Contents lists available at ScienceDirect

Games and Economic Behavior

Partners or rivals? Strategies for the iterated prisoner's dilemma ☆

Christian Hilbe [a,*], Arne Traulsen [b], Karl Sigmund [c,d]

[a] *Program for Evolutionary Dynamics, Harvard University, Cambridge, MA, United States*
[b] *Department of Evolutionary Theory, Max-Planck Institute for Evolutionary Biology, 24306 Plön, Germany*
[c] *Faculty of Mathematics, University of Vienna, Nordbergstrasse 15, 1090 Vienna, Austria*
[d] *International Institute for Applied Systems Analysis, Schlossplatz 1, 2361 Laxenburg, Austria*

ARTICLE INFO

ABSTRACT

Within the class of memory-one strategies for the iterated Prisoner's Dilemma, we characterize partner strategies, competitive strategies and zero-determinant strategies. If a player uses a partner strategy, both players can fairly share the social optimum; but a co-player preferring an unfair solution will be penalized by obtaining a reduced payoff. A player using a competitive strategy never obtains less than the co-player. A player using a zero-determinant strategy unilaterally enforces a linear relation between the two players' payoffs. These properties hold for every strategy used by the co-player, whether memory-one or not.

1. Introduction

In a one-shot Prisoner's Dilemma (PD) game, the two players have to choose between C and D (to cooperate resp. to defect). Following the notation in Rapoport and Chammah (1965), the payoff matrix is given by

$$
\begin{array}{c|c|c|}
 & C & D \\
\hline
C & R,R & S,T \\
\hline
D & T,S & P,P \\
\hline
\end{array}
\tag{1}
$$

in which the four payoff variables represent the reward for mutual cooperation R, the sucker's payoff S, the temptation to defect T, and the punishment for mutual defection P. Payoffs satisfy the inequalities $T > R > P > S$, such that defection is a dominant strategy, but mutual cooperation is preferred over mutual defection. In addition to these inequalities, we shall also assume that

* Corresponding author.
  E-mail addresses: hilbe@fas.harvard.edu (C. Hilbe), traulsen@evolbio.mpg.de (A. Traulsen), karl.sigmund@univie.ac.at (K. Sigmund).

$$2R > T + S, \tag{2}$$

such that mutual cooperation is unanimously preferred from a group perspective.[1] In such cases, experimental evidence suggests that many players want to achieve conditional cooperation. They are willing to play C, provided the co-player also plays C (see, e.g., Fehr and Fischbacher, 2003; Yamagishi et al., 2005). However, short of a commitment device, this cannot be ensured. Thus, players either have to trust their co-player, or else use their dominant strategy.

The situation is different for an iterated PD game (IPD). Diverse 'folk theorems' state that any feasible and individually rational outcome can be sustained as an equilibrium if the probability δ of a further round is sufficiently large. Such outcomes can be enforced in various ways, and under a wide range of circumstances (see, e.g., Friedman, 1971; Aumann, 1981; Aumann and Shapley, 1994; Fudenberg and Maskin, 1986; Kalai, 1990; Myerson, 1991; Mailath and Olszewski, 2011). If subjects can make binding commitments ahead of the game, then analogous results can be obtained even for one-shot games (Kalai et al., 2010).

Experimental research has uncovered considerable heterogeneity in human social preferences (Colman, 1995; Kagel and Roth, 1997; Camerer, 2003), and a similar variety can be found among the strategies that are played in the IPD (Milinski and Wedekind, 1998; Dal Bó and Fréchette, 2011; Fudenberg et al., 2012). Players who in the one-shot game would opt for conditional cooperation should be willing to engage in 'partner' strategies. Such strategies aim for an average payoff R per round, which necessarily provides the same payoff R for the co-player; should the co-player not go along, however, then the co-player's payoff will be less than R. Thus, a partner strategy appeals to the co-player's self-interest in order to further the own self-interest. It is fair, and provides an incentive for the co-player to also be fair. In contrast, some players tend to view their co-player as a rival, rather than a partner. The main purpose, for such competitive players, is to do better, or at least as well as the other player. A preference for dominating the co-player is particularly likely in the context of a game, which often has antagonistic connotations.

