( R \) because \( \ell_\ell \) is a protection line. This implies (5.4) and so the result follows from Proposition 5.2.

Since \( \xi_{\ell}(0) = 0 \) implies \( \xi_{\ell}(t) = 0 \) for all \( t \) the best stability result we can hope for is that every solution with \( \xi_{\ell}(0) > 0 \) converges to fixation at \( i^* \). We will call this **global stability**.

As an illustration we describe a very special case.

**Theorem 5.7.** Assume that \( \ell_{i^*} \) is a protection line. If for every \( j \in I \setminus \{i^*\} \), \( \ell_j \) is a horizontal line \( \{s_j = C_j\} \) with \( P \leq C_j < R \), then for every \( j \in I \setminus \{i^*\} \), \( i^* \) weakly dominates \( j \) and so the system exhibits **global stability**.

**Proof.** Because \( \ell_{i^*} \) is a protection line, we have, as in Theorem 5.6, \( A_{jj,i^*} < A_{i^*,i^*} = R \). For any \( k \in I \setminus \{i^*\} \), \( C_k = A_{jk} = A_{i^*k} \) for all \( j \in I \) and weak domination, (5.9), follows. A fortiori, \( i^* \) weakly dominates the sequence \( \{j_1, \ldots, j_n\} \) and the result follows from Corollary 5.5.

We will show that we achieve global stability if \( \pi_{i^*} \) is good, \((R,R) \not\in \ell_j \) for any \( j \in I \setminus \{i^*\} \) and, in addition, all the lines \( \ell_j \) have positive slope. This requires a bit of geometry.

**Lemma 5.8.** Assume that \( \ell_{i^*} \) is a protection line, \((R,R) \not\in \ell_k \) for any \( k \in I \setminus \{i^*\} \) and that \( \ell_k \) has non-negative slope for every \( k \in I \). If for some \( i \in I \) the segment \( \ell_i \cap S \) lies below \( \ell_{i^*} \) then \( i^* \) weakly dominates \( i \).

**Proof.** As usual \( A_{ii^*} < A_{i^*,i^*} = R \). For any \( k \in I \) the line \( \text{Switch}(\ell_k) \) is either vertical or has positive slope. Let \( \tilde{V}, V^* \) be the intersection points of \( \text{Switch}(\ell_k) \cap \ell_i \) and \( \text{Switch}(\ell_k) \cap \ell_{i^*} \), respectively. If \( \text{Switch}(\ell_k) \) is vertical then the \( X \) coordinates of \( \tilde{V} \) and \( V^* \) are equal. If \( \text{Switch}(\ell_k) \) has positive slope then \( V^* \) is above and to the right of \( \tilde{V} \) and so has a larger \( X \) coordinate. Thus, \( A_{ik} \leq A_{i^*k} \), proving weak domination.
Now we assume that the slope of $\ell_i < 1$, i.e. $\ell_i$ is not the diagonal. Let $V$ be the point of intersection $\ell_i \cap \{(S, T), (P, \bar{P})\}$, where $\bar{P} = \min \{P, \frac{1}{2}(T + S)\}$. Thus, $\ell_i \cap S \cap \{(R, R) = [V, (R, R)] \}$ and the entire half-open segment lies above the diagonal. Now let $\ell_j$ be a separation line which does not contain $(R, R)$. So it contains a point $A \in ((R, R), (T, S)]$. Let $B$ be the intersection point $\ell_j \cap \{(S, T), (P, \bar{P})\}$. If $B$ lies below $\ell_i$, then the entire segment $\ell_j \cap S$ lies below $\ell_i$. Otherwise, $B \in [V, (S, T)]$ and this is the situation we wish to examine.

Since $A$ is below $\ell_i$, and $B$ is on or above $\ell_i$, it follows that $\ell_j$ intersects $\ell_i$ at a point $V^j$ of $S$ with $V_X^j$ its $X$ coordinate. Notice that the portion of $\ell_j$ to the right of $V^j$ lies below $\ell_i$. In any case, $\text{Switch}(\ell_j)$ intersects $\ell_i$ at a point $W^j$ with $X$ coordinate $W_X^j$.

