Ethan Akin

The iterated Prisoner's Dilemma: good strategies and their dynamics

Abstract: For the iterated Prisoner's Dilemma, there exist Markov strategies that solve the problem when we restrict attention to the long-term average payoff. When used by both players, these assure the cooperative payoff for each of them. Neither player can benefit by moving unilaterally any other strategy, i.e., these are Nash equilibria. In addition, if a player uses instead an alternative that decreases the opponent's payoff below the cooperative level, then his own payoff is decreased as well. Thus, if we limit attention to the long-term payoff, these *good strategies* effectively stabilize cooperative behavior. We characterize these good strategies and analyze their role in evolutionary dynamics.

Keywords: Prisoner's Dilemma, stable cooperative behavior, iterated play, Markov strategies, zero-determinant strategies, Press–Dyson equations, evolutionary game dynamics.

1 The iterated Prisoner's Dilemma

The *Prisoner's Dilemma* is a two-person game that provides a simple model of a disturbing social phenomenon. It is a symmetric game in which each of the two players, X and Y, has a choice between two strategies, c and d. Thus, there are four outcomes that we list in the order: cc, cd, dc, dd, where, for example, cd is the outcome when X plays c and Y plays d. Each then receives a payoff. The following 2×2 chart describes the payoff to the X player. The transpose is the Y payoff.

X/Y	c	d
c	R	S
d	T	P

(1.1)

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

$$\mathbf{S}_X = (R, S, T, P) \quad \text{and} \quad \mathbf{S}_Y = (R, T, S, P). \tag{1.2}$$

Davis [6] and Straffin [17] provide clear introductory discussions of the elements of game theory.

Either player can use a *mixed strategy*, randomizing by choosing c with probability p_c and d with the complementary probability $1 - p_c$.

A probability *distribution* $\mathbf{v}$ on the set of outcomes is a non-negative vector with unit sum, indexed by the four states. That is, $v_i \geq 0$ for $i = 1, \ldots, 4$ and the dot product $\langle \mathbf{v} \cdot \mathbf{1} \rangle = 1$. For example, v_2 is the probability that X played c and Y played d. In particular, $v_1 + v_2$ is the probability X played c. With respect to $\mathbf{v}$, the expected payoffs to X and Y, denoted s_X and s_Y, are the dot products with the corresponding payoff vectors:

$$s_X = \langle \mathbf{v} \cdot \mathbf{S}_X \rangle \quad \text{and} \quad s_Y = \langle \mathbf{v} \cdot \mathbf{S}_Y \rangle. \tag{1.3}$$

The payoffs are assumed to satisfy

$$T > R > P > S \quad \text{and} \quad 2R > T + S. \tag{1.4}$$

We will later use the following easy consequence of these inequalities.

Proposition 1.1. *If* $\mathbf{v}$ *is a distribution, then the associated expected payoffs to the two players, as defined by Equation (1.3), satisfy the following equation:*

$$s_Y - s_X = (v_2 - v_3)(T - S). \tag{1.5}$$

So we have $s_Y = s_X$ *iff* $v_2 = v_3$.

In addition,

$$\frac{1}{2}(s_Y + s_X) \leq R, \tag{1.6}$$

with equality iff $\mathbf{v} = (1, 0, 0, 0)$. *Hence, the following statements are equivalent.*

(i) $1/2(s_Y + s_X) = R$,

(ii) $v_1 = 1$,

(iii) $s_Y = s_X = R$.

Proof. Dot $\mathbf{v}$ with $\mathbf{S}_Y - \mathbf{S}_X = (0, T - S, S - T, 0)$ and with $(1/2)(\mathbf{S}_Y + \mathbf{S}_X) = (R, (1/2)(T + S), (1/2)(T + S), P)$. Observe that R is the maximum entry of the latter. □

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.

This is where it gets ugly. In game theory lingo, 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 Array (1.1), 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. Having firmly agreed to cooperate, X hopes that Y will stick to the agreement because X can then obtain the large payoff *T* by defecting. Furthermore, if he were not to play *d*, then he risks getting *S* when Y defects. All the more reason to defect, as X realizes Y is thinking the same thing.

The payoffs are often stated in money amounts or in years reduced from a prison sentence (the original "prisoner" version), but it is important to understand that the payoffs are really in units of *utility*. That is, the ordering in Equation (1.4) is assumed to describe the order of desirability of the various outcomes to each player when all the consequences of each outcome are taken into account. Thus, if X is induced to feel guilty at the *dc* outcome, then the payoff to X of that outcome is reduced. Adjusting the payoffs is the classic way of stabilizing cooperative behavior. Suppose prisoner X walks out of prison, free after defecting, having consigned Y, who played *c*, to a 20-year sentence. Colleagues of Y might well do X some serious damage. Anticipation of such an event considerably reduces the desirability of *dc* for X, perhaps to well below *R*. If X and Y each have threatening friends, then it is reasonable for each to expect that a prior agreement to play *cc* will stand and so they each receive *R*. However, in terms of utility this is no longer a Prisoner's Dilemma. In the book that originated modern game theory, Von Neumann and Morgenstern [19], the authors developed an axiomatic theory of utility that allows us to make sense of such arithmetic relationships as the second inequality in Equation (1.4). We won't consider this further, but the reader should remember that the payoffs are numerical measurements of desirability.

This two-person collapse of cooperation can be regarded as a simple model of what Garret Hardin [7] calls *the tragedy of the commons*. This is a similar sort of collapse of mutually beneficial cooperation on a multi person scale.

In the search for a way to avert this tragedy, 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.

Robert Axelrod devised a tournament in which submitted computer programs played against one another. Each program played a fixed, but unknown, number of rounds against each of the competing programs, and the resulting payoffs were summed. The results are described and analyzed in his landmark book [4]. The winning program, Tit-for-Tat, submitted by game theorist Anatol Rapaport, consists, after initial cooperation, in playing in each round the strategy used by the opponent in the previous round. A second tournament yielded the same winner. Axelrod extracted some interesting rules of thumb from Tit-for-Tat and applied these to some historical examples.

At around the same time, game theory was being introduced by John Maynard Smith into biology in order to study problems in the evolution of behavior. Maynard Smith [11] and Sigmund [13] provide good surveys of the early work. Tournament play for games, which has been widely explored since, exactly simulates the dynamics examined in this growing field of evolutionary game theory. However, the tournament/evolutionary viewpoint changes the problem in a subtle way. In evolutionary game theory, what matters is how a player is doing as compared with the competing players. Consider this with just two players and suppose they are currently considering strategies with the same payoff to each. Comparing outcomes, Y would reject a move to a strategy where she does better, but which allows X to do still better than she. That this sort of *altruism* is selected against is a major problem in the theory of evolution. However, in classical game theory, the payoffs are in utilities. Y simply desires to obtain the highest absolute payoff. The payoffs to her opponent are irrelevant, except as data to predict X's choice of strategy. It is the classical problem that we will mainly consider, although we will return to evolutionary dynamics in the last section.

I am not competent to summarize the immense literature devoted to these matters. I recommend the excellent book-length treatments of Hofbauer and Sigmund [9] and Nowak [12] and Sigmund [14]. The latter two discuss the Markov approach that we now examine.

The choice of play for the first round is the *initial play*. A *strategy* is a choice of initial play together with what we will call a *plan*: a choice of play, after the first round, to respond to any possible past history of outcomes in the previous rounds. A memory-one plan bases its response entirely on outcome of the previous round. The Tit-for-Tat plan (hereafter, just TFT) is an example of a memory-one plan.

With the outcomes listed in order as cc, cd, dc, dd, a memory one plan for X is a vector $\mathbf{p} = (p_1, p_2, p_3, p_4) = (p_{cc}, p_{cd}, p_{dc}, p_{dd})$, where p_z is the probability of playing c when the outcome z occurred in the previous round. If Y uses the memory-one plan $\mathbf{q} = (q_1, q_2, q_3, q_4)$, 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

$$\mathbf{M} = \begin{pmatrix} p_1q_1 & p_1(1-q_1) & (1-p_1)q_1 & (1-p_1)(1-q_1) \\ p_2q_3 & p_2(1-q_3) & (1-p_2)q_3 & (1-p_2)(1-q_3) \\ p_3q_2 & p_3(1-q_2) & (1-p_3)q_2 & (1-p_3)(1-q_2) \\ p_4q_4 & p_4(1-q_4) & (1-p_4)q_4 & (1-p_4)(1-q_4) \end{pmatrix}. \tag{1.7}$$

We use the switch in numbering from the Y strategy $\mathbf{q}$ to the Y response vector 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 $\mathbf{p} = \mathbf{q} = (1, 0, 1, 0)$, but the response vector for Y is $(1, 1, 0, 0)$. The plan *Repeat* is given by $\mathbf{p} = \mathbf{q} = (1, 1, 0, 0)$ with the response vector for Y equal to $(1, 0, 1, 0)$. This plan just repeats the previous play, regardless of what the opponent did.

We describe some elementary facts about finite Markov chains, see, e.g., Chapter 2 of Karlin and Taylor [10].

A Markov matrix like **M** is a non-negative matrix with row sums equal to 1. Thus, the vector **1** is a right eigenvector with eigenvalue 1. For such a matrix, we can represent the associated Markov chain as movement along a directed graph with vertices the states, in this case, cc, cd, dc, dd, and with a directed edge from the i-th state z_i to the j-th state z_j when $\mathbf{M}_{ij} > 0$, that is, when we can move from z_i to z_j with positive probability. In particular, there is an edge from z_i to itself iff the diagonal entry $\mathbf{M}_{ii}$ is positive.

A *path* in the graph is a state sequence $z^1, \ldots, z^n$ with $n > 1$ such that there is an edge from z^i to z^{i+1} for $i = 1, \ldots, n-1$. A set of states I is called a *closed set* when no path that begins in I can exit I. For example, the entire set of states is closed and for any z the set of states accessible via a path that begins at z is a closed set. I is closed iff $\mathbf{M}_{ij} = 0$ whenever $z_i \in I$ and $z_j \notin I$. In particular, when we restrict the chain to a closed set I, the associated submatrix of **M** still has row sums equal to 1. A minimal, nonempty, closed set of states is called a *terminal set*. A state is called *recurrent* when it lies in some terminal set and *transient* when it does not. The following facts are easy to check.

- A nonempty, closed set of states I is terminal iff for all $z_i, z_j \in I$, there exists a path from z_i to z_j.
- If I is a terminal set and $z_j \in I$, then there exists $z_i \in I$ with an edge from z_i to z_j.
- Distinct terminal sets are disjoint.
- Any nonempty, closed set contains at least one terminal set.
- From any transient state, there is a path into some terminal set.

Suppose we are given an initial distribution $\mathbf{v}^1$, describing the outcome of the first round of play. The Markov process evolves in discrete time via the equation

$$\mathbf{v}^{n+1} = \mathbf{v}^n \cdot \mathbf{M}, \tag{1.8}$$

where we regard the distributions as row vectors.

In our game context, the initial distribution is given by the initial plays, pure or mixed, of the two players. If X uses initial probability p_c and Y uses q_c, then

$$\mathbf{v}^1 = (p_c q_c, p_c(1-q_c), (1-p_c)q_c, (1-p_c)(1-q_c)). \tag{1.9}$$

Thus, v_i^n is the probability that outcome z_i occurs on the n-th round of play. A distribution **v** is *stationary* when it satisfies $\mathbf{vM} = \mathbf{v}$. That is, it is a left eigenvector with eigenvalue 1. From Perron–Frobenius theory (see, e.g., Appendix 2 of [10]), it follows that if I is a terminal set, then there is a unique stationary distribution **v** with $v_i > 0$ iff $i \in I$. That is, the *support* of **v** is exactly I. In particular, if the eigenspace of **M** associated with the eigenvalue 1 is one-dimensional, then there is a unique stationary distribution, and so a unique terminal set that is the support of the stationary distribution. The converse is also true and any stationary distribution **v** is a mixture of the $\mathbf{v}_J$'s, where $\mathbf{v}_J$ is supported on the terminal set J. This follows from the fact that any stationary distribution $\mathbf{v}$ satisfies $v_i = 0$ for all transient states z_i and so is supported on the set of recurrent states. On the recurrent states, the matrix $\mathbf{M}$ is block diagonal. Hence, the following are equivalent in our 4×4 case:

- There is a unique terminal set of states for the process associated with M.
- There is a unique stationary distribution vector for M.
- The matrix $M' = M - I$ has rank 3.

We will call $\mathbf{M}$ *convergent* when these conditions hold. For example, when all of the probabilities of $\mathbf{p}$ and $\mathbf{q}$ lie strictly between 0 and 1, then all the entries of $\mathbf{M}$ given by Equation (1.7) are positive and so the entire set of states is the unique terminal state and the positive matrix $\mathbf{M}$ is convergent.

The sequence of the Cesaro averages $\{1/n\Sigma_{i=1}^n \mathbf{v}^i\}$ of the outcome distributions always converges to some stationary distribution $\mathbf{v}$. That is,

$$\lim_{n\to\infty} \frac{1}{n}\Sigma_{k=1}^n \mathbf{v}^k = \mathbf{v}. \tag{1.10}$$

Hence, using the payoff vectors from Equation (1.2), the long-run average payoffs for X and Y converge to s_X and s_Y of Equation (1.3) with $\mathbf{v}$ this limiting stationary distribution.