The aim of the present manuscript is two-fold: first, we are going to characterize all memory-one strategies which are either competitive, or partner strategies in the sense above. We emphasize that players using such strategies can enforce their preferences against all comers, since we impose no restrictions on the strategies used by the co-players. For partner strategies, the corresponding results in the limiting case of an IPD without discounting have been obtained within the last two years (Akin, 2013; Stewart and Plotkin, 2013, 2014). Here, we are going to extend the theory by allowing for discount factors $\delta \leq 1$ (which may either be interpreted as the constant continuation probability of having another round, or as the players' common discount rate on future payoff streams). The recent progress was stimulated by the unexpected discovery of so-called zero-determinant (ZD) strategies, a class of memory-one strategies enforcing a linear relationship between the payoffs of the two players, irrespective of the co-player's strategy (Press and Dyson, 2012). In particular, ZD strategies can fix the co-player's payoff to an arbitrary value between P and R; or ensure that the own 'surplus' (over the maximin value P) is twice as large as the co-player's surplus; etc. Also for ZD strategies, we are going to extend the theory to the case when future payoffs are discounted, and $\delta \leq 1$.

The nature of our results is somewhat different from usual treatments of repeated games. Our article does not focus on equilibrium behavior (in particular, we do not aim to explore which payoffs rational players can achieve). Instead, we define some interesting properties that a player's strategy may have (e.g., being competitive); and then we are going to characterize all memory-one strategies that have the respective property (independent of whether such a strategy can be sustained as an equilibrium). Thereby, we do not make any assumptions on the behaviors of the co-players (e.g., we do not require them to play best responses, or to follow a predefined equilibrium path). Nevertheless, there are natural connections between several of the described strategy classes and equilibrium behavior, and in that case we will discuss these connections in detail.

In the discussion, we will briefly review the previous development, and in particular the relevant findings in evolutionary game theory (Stewart and Plotkin, 2012, 2013, 2014; Adami and Hintze, 2013; Hilbe et al., 2013a, 2013b; Akin, 2013; Szolnoki and Perc, 2014a, 2014b; Wu and Rong, 2014). In a nutshell, these findings say that in the context of populations of adapting players, partner strategies do well, whereas competitive strategies fare poorly.

2. A fundamental lemma on mean distributions

We consider the standard setup of an IPD with perfect monitoring. In each round, the two players choose whether to cooperate or to defect. That is, they choose an action from their respective action set $A_i = \{C, D\}$, with $i \in \{I, II\}$. The

---

[1] Whereas this additional constraint is rather uncommon in economics, it is fairly common in psychology and in the evolutionary game theory literature (e.g., Rapoport and Chammah, 1965; Axelrod, 1984). Inequality (2) rules out some additional complications that arise when players need to coordinate on different actions to obtain the social optimum. The difficulties become more apparent in the repeated game. As part of our analysis, we wish to characterize strategies which only depend on the decisions of the last round (so-called *memory-one strategies*), and which enforce a fair and efficient outcome (which will be referred to as *partner strategies*). When $2R < T + S$, efficiency requires players to alternate between cooperation and defection, and to establish an equilibrium path (CD, DC, CD, DC, ...). Now problems can arise when players observe a round with mutual defection, since a memory-one player is unable to determine who of the two players deviated from the equilibrium path (or in which stage of a possible punishment phase the players are). These issues can be circumvented when players are allowed to have longer memory (Mailath and Olszewski, 2011), or when the action space is *rich* (e.g., when the action space is a convex set, see Barlo et al., 2009). Herein, we will neglect these additional complications by focusing on games that satisfy the constraint (2). However, we note that inequality (2) is only used in proofs of results pertaining to partner strategies (Lemma 2 and Proposition 1). All other results presented in this manuscript (Lemma 1 and the Propositions 2–6) are independent of this condition.

outcome of a given round t can then be described by an action profile $a^t \in A = A_I \times A_{II}$. After each round, both players observe the chosen action profile, and they receive the respective payoffs as specified in the payoff matrix (1). The round t history is a vector $h^t = (a^0, a^1, \ldots, a^{t-1}) \in A^t$, and the set of possible histories is the union $\mathcal{H} = \cup_{t=0}^{\infty} A^t$, with the initial history A^0 being defined as the null set $A^0 = \{\emptyset\}$. A strategy for player i is a rule that tells the player how to act after any possible history; that is, a strategy is a map $\sigma_i : \mathcal{H} \to \Delta(A_i)$, where $\Delta(A_i)$ denotes the set of probability distributions over the action set A_i.[2]