**Lemma 5.9.** If $\ell_j$ has non-negative slope then $V_X^j < W_X^j$.

**Proof.** The lines $\ell_j$ and $\text{Switch}(\ell_j)$ meet the diagonal at a common point $(Q, Q) = \ell_j \cap \text{Switch}(\ell_j)$. To the right of $\{s_X = Q\}$ the line $\ell_j$ lies below the diagonal, because $A$ is below the diagonal. On the other hand, all of $\ell_i$ and $\text{Switch}(\ell_j)$ lies above the diagonal. Hence, $V_X^j < Q$. Similarly, $\text{Switch}(\ell_j)$ intersects $\ell_i$ above the diagonal. Since $\text{Switch}(\ell_j)$ is either vertical or has positive slope, it follows that $Q \leq W_X^j$.

From this we obtain the main result of this section.

**Theorem 5.10.** Let $\{\pi_i : i \in I\}$ be a finite indexed collection of simple Smale plans with $\ell_i$ the separation line for $\pi_i$. Assume that for some $i^* \in I$, $\ell_{i^*}$ is a line through $(R, R)$ with slope strictly between 0 and 1 and $(R, R) \notin \ell_j$ for any $j \in I \setminus \{i^*\}$. If $\ell_i$ has non-negative slope for all $i \in I$ and, in addition, $\ell_i \cap S$ lies below $\ell_i$, for those $i \in I$ with $\ell_i$ horizontal, then fixation at $i^*$ is a globally stable equilibrium. That is, if $\xi_{i^*}(0) > 0$ then $\lim_{t \to \infty} \xi_{i^*}(t) = 1$.

**Proof.** We choose a numbering of the $n$ elements of $I \setminus \{i^*\}$ by letting $j_1, \ldots, j_m$ with $0 \leq m \leq n$ so that $\ell_j \cap S$ lies below $\ell_i$, if and only if $j = j_p$ for some $p \leq m$. If no such exist then $m = 0$ and the set is empty.

For the remaining $\ell_j$’s the slope is positive and the numbers $V_X^j$ and $W_X^j$ are defined as above. Number them so that $V_X^{j_p} \leq V_X^{j_{p+1}}$ for $m < p < n$.

By Corollary 5.5 it suffices to show that $i^*$ weakly dominates the sequence \{\(j_1, \ldots, j_n\)\}.

To begin with $i^*$ weakly dominates each $j_p$ for $p \leq m$ by Lemma 5.8.

We must show that if $m < p \leq n$ then $i^*$ dominates $j_p$ in $\{i^*, j_p, j_{p+1}, \ldots, j_n\}$. As before, $A_{j_p, i^*} < A_{i^*, j_p}$. Now let $k \in \{j_p, j_{p+1}, \ldots, j_n\}$.

Because of the chosen numbering and Lemma 5.9 we have $V_X^{j_p} \leq V_X^k < W_X^k$. That is, intersection point $W^k$ of $\text{Switch}(\ell_k) \cap \ell_i$ lies to the right of $V_X^{j_p}$. The slope of $\text{Switch}(\ell_k)$ is greater than 1 and the slope of $\ell_i$ is less than 1. Hence, $\text{Switch}(\ell_k)$ is above $\ell_i$, to the right of $W^k$ and below $\ell_i$, to the left. It follows that $\text{Switch}(\ell_k)$ intersects the vertical line $\{s_X = V_X^{j_p}\}$ below $V_X^{j_p}$ and so below the line $\ell_{j_p}$, because $V_X^{j_p}$ lies on $\ell_{j_p}$. Again $\text{Switch}(\ell_k)$ has slope greater than 1 and $\ell_{j_p}$ has slope less than one. So $\text{Switch}(\ell_k)$ intersects $\ell_{j_p}$ to the right of this vertical line. Right of this $V_X^{j_p}$ vertical line, $\ell_{j_p}$ lies below $\ell_{i^*}$. As in Lemma 5.8 the intersection point $\text{Switch}(\ell_k) \cap \ell_{j_p}$ lies below and to the right of $W^k = \text{Switch}(\ell_k) \cap \ell_i$. That is, $A_{j_p, k} < A_{i^*, k}$. Thus, $i^*$ dominates $j_p$ in $\{i^*, j_p, j_{p+1}, \ldots, j_n\}$, as required. \qed
6. **Variations.** Following Smale we consider alternative weighting schemes.