When $\mathbf{M}$ is convergent, the limit $\mathbf{v}$ is the unique stationary distribution and so the average payoffs are independent of the initial distribution. In the non-convergent case, the long-term payoffs depend on the initial distribution. Suppose there are exactly two terminal sets, I and J, with stationary distribution vectors $\mathbf{v}_I$ and $\mathbf{v}_J$, supported on I and J, respectively. For any initial distribution $\mathbf{v}^1$, there are probabilities p_I and $p_J = 1 - p_I$ of entering into, and so terminating in, I or J, respectively. In that case, the limit of the Cesaro averages sequence for $\{\mathbf{v}^n\}$ is given by

$$\mathbf{v} = p_I\mathbf{v}_I + p_J\mathbf{v}_J, \tag{1.11}$$

and the limits of the average payoffs are given by Equation (1.3) with this distribution $\mathbf{v}$. This extends in the obvious way when there are more terminal sets.

When Y responds to the memory-one plan $\mathbf{p}$ with a memory-one plan $\mathbf{q}$, we have the *Markov case* as above. We will also want to see how a memory-one plan $\mathbf{p}$ for X fares against a not necessarily memory-one response by Y. We will call such a response pattern a *general plan* to emphasize that it need not be memory-one. That is, a general plan is a choice of response, pure or mixed, for any sequence of previous outcomes. Hereafter, unless we use the expression "general plan," we will assume a plan is memory-one.

If Y uses a general plan, then the sequence of Cesaro averages need not converge. We will call any limit point of the sequence an associated *limit distribution*. We will call s_X and s_Y, given by Equation (1.3) with such a limit distribution $\mathbf{v}$, the *expected payoffs* associated with $\mathbf{v}$.

Call a plan **p** *agreeable* when $p_1 = 1$ and *firm* when $p_4 = 0$. That is, an agreeable plan always responds to *cc* with *c* and a firm plan always responds to *dd* with *d*. If both **p** and **q** are agreeable, then $\{cc\}$ is a terminal set for the Markov matrix **M** given by Equation (1.7) and so $\mathbf{v} = (1, 0, 0, 0)$ is a stationary distribution with fixation at *cc*. If both **p** and **q** are firm, then $\{dd\}$ is a terminal set for **M** and $\mathbf{v} = (0, 0, 0, 1)$ is a stationary distribution with fixation at *dd*. Any convex combination of agreeable plans (or firm plans) is agreeable (respectively, firm).

An agreeable plan together with initial cooperation is called an *agreeable strategy*.

The plans TFT $= (1, 0, 1, 0)$ and Repeat $= (1, 1, 0, 0)$ are each agreeable and firm. The same is true for any mixture of these. If both X and Y use TFT, then the outcome is determined by the initial play. Initial outcomes *cc* and *dd* lead to immediate fixation. Either *cd* or *dc* results in period 2 alternation between these two states. Thus, $\{cd, dc\}$ is another terminal set with stationary distribution $(0, 1/2, 1/2, 0)$. If $a \cdot \text{TFT} + (1 - a)$Repeat is used instead by either player (with $0 < a < 1$), then eventually fixation at *cc* or *dd* results. There are then only two terminal sets instead of three. The period 2 alternation described above illustrates why we needed the Cesaro limit, i.e., the limit of averages, in Equation (1.10) rather than the limit per se.

Because so much work had been done on this Markov model, the exciting new ideas in Press and Dyson [15] took people by surprise. They have inspired a number of responses, e.g., Stewart and Plotkin [16] and especially, Hilbe, Nowak, and Sigmund [8]. I would here like to express my gratitude to Karl Sigmund whose kind, but firm, criticism of the initial draft directed me to this recent work. The result is both a substantive and expository improvement.

Our purpose here is to use these new ideas to characterize the plans that are good in the following sense.

Definition 1.2. A plan **p** for X is called *good* if it is agreeable and if for any general plan chosen by Y against it and any associated limit distribution, the expected payoffs satisfy

$$s_Y \geq R \implies s_Y = s_X = R. \tag{1.12}$$

The plan is called *of Nash type* if it is agreeable and if the expected payoffs against any Y general plan satisfy

$$s_Y \geq R \implies s_Y = R. \tag{1.13}$$

A *good strategy* is a good plan together with initial cooperation.

By Proposition 1.1, $s_Y = s_X = R$ iff the associated limit distribution is $(1, 0, 0, 0)$. In the memory-one case, $(1, 0, 0, 0)$ is a stationary distribution iff both plans are agreeable. It is the unique stationary distribution iff, in addition, the matrix **M** is convergent. If **p** is not agreeable, then Equation (1.12) can be vacuously true. For example, if X plays All$D = (0, 0, 0, 0)$, then for any Y response $P \geq s_Y$ and the implication is true.

When both players use agreeable strategies, i.e., agreeable plans with initial cooperation, then the joint cooperative payoff is achieved. The pair of strategies is a Nash equilibrium exactly when the two plans are of Nash type. That is, both players receive R and neither player can do better by playing an alternative plan. A good plan is of Nash type, but more is true. We will see that with a Nash equilibrium it is possible that Y can play an alternative that still yields R for herself but with the payoff to X smaller than R. That is, Y has no incentive to play so as to reach the joint cooperative payoff. On the other hand, if X uses a good plan, then the only responses for Y that obtain R for her also yield R for X.

The strategy Repeat $= (1, 1, 0, 0)$ is an agreeable plan that is not of Nash type. If both players use Repeat, then the initial outcome repeats forever. If the initial outcome is cd, then $s_Y = T$ and $s_X = S$.

For a plan $\mathbf{p}$, we define the X *Press–Dyson vector* $\tilde{\mathbf{p}} = \mathbf{p} - \mathbf{e}_{12}$, where $\mathbf{e}_{12} = (1, 1, 0, 0)$. Considering the utility of the following result of Hilbe, Nowak and Sigmund, its proof, taken from Appendix A of [8], is remarkably simple.

Theorem 1.3. *Assume that* X *uses the plan* $\mathbf{p}$ *with* X *Press–Dyson vector* $\tilde{\mathbf{p}}$. *If the initial plays and the general plan of* Y *yields the sequence of distributions* $\{\mathbf{v}^n\}$, *then*

$$\lim_{n\to\infty} \frac{1}{n} \Sigma_{k=1}^{n} \langle \mathbf{v}^k \cdot \tilde{\mathbf{p}} \rangle = 0,$$
$$\text{and so} \quad \langle \mathbf{v} \cdot \tilde{\mathbf{p}} \rangle = v_1\tilde{p}_1 + v_2\tilde{p}_2 + v_3\tilde{p}_3 + v_4\tilde{p}_4 = 0 \tag{1.14}$$

for any associated limit distribution $\mathbf{v}$.

Proof. Let $v_{12}^n = v_1^n + v_2^n$, the probability that either cc or cd is the outcome in the n-th round of play. That is, $v_{12}^n = \langle \mathbf{v}^n \cdot \mathbf{e}_{12} \rangle$ is the probability that X played c in the n-th round. On the other hand, since X is using the plan $\mathbf{p}$, p_i is the conditional probability that X plays c in the next round, given outcome z_i in the current round. Thus, $\langle \mathbf{v}^n \cdot \mathbf{p} \rangle$ is the probability that X plays c in the $(n+1)^{st}$ round, i.e., it is v_{12}^{n+1}. Hence, $v_{12}^{n+1} - v_{12}^n = \langle \mathbf{v}^n \cdot \tilde{\mathbf{p}} \rangle$. The sum telescopes to yield

$$v_{12}^{n+1} - v_{12}^1 = \Sigma_{k=1}^{n} \langle \mathbf{v}^k \cdot \tilde{\mathbf{p}} \rangle. \tag{1.15}$$

As the left side has absolute value at most 1, Limit (1.14) follows. If a subsequence of the Cesaro averages converges to $\mathbf{v}$, then $\langle \mathbf{v} \cdot \tilde{\mathbf{p}} \rangle = 0$ by continuity of the dot product. □

To illustrate the use of this result, we examine $TFT = (1, 0, 1, 0)$ and another plan that has been labeled in the literature Grim $= (1, 0, 0, 0)$. We consider mixtures of each with Repeat $= (1, 1, 0, 0)$.

Corollary 1.4. *Let* $1 \geq a > 0$.

(a) *The plan* $\mathbf{p} = a\text{TFT} + (1-a)\text{Repeat}$ *is a good plan with* $s_Y = s_X$ *for any limiting distribution.*

(b) *The plan* $\mathbf{p} = a\text{Grim} + (1-a)\text{Repeat}$ *is good.*

Proof. (a) In this case, $\tilde{\mathbf{p}} = a(0,-1,1,0)$ and so Equation (1.14) implies that $v_2 = v_3$. Thus, $s_Y = s_X$. From this Equation (1.12) follows from Proposition 1.1.

(b) Now $\tilde{\mathbf{p}} = a(0,-1,0,0)$ and so Equation (1.14) implies that $v_2 = 0$. Thus, $s_Y = v_1R + v_3S + v_4P$ and this is less than R unless $v_3 = v_4 = 0$ and $v_1 = 1$. When $v_1 = 1$, $s_Y = s_X = R$, proving Equation (1.12). □

In the next section, we will prove the following characterization of the good plans.

Theorem 1.5. *Let* $\mathbf{p} = (p_1, p_2, p_3, p_4)$ *be an agreeable plan other than* Repeat. *That is,* $p_1 = 1$ *but* $\mathbf{p} \neq (1,1,0,0)$.

The plan $\mathbf{p}$ *is of Nash type iff the following inequalities hold.*

$$\frac{T-R}{R-S} \cdot p_3 \leq (1-p_2) \quad \textit{and} \quad \frac{T-R}{R-P} \cdot p_4 \leq (1-p_2). \tag{1.16}$$

The plan $\mathbf{p}$ *is good iff, in addition, both inequalities are strict.*

Corollary 1.6. *In the compact convex set of agreeable plans, the set* $\{\mathbf{p}$ *equals* Repeat *or is of Nash type*$\}$ *is a closed convex set with interior set of good plans.*

Proof. The X Press–Dyson vectors form a cube and the agreeable plans are the intersection with the subspace $\tilde{p}_1 = 0$. We then intersect with the half-spaces defined by

$$\frac{T-R}{R-S}\tilde{p}_3 + \tilde{p}_2 \leq 0 \quad \text{and} \quad \frac{T-R}{R-P}\tilde{p}_4 + \tilde{p}_2 \leq 0. \tag{1.17}$$

The result is a closed convex set with interior given by the strict inequalities. Notice that these conditions are preserved by multiplication by a positive constant $a \leq 1$ or by any larger constant so long as $a\tilde{\mathbf{p}}$ remains in the cube. Hence, Repeat with $\tilde{\mathbf{p}} = 0$ is on the boundary. □

It is easy to compute that

$$\det \begin{pmatrix} R & R & 1 & 0 \\ S & T & 1 & 1 \\ T & S & 1 & 1 \\ P & P & 1 & 0 \end{pmatrix} = -2(R-P)(T-S). \tag{1.18}$$

Hence, with $\mathbf{e}_{23} = (0,1,1,0)$, we can use $\{\mathbf{S}_X, \mathbf{S}_Y, \mathbf{1}, \mathbf{e}_{23}\}$ as a basis for $\mathbb{R}^4$. For a distribution vector $\mathbf{v}$, we will write v_{23} for $v_2 + v_3 = \langle \mathbf{v} \cdot \mathbf{e}_{23} \rangle$. From Theorem 1.3, we immediately obtain the following.

Theorem 1.7. *If* $\mathbf{p}$ *is a plan whose X Press–Dyson vector* $\tilde{\mathbf{p}} = \alpha\mathbf{S}_X + \beta\mathbf{S}_Y + \gamma\mathbf{1} + \delta\mathbf{e}_{23}$ *and* $\mathbf{v}$ *is a limit distribution when Y plays some general plan against* $\mathbf{p}$*, then the average payoffs satisfy the following* Press–Dyson equation:

$$\alpha s_X + \beta s_Y + \gamma + \delta v_{23} = 0. \tag{1.19}$$

The most convenient cases to study occur when $\delta = 0$. Press and Dyson called such a plan a *zero-determinant strategy* (hereafter ZDS) because of an ingenious determinant argument leading to Equation (1.19). We have used Theorem 1.3 of Hilbe–Nowak–Sigmund instead.

This representation yields a simple description of the good plans.

Theorem 1.8. *Assume that* $\mathbf{p} = (p_1, p_2, p_3, p_4)$ *is an agreeable plan with* X *Press–Dyson vector* $\tilde{\mathbf{p}} = \alpha\mathbf{S}_X + \beta\mathbf{S}_Y + \gamma\mathbf{1} + \delta\mathbf{e}_{23}$. *Assume that* $\mathbf{p}$ *is not* Repeat, *i.e.*, $(\alpha, \beta, \gamma, \delta) \neq (0, 0, 0, 0)$. *The plan* $\mathbf{p}$ *is of Nash type iff*

$$\max\left(\frac{\delta}{(T-S)}, \frac{\delta}{(2R-(T+S))}\right) \leq \alpha. \tag{1.20}$$

The plan $\mathbf{p}$ *is good iff, in addition, the inequality is strict.*

Remark: Observe that $T - S > 2R - (T + S) > 0$. It follows that if $\delta \leq 0$, then $\mathbf{p}$ is good iff $\delta/(T - S) < \alpha$. On the other hand, if $\delta > 0$, then $\mathbf{p}$ is good iff $\delta/(2R - (T + S)) < \alpha$.