For given strategies of the two players, let $v_a(t)$ denote the probability that the resulting action profile played in round t is $a \in \{CC, CD, DC, DD\}$. For convenience, we use the following vector notation:

$$\mathbf{v}(t) = \big(v_{CC}(t), v_{CD}(t), v_{DC}(t), v_{DD}(t)\big)$$
$$\mathbf{g_I} = \big(R, S, T, P\big)$$
$$\mathbf{g_{II}} = \big(R, T, S, P\big) \tag{3}$$

Using this notation, we can write the players' expected payoffs in round t as $\pi_I(t) = \mathbf{g_I} \cdot \mathbf{v}(t)$ and $\pi_{II}(t) = \mathbf{g_{II}} \cdot \mathbf{v}(t)$. For a discount factor $\delta < 1$, the expected payoffs of the repeated game can then be defined by the Abelian means

$$\pi_I = (1 - \delta) \sum_{t=0}^{\infty} \delta^t \pi_I(t) = \mathbf{g_I} \cdot \mathbf{v}, \tag{4}$$

and similarly $\pi_{II} = \mathbf{g_{II}} \cdot \mathbf{v}$, where $\mathbf{v} = \big(v_{CC}, v_{CD}, v_{DC}, v_{DD}\big)$ refers to the (Abelian) mean distribution

$$\mathbf{v} = (1 - \delta) \sum_{t=0}^{\infty} \delta^t \mathbf{v}(t). \tag{5}$$

In the limiting case $\delta = 1$, the payoff per round is given by the Cesaro mean

$$\pi_I = \lim_{\tau \to \infty} \frac{1}{\tau + 1} \sum_{t=0}^{\tau} \pi_I(t) \tag{6}$$

(if this limit exists), and a similar expression for π_{II}.[3] A theorem by Frobenius states that if the Cesaro mean exists, it is the limit of the Abelian mean, for $\delta \nearrow 1$.

In the following, we will sometimes focus on players who only take the decisions in the previous round into account.

Definition 1. A strategy σ is a memory-one strategy if $\sigma(h^t) = \sigma(\tilde{h}^{t'})$ for all histories $h^t = (a^0, \ldots, a^{t-1})$ and $\tilde{h}^{t'} = (\tilde{a}^0, \ldots, \tilde{a}^{t'-1})$ with $t, t' \geq 1$ and $a^{t-1} = \tilde{a}^{t'-1}$.

For rounds $t \geq 1$, the move of a memory-one player is therefore solely determined by the action profile played in the previous round (in particular, we note that such players do not condition their behavior on the round number, as sometimes considered in models of bounded recall, e.g. Mailath and Olszewski, 2011). Such memory-one strategies can be written as a 5-tuple $\mathbf{p} = (p_{CC}, p_{CD}, p_{DC}, p_{DD}; p_0)$. The element p_0 denotes the probability to cooperate in the initial round. The continuation vector $\tilde{\mathbf{p}} := (p_{CC}, p_{CD}, p_{DC}, p_{DD})$ denotes the conditional probabilities to cooperate in rounds $t \geq 1$, depending on the outcome of the previous round (slightly abusing notation, we let the first letter in the subscript refer to the player's own action in the previous round, and the second letter to the co-player's action. Using this convention, we ensure that the interpretation of a memory-one strategy does not depend on whether the player acts as player I or as player II, see Nowak and Sigmund, 1995). Examples of memory-one strategies include $AllD = (0, 0, 0, 0; 0)$, Tit For Tat $(1, 0, 1, 0; 1)$, or Win-Stay, Lose-Shift $(1, 0, 0, 1; 1)$, see Sigmund (2010) for a comprehensive discussion.

When both players apply a memory-one strategy, the resulting mean distribution $\mathbf{v}$ can be calculated explicitly (Nowak and Sigmund, 1995): if player I uses the memory-one strategy $\mathbf{p} = (p_{CC}, p_{CD}, p_{DC}, p_{DD}; p_0)$ against a player II with memory-one strategy $\mathbf{q} = (q_{CC}, q_{CD}, q_{DC}, q_{DD}; q_0)$, then

$$\mathbf{v} = (1 - \delta)\mathbf{v}(0) \cdot (I_4 - \delta M)^{-1}, \tag{7}$$

where $\mathbf{v}(0) = \big(p_0 q_0, \ p_0(1 - q_0), \ (1 - p_0)q_0, \ (1 - p_0)(1 - q_0)\big)$ is the initial distribution, I_4 is the 4×4 identity matrix, and M is the transition matrix of the Markov chain,

---

[2] Strictly speaking, this means that we are considering *behavior strategies*, see Section 2.1.3 of Mailath and Samuelson (2006).