Let \( \{w_1, w_2, \ldots \} \) be an infinite sequence of positive numbers. Let \( W_N = \sum_{k=1}^{N} w_k \) and \( \Delta_N = \sum_{k=1}^{N} |w_{k+1} - w_k| \). Consider the conditions:

- (Weight Condition 1) \( \lim_{N \to \infty} \frac{w_N}{W_N} = 0 \).
- (Weight Condition 2) \( \lim_{N \to \infty} W_N = \infty \).
- (Weight Condition 3) \( \lim_{N \to \infty} \frac{\Delta_N}{W_N} = 0 \).

**Lemma 6.1.** Condition (3) implies Conditions (1) and (2). If the sequence \( \{w_n\} \) is monotonically non-increasing or non-decreasing, then Conditions (1) and (2) imply Condition (3).

**Proof.** \( \Delta_N \geq |w_{N+1} - w_1| \). So (3) implies \( |w_{N+1} - w_1| \to 0 \). If \( |w_{n+1} - w_1| \leq \frac{1}{2} w_1 \) infinitely often then \( w_{n+1} > \frac{1}{2} w_1 \) infinitely often and so the increasing sequence \( \{W_N\} \) is unbounded, implying (2). Otherwise, eventually \( |w_{n+1} - w_1| > \frac{1}{2} w_1 \) and so \( \frac{w_{n+1}}{W_N} \to 0 \) which implies (2). Then \( w_{n+1} \leq w_1 + |w_{n+1} - w_1| \) and so \( \frac{w_{N+1}}{W_{N+1}} < \frac{w_{N+1}}{W_N} \to 0 \) which is (1).

Condition (1) implies \( \frac{W_N}{W_{N+1}} = 1 - \frac{w_{N+1}}{W_{N+1}} \to 1 \). Hence, (1) implies \( \frac{w_{N+1}}{W_N} \to 0 \). Condition (2) implies \( \frac{W_{N+1}}{W_N} \to 0 \).

If the sequence \( \{w_n\} \) is monotone then the sum defining \( \Delta_N \) telescopes to yield \( \Delta_N = |w_{N+1} - w_1| \). So in that case (1) and (2) imply (3).

If the sequence is non-increasing then (1) certainly holds and (2) says that the sequence does not decrease so fast that the associated series converges. If the sequence is non-decreasing then (2) certainly holds and (1) says that the sequence does not increase too fast. For example, if \( w_{N+1} \geq c W_N \) then \( \frac{w_{N+1}}{W_{N+1}} \geq \frac{c}{1+c} \).

The initial averaging procedure that we used had \( w_n = 1 \) for all \( n \) and so \( W_N = N \).

Since the averaging procedure uses ratios we may multiply by a positive constant and so assume \( w_1 = 1 \) and hence \( W_N \geq 1 \) for all \( N \).

Now assume that \( \{w_n\} \) is a positive sequence with \( w_1 = 1 \) and Conditions (1) and (2) hold.

We replace our previous averaging of the payoff sequence in \((2.4)\) to define

\[
S_N^N = \frac{1}{W_N} \sum_{k=1}^{N} w_k S_k.
\]  

(6.1)

We obtain the analogues of \((2.5)\) and \((2.6)\).

\[
S_{N+1}^N = \frac{W_N}{W_{N+1}} S_N + \frac{W_{N+1}}{W_{N+1}} S_{N+1}.
\]  

(6.2)

and so

\[
S_{N+1}^N - S_N^N = \frac{W_{N+1}}{W_{N+1}} (S_{N+1}^N - S_N^N).
\]  

(6.3)