In the next section, we will investigate the geometry of the $\{\mathbf{S}_X, \mathbf{S}_Y, \mathbf{1}, \mathbf{e}_{23}\}$ decomposition of the Press–Dyson vectors and prove the theorems.

2 Good plans and the Press–Dyson decomposition

We begin by normalizing the payoffs. We can add to all a common number and multiply all by a common positive number without changing the relationship between the various strategies. We subtract S and divide by $T - S$. So from now on we will assume that $T = 1$ and $S = 0$.

The payoff vectors of Equation (1.2) are then given by

$$\mathbf{S}_X = (R, 0, 1, P), \quad \mathbf{S}_Y = (R, 1, 0, P), \tag{2.1}$$

and from Equation (1.4), we have

$$1 > R > \frac{1}{2}, \quad \text{and} \quad R > P > 0. \tag{2.2}$$

After normalization Theorem 1.5 becomes the following.

Theorem 2.1. *Let* $\mathbf{p} = (p_1, p_2, p_3, p_4)$ *be an agreeable plan other than* Repeat. *That is,* $p_1 = 1$ *but* $\mathbf{p} \neq (1, 1, 0, 0)$.

The plan **p** *is of Nash type iff the following inequalities hold.*

$$\frac{1-R}{R} \cdot p_3 \leq (1-p_2) \quad \textit{and} \quad \frac{1-R}{R-P} \cdot p_4 \leq (1-p_2). \tag{2.3}$$

The plan **p** *is good iff, in addition, both inequalities are strict.*

Proof. We first eliminate the possibility $p_2 = 1$. If $1 - p_2 = 0$, then the inequalities would yield $p_3 = p_4 = 0$ and so $\mathbf{p} =$ Repeat, which we have excluded. On the other hand, if $p_2 = 1$, then $\mathbf{p} = (1, 1, p_3, p_4)$. If against this Y plays All$D = (0, 0, 0, 0)$, then $\{cd\}$ is a terminal set with stationary distribution $(0, 1, 0, 0)$ and so with $s_Y = 1$ and $s_X = 0$. Hence, **p** is not of Nash type. Thus, if $p_2 = 1$, then neither is **p** of Nash type nor do the inequalities hold for it. We now assume $1 - p_2 > 0$.

Observe that

$$\begin{aligned} s_Y - R &= (v_1 R + v_2 + v_4 P) - (v_1 R + v_2 R + v_3 R + v_4 R) \\ &= v_2(1-R) - v_3 R - v_4(R-P). \end{aligned} \tag{2.4}$$

Hence, multiplying by the positive quantity $(1 - p_2)$, we have

$$s_Y >= R \iff (1-p_2)v_2(1-R) >= v_3(1-p_2)R + v_4(1-p_2)(R-P), \tag{2.5}$$

where this notation means that the inequalities are equivalent and the equations are equivalent.

Since $\tilde{p}_1 = 0$, Equation (1.14) implies $v_2\tilde{p}_2 + v_3\tilde{p}_3 + v_4\tilde{p}_4 = 0$ and so $(1-p_2)v_2 = v_3 p_3 + v_4 p_4$. Substituting in the above inequality and collecting terms we get

$$\begin{aligned} &s_Y >= R \iff Av_3 >= Bv_4, \text{ with} \\ &A = [p_3(1-R) - (1-p_2)R] \quad \text{and} \quad B = [(1-p_2)(R-P) - p_4(1-R)]. \end{aligned} \tag{2.6}$$

Observe that the inequalities of Equation (2.3) are equivalent to $A \leq 0$ and $B \geq 0$. The proof is completed by using a sequence of little cases.

Case (i) $A = 0$, $B = 0$: In this case, $Av_3 = Bv_4$ holds for any strategy for Y. So for any Y strategy, $s_Y = R$ and **p** is of Nash type. If Y chooses a plan that is not agreeable, then $\{cc\}$ is not a closed set of states and so $v_1 \neq 1$. From Proposition 1.1, $s_X < R$ and so **p** is not good.

Case (ii) $A < 0$, $B = 0$: The inequality $Av_3 \geq Bv_4$ holds iff $v_3 = 0$. If $v_3 = 0$, then $Av_3 = Bv_4$ and so $s_Y = R$. Thus, **p** is Nash.

Case (iia) $B \leq 0$, any A: Assume Y chooses a plan that is not agreeable and is such that $v_3 = 0$. For example, if Y plays All$D = (0, 0, 0, 0)$, then no state moves to dc. With such a Y choice, $Av_3 \geq Bv_4$ and so $s_Y \geq R$. As above, $v_1 \neq 1$ because the Y plan is not agreeable. Again $s_X < R$ and **p** is not good. Furthermore, $v_3 = 0$, $v_1 < 1$, $p_2 < 1$, and $(1-p_2)v_2 = v_4 p_4$ imply that $v_4 > 0$. So if $B < 0$, then $Av_3 > Bv_4$ and so $s_Y > R$. Hence, **p** is not Nash when $B < 0$.

Case (iii) $A = 0$, $B > 0$: The inequality $Av_3 \geq Bv_4$ holds iff $v_4 = 0$. If $v_4 = 0$, then $Av_3 = Bv_4$ and $s_Y = R$. Thus, **p** is Nash.

Case (iiia) $A \geq 0$, any B: Assume Y chooses a plan that is not agreeable and is such that $v_4 = 0$. For example, if Y plays $(0, 1, 1, 1)$, then no state moves to dd. With such a Y choice, $Av_3 \geq Bv_4$ and so $s_Y \geq R$. As before, $v_1 \neq 1$ implies $s_X < R$ and the plan is not good. Furthermore, $v_4 = 0$, $v_1 < 1$, $p_2 < 1$, and $(1 - p_2)v_2 = v_3p_3$ imply that $v_3 > 0$. So if $A > 0$, then $Av_3 > Bv_4$ and so $s_Y > R$. Hence, $\mathbf{p}$ is not Nash when $A > 0$.

Case (iv) $A < 0$, $B > 0$: The inequality $Av_3 \geq Bv_4$ implies $v_3, v_4 = 0$. So $(1 - p_2)v_2 = v_3p_3 + v_4p_4 = 0$. Since $p_2 < 1$, $v_2 = 0$. Hence, $v_1 = 1$. That is, $s_Y \geq R$ implies $s_Y = s_X = R$ and so $\mathbf{p}$ is good. □

Remarks:

(a) Since $1 > R > 1/2$, it is always true that $(1 - R)/R < 1$. On the other hand, $(1 - R)/(R - P)$ can be greater than 1 and the second inequality requires $p_4 \leq (R - P)/(1 - R)$. In particular, if $p_2 = 0$, then the plan is good iff $p_4 < (R - P)/(1 - R)$. For example, the plan $(1, 0, 0, 1)$ is, in the literature, labeled *Pavlov*, or *WinStay, LoseShift*. This plan always satisfies the first inequality strictly, but it satisfies the second strictly, and so is good, iff $1 - R < R - P$.

(b) In Case (i) of the proof, the payoff $s_Y = R$ is determined by $\mathbf{p}$ independent of the choice of strategy for Y. In general, plans that fix the opponent's payoff in this way were described by Press and Dyson [15] and, earlier, by Boerlijst, Nowak, and Sigmund [5], where they are called *equalizer strategies*. The agreeable equalizer plans have $\tilde{\mathbf{p}} = a(0, -(1 - R)/R, 1, (R - P)/R)$ with $1 \geq a > 0$.

Christian Hilbe suggests a nice interpretation of the above results.

Corollary 2.2. *Let* $\mathbf{p}$ *be an agreeable plan with* $p_2 < 1$.

(a) *If* $\mathbf{p}$ *is good, then using any plan* $\mathbf{q}$ *that is not agreeable forces* Y *to get a payoff* $s_Y < R$.

(b) *If* $\mathbf{p}$ *is not good, then by using at least one of the two plans* $\mathbf{q} = (0, 0, 0, 0)$ *or* $\mathbf{q} = (0, 1, 1, 1)$, Y *can certainly obtain a payoff* $s_Y \geq R$, *and force* X *to get a payoff* $s_X < R$.

(c) *If* $\mathbf{p}$ *is not Nash, then by using at least one of the two plans* $\mathbf{q} = (0, 0, 0, 0)$ *or* $\mathbf{q} = (0, 1, 1, 1)$, *Y can certainly obtain a payoff* $s_Y > R$, *and force* X *to get a payoff* $s_X < R$.

Proof. (a) If $\mathbf{p}$ is good, then $s_Y \geq R$ implies $s_Y = s_X = R$, which requires $v = (1, 0, 0, 0)$. This is only stationary when $\mathbf{q}$ as well as $\mathbf{p}$ is agreeable.
(b) and (c) follow from the analysis of cases in the above proof. □

Remark: If $p_2 = p_1 = 1$, then the plan $\mathbf{p}$ is not Nash. As observed in the proof of Theorem 2.1, if Y plays $\mathbf{q} = (0, 0, 0, 0)$, then cd is a terminal set with stationary distribution

$\mathbf{v} = (0, 1, 0, 0)$ and so with $s_Y = 1$, $s_X = 0$. However, if, in addition, $p_4 = 0$, e.g., if X uses Repeat, then dd is also a terminal set. Thus, if X plays $\mathbf{p}$ with $1-p_4 = p_2 = p_1 = 1$ and Y always defects, then fixation occurs immediately at either cd with $s_Y = 1$ and $s_X = 0$, or else at dd with $s_Y = s_X = P$. The result is determined by the initial play of X.

We now consider the Press–Dyson representation, using the normalized payoff vectors of Equation (2.1). If $\tilde{\mathbf{p}} = \alpha \mathbf{S}_X + \beta \mathbf{S}_Y + \gamma \mathbf{1} + \delta \mathbf{e}_{23}$ is the X Press–Dyson vector of a plan $\mathbf{p}$, then it must satisfy two sorts of constraints.

The *sign constraints* require that the first two entries be non-positive and the last two be non-negative. That is,

$$\begin{aligned} &(\alpha + \beta)R + \gamma \le 0, \\ &\beta + \gamma + \delta \le 0, \\ &\alpha + \gamma + \delta \ge 0, \\ &(\alpha + \beta)P + \gamma \ge 0. \end{aligned} \tag{2.7}$$

Lemma 2.3. *If $\tilde{\mathbf{p}} = \alpha \mathbf{S}_X + \beta \mathbf{S}_Y + \gamma \mathbf{1} + \delta \mathbf{e}_{23}$ satisfies the sign constraints, then*

$$\begin{aligned} &\alpha + \beta \le 0 \quad \textit{and} \quad \gamma \ge 0, \\ &\alpha + \beta = 0 \Leftrightarrow \gamma = 0. \end{aligned} \tag{2.8}$$

Proof. Subtracting the fourth inequality from the first we see that $(\alpha + \beta)(R - P) \le 0$ and so $R - P > 0$ implies $\alpha + \beta \le 0$. Then the fourth inequality and $P > 0$ imply $\gamma \ge 0$. The first and fourth imply $\alpha + \beta = 0$ iff $\gamma = 0$. □

Remark: Notice that both $\tilde{p}_1$ and $\tilde{p}_4$ vanish iff $\alpha + \beta = \gamma = 0$. These are the cases when plan $\mathbf{p}$ is both agreeable and firm.

In addition, the entries of an X Press–Dyson vector have absolute value at most 1. These are the *size constraints*. If a vector satisfies the sign constraints, then, multiplying by a sufficiently small positive number, we obtain the size constraints as well. Any vector in $\mathbb{R}^4$ that satisfies both the sign and the size constraints is an X Press–Dyson vector. Call $\mathbf{p}$ a *top plan* if $|\tilde{p}_i| = 1$ for some i. For any plan $\mathbf{p}$, other than Repeat, which has X Press–Dyson vector $\mathbf{0}$, $\mathbf{p} = a(\mathbf{p}^t) + (1 - a)$Repeat for a unique top plan $\mathbf{p}^t$ and a unique positive $a \le 1$. Equivalently, $\tilde{\mathbf{p}} = a\tilde{\mathbf{p}}^t$.

Observe that $\mathbf{p}$ is agreeable iff $\tilde{p}_1 = 0$ and so iff $(\alpha + \beta)R + \gamma = 0$. In that case, $\beta = -\alpha - \gamma R^{-1}$. Substituting into Equation (1.19), we obtain the following corollary of Theorem 1.7.

Corollary 2.4. *If $\mathbf{p}$ is an agreeable plan with X Press–Dyson vector $\tilde{\mathbf{p}} = \alpha \mathbf{S}_X + \beta \mathbf{S}_Y + \gamma \mathbf{1} + \delta \mathbf{e}_{23}$, then the payoffs with any limit distribution satisfy the following version of the* Press–Dyson equation.

$$\gamma R^{-1} s_Y + \alpha(s_Y - s_X) - \delta v_{23} = \gamma. \tag{2.9}$$

Now we justify the description in Theorem 1.8. Notice that if we label by S^0_X and S^0_Y our original payoff vectors before normalization, then $\mathbf{S}^0_X = (T-S)\mathbf{S}_X + S\mathbf{1}$, $S^0_Y = (T-S)\mathbf{S}_Y + S\mathbf{1}$ and so if $(\alpha, \beta, \gamma, \delta)$ are the coordinates of $\tilde{\mathbf{p}}$ with respect to the basis $\{\mathbf{S}_X, \mathbf{S}_Y, \mathbf{1}, \mathbf{e}_{23}\}$, then $(\alpha^0, \beta^0, \gamma^0, \delta^0) = (\alpha/(T-S), \beta/(T-S), \gamma - (\alpha+\beta)S/(T-S), \delta)$ are the coordinates with respect to $\{S^0_X, S^0_Y, \mathbf{1}, \mathbf{e}_{23}\}$. In particular, $\alpha >= k\delta$ iff $\alpha^0 >= k\delta^0/(T-S)$ for any k. Furthermore, the constant $k = (T-S)/(2R-(T+S))$ is independent of normalization. So it suffices to prove the normalized version of the theorem, which is the following.