[3] This definition of payoffs for $\delta = 1$ is common in evolutionary game theory (e.g. Sigmund, 2010), whereas the equilibrium literature usually takes the lim inf of average payoffs to ensure that payoffs are always defined. Obviously, if the limit in (6) exists, the two definitions coincide. In the evolutionary literature, the strategy space is typically restricted (for example to memory-one strategies), which often guarantees the existence of the limit. Here, we have chosen the definition (6) to be consistent with the previous literature on ZD strategies in repeated games without discounting, e.g. Press and Dyson (2012) and Akin (2013).

$$M = \begin{pmatrix} p_{CC}q_{CC} & p_{CC}(1-q_{CC}) & (1-p_{CC})q_{CC} & (1-p_{CC})(1-q_{CC}) \\ p_{CD}q_{DC} & p_{CD}(1-q_{DC}) & (1-p_{CD})q_{DC} & (1-p_{CD})(1-q_{DC}) \\ p_{DC}q_{CD} & p_{DC}(1-q_{CD}) & (1-p_{DC})q_{CD} & (1-p_{DC})(1-q_{CD}) \\ p_{DD}q_{DD} & p_{DD}(1-q_{DD}) & (1-p_{DD})q_{DD} & (1-p_{DD})(1-q_{DD}) \end{pmatrix}. \tag{8}$$

But even if only one of the players is using a memory-one strategy $\mathbf{p}$, there is still a powerful relationship between $\mathbf{p}$ and the resulting mean distribution $\mathbf{v}$.

Lemma 1. *Suppose player I applies a memory-one strategy $\mathbf{p}$, and let the strategy of player II be arbitrary, but fixed.*

(i) *In the case with discounting ($\delta < 1$), let $\mathbf{v}$ denote the mean distribution of the repeated game. Then*

$$(\delta p_{CC} - 1)v_{CC} + (\delta p_{CD} - 1)v_{CD} + \delta p_{DC}v_{DC} + \delta p_{DD}v_{DD} = -(1-\delta)p_0, \tag{9}$$

or in vector notation, $(\delta\tilde{\mathbf{p}} - \mathbf{g_0}) \cdot \mathbf{v} = -(1-\delta)p_0$, where $g_0 = (1, 1, 0, 0)$.

(ii) *In the case without discounting, we have*

$$\lim_{\tau \to \infty} \frac{1}{\tau + 1} \sum_{t=0}^{\tau} (\tilde{\mathbf{p}} - \mathbf{g_0}) \cdot \mathbf{v}(t) = 0. \tag{10}$$

In particular, if the Cesaro mean distribution $\mathbf{v}$ exists, $(\tilde{\mathbf{p}} - \mathbf{g_0}) \cdot \mathbf{v} = 0$.

Proof. Suppose $\delta < 1$, and let $q_I(t)$ denote the probability that player I cooperates in round t. Then $q_I(t) = \mathbf{g_0} \cdot \mathbf{v}(t)$ and $q_I(t+1) = \tilde{\mathbf{p}} \cdot \mathbf{v}(t)$. It follows that $w(t) := \delta q_I(t+1) - q_I(t)$ is given by

$$w(t) = (\delta\tilde{\mathbf{p}} - \mathbf{g_0}) \cdot \mathbf{v}(t). \tag{11}$$

Multiplying each $w(t)$ by $(1-\delta)\delta^t$ and summing up over $t = 0, \ldots, \tau$ yields

$$(1-\delta)\sum_{t=0}^{\tau}\delta^t w(t) = (1-\delta)\big(\delta q_I(1) - q_I(0) + \delta^2 q_I(2) - \delta q_I(1)\ldots\big)$$
$$= (1-\delta)\delta^{\tau+1}q_I(\tau+1) - (1-\delta)q_I(0) \to -(1-\delta)p_0. \tag{12}$$

On the other hand, due to Eq. (11),

$$(1-\delta)\sum_{t=0}^{\tau}\delta^t w(t) = (1-\delta)\sum_{t=0}^{\tau}\delta^t(\delta\tilde{\mathbf{p}} - \mathbf{g_0}) \cdot \mathbf{v}(t) \to (\delta\tilde{\mathbf{p}} - \mathbf{g_0}) \cdot \mathbf{v} \tag{13}$$