By Condition (1) \((6.1)\) implies that \( ||S_{N+1}^N - S_N^N|| \to 0 \) and so the limit point set is connected as before. However, the crucial fact is \((6.2)\) which says that \( S_{N+1}^N \) is on the segment \([S_{N+1}^N, S_N^N]\) with the weight on \( S_N^N \) approaching 1 as \( N \to \infty \). Consequently, all of the linear estimates for Smale plans go through as before. The only change is that the numerical estimates \( MN^*/N \) are replaced by \( MW_{N^*}/W_N \) which tends to 0 as \( N \to \infty \) by Condition (2). In particular, when two non-extreme, simple Smale plans compete we obtain convergence to the intersection point regardless of the averaging procedure.
It is the similar result for Markov plans that requires Condition (3).
Suppose that $\mathbf{M}$ is the Markov matrix when $X$ plays $p$ and $Y$ plays $q$. Let $v^1$ be the initial distribution and $v^{n+1} = v^n\mathbf{M}$, the distribution after round $n+1$. Define

$$v^N = \frac{1}{W_N} \sum_{k=1}^{N} w_k v^k.$$  

(6.4)

It follows that

$$v^N\mathbf{M} = \frac{1}{W_N} \sum_{k=1}^{N} w_k v^{k+1} = \frac{1}{W_N} \sum_{k=2}^{N+1} w_{k-1} v^k.$$  

(6.5)

Since the length of a distribution is at most 1 we have that

$$||v^N - v^N\mathbf{M}|| \leq \frac{w_1 + w_{N+1} + \Delta_N}{W_N}.$$  

(6.6)

From Condition (3) it follows that any limit point of the sequence $\{v^N\}$ is a stationary distribution. In particular if there is a unique terminal set and so a unique stationary distribution $v$ then $\{v^N\}$ converges to $v$. If $J$ is one of several terminal sets then with probability $p_J$, depending only on the initial distribution, $v^1$, the sequence of outcomes enters $J$. The conditional distributions assuming entrance into $J$ then converge to the unique stationary distribution on $J$.

In contrast with all this, there is another sort of natural averaging which does not work. Suppose we use

$$s^N = \frac{1}{W_N} \sum_{k=1}^{N} w_k S^{N+1-k}.$$  

(6.7)

With Condition (3) one can still show that $||s^{N+1} - s^N|| \to 0$, but this time $s^{N+1}$ is not on the segment $[S^{N+1}, s^N]$ except when all the $w_n$ are equal, in which case the two sorts of averaging agree (This is the original $w_n = 1$ for all $n$ case). So the results from the first section will not carry over.

The other variation to consider is an asymmetric version of the Prisoner’s Dilemma with payoffs given by

$$X \setminus Y | \begin{array}{c|c|c}
  c & d \\
  \hline
  c & (R_X, R_Y) & (S_X, T_Y) \\
  d & (T_X, S_Y) & (P_X, P_Y) \\
\end{array}$$  

(6.8)

and with inequalities for $X$ and for $Y$ analogous to those of (2.3).

This is a real issue because in the classic version of the Prisoner’s Dilemma the payoffs are not in units of dollars, time reduced from a prison sentence or population fitness, but in terms of utility and there is no reason that the two players would have the same Von Neumann-Morgenstern utility functions.

At first glance, there is no problem. In [4] the good Markov strategies are characterized for the asymmetric case. In [18] Smale points out that the theory will work the same way for the asymmetric case. Now one must describe separate Smale strategies for $Y$, rather than using $\pi \circ Switch$, but as he indicates the mathematics is essentially the same.

There is, however, an underlying philosophical problem. In [4] the inequalities for a good plan for $X$ use the payoffs for $Y$, which, in theory, $X$ does not know. In the Markov case, this is not too bad because only a rough estimate is needed to ensure that the strategy is good.
In the Smale case, the running averages use the payoffs to both players. Perhaps the best way to proceed would be to begin again and operate, not in the two dimensional convex set generated by the payoff pairs but in the three dimensional simplex of outcomes. That is, let

\[ e_{cc} = (1, 0, 0, 0), \quad e_{cd} = (0, 1, 0, 0), \quad e_{dc} = (0, 0, 1, 0), \quad e_{dd} = (0, 0, 0, 1) \]

(6.9)

The convex hull \( S' \) with these vertices is the simplex of distributions on the four outcomes. The data we use from the sequence of outcomes \( \{o_1, \ldots, o_N\} \) is the frequency of past outcomes:

\[ s^N = \frac{1}{N} \sum_{k=1}^{N} e_{o_k}. \]

(6.10)

so that, analogous with (2.5)

\[ s^{N+1} = \frac{1}{N+1} o_{N+1} + \frac{N}{N+1} s^N. \]

(6.11)

A plan for \( X \) is then a map \( \pi : S' \to [0, 1] \) with \( \pi(s) \) the probability of cooperating in response to position \( s \). So a pure strategy plan, of the sort Smale uses would be a map \( \pi : S' \to \{0, 1\} \).