Theorem 2.5. *Assume that* $\mathbf{p} = (p_1, p_2, p_3, p_4)$ *is an agreeable plan with X Press–Dyson vector* $\tilde{\mathbf{p}} = \alpha\mathbf{S}_X + \beta\mathbf{S}_Y + \gamma\mathbf{1} + \delta\mathbf{e}_{23}$. *Assume that* $\mathbf{p}$ *is not* Repeat, *i.e.*, $(\alpha, \beta, \gamma, \delta) \neq (0,0,0,0)$. *The plan* $\mathbf{p}$ *is of Nash type iff*

$$\max(\delta, (2R-1)^{-1}\delta) \leq \alpha. \tag{2.10}$$

The plan $\mathbf{p}$ *is good iff, in addition, the inequality is strict.*

Proof. Since $\beta = -\alpha - \gamma R^{-1}$, we have

$$\begin{aligned}(1-p_2) &= -\tilde{p}_2 = -\beta - \gamma - \delta = \alpha + \frac{1-R}{R}\gamma - \delta,\\ p_3 &= \tilde{p}_3 = \alpha + \gamma + \delta, \quad p_4 = \tilde{p}_4 = \frac{R-P}{R}\gamma.\end{aligned} \tag{2.11}$$

The inequality $(1-R)p_3 \leq R(1-p_2)$ becomes $(1-R)(\alpha+\gamma+\delta) \leq R\alpha + (1-R)\gamma - R\delta$. This reduces to $\delta \leq (2R-1)\alpha$. Similarly, the inequality $(1-R)p_4 \leq (R-P)(1-p_2)$ reduces to $\delta \leq \alpha$. □

Remarks: (a) Thus, when $\delta \leq 0$, $\mathbf{p}$ is good iff $\delta < \alpha$. When $\delta > 0$, $\mathbf{p}$ is good iff $\delta/(2R-1) < \alpha$.

(b) From the proof we see that the equalizer case, when both inequalities of Equation (2.3) are equations, occurs when $\delta = \alpha = (2R-1)^{-1}\delta$. Since $2R-1 < 1$, this reduces to $0 = \delta = \alpha$.

In the ZDS case, when $\delta = 0$, we can rewrite Equation (2.9) as

$$\kappa \cdot (s_X - R) = s_Y - R \tag{2.12}$$

with $\kappa = \alpha R/(\gamma + \alpha R)$. Thus, the condition $\alpha > 0$ is equivalent to $0 < \kappa \leq 1$. In [8], these plans are introduced and called *complier strategies*. The equation and the condition $\kappa > 0$ make it clear that such plans are good. In addition, if $s_Y < R$, then it follows that $s_X \leq s_Y$ with strict inequality when $\gamma > 0$ and so $\kappa < 1$. The strategy ZGTFT-2 analyzed in Stewart and Plotkin [16] is an example of a complier plan. When X plays a complier plan, then either both s_X and s_Y are equal to R or else both are below R. This is not true for good plans in general. If X plays the good plan Grim $= (1,0,0,0)$ and Y plays $(0,1,1,1)$, then fixation at *dc* occurs with $v = (0,0,1,0)$ and so with $s_Y = 0$ ($<R$ as required by Corollary 2.2), but with $s_X = 1 > R$.

Let us look at the geometry of the Press–Dyson representation.

We begin with the *exceptional plans* that are defined by $\gamma = \alpha + \beta = 0$. The sign constraints yield $\alpha = -\beta \geq |\delta|$ and $\tilde{\mathbf{p}} = (0, \delta - \alpha, \delta + \alpha, 0)$. As remarked after Lemma 2.3, the exceptional plans are exactly those plans that are both agreeable and firm. In the xy plane with $x = \tilde{p}_2$ and $y = \tilde{p}_3$, they form a square with vertices: Repeat($\tilde{\mathbf{p}} = (0,0,0,0)$), Grim($\tilde{\mathbf{p}} = (0,-1,0,0)$), TFT ($\tilde{\mathbf{p}} = (0,-1,1,0)$), and what we will call Lame ($\tilde{\mathbf{p}} = (0,0,1,0)$).

Thus, Lame $= (1,1,1,0)$. The top plans consist of the segment that connects TFT with Grim together with the segment that connects TFT with Lame.

On the Grim-TFT segment, $\delta - \alpha = -1$ and $0 \leq \delta + \alpha \leq 1$. That is, $\delta = \alpha - 1$ and $-1/2 \leq \delta \leq 0$. By Theorem 2.5, the plans in the triangle with vertices Grim, TFT, and Repeat are all good except for Repeat itself.

On the Lame-TFT segment, $-1 \leq \delta - \alpha \leq 0$ and $\delta + \alpha = 1$. That is, $\delta = 1 - \alpha$ and $1/2 \geq \delta \geq 0$. Such a plan is the mixture tTFT $+ (1-t)$Lame $= (1, 1-t, 1, 0)$ with $t = 2\alpha - 1$. By Theorem 2.1, this plan is good iff $t > (1-R)/R$. The plan on the TFT-Lame segment with $t = (1-R)/R$, and so with $2\alpha = R^{-1}$, we will call Edge $= (1, (2R-1)/R, 1, 0)$. The plans in the TFT-Edge-Repeat triangle that are not on the Edge-Repeat side are good plans. The plans in the complementary Edge-Lame-Repeat triangle are not good.

Now assume $\gamma > 0$, and define

$$\bar{\alpha} = \alpha/\gamma, \quad \bar{\beta} = \beta/\gamma, \quad \bar{\delta} = \delta/\gamma. \tag{2.13}$$

with the sign constraints

$$\begin{aligned} -P^{-1} &\leq \bar{\alpha} + \bar{\beta} \leq -R^{-1}, \\ \bar{\beta} &\leq -1 - \bar{\delta} \leq \bar{\alpha}. \end{aligned} \tag{2.14}$$

For any triple $(\bar{\alpha}, \bar{\beta}, \bar{\delta})$ that satisfies these inequalities, we obtain an X Press–Dyson vector $\tilde{\mathbf{p}} = \alpha \mathbf{S}_X + \beta \mathbf{S}_Y + \gamma \mathbf{1} + \delta \mathbf{e}_{23}$ that satisfies the size constraints as well by using $(\alpha, \beta, \gamma, \delta) = \gamma \cdot (\bar{\alpha}, \bar{\beta}, 1, \bar{\delta})$ with $\gamma > 0$ small enough. When we use the largest value of γ such that the size constraints hold, we obtain the top plan associated with $(\bar{\alpha}, \bar{\beta}, \bar{\delta})$. The others are mixtures of the top plan with Repeat. For a plan with this triple, the Press–Dyson equation (1.19) becomes

$$\bar{\alpha} s_X + \bar{\beta} s_Y + \bar{\delta} v_{23} + 1 = 0. \tag{2.15}$$

The points $(x, y) = (\bar{\alpha}, \bar{\beta})$ lie in the *Strategy Strip*. This consists of the points of the xy plane with $y \leq x$ and that lie on or below the line $x + y = -R^{-1}$ and on or above the line $x + y = -P^{-1}$. Then $\bar{\delta}$ must satisfy $-1 - x \leq \bar{\delta} \leq -1 - y$. Alternatively, we can fix $\bar{\delta}$ to be arbitrary and intersect the Strategy Strip with the fourth quadrant when the origin is at $(-1-\bar{\delta}, -1-\bar{\delta})$, i.e., the points with $y \leq -1 - \bar{\delta} \leq x$.

Together with the exceptional plans those with $(\bar{\alpha}, \bar{\beta})$ on the line $x + y = -R^{-1}$ are exactly the agreeable plans. Together with the exceptional plans those on the line $x + y = -P^{-1}$ are exactly the firm plans.

Let us look at the good ZDS's, i.e., the good plans with $\delta = 0$. In the exceptional case with $\gamma = 0$, the top good plan is TFT. When $\delta = 0$ and $\gamma > 0$, the good plans are those that satisfy $\bar{\alpha} + \bar{\beta} = -R^{-1}$ and $\bar{\alpha} > 0$. As mentioned above, these are the complier plans.

Proposition 2.6. *Given $\bar{\alpha} > 0$, the associated agreeable ZDS top plan is given by*

$$\mathbf{p} = \left(1, \frac{2R-1}{R(\bar{\alpha}+1)}, 1, \frac{R-P}{R(\bar{\alpha}+1)}\right). \tag{2.16}$$

Proof. The agreeable plan $\mathbf{p}$ with $\gamma, \bar{\alpha} > 0$ and $\bar{\delta} = 0$ has X Press–Dyson vector

$$\tilde{\mathbf{p}} = \left(0, -\gamma(\bar{\alpha} + R^{-1} - 1), \gamma(\bar{\alpha}+1), \gamma(1 - P \cdot R^{-1})\right). \tag{2.17}$$

With $\bar{\alpha}$ fixed, the largest value for γ so that the size constraints hold is $(\bar{\alpha}+1)^{-1}$. This easily yields Equation (2.16) for the top plan. □

When $\bar{\delta} = 0$, the vertical line $\bar{\alpha} = 0$ intersects the strip in points whose plans are all the *equalizers*, as discussed by Press and Dyson [15] and by Boerlijst, Nowak, and Sigmund [5]. Observe that with $\bar{\delta} = 0$, and $\bar{\alpha} = 0$, the Press–Dyson equation (2.15) becomes $\bar{\beta} s_Y + 1 = 0$, and so $s_Y = -\bar{\beta}^{-1}$ regardless of the choice of strategy for Y. The agreeable case has $\bar{\beta} = -R^{-1}$. The vertical line of equalizers cuts the line of agreeable plans, separating it into the unbounded ray with good plans and the segment with plans that are not even of Nash type.

Finally, we call a plan $\mathbf{p}$ *generous* when $p_2 > 0$ and $p_4 > 0$. That is, whenever Y defects, there is a positive probability that X will cooperate. The complier plans given by Equation (2.16) are generous.

Proposition 2.7. *Assume that X plays $\mathbf{p}$, a generous plan of Nash type. If Y plays plan $\mathbf{q}$ of Nash type and either (i) $\mathbf{q}$ is generous or (ii) $q_3 + q_4 > 0$, then $\{cc\}$ is the unique terminal set for the associated Markov matrix $\mathbf{M}$. Thus, $\mathbf{M}$ is convergent.*

Proof. Since $\mathbf{p}$ and $\mathbf{q}$ are both agreeable, $\{cc\}$ is a terminal set for $\mathbf{M}$.

Since $\mathbf{p}$ is of Nash type, it is not Repeat and so Equation (2.3) implies that $p_2 < 1$.

For the first case, we prove that if $p_1 = 1, p_2 < 1, p_4 > 0$ and $\mathbf{q}$ satisfies analogous conditions and not both $\mathbf{p}$ and $\mathbf{q}$ are of the form $(1, 0, 1, a)$, then $\mathbf{M}$ is convergent.

Recall that Y responds to *cd* using q_3 and to *dc* using q_2.

The assumptions $p_4, q_4 > 0$ imply that there is an edge from *dd* to *cc*, and so that *dd* is transient. There is an edge from *dc* to *dd* if $p_3 < 1$ since $q_2 < 1$. If $p_3 = 1$ and $q_2 > 0$, then there is an edge to *cc*. There remains the case that $p_3 = 1, q_2 = 0$ with the only edge from *dc* going to *cd*. Similarly, there is an edge from *cd* to either *dd* or *cc* except when $p_2 = 0, q_3 = 1$. Thus, the only case when $\mathbf{M}$ is not convergent is when both $\mathbf{p}$ and $\mathbf{q}$ are of the form $(1, 0, 1, a)$. In that case, $\{cd, dc\}$ is an additional terminal set. In particular, if either p_2 or q_2 is positive, then $\{cc\}$ is the only terminal set. This completes case (i). It also shows that if $\mathbf{p}$ is generous and $q_4 > 0$, then $\mathbf{M}$ is convergent.

To complete case (ii), we assume that $\mathbf{p}$ is generous and $q_3 > 0$. Since $p_2 > 0$ and $q_3 > 0$, there is an edge from cd to cc and so cd is transient. Since $p_4 > 0$, there is an edge from dd either to cc or to cd and so dd is transient. Finally, $q_2 < 1$ implies there is an edge from dc to cd or to dd. Thus, dc is transient as well. □

This result indicates the advantage which the good plans that are generous have over the good exceptional plans like Grim and TFT. The latter are firm as well as agreeable. Playing them against each other yields a non-convergent matrix with both $\{cc\}$ and $\{dd\}$ as terminal sets. Initial cooperation does lead to immediate fixation at cc, but an error might move the sequence of outcomes on a path leading to another terminal set. When generous good plans are used against each other, $\{cc\}$ is the unique terminal set. Eventual fixation at cc occurs whatever the initial distribution is, and if an error occurs, then the strategies move the successive outcomes along a path that returns to cc. It is easy to compute the expected number of steps T_z from transient state z to cc.

$$T_z = 1 + \Sigma_{z'} p_{zz'} T_{z'}, \tag{2.18}$$

where we sum over the three transient states and $p_{zz'}$ is the probability of moving along an edge from z to z'. Thus, with $\mathbf{M}' = \mathbf{M} - I$, we obtain the formula for the vector $\mathbf{T} = (T_2, T_3, T_4)$:

$$\mathbf{M}'_t \cdot \mathbf{T} = -\mathbf{1}, \tag{2.19}$$

where $\mathbf{M}'_t$ is the invertible 3×3 matrix obtained from $\mathbf{M}'$ by omitting the first row and column.