As both limits need to coincide, we have confirmed Eq. (9). For the case without discounting, an analogous calculation as in Eq. (12) yields

$$\frac{1}{\tau + 1} \sum_{t=0}^{\tau} w(t) \to 0, \tag{14}$$

whereas Eq. (13) becomes

$$\frac{1}{\tau + 1} \sum_{t=0}^{\tau} w(t) = \frac{1}{\tau + 1} \sum_{t=0}^{\tau} (\tilde{\mathbf{p}} - \mathbf{g_0}) \cdot \mathbf{v}(t). \tag{15}$$

It follows that the limit of $\frac{1}{\tau+1}\sum_{t=0}^{\tau}(\tilde{\mathbf{p}} - \mathbf{g_0}) \cdot \mathbf{v}(t)$ for $\tau \to \infty$ exists and equals zero. $\square$

It is worthwhile to stress the generality of Lemma 1: it neither makes any assumption on the strategy used by the co-player, nor does it depend on the specific payoff constraints of a prisoner's dilemma. In the limiting case $\delta = 1$, Lemma 1 allows a geometric interpretation: the mean distribution $\mathbf{v}$ (if it exists) is orthogonal to $\tilde{\mathbf{p}} - \mathbf{g_0}$ (see Akin, 2013).

3. Partner strategies and competitive strategies

Definition 2. A player's strategy is *nice*, if the player is never the first to defect. A player's strategy is *cautious* if the player is never the first to cooperate.

For memory-one strategies, nice strategies fulfill $p_0 = p_{CC} = 1$, and cautious strategies $p_0 = p_{DD} = 0$. As an example, the strategy *TFT* $(1, 0, 1, 0; 1)$ is nice, whereas the defector's strategy *AllD* $(0, 0, 0, 0; 0)$ is cautious.
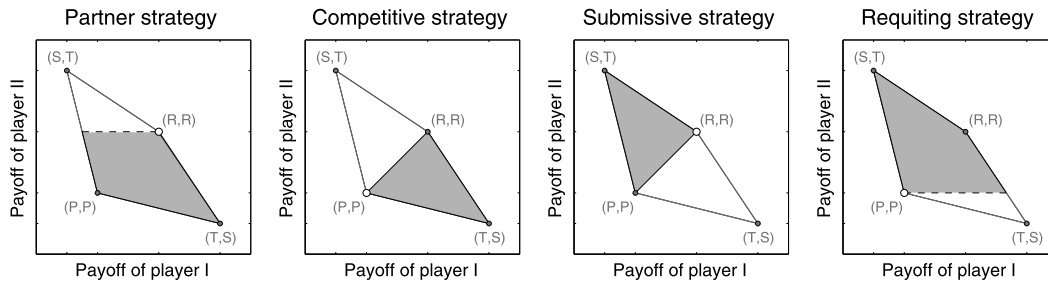
Fig. 1. Schematic representation of partner strategies, competitive strategies, submissive strategies and requiting strategies. The grey-shaded area depicts the set of possible payoff pairs when player I adopts a strategy of the respective strategy class. The white dot represents the payoff that player I gets against a co-player using the same strategy.

Lemma 2. *If $2R > T + S$, then payoffs satisfy $\pi_I + \pi_{II} \geq 2R$ if and only if $\pi_I = \pi_{II} = R$ (which for $\delta < 1$ holds if and only if both players are nice). Similarly, if $2P < T + S$, then $\pi_I + \pi_{II} \leq 2P$ if and only if $\pi_I = \pi_{II} = P$ (which for $\delta < 1$ is equivalent to both players being cautious).*

Proof. Due to Eq. (4), $\pi_I + \pi_{II} = (\mathbf{g_I} + \mathbf{g_{II}}) \cdot \mathbf{v} = (2R, T + S, T + S, 2P) \cdot \mathbf{v}$. As $2R > T + S$, the inequality $\pi_I + \pi_{II} \geq 2R$ implies $v_{CC} = 1$. For $\delta < 1$, this requires both players to cooperate in every round (if $\delta = 1$, it only requires the players to cooperate in almost every round). Similarly, for a prisoner's dilemma with $2P < T + S$, the inequality $\pi_I + \pi_{II} \leq 2P$ implies $v_{DD} = 1$. □

Definition 3.