Linear results analogous to those of Section 2 can then be carried over. Nonetheless, determining what is a good plan would still require some estimate of the opponent’s payoffs. This is a task for another day.

REFERENCES

[1] K. Abhyankar, Smale strategies for Prisoner’s Dilemma type games, in Algebra, Arithmetic and Geometry with Applications; Papers from Shreeram S. Abhyankar’s 70th Birthday Conference (eds. C. Christensen et al.), Springer-Verlag, Berlin, (2004), 45–48.

[2] E. Akin, The differential geometry of population genetics and evolutionary games, in Mathematical and Statistical Developments of Evolutionary Theory (ed. S. Lessard), Kluwer, Dordrecht, 229 (1990), 1–93.

[3] E. Akin, The iterated prisoner’s dilemma: Good strategies and their dynamics, Ergodic theory, De Gruyter, Berlin, (2016), 77–107. arXiv:1211.0969v3.

[4] E. Akin, What you gotta know to play good in the Iterated Prisoner’s Dilemma, Games, 6 (2015), 175–190.

[5] E. Akin, The Iterated Prisoner’s Dilemma: Good strategies and their dynamics, in Ergodic Theory, Advances in Dynamical Systems (ed. I. Assani), De Gruyter, Berlin, (2016), 77–107.

[6] R. Axelrod, The Evolution of Cooperation, Basic Books, New York, NY, 1984.

[7] K. Behrstock, M. Benaim and M. Hirsch, Smale strategies for network Prisoner’s Dilemma Games, Journal of Dynamics and Games, 2 (2015), 141–155.

[8] M. Benaim and M. Hirsch, Stochastic adaptive behavior for Prisoner’s Dilemma, unpublished manuscript.

[9] M. Boerlijst, M. Nowak and K. Sigmund, Equal pay for all prisoners, Amer. Math. Monthly, 104 (1997), 303–305.

[10] C. Hilbe, M. Nowak and K. Sigmund, The evolution of extortion in Iterated Prisoner’s Dilemma games, Proceedings of the National Academy of Sciences, 110 (2013), 6913–6918.

[11] J. Hofbauer and K. Sigmund, Evolutionary Games and Population Dynamics, Cambridge Univ. Press, Cambridge, UK, 1998.

[12] G. Kendall, X. Yao and S. W. Chong (eds.), The Iterated Prisoner’s Dilemma, 20 Years On, Advances in Natural Computation vol.4, World Scientific, Singapore, 2007.

[13] J. Maynard Smith, Evolution and the Theory of Games, Cambridge Univ. Press, Cambridge, UK, 1982.

[14] M. Nowak, Evolutionary Dynamics, Harvard Univ. Press, Cambridge, MA, 2006.

[15] W. Press and F. Dyson, Iterated Prisoner’s Dilemma contains strategies that dominate any evolutionary opponent, Birds and Frogs, (2015), 329–341.

[16] K. Sigmund, Games of Life, Oxford Univ. Press, Oxford, UK, 1993.
[17] K. Sigmund, *The Calculus of Selfishness*, Princeton Univ. Press, Princeton, NJ, 2010.
[18] S. Smale, The Prisoner’s Dilemma and dynamical systems associated to non-cooperative games, *Econometrica*, 48 (1980), 1617–1634.
[19] A. Stewart and J. Plotkin, Extortion and cooperation in the Prisoner’s Dilemma, *Proceedings of the National Academy of Sciences*, 109 (2012), 10134–10135.
[20] A. Stewart and J. Plotkin, From extortion to generosity, evolution in the Iterated Prisoner’s Dilemma, *Proceedings of the National Academy of Sciences*, 110 (2013), 15348–15353.
[21] P. Taylor and L. Jonker, Evolutionarily stable strategies and game dynamics, *Math. Biosciences*, 40 (1978), 145–156.

Received October 2016; revised March 2017.

E-mail address: ethanakin@earthlink.net