Consider the case when X and Y both use the plan given by Equation (2.16), so that $\mathbf{p} = \mathbf{q} = (1, p_2, 1, p_4)$. The only edges coming from cd connect with cc or with dc and similarly for the edges from dc. Symmetry will imply that $T_{cd} = T_{dc}$. So with T this common value we obtain from Equation (2.18) $T = 1 + (1 - p_2)T$. Hence, from Equation (2.16), we get

$$T = T_{cd} = T_{dc} = \frac{1}{p_2} = \frac{\bar{\alpha} + 1}{2 - R^{-1}}. \tag{2.20}$$

Thus, the closer the plan is to the equalizer plan with $\bar{\alpha} = 0$, the shorter the expected recovery time from an error leading to a dc or cd outcome. From Equation (2.18), one can see that

$$T_{dd} = 1 + 2p_4(1 - p_4) \cdot T + (1 - p_4)^2 \cdot T_{dd}. \tag{2.21}$$

We won't examine this further as arriving at dd from cc implies errors on the part of both players.

Of course, one might regard such departures from cooperation not as noise or error but as ploys. Y might try a rare move to *cd* in order to pick up the temptation payoff for defection as an occasional bonus. But if this is strategy rather than error, it means that Y is departing from the good plan to one with q_1 a bit less than 1. Corollary 2.2(a) implies that Y loses by executing such a ploy.

3 Competing zero-determinant strategies

We now examine the ZDS's in more detail. Recall that a plan $\mathbf{p}$ is a ZDS when $\delta = 0$ in the Press–Dyson decomposition of the X Press–Dyson vector $\tilde{\mathbf{p}} = \mathbf{p} - \mathbf{e}_{12}$. With Normalization (2.1), the inverse matrix of $(\mathbf{S}_X\ \mathbf{S}_Y\ \mathbf{1}\ \mathbf{e}_{23})$ is

$$\frac{-1}{2(R-P)}\begin{pmatrix} -1 & R-P & P-R & 1 \\ -1 & P-R & R-P & 1 \\ 2P & 0 & 0 & -2R \\ 1-2P & P-R & P-R & 2R-1 \end{pmatrix} \tag{3.1}$$

and so if $\tilde{\mathbf{p}} = \alpha\mathbf{S}_X + \beta\mathbf{S}_Y + \gamma\mathbf{1} + \delta\mathbf{e}_{23}$,

$$2(R-P)\delta = (2P-1)\tilde{p}_1 + (R-P)(\tilde{p}_2 + \tilde{p}_3) - (2R-1)\tilde{p}_4. \tag{3.2}$$

Thus, for example, if $R + P = 1$, both All$D$ with $\tilde{\mathbf{p}} = (-1,-1,0,0)$ and All$C$ with $\tilde{\mathbf{p}} = (0,0,1,1)$ are ZDS.

The exceptional ZDSs, which have $\gamma = 0$ as well as $\delta = 0$, are mixtures of TFT and Repeat. Otherwise, $\gamma > 0$ and we can write $\tilde{\mathbf{p}} = \gamma(\bar{\alpha}\mathbf{S}_X + \bar{\beta}\mathbf{S}_Y + \mathbf{1})$. When $(\bar{\alpha}, \bar{\beta})$ lies in the ZDS strip defined by

$$\text{ZDS strip} = \{(x,y) : x \ge -1 \ge y \quad \text{and} \quad -R^{-1} \ge x + y \ge -P^{-1}\}, \tag{3.3}$$

then the sign constraints are satisfied. The size constraints hold as well when $\gamma > 0$ is small enough. For Z with $P \le Z \le R$ the intersection of the ZDS strip with the line $x + y = -Z^{-1}$ is a *value line* in the strip.

Lemma 3.1. *Assume that* $(\bar{\alpha}, \bar{\beta})$ *in the ZDS strip, with* $\bar{\alpha} + \bar{\beta} = -Z^{-1}$. *We then have* $-\bar{\beta} \ge \max(1, |\bar{\alpha}|)$ *and* $-\bar{\beta} = |\bar{\alpha}|$ *iff* $\bar{\alpha} = \bar{\beta} = -1$. *If* $(\bar{a}, \bar{b})$ *is also in the strip, then* $D = \bar{\beta}\bar{b} - \bar{\alpha}\bar{a} \ge 0$ *with equality iff* $\bar{\alpha} = \bar{\beta} = \bar{a} = \bar{b} = -1$.

Proof. By definition of Z, $-\bar{\beta} = \bar{\alpha} + Z^{-1} > \bar{\alpha}$. Also, the sign constraints imply $-\bar{\beta} \ge 1 \ge -\bar{\alpha}$, and so $-\bar{\beta} \ge -\bar{\alpha}$ with equality iff $\bar{\alpha} = \bar{\beta} = -1$. $D \ge (-\bar{\beta})(-\bar{b}) - |\bar{\alpha}||\bar{a}| \ge 0$ and the inequality is strict unless $\bar{\alpha} = \bar{\beta} = \bar{a} = \bar{b} = -1$. □

Remark: Because $R > 1/2$ it is always true that $-R^{-1} > -2$, but $-2 \ge -P^{-1}$ iff $1/2 \ge P$. Hence, $(-1,-1)$ is in the ZDS strip iff $1/2 \ge P$.

For a ZDS, we can usefully transform the Press–Dyson equation (2.15).

Proposition 3.2. *Assume that* X *uses plan* $\mathbf{p}$ *with* X *Press–Dyson vector* $\tilde{\mathbf{p}} = \gamma(\bar{\alpha}\mathbf{S}_X + \bar{\beta}\mathbf{S}_Y + \mathbf{1})$, $\gamma > 0$. *Let* $-Z^{-1} = \bar{\alpha} + \bar{\beta}$, *so that* $P \le Z \le R$.

For any general plan played by Y,

$$\bar{\alpha}Z(s_X - s_Y) = (s_Y - Z). \tag{3.4}$$

If $\kappa = \bar{\alpha}Z/(1 + \bar{\alpha}Z)$, *then* $1 > \kappa$ *and* κ *has the same sign as* $\bar{\alpha}$. *For any general plan played by* Y,

$$\kappa(s_X - Z) = (s_Y - Z). \tag{3.5}$$

Proof. Notice that $1 + \bar{\alpha}Z = -\bar{\beta}Z \ge Z \ge P > 0$. Multiplying Equation (2.15) by Z and substituting for $\bar{\beta}Z$ easily yields Equation (3.4) and then Equation (3.5). □

If $\bar{\alpha} = 0$, which is the equalizer case, $s_Y = Z$ and s_X is undetermined. When $\bar{\alpha} > 0$, the payoffs s_X and s_Y are on the same side of Z, while they are on opposite sides when $\bar{\alpha} < 0$. To be precise, we have the following.

Corollary 3.3. *Assume that* X *uses a plan* $\mathbf{p}$ *with* X *Press–Dyson vector* $\tilde{\mathbf{p}} = \gamma(\bar{\alpha}\mathbf{S}_X + \bar{\beta}\mathbf{S}_Y + \mathbf{1})$, $\gamma > 0$. *Let* $-Z^{-1} = \bar{\alpha} + \bar{\beta}$. *Assume that* Y *uses an arbitrary general plan.*

(a) *If* $\bar{\alpha} = 0$, *then* $s_Y = Z$. *If* $\bar{\alpha} \ne 0$, *then the following are equivalent*

(i) $s_Y = s_X$,

(ii) $s_Y = Z$,

(iii) $s_X = Z$.

(b) *If* $s_Y > s_X$, *then*

$$\begin{cases} \bar{\alpha} > 0, \Rightarrow\ Z > s_Y > s_X, \\ \bar{\alpha} = 0, \Rightarrow\ Z = s_Y > s_X, \\ \bar{\alpha} < 0, \Rightarrow\ s_Y > Z > s_X. \end{cases} \tag{3.6}$$

(c) *If* $s_X > s_Y$, *then*

$$\begin{cases} \bar{\alpha} > 0, \Rightarrow\ s_X > s_Y > Z, \\ \bar{\alpha} = 0, \Rightarrow\ s_X > s_Y = Z, \\ \bar{\alpha} < 0, \Rightarrow\ s_X > Z > s_Y. \end{cases} \tag{3.7}$$

Proof. (a) If $\bar{\alpha} = 0$, then $s_Y = Z$ by Equation (3.4). If $\bar{\alpha} \ne 0$, then (i) $\Leftrightarrow$ (ii) by Equation (3.4) and (ii) $\Leftrightarrow$ (iii) by Equation (3.5).

(b), (c) If $\bar{\alpha} \ne 0$, then by Equation (3.4) $s_Y - Z$ has the same sign as that of $\bar{\alpha}(s_X - s_Y)$. □

For $Z = R$, Equation (3.5) is Equation (2.12). When $\bar{\alpha} > 0$, these are the complier strategies, i.e., the generous, good plans described in Proposition 2.6.

For $Z = P, \bar{\alpha} > 0$, the plans are firm. These were considered by Press and Dyson who called them *extortion strategies*. The name comes from the observation that whenever Y chooses a strategy so that her payoff is above P, the bonus beyond P is divided between X and Y in a ratio of $1 : \kappa$. They point out that the best reply against such an extortion play by X is for Y is to play All$C = (1, 1, 1, 1)$, which gives X a payoff above R. At first glance, it seems hard to escape from this coercive effect. I believe that the answer is for Y to play a generous good strategy like the compliers above. With repeated play, each player receives enough data to estimate statistically the strategy used by the opponent. Y's good strategy represents a credible invitation for X to switch to an agreeable plan and receive R, or else be locked below R. Hence, it undercuts the threat from X to remain extortionate.

In order to compute what happens when both players use a ZDS, we need to examine the symmetry between the two players. Let Switch : $\mathbb{R}^4 \to \mathbb{R}^4$ be defined by Switch$(x_1, x_2, x_3, x_4) = (x_1, x_3, x_2, x_4)$. Notice that Switch interchanges the vectors $\mathbf{S}_X$ and $\mathbf{S}_Y$. If X uses $\mathbf{p}$ and Y uses $\mathbf{q}$, then recall that the response vectors used to build the Markov matrix $\mathbf{M}$ are $\mathbf{p}$ and Switch($\mathbf{q}$). Now suppose that the two players exchange plans so that X uses $\mathbf{q}$ and Y uses $\mathbf{p}$. Then the X response is $\mathbf{q}$ = Switch(Switch($\mathbf{q}$)) and the Y response is Switch($\mathbf{p}$). Hence, the new Markov matrix is obtained by transposing both the second and third rows and the second and third columns. It follows that if $\mathbf{v}$ was a stationary vector for $\mathbf{M}$, then Switch($\mathbf{v}$) is a stationary vector for the new matrix. Hence, Theorem 1.3 applied to the X Press–Dyson vector $\tilde{\mathbf{q}}$ implies that $0 = \langle \text{Switch}(\mathbf{v}) \cdot \tilde{\mathbf{q}} \rangle = \langle \mathbf{v} \cdot \text{Switch}(\tilde{\mathbf{q}}) \rangle$. Furthermore, if $\tilde{\mathbf{q}} = a\mathbf{S}_X + b\mathbf{S}_Y + g\mathbf{1} + \delta\mathbf{e}_{23}$, then $\text{Switch}(\tilde{\mathbf{q}}) = b\mathbf{S}_X + a\mathbf{S}_Y + g\mathbf{1} + \delta\mathbf{e}_{23}$.

For a plan $\mathbf{q}$, we define Y *Press–Dyson vector* $\tilde{\tilde{\mathbf{q}}} = \text{Switch}(\tilde{\mathbf{q}}) = \text{Switch}(\mathbf{q}) - \mathbf{e}_{13}$, where $\mathbf{e}_{13} = (1, 0, 1, 0)$. For any general plan for X and any limiting distribution $\mathbf{v}$ when Y uses $\mathbf{q}$ we have $\langle \mathbf{v} \cdot \tilde{\tilde{\mathbf{q}}} \rangle = 0$. The plan $\mathbf{q}$ is a ZDS associated with $(\bar{a}, \bar{b})$ in the ZDS strip when $\tilde{\tilde{\mathbf{q}}} = g(\bar{b}\mathbf{S}_X + \bar{a}\mathbf{S}_Y + \mathbf{1})$ with some $g > 0$.