(i) A *partner strategy* for player I is a nice strategy such that, irrespective of the co-player's strategy,

$$\pi_I < R \quad \Rightarrow \quad \pi_{II} < R. \tag{16}$$

(ii) A *competitive strategy* for player I is a strategy such that, irrespective of the co-player's strategy,

$$\pi_I \geq \pi_{II}. \tag{17}$$

Fig. 1 gives a schematic illustration of these two strategy classes. The definition of partner strategies implies that these strategies are best replies to themselves, and thus they are Nash equilibria. Even more, because condition (16) is equivalent to $(\pi_{II} \geq R) \Rightarrow (\pi_I \geq R)$, we can conclude due to Lemma 2 that $(\pi_{II} \geq R) \Rightarrow (\pi_I = \pi_{II} = R)$. Thus, no matter *which* best reply the co-player applies, a player with a partner strategy will always obtain the mutual cooperation payoff R.

On the other hand, players with a competitive strategy always obtain at least the co-player's payoff. It is easy to see that for $\delta < 1$ a competitive strategy needs to be cautious (otherwise the focal player would be outcompeted by an *AllD*-player). In the limiting case $\delta = 1$, competitiveness is closely related to the concept of being unbeatable, as introduced by Duersch et al. (2012). A strategy for player I is unbeatable, if against any co-player and for any number of rounds, the payoff differential $\sum_{t=0}^{\tau} \left(\pi_{II}(t) - \pi_I(t) \right)$ is bounded from above (in particular, if the average payoffs per round converge to π_I and π_{II}, then $\pi_I \geq \pi_{II}$).

Proposition 1. *For a player I with a nice memory-one strategy $\mathbf{p}$, the following are equivalent:*

(i) *$\mathbf{p}$ is a partner strategy;*
(ii) *If the co-player uses either AllD or the strategy $(0, 1, 1, 1; 0)$, then $\pi_{II} < R$;*
(iii) *The two inequalities $B_1 < 0$ and $B_2 < 0$ hold, with*

$$B_1 = \delta(T - R)p_{DD} - \delta(R - P)(1 - p_{CD}) + (1 - \delta)(T - R)$$
$$B_2 = \delta(T - R)p_{DC} - \delta(R - S)(1 - p_{CD}) + (1 - \delta)(T - R). \tag{18}$$

Proof.

(i) $\Rightarrow$ (ii) Assume to the contrary that $\pi_{II} \geq R$. Then the definition of partner strategies implies that $\pi_I = \pi_{II} = R$. Since all players use memory-one strategies, this would require that everyone cooperates after mutual cooperation, which is neither true for $AllD = (0, 0, 0, 0; 0)$ nor for the strategy $(0, 1, 1, 1; 0)$.

(ii) $\Rightarrow$ (iii) Against a player using a nice memory-one strategy $\mathbf{p}$ (with $p_0 = p_{CC} = 1$), the payoff of an *AllD* co-player is given by

$$\hat{\pi}_{II} = \frac{(1-\delta)T + \delta P - \delta P p_{CD} + \delta T p_{DD}}{1 + \delta(p_{DD} - p_{CD})}. \tag{19}$$

We note that this payoff is also defined when $\delta = 1$, because $p_{CD} < 1$ (otherwise **p** would satisfy $p_0 = p_{CC} = p_{CD} = 1$, and player *I* would always cooperate. In that case, an *AllD* co-player would receive $T > R$, which is ruled out by (ii)). Elementary algebra yields

$$B_1 = \big(1 + \delta(p_{DD} - p_{CD})\big)\big(\hat{\pi}_{II} - R\big), \tag{20}$$

In particular, B_1 has the same sign as $\hat{\pi}_{II} - R$. On the other hand, if the co-player uses the strategy (0,1,1,1;0), the co-player's payoff is

$$\tilde{\pi}_{II} = \frac{(1-\delta)T + \delta S + \delta\big((1-\delta)R - S\big)p_{CD} + \delta(T + \delta R)p_{DC}}{1 + \delta^2(p_{DC} - p_{CD}) + \delta p_{DC}}, \tag{21}$$

and

$$B_2 = \big(1 + \delta^2(p_{DC} - p_{CD}) + \delta p_{DC}\big)\big(\tilde{\pi}_{II} - R\big). \tag{22}$$

Therefore, B_2 has the same sign as $\tilde{\pi}_{II} - R$.