Now we compute what happens when X and Y use ZDS plans associated, respectively, with points $(\bar{\alpha}, \bar{\beta})$ and $(\bar{a}, \bar{b})$ in the ZDS strip. This means that for some $\gamma > 0, g > 0$, $\tilde{\mathbf{p}} = \gamma(\bar{\alpha}\mathbf{S}_X + \bar{\beta}\mathbf{S}_Y + \mathbf{1})$ and $\tilde{\tilde{\mathbf{q}}} = g(\bar{b}\mathbf{S}_X + \bar{a}\mathbf{S}_Y + \mathbf{1})$. We obtain two Press–Dyson equations that hold simultaneously

$$\begin{aligned} \bar{\alpha}s_X + \bar{\beta}s_Y &= -1, \\ \bar{b}s_X + \bar{a}s_Y &= -1. \end{aligned} \tag{3.8}$$

If $\bar{\alpha} = \bar{\beta} = \bar{a} = \bar{b} = -1$, which we will call a Vertex plan $= \gamma(2(1-R), 1, 0, 1-2P)$, then the two equations are the same. Following the Remark after Lemma 3.1, a Vertex plan can occur only when $P \leq 1/2$. Clearly, $\{cd\}$ and $\{dc\}$ are both terminal sets when both players use a Vertex plan and so the payoffs depend upon the initial plays. If the two players use the same initial play as well as the same plan, then $s_X = s_Y$ and the single equation of (3.8) yields $s_X = s_Y = 1/2$.

Otherwise, Lemma 3.1 implies that the determinant $D = \bar{\beta}\bar{b} - \bar{\alpha}\bar{a}$ is positive, and by Cramer's rule, we get

$$\begin{aligned} &s_X = D^{-1}(\bar{a} - \bar{\beta}), \quad s_Y = D^{-1}(\bar{\alpha} - \bar{b}), \\ &\text{and so} \quad s_Y - s_X = D^{-1}[(\bar{\alpha} + \bar{\beta}) - (\bar{a} + \bar{b})]. \end{aligned} \tag{3.9}$$

Notice that s_X and s_Y are independent of γ and g.

Thus, when both X and Y use ZDS plans from the ZDS strip, these long-term payoffs depend only on the plans and so the results are independent of the choice of initial plays.

Proposition 3.4. *Assume that* $\tilde{\mathbf{p}} = \gamma(\bar{\alpha}\mathbf{S}_X + \bar{\beta}\mathbf{S}_Y + \mathbf{1})$ *and* $\tilde{\tilde{\mathbf{q}}} = g(\bar{b}\mathbf{S}_X + \bar{a}\mathbf{S}_Y + \mathbf{1})$. *Let* $\bar{\alpha} + \bar{\beta} = -Z_X^{-1}$ *and* $\bar{a} + \bar{b} = -Z_Y^{-1}$. *Assume that* $(-1, -1)$ *is not equal to both* $(\bar{\alpha}, \bar{\beta})$ *and* $(\bar{a}, \bar{b})$.

(a) *The points* $(\bar{\alpha}, \bar{\beta}), (\bar{a}, \bar{b})$ *lie on the same value line* $x + y = -Z^{-1}$, *i.e.,* $Z_X = Z_Y$, *iff* $s_X = s_Y$. *In that case,* $Z_X = s_X = s_Y = Z_Y$.

(b) $s_Y > s_X$ *iff* $Z_X > Z_Y$.

(c) *Assume* $Z_X > Z_Y$. *The following implications hold.*

$$\begin{cases} \bar{\alpha} > 0, \Rightarrow \; Z_X > s_Y > s_X. \\ \bar{\alpha} = 0, \Rightarrow \; Z_X = s_Y > s_X. \\ \bar{\alpha} < 0, \Rightarrow \; s_Y > Z_X > s_X. \end{cases}$$
$$\begin{cases} \bar{a} > 0, \Rightarrow \; s_Y > s_X > Z_Y. \\ \bar{a} = 0, \Rightarrow \; s_Y > s_X = Z_Y. \\ \bar{a} < 0, \Rightarrow \; s_Y > Z_Y > s_X. \end{cases} \tag{3.10}$$

Proof. We are excluding by assumption the case when both players use Vertex plans and so we have $D > 0$.

(a) Assume $Z_X = Z_Y$. From Equation (3.9), we see that $s_Y - s_X = 0$.
When $s_X = s_Y$, Corollary 3.3(a) implies that $s_X = s_Y = Z_X$. By using the XY symmetry, we see that the common value is Z_Y as well. Hence, $Z_X = Z_Y$ and the points lie on the same line.

(b) Since $D > 0$, (b) follows from Equation (3.9).

(c) From (b), $s_Y - s_X > 0$. The first part follows from Equation (3.6) with $Z = Z_X$. The second follows from Equation (3.7) by using the XY symmetry with $\bar{\alpha}$, $\bar{\beta}$, Z replaced by $\bar{a}$, $\bar{b}$, Z_Y. □

Remark: If both players use a Vertex plan and the same initial play, then (a) holds with $Z_X = Z_Y = 1/2$.

4 Dynamics among zero-determinant strategies

In this section, we move beyond the classical question that motivated our original interest in good strategies. We consider now the evolutionary dynamics among memory one strategies. We follow Chapter 9 of Hofbauer and Sigmund [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 $\mathcal{I}$. When players X and Y use strategies with index $i, j \in \mathcal{I}$, respectively, the payoff to player X is given by A_{ij} and the payoff to Y by 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 π_i denote the ratio of the number of i players to the total population. The frequency vector $\{\pi_i\}$ lives in the unit simplex $\Delta \subset \mathbb{R}^{\mathcal{I}}$, i.e., the entries are non-negative and sum to 1. The vertex $v(i)$ associated with $i \in \mathcal{I}$ corresponds to a population consisting entirely of i players. We assume the population is large so that we can regard π 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 ρ of the members of the population. So the i player is replaced by $1 + (\rho + A_{ij})dt$ i players. Averaging over the current population distribution, the expected relative reproductive rate for the subpopulation of i players is $\rho + A_{i\pi}$, where

$$\begin{aligned} A_{i\pi} &= \Sigma_{j \in \mathcal{I}} \pi_j A_{ij} \quad \text{and} \\ A_{\pi\pi} &= \Sigma_{i \in \mathcal{I}} \pi_i A_{i\pi} = \Sigma_{i,j \in \mathcal{I}} \pi_i \pi_j A_{ij}. \end{aligned} \tag{4.1}$$

The resulting dynamical system on Δ is given by the *Taylor–Jonker game dynamics equations* introduced by Taylor and Jonker [18].

$$\frac{d\pi_i}{dt} = \pi_i (A_{i\pi} - A_{\pi\pi}). \tag{4.2}$$

This system is an example of the *replicator equations* studied in great detail by Hofbauer and Sigmund [9].

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

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

Each nonempty subset $\mathcal{J}$ of $\mathcal{I}$ determines the *face* $\Delta_{\mathcal{J}}$ of the simplex consisting of those $\pi \in \Delta$ such that $\pi_i = 0$ for all $i \notin \mathcal{J}$. Each face of the simplex is invariant because $\pi_i = 0$ implies that $d\pi_i/dt = 0$. In particular, for each $i \in \mathcal{I}$, the vertex $v(i)$, which represents fixation at the i strategy, is an equilibrium. In general, π is an equilibrium when, for all $i, j \in \mathcal{I}$, π_i, $\pi_j > 0$ imply $A_{i\pi} = A_{j\pi}$. This implies that $A_{i\pi} = A_{\pi\pi}$ for all i such that $\pi_i > 0$. That is, for all i in the *support* of π.

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

$$\omega(\pi) = \bigcap_{t>0} \overline{\{\pi(s) : s \geq t\}}. \tag{4.3}$$

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

Definition 4.1. We call a strategy i^* an *evolutionarily stable strategy* (hereafter, an ESS) when

$$A_{ji^*} < A_{i^*i^*} \quad \textit{for all } j \neq i^* \textit{ in } \mathcal{I}. \tag{4.4}$$

We call a strategy i^* an *evolutionarily unstable strategy* (hereafter, an EUS) when

$$A_{ji^*} > A_{i^*i^*} \quad \textit{for all } j \neq i^* \textit{ in } \mathcal{I}. \tag{4.5}$$

The ESS condition above is really a special case of a more general notion, see page 63 of [9], and is referred to there as a *strict Nash equilibrium*. We will not need the generalization and we use the term to avoid confusion with the strategies of Nash type considered in the previous sections.

Proposition 4.2. *If i^* is an ESS, then the vertex $v(i^*)$ is an attractor, i.e., a locally stable equilibrium, for System (4.2). In fact, there exists $\epsilon > 0$ such that*

$$1 > \pi_{i^*} \geq 1 - \epsilon \implies \frac{d\pi_{i^*}}{dt} > 0. \tag{4.6}$$

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

If i^* is an EUS, then the vertex $v(i^*)$ is a repellor, i.e., a locally unstable equilibrium, for System (4.2). In fact, there exists $\epsilon > 0$ such that

$$1 > \pi_{i^*} \geq 1 - \epsilon \implies \frac{d\pi_{i^*}}{dt} < 0. \tag{4.7}$$

Thus, near the equilibrium $v(i^*)$, $\pi_{i^*}(t)$ decreases monotonically, until the system enters, and then remains in the region where $\pi_{i^*} < 1 - \epsilon$.

Proof. When i^* is an ESS, $A_{i^*i^*} > A_{ji^*}$ for all $j \neq i^*$. It then follows for $\epsilon > 0$ sufficiently small that $\pi_{i^*} \geq 1-\epsilon$ implies $A_{i^*\pi} > A_{j\pi}$ for all $j \neq i^*$. If also $1 > \pi_{i^*}$, then $A_{i^*\pi} > A_{\pi\pi}$. So Equation (4.2) implies Equation (4.6).

The EUS case is similar. Notice that no solution path can cross $\Delta \cap \{\pi_{i^*} = 1 - \epsilon\}$ from $\{\pi_{i^*} < 1 - \epsilon\}$. □

Definition 4.3. 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

$$A_{jk} \leq A_{ik} \quad \text{for all } k \in \mathcal{J}, \tag{4.8}$$

and the inequality is strict for either $k = i$ or $k = j$. If the inequalities are strict for all k, then we say that i *dominates* j in $\mathcal{J}$.

We say that $i \in \mathcal{J}$ dominates a sequence $\{j_1, \ldots, j_n\}$ in $\mathcal{J}$ when i dominates j_1 in $\mathcal{J}$ and for $p = 2, \ldots, n$, i dominates j_p in $\mathcal{J} \setminus \{j_1, \ldots, j_{p-1}\}$.

When $\mathcal{J}$ equals all of $\mathcal{I}$, we will omit the phrase "in $\mathcal{J}$."

For $i, j \in \mathcal{I}$, define the set Q_{ij} and on it the real valued function L_{ij} by

$$\begin{aligned} Q_{ij} &= \{\pi \in \Delta : \pi_i, \pi_j > 0\} \\ L_{ij}(\pi) &= \ln(\pi_i) - \ln(\pi_j). \end{aligned} \tag{4.9}$$

Lemma 4.4. (a) *If i weakly dominates j, then $dL_{ij}/dt > 0$ on the set Q_{ij}.*
(b) *If i dominates j in $\mathcal{J}$, then there exists $\epsilon > 0$ such that $dL_{ij}/dt > 0$ on the set $Q_{ij} \cap \{\pi \in \Delta : \Sigma_{k\notin\mathcal{J}}\pi_k \leq \epsilon\}$.*

Proof. Observe that

$$dL_{ij}/dt = A_{i\pi} - A_{j\pi} = \Sigma_{k\in\mathcal{I}}\pi_k(A_{ik} - A_{jk}) \tag{4.10}$$

(a) Since $\pi_i, \pi_j > 0$ in Q_{ij} and $A_{ik} - A_{jk} \geq 0$ for all k with strict inequality for $k = i$ or $k = j$, it follows that the derivative is positive.
(b) Define

$$\begin{aligned} m &= \min\{A_{ik} - A_{jk} : k \in \mathcal{J}\} > 0, \\ M &= \max\{|A_{ik} - A_{jk}| : k \notin \mathcal{J}\}, \\ \pi_{\mathcal{J}} &= \Sigma_{k\in\mathcal{J}}\pi_k, \\ \pi_{k|\mathcal{J}} &= \pi_k/\pi_{\mathcal{J}} \quad \text{for } k \in \mathcal{J}. \end{aligned} \tag{4.11}$$

Observe that $\Sigma_{k\notin\mathcal{J}}\pi_k = 1 - \pi_{\mathcal{J}}$.

For any $\pi \in Q_{ij}$

$$\begin{aligned} A_{i\pi} - A_{j\pi} &= \pi_{\mathcal{J}}\Sigma_{k\in\mathcal{J}}\pi_{k|\mathcal{J}}(A_{ik} - A_{jk}) + \Sigma_{k\notin\mathcal{J}}\pi_k(A_{ik} - A_{jk}) \\ &\geq \pi_{\mathcal{J}}m - (1 - \pi_{\mathcal{J}})M. \end{aligned} \tag{4.12}$$

So if ϵ is chosen with $0 < \epsilon < m/(m + M)$, then $A_{i\pi} - A_{j\pi} > 0$ when $\pi \in Q_{ij} \cap \{\pi \in \Delta : (1 - \pi_{\mathcal{J}}) \leq \epsilon\}$. □

Lemma 4.5. *If $\pi(t)$ is a solution path with $\pi(0) \in Q_{ij}$ and there exists $T \in \mathbb{R}$ such that $dL_{ij}/dt > 0$ on the set $Q_{ij} \cap \overline{\{\pi(t) : t \geq T\}}$, then*

$$\text{Lim}_{t\to\infty}\pi_j(t) = 0. \tag{4.13}$$

Proof. By assumption, $L_{ij}(\pi(t))$ is a strictly increasing function of t for $t \geq T$. Thus, as a t tends to infinity, $L_{ij}(\pi(t))$ approaches $\ell = \sup\{L_{ij}(\pi(t)) : t \geq T\}$ with $L_{ij}(\pi(T)) < \ell \leq +\infty$.