(iii) $\Rightarrow$ (i) Suppose that $B_1 < 0$ and $B_2 < 0$, and that $\pi_{II} \geq R$. We need to show that $\pi_{II} = \pi_I = R$. As $\pi_{II} - R = (\mathbf{g_{II}} - R\mathbf{1}) \cdot \mathbf{v}$, with $\mathbf{1} = (1, 1, 1, 1)$, we note that $\pi_{II} \geq R$ is equivalent to

$$(T - R)v_{CD} - (R - S)v_{DC} - (R - P)v_{DD} \geq 0. \tag{23}$$

Using the linear equations $\mathbf{1} \cdot \mathbf{v} = 1$ and $(\delta\tilde{\mathbf{p}} - \mathbf{g_0}) \cdot \mathbf{v} = -(1 - \delta)$ (Lemma 1 with $p_0 = p_{CC} = 1$, since the memory-one strategy is nice), we calculate v_{CD} as a function of v_{DC} and v_{DD}:

$$v_{CD} = \frac{\big(1 - (1 - p_{DC})\delta)\big)v_{DC} + \big(1 - (1 - p_{DD})\delta\big)v_{DD}}{(1 - p_{CD})\delta}. \tag{24}$$

The denominator of v_{CD} is positive, as $B_1 < 0$ implies $p_{CD} < 1$. Plugging (24) into (23) and multiplying both sides with $(1 - p_{CD})\delta$ shows that $\pi_{II} \geq R$ if and only if

$$B_2 v_{DC} + B_1 v_{DD} \geq 0, \tag{25}$$

with B_1 and B_2 as defined in (18). Thus, the assumptions $B_1 < 0$ and $B_2 < 0$ indicate that $v_{DC} = v_{DD} = 0$, and by (24) that $v_{CD} = 0$. We conclude that $v_{CC} = 1$, and therefore $\pi_I = \pi_{II} = R$, i.e., **p** is a partner strategy. □

For example, TFT is a partner strategy if and only if $\delta > \frac{T-R}{T-P}$ and $\delta > \frac{T-R}{R-S}$, whereas WSLS is a partner strategy if and only if $\delta > \frac{T-R}{R-P}$ and $\delta > \frac{T-R}{T-S}$, which is a sharper condition. In analogy to the definition of partner strategies, one may define a *mild partner strategy* for player *I* as a nice strategy such that $\pi_I \leq R$ implies $\pi_{II} \leq R$, irrespective of the co-player's strategy.[4] For memory-one strategies, the characterization of mild partner strategies is analogous to the characterization of partner strategies (only the strict inequalities in Proposition 1 need to be replaced by weak inequalities).

Proposition 1 also provides an interesting connection to the folk theorems. The existence of an equilibrium with individually rational payoffs (π_I, π_{II}) in the IPD is typically shown by applying trigger strategies – any deviation from the equilibrium path is punished with relentless defection (as for example in Friedman, 1971).[5] The following Corollary states that trigger strategies are, in some sense, the most effective means to enforce a cooperative equilibrium in the IPD.

Corollary 1. *For a given prisoner's dilemma and a given continuation probability δ, there exists a memory-one partner strategy if and only if the trigger strategy Grim = (1, 0, 0, 0; 1) is a partner strategy.*

Proof. The two quantities B_1 and B_2 in Proposition 1 are minimal for $p_{CD} = p_{DC} = p_{DD} = 0$. Thus, if there is a memory-one strategy that meets the inequalities $B_1 < 0$ and $B_2 < 0$, then the corresponding inequalities are also met by Grim. □

From Corollary 1, we may also conclude that partner strategies exist if and only if $\delta > \frac{T-R}{T-P}$ (this condition for the existence of fully cooperative equilibria has been previously derived by Roth and Murnighan, 1978; Stahl, 1991).

Let us next give a characterization of competitive memory-one strategies:

---

[4] Equivalently, one may define mild partner strategies as nice strategies such that $\pi_{II} > R$ implies $\pi_I > R$. We note that if the premise was true and $\pi_{II} > R$, then total payoffs would exceed $2R$, which is ruled out by Lemma 2. We conclude that mild partner strategies enforce $\pi_{II} \leq R$. That is, mild partner strategies are exactly those nice strategies that support mutual cooperation in a Nash equilibrium.

[5] For arbitrary stage games, trigger strategies support all outcomes that Pareto dominate a Nash equilibrium of the stage game. To support any individually rational outcome in a perfect equilibrium, players may have to use "stick and carrot" strategies instead, which punish deviations only for a finite number of rounds, see Fudenberg and Maskin (1986).
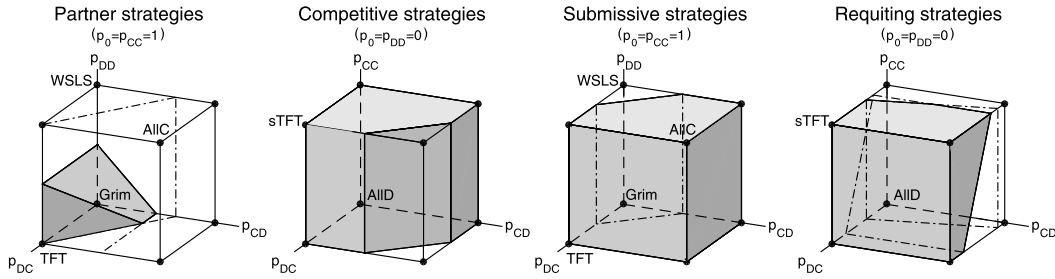
Fig. 2. The space of partner strategies, competitive strategies, submissive strategies and requiting strategies. Each grey block represents the set of strategies that fulfill the respective constraints in Propositions 1–4. For this representation, the continuation probability was set to $\delta = 2/3$, using the payoff values in Axelrod (1984), i.e. $T = 5$, $R = 3$, $P = 1$, $S = 0$. The depicted pure strategies are: TFT $= (1,0,1,0;1)$, Grim $= (1,0,0,0;1)$, Win-stay lose-shift: WSLS $= (1,0,0,1;1)$, AllC $= (1,1,1,1;1)$, AllD $= (0,0,0,0;0)$ and suspicious Tit For Tat: sTFT $= (1,0,1,0;0)$.

Proposition 2. *Suppose player I applies the memory-one strategy* $\mathbf{p}$*. Then the following are equivalent:*

  (i) $\mathbf{p}$ *is competitive.*
 (ii) *If the co-player uses either AllD or the strategy* $(0, 0, 0, 1; 0)$*, then* $\pi_I \geq \pi_{II}$.
(iii) *The entries of* $\mathbf{p}$ *satisfy* $p_0 = p_P = 0$ *and* $\delta(p_{CD} + p_{DC}) \leq 1$.

Proof.

(i) $\Rightarrow$ (ii) Follows immediately from the definition, a competitive strategy yields $\pi_I \geq \pi_{II}$ against any co-player.
(ii) $\Rightarrow$ (iii) If player *II* applies *AllD*, then an explicit calculation of payoffs yields

$$\pi_I - \pi_{II} = -\frac{(T - S)\big((1 - \delta)p_0 + \delta p_{DD}\big)}{1 + \delta(p_{DD} - p_{CD})}, \tag{26}$$

and thus $\pi_I \geq \pi_{II}$ implies $p_0 = p_{DD} = 0$. Similarly, if player *II* applies the strategy $(0, 0, 0, 1; 0)$, we obtain (using $p_0 = p_{DD} = 0$):

$$\pi_I - \pi_{II} = \frac{\delta\big(T - S\big)\big(1 - \delta(p_{CD} + p_{DC})\big)}{1 + \delta\big(1 - (1 + \delta)p_{CD} + \delta p_{DC}\big)}. \tag{27}$$

This is non-negative if and only if $\delta(p_{CD} + p_{DC}) \leq 1$.
(iii) $\Rightarrow$ (i) By Lemma 1 and as $p_{DD} = p_0 = 0$,

$$\delta p_{DC} v_{DC} = (1 - \delta p_{CC})v_{CC} + (1 - \delta p_{CD})v_{CD}. \tag{28}$$

Using the inequality $\delta p_{DC} \leq 1 - \delta p_{CD}$, this leads to

$$(1 - \delta p_{CD})v_{DC} \geq (1 - \delta p_{CC})v_{CC} + (1 - \delta p_{CD})v_{CD}, \tag{29}$$

or equivalently $(1 - \delta p_{CD})(v_{DC} - v_{CD}) \geq (1 - \delta p_{CC})v_{CC}$. This implies $v_{DC} \geq v_{CD}$. As a consequence, $\pi_I - \pi_{II} = (