We must prove that $\pi_j = 0$ on the omega limit set. Assume instead that $\pi^* \in \omega(\pi(0))$ with $\pi^*_j > 0$. If π^*_i were 0, then $L_{ij}(\pi(t))$ would not be bounded below on $\{\pi(t) : t \geq T\}$. Hence, π^* lies in Q_{ij} with $\ell = L_{ij}(\pi^*) < \infty$. So on the invariant set $\omega(\pi(0)) \cap Q_{ij}$, which contains π^* and so is nonempty, L_{ij} would be constantly $\ell < \infty$. Since this set is invariant, dL_{ij}/dt would equal zero. This contradicts our assumption that the derivative is positive on $\omega(\pi(0)) \cap Q_{ij}$. □

Proposition 4.6. *For $i \in \mathcal{I}$, let $\pi(t)$ be a solution path with $\pi_i(0) > 0$*
(a) *If i weakly dominates j, then* $\lim_{t\to\infty} \pi_j(t) = 0$.
(b) *If i dominates the sequence* $\{j_1, \ldots, j_n\}$ *then for* $j = j_1, \ldots, j_n$, $\lim_{t\to\infty} \pi_j(t) = 0$.

Proof. (a) If $\pi_j(0) = 0$, then $\pi_j(t) = 0$ for all t and so the limit is 0. Hence, we may assume $\pi_j(0) > 0$ and so that $\pi(0) \in Q_{ij}$. By Lemma 4.4(a), $dL_{ij}/dt > 0$ on Q_{ij} and so Lemma 4.5 implies $\mathrm{Lim}_{t\to\infty}\pi_j(t) = 0$.
(b) We prove the result by induction on n.
By part (a) $\lim_{t\to\infty} \pi_j(t) = 0$, for $j = j_1$.
Now assume the limit result is true for $j = j_1, \ldots, j_{p-1}$ with $1 < p \leq n$. We prove the result for $j = j_p$.
Let $\mathcal{J} = \mathcal{I} \setminus \{j_1, \ldots, j_{p-1}\}$. By assumption, i dominates j_p in $\mathcal{J}$. Hence, with $j = j_p$, Lemma 4.4(b) implies that there exists $\epsilon > 0$ such that $dL_{ij}/dt > 0$ on the set $Q_{ij} \cap \{\pi \in \Delta : \Sigma_{k\notin\mathcal{J}}\pi_k \leq \epsilon\}$.
By induction hypothesis, there exists T such that $\Sigma_{k\notin\mathcal{J}}\pi_k(t) \leq \epsilon$ for all $t \geq T$. Hence, $\overline{\{\pi(t) : t \geq T\}} \subset \{\pi : \Sigma_{k\notin\mathcal{J}}\pi_k(t) \leq \epsilon\}$.
As in part (a), we can assume $\pi \in Q_{ij}$ and then apply Lemma 4.5 to conclude $\lim_{t\to\infty} \pi_j(t) = 0$. This completes the inductive step. □

Now we specialize to the iterated Prisoner's Dilemma. By a *strategy*, we will mean a plan **p** together with an initial play, pure or mixed. Recall that a good (or agreeable) strategy is a good (respectively agreeable) plan together with initial cooperation.

To apply the Taylor–Jonker dynamics to our case, we suppose that $\mathcal{I}$ indexes a finite collection of strategies. We then use

$$A_{ij} = s_X \text{ so that } A_{ji} = s_Y. \tag{4.14}$$

That is, when the X player uses the *i* strategy and the Y player uses the *j* strategy, the players receive the payoffs s_X and s_Y, respectively, as additions to their reproductive rate. When the associated Markov matrix is convergent, there is a unique terminal set, and the long-term payoffs, s_X, s_Y depend only on the plans and not on the initial plays.

Theorem 4.7. *Let $\mathcal{I}$ index a finite set of strategies for the iterated Prisoner's Dilemma. Suppose that associated with $i^* \in \mathcal{I}$ is a good strategy $\mathbf{p}^{i*}$. If for no other $j \in \mathcal{I}$ is the plan $\mathbf{p}^j$ agreeable, then i^* is an ESS for the associated game $\{A_{ij} : i, j \in \mathcal{I}\}$ and so the vertex $v(i^*)$ is an attractor for the dynamic.*

Proof. Since i^* is associated with an agreeable strategy, $A_{i^*i^*} = R$. Since $\mathbf{p}^{i^*}$ is good and $\mathbf{p}^j$ is not agreeable for $j \neq i^*$, it follows from Corollary 2.2(a) that $A_{ji^*} < R$ for $j \neq i^*$. Thus, i^* is an ESS. □

There are other cases of ESS that are far from good.

Lemma 4.8. (a) *Assume that* X *uses a plan* $\mathbf{p} = (p_1, p_2, 0, 0)$ *with* $p_1, p_2 < 1$. *If* Y *uses any plan* $\mathbf{q}$ *that is not firm, then*

$$s_Y < P < s_X. \tag{4.15}$$

(b) *Assume that* X *uses a plan* $\mathbf{p} = (1, p_2, 0, 0)$ *with* $p_2 < 1$. *If* Y *uses any plan* $\mathbf{q}$ *that is neither firm nor agreeable, then Equation (4.15) holds.*

(c) *Assume* $P < 1/2$ *and that* X *uses a firm, non-exceptional ZDS with* $\bar{\alpha} < 0$. *If* Y *uses any plan* $\mathbf{q}$ *that is not firm, then Equation (4.15) holds.*

Proof. (a) and (b) Since $p_3 = p_4 = 0$, the set $\{dc, dd\}$ is closed. If $q_4 > 0$, then $\{dd\}$ is not closed and so is not a terminal set.

(a) Since $p_2 < 1$, there is an edge from *cd* to either *dc* or *dd*. Hence, *cd* is transient. Similarly, $p_1 < 1$ implies that *cc* is transient. Hence, for any stationary distribution $\mathbf{v}$, $v_1 = v_2 = 0$. Since $\mathbf{q}$ is not firm, $q_4 > 0$ and so $v_4 < 1$. Hence, $s_Y = v_4 P < P$ and $s_X = v_3 + v_4 P = (1 - v_4) + v_4 P > P$.

(b) As before, $p_2 < 1$ implies that *cd* is transient. Now $\mathbf{q}$ is not agreeable and so $q_1 < 1$. This implies that there is an edge from *cc* to the transient state *cd* and so *cc* is transient. The proof is completed as in (a).

(c) Because $P < 1/2$, the smallest entry in $1/2(\mathbf{S}_X + \mathbf{S}_Y)$ is P and so $1/2(s_X + s_Y) \leq P$ can only happen when $v_4 = 1$, which implies that $s_X = s_Y = P$. This requires that Y play a firm plan so that $\{dd\}$ is a terminal set. Compare Proposition 1.1.

From Equation (3.5) we see that with $\bar{\alpha} + \bar{\beta} = -Z^{-1}$ and $\kappa = \bar{\alpha}Z/(1 + \bar{\alpha}Z)$

$$\frac{1}{2}(1 + \kappa)(s_X - Z) = \left(\frac{1}{2}(s_X + s_Y) - Z\right). \tag{4.16}$$

When $Z = P$, $P < 1/2$ and $-1 \geq \bar{\alpha}$ imply that $(1 + \kappa) = (1 + 2\bar{\alpha}P)/(1 + \bar{\alpha}P) > 0$. Hence, $s_X \leq P$ implies that $1/2(s_X + s_Y) \leq P$. Since the Y plan is not firm, this does not happen. Hence, $s_X > P$. Since $\kappa < 0$, Equation (3.5) implies that $s_Y < P$. □

Theorem 4.9. *Let* $\mathcal{I}$ *index a finite set of strategies for the iterated Prisoner's Dilemma.*

(a) *Suppose that associated with* $i^* \in \mathcal{I}$ *is a plan* $\mathbf{p}^{i*} = (p_1, p_2, 0, 0)$ *with* $p_1, p_2 < 1$ *together with any initial play. If for no other* $j \in \mathcal{I}$ *is the plan* $\mathbf{p}^j$ *firm, then* i^* *is an ESS for the associated game* $\{A_{ij} : i, j \in \mathcal{I}\}$.

(b) *Suppose that associated with* $i^* \in \mathcal{I}$ *is a plan* $\mathbf{p}^{i*} = (1, p_2, 0, 0)$ *with* $p_2 < 1$ *together with any initial play. If for no other* $j \in \mathcal{I}$ *is the plan* $\mathbf{p}^j$ *either agreeable or firm, then* i^* *is an ESS for the associated game* $\{A_{ij} : i, j \in \mathcal{I}\}$.

(c) *Assume that* $P < 1/2$. *Suppose that associated with* $i^* \in \mathcal{I}$ *is a firm, non-exceptional ZDS with* $\bar{\alpha} < 0$ *together with any initial play. If for no other* $j \in \mathcal{I}$ *is the plan* $\mathbf{p}^j$ *firm, then* i^* *is an ESS for the associated game* $\{A_{ij} : i, j \in \mathcal{I}\}$.

Proof. (a) If both players use p^{i^*}, then there is an edge from *dc* to *dd* and so *dc*, *cd*, and *cc* are all transient. Thus, $\{dd\}$ is the unique terminal set and so

$A_{i^*i^*} = P$ regardless of the initial plays. By Lemma 4.8(a), $A_{ji^*} < P$ for all $j \neq i^*$.

(b) If both players use p^{i^*}, then there are edges from cd to dd and from dc to dd. The two terminal sets are $\{cc\}$ and $\{dd\}$. Hence, $R \geq A_{i^*i^*} \geq P$. This time Lemma 4.8(b) implies that $A_{ji^*} < P$ for any $j \neq i^*$.

(c) $A_{i^*i^*} = Z_{i^*} = P$ since the i^* plan is firm. Lemma 4.8(c) implies that $A_{ji^*} < P$ for any $j \neq i^*$. □

Thus, $p^{i^*} = \text{All}D = (0, 0, 0, 0)$ with any initial play is an ESS when played against plans that are not firm. If $p^{i^*} = \text{Grim} = (1, 0, 0, 0)$, then with any initial play i^* is an ESS when played against strategies that are neither agreeable nor firm.

At the other extreme, we have the following.

Theorem 4.10. *Let $\mathcal{I}$ index a finite set of strategies for the iterated Prisoner's Dilemma. Assume that $P < 1/2$. Suppose that associated with $i^* \in \mathcal{I}$ is an extortionate plan $\mathbf{p}^{i*}$ together with initial defection. That is, $\mathbf{p}^{i^*}$ is a firm ZDS with $\bar{\alpha} > 0$. If for no other $j \in \mathcal{I}$ is the plan $\mathbf{p}^j$ firm, then i^* is an EUS for the associated game $\{A_{ij} : i, j \in \mathcal{I}\}$ and so the vertex $v(i^*)$ is a repellor for the dynamic.*

Proof. Because $P < 1/2$, the smallest entry in $1/2(\mathbf{S}_\text{X} + \mathbf{S}_\text{Y})$ is P and so $s_\text{X}, s_\text{Y} \leq P$ implies $s_\text{X} = s_\text{Y} = P$ and this can only happen when $v_4 = 1$, which requires that Y play a firm plan so that $\{dd\}$ is a terminal set. Compare Proposition 1.1.

Since the i^* strategy is firm with initial defection, $A_{i^*i^*} = P$.

If Y uses any plan that is not firm then Equation (3.5) with $z = P$ and $\bar{\alpha} > 0$ shows that if $s_\text{Y} \leq P$ then $s_\text{X} \leq P$ as well. Because $P < 1/2$ this can only happen when $s_\text{X} = s_\text{Y} = P$ and $v_4 = 1$. But the Y plan is not firm. It follows that $s_\text{Y} > P$. Thus, for any $j \neq i^*$, $A_{ji^*} > A_{i^*i^*}$. This says that strategy i^* is an EUS. □

We now specialize to the case when all the strategies indexed by $\mathcal{I}$ are ZDS s with the exceptional strategies excluded. We can thus regard $\mathcal{I}$ as listing a finite set of points $(\bar{\alpha}_i, \bar{\beta}_i)$ in the ZDS strip and, except for the Vertex plans, we may disregard the initial plays. We define $Z_i = -(\bar{\alpha}_i + \bar{\beta}_i)^{-1}$. That is, the point $(\bar{\alpha}_i, \bar{\beta}_i)$ lies on the value line $x + y = -(Z_i)^{-1}$.

X uses $\mathbf{p}$ associated with $(\bar{\alpha}_i, \bar{\beta}_i)$ when $\tilde{\mathbf{p}} = \gamma_i(\bar{\alpha}_i\mathbf{S}_\text{X} + \bar{\beta}_i\mathbf{S}_\text{Y} + \mathbf{1})$ and Y uses $\mathbf{q}$ associated with $(\bar{\alpha}_j, \bar{\beta}_j)$ when $\tilde{\bar{\mathbf{q}}} = \gamma_j(\bar{\beta}_j\mathbf{S}_\text{X} + \bar{\alpha}_j\mathbf{S}_\text{Y} + \mathbf{1})$ for some $\gamma_i, \gamma_j > 0$. Notice the XY switch.

If both players use a Vertex plans with the same initial plays, then $(\bar{\alpha}_i, \bar{\beta}_i) = (-1, -1)$ and $A_{ii} = 1/2 = Z_i$. Recall that $(-1, -1)$ lies in the ZDS strip iff $P \leq 1/2$.

Otherwise, we apply Equation (3.8) with $(\bar{\alpha}, \bar{\beta}) = (\bar{\alpha}_i, \bar{\beta}_i)$ and $(\bar{a}, \bar{b}) = (\bar{\alpha}_j, \bar{\beta}_j)$. Then from Equation (3.9) we get, for $i \neq j$

$$\begin{aligned} A_{ij} &= s_\text{X} = K_{ij}(\bar{\alpha}_j - \bar{\beta}_i) \\ \text{with} \quad K_{ij} &= K_{ji} = (\bar{\beta}_i\bar{\beta}_j - \bar{\alpha}_i\bar{\alpha}_j)^{-1} > 0. \end{aligned} \tag{4.17}$$

Note that these payoffs are independent of the choice of γ_i, γ_j as well as the initial plays.

By Proposition 3.4(a)

$$A_{ii} = Z_i \quad \text{for all } i \in \mathcal{I}. \tag{4.18}$$

We begin with some degenerate cases. For convenience, we exclude the Vertex plans.

First, if all of the points $(\bar{\alpha}_i, \bar{\beta}_i)$ lie on the same value line $x + y = -Z^{-1}$, i.e., all the Z_is are equal, then by Proposition 3.4(a) $A_{ij} = Z$ for all i, j and so $d\pi/dt = 0$ and every population distribution is an equilibrium. In general, if for two strategies i, j $A_{ij} = A_{ji} = Z$, then by Proposition 3.4(a) both points lie on $x + y = -Z^{-1}$ and it follows that $A_{ii} = A_{jj} = Z$ as well. In general, if $\mathcal{I}_Z = \{i : Z_i = Z\}$ contains more than one $i \in I$, then the dynamics is degenerate on the face $\Delta_{\mathcal{I}_Z}$ of the simplex.

Second, if all of the points satisfy $\bar{\alpha}_i = 0$, then all the strategies are equalizer strategies. In this case, the payoff matrix need not be constant but A_{ij} depends only on j. This implies that for all i $A_{i\pi} = A_{\pi\pi}$ and so again $d\pi/dt = 0$ and every population distribution is an equilibrium.

We will now see that the line $\bar{\alpha} = 0$ separates different interesting dynamic behaviors.

Theorem 4.11. *Let $\mathcal{I}$ index a set of non-exceptional ZDS plans. Thus, each $i \in \mathcal{I}$ is associated with a point $(\bar{\alpha}_i, \bar{\beta}_i)$ in the ZDS strip and $\bar{\alpha}_i + \bar{\beta}_i = -(Z_i)^{-1}$.*

Assume either

Case (+): $\bar{\alpha}_i > 0$ for all $i \in \mathcal{I}$ and for some $i^* \in \mathcal{I}$, $Z_{i^*} > Z_j$ for all $j \neq i^*$;

or

Case (−): $\bar{\alpha}_i < 0$ for all $i \in \mathcal{I}$ and for some $i^* \in \mathcal{I}$, $Z_{i^*} < Z_j$ for all $j \neq i^*$.

The strategy i^* is an ESS, and if $\pi_{i^*}(0) > 0$, then the solution path converges to the vertex $v(i^*)$.

Proof. List the strategies $j_1, \dots, j_n$ of $\mathcal{I} \setminus \{i^*\}$ so that in Case(+) $Z_{j_1} \leq Z_{j_2} \leq \cdots \leq Z_{j_n} < Z_{i^*}$ and in Case(−) $Z_{j_1} \geq Z_{j_2} \geq \cdots \geq Z_{j_n} > Z_{i^*}$. For both cases, we apply Proposition 3.4. It first implies that if $Z_i = Z_j$ then

$$A_{ii} = Z_i = A_{ji} = A_{ij} = Z_j = A_{jj}. \tag{4.19}$$

Case(+): If $Z_i > Z_j$, then, because $\bar{\alpha}_i, \bar{\alpha}_j > 0$, Proposition 3.4 implies that

$$A_{ii} = Z_i > A_{ji} > A_{ij} > Z_j = A_{jj}. \tag{4.20}$$

Hence, if $Z_i > Z_k \geq Z_j$, then $A_{ii} > A_{ji}$ and $A_{ik} > A_{kk} \geq A_{jk}$.

It follows that i^* dominates the sequence $\{j_1, \dots, j_n\}$. Hence, Proposition 4.6(b) implies that $\mathrm{Lim}_{t\to\infty}\pi_j(t) = 0$ for $j = j_1, \dots, j_n$ when $\pi_{i^*}(0) > 0$. Consequently, $\pi_{i^*}(t) = 1 - \Sigma_{p=1}^{n}\pi_{j_p}(t)$ tends to 1. That is, $\pi(t)$ converges to $v(i^*)$.

Case(−): If $Z_i < Z_j$, then, because $\bar{\alpha}_i, \bar{\alpha}_j < 0$, Proposition 3.4 implies that

$$A_{ij} > Z_j = A_{jj} > A_{ii} = Z_i > A_{ji}. \tag{4.21}$$

It again follows that i^* dominates the sequence $\{j_1, \ldots, j_n\}$ and convergence to $v(i^*)$ again follows from Proposition 3.4.

In both cases, it is clear that i^* is an ESS. □

Thus, when only $\bar{\alpha} > 0$ ZDS plans are competing with one another, the ones on the highest value line win. Among $\bar{\alpha} < 0$ ZDS plans, the ones on the lowest value line win.

The local stability of an ESS good strategy will not be global when both signs occur. To illustrate this, consider the case of two strategies indexed by $\mathcal{I} = \{1, 2\}$. Letting $w = \pi_1$, it is an easy exercise to show that Equation (4.2) reduces to

$$\frac{dw}{dt} = w(1 - w)[(A_{11} - A_{21})w + (A_{12} - A_{22})(1 - w)]. \tag{4.22}$$

Proposition 4.12. *Assume that $Z_1 > Z_2$ and that $\bar{\alpha}_1 \cdot \bar{\alpha}_2 < 0$. There is an equilibrium population $\pi^* = (w^*, (1 - w^*))$ that contains both strategies with*

$$w^*/(1 - w^*) = (A_{22} - A_{12})/(A_{11} - A_{21}). \tag{4.23}$$

This equilibrium is stable if $\bar{\alpha}_1 < 0$ and is unstable if $\bar{\alpha}_1 > 0$.

Proof. If $\bar{\alpha}_1 < 0$ and $\bar{\alpha}_2 > 0$, then Proposition 3.4 implies that $A_{11} - A_{21} = Z_1 - A_{21} < 0$ and $A_{12} - A_{22} = A_{12} - Z_2 > 0$. Reversing the signs reverses the inequalities. The result then easily follows from Equation (4.22). Just graph the linear function of w in the brackets and observe where the result is positive or negative. □

Question 4.13. Suppose we restrict to the case where $\mathcal{I}$ indexes ZDS's lying on different value lines to avoid degeneracies. We ask:

- How large a population can coexist? If N is the size of $\mathcal{I}$, the number of competing strategies, then for what N do there exist examples with an interior equilibrium, that is, an equilibrium π such that $\pi_i > 0$ for all $i \in \mathcal{I}$? When is there a locally stable interior equilibrium? For how large an N can *permanence* occur (see Section 3 of [9]), that is, the boundary of Δ be a repellor? The *Brouwer fixed point theorem* implies that such a permanent system always admits an interior equilibrium. When an interior equilibrium does not exist, there is always some sort of dominance among the mixed strategies of the game $\{A_{ij}\}$. See [1] and [3].
- Can there exist a stable, closed invariant set containing no equilibria, e.g., a stable limit cycle?

There is alternative version of the dynamics that explicitly considers for X not the payoff s_X but the advantage that X has over Y. That is, the addition to the growth rate is given not by s_X but by the difference $s_X - s_Y$. This amounts to replacing A_{ij} by the anti-symmetric matrix $S_{ij} = A_{ij} - A_{ji}$ so that the game becomes zero-sum. In this case, we define $\xi_i = -Z_i^{-1} = \bar{\alpha}_i + \bar{\beta}_i$. Thus, ξ_i varies in the interval $[-P^{-1}, -R^{-1}]$. Define

$\xi_\pi = \Sigma_{i \in \mathcal{I}} \pi_i \xi_i$. From Equation (4.17), we get

$$S_{ij} = K_{ij}(\xi_j - \xi_i), \tag{4.24}$$

where we let $K_{ii} = 1$ for all $i \in \mathcal{I}$.

Since $\{S_{ij}\}$ is antisymmetric, $S_{\pi\pi} = 0$.

For this system, the behavior is always like the $\bar{\alpha} < 0$ case for the previous system.

Theorem 4.14. *Let $\mathcal{I}$ index a finite list of non-exceptional ZDS strategies, with at most one using a Vertex plan. For the system with*

$$\frac{d\pi_i}{dt} = \pi_i(S_{i\pi} - S_{\pi\pi}) = \pi_i S_{i\pi}, \tag{4.25}$$

we have

$$\begin{aligned} &\frac{d\xi_\pi}{dt} \leq 0, \\ &\textit{with equality iff} \quad \pi_i, \pi_j > 0 \Longrightarrow \xi_i = \xi_j. \end{aligned} \tag{4.26}$$

Assume now that $Z_{i^*} < Z_j$, or equivalently $\xi_{i^*} < \xi_j$ for all $j \neq i^*$. The strategy i^* is an ESS, and if $\pi_{i^*}(0) > 0$, then the solution path converges to the vertex $v(i^*)$.

Proof. Because K_{ij} is symmetric and positive, $d\xi_\pi/dt$ equals

$$\begin{aligned} &- \Sigma_{i,j \in \mathcal{I}} \pi_i \pi_j K_{ij} \xi_i (\xi_i - \xi_j) \\ &= -\frac{1}{2} \left[\Sigma_{i,j \in \mathcal{I}} \pi_i \pi_j K_{ij} \xi_i (\xi_i - \xi_j) - \Sigma_{i,j \in \mathcal{I}} \pi_j \pi_j K_{ij} \xi_j (\xi_i - \xi_j) \right] \\ &= -\frac{1}{2} \Sigma_{i,j \in \mathcal{I}} \pi_i \pi_j K_{ij} (\xi_i - \xi_j)^2 \leq 0. \end{aligned} \tag{4.27}$$

Equality holds iff $\pi_i \pi_j (\xi_i - \xi_j)^2 = 0$ for all $i, j \in \mathcal{I}$. That is, when $\xi_i = \xi_j$ for all i, j with $\pi_i, \pi_j > 0$.

If $\xi_i < \xi_j$, then

$$S_{ij} > 0 = S_{jj} = S_{ii} > S_{ji}. \tag{4.28}$$

If $\xi_i < \xi_k \leq \xi_j$, then $S_{ik} > S_{kk} \geq S_{jk}$. Let $\{j_1, \ldots, j_n\}$ list $\mathcal{I} \setminus \{i^*\}$ with $\xi_{j_1} \geq \cdots \geq \xi_{j_n}$. As in Case(−) of Theorem 4.11, it follows that i^* dominates the sequence $\{j_1, \ldots, j_n\}$. If $\pi_{i^*}(0) > 0$, then $\pi(t)$ converges to $v(i^*)$ by Proposition 3.4. □

References

[1] E. Akin, Domination or equilibrium, *Math. Biosciences* **50** (1980), 239–250.

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

[3] E. Akin and J. Hofbauer, Recurrence of the unfit, *Math. Biosciences* **61** (1982), 51–63.

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

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

[6] M. D. Davis, *Game Theory: A Nontechnical Introduction*, Dover Publications, Mineola, NY, 1983.

[7] G. Hardin, The tragedy of the commons, *Science* **162** (1968), 1243–1248.

[8] C. Hilbe, M. Nowak and K. Sigmund, The evolution of extortion in iterated Prisoner's Dilemma games, *PNAS* **110** (2013), no. 17, 6913–6918.

[9] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*, Cambridge University Press, Cambridge, UK, 1998.

[10] S. Karlin and H. Taylor, *A First Course in Stochastic Processes*, 2nd edition, Academic Press, New York, NY, 1975.

[11] J. Maynard Smith, *Evolution and the Theory of Games*, Cambridge University Press, Cambridge, UK, 1982.

[12] M. Nowak, *Evolutionary Dynamics*, Harvard University Press, Cambridge, MA, 2006.

[13] K. Sigmund, *Games of life*, Oxford University Press, Oxford, UK, 1993.

[14] K. Sigmund, *The Calculus of Selfishness*, Princeton University Press, Princeton, NJ, 2010.

[15] W. Press and F. Dyson, Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent, *PNAS* **109** (2012), no. 26, 10409–10413.

[16] A. Stewart and J. Plotkin, Extortion and cooperation in the Prisoner's Dilemma, *PNAS* **109** (2012), no. 26, 10134–10135.

[17] P. D. Straffin, *Game Theory and Strategy*, Mathematical Association of America, Washington, DC, 1993.

[18] P. Taylor and L. Jonker, Evolutionarily stable strategies and game dynamics, *Math. Biosciences* **40** (1978), 145–156.

[19] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, NJ, 1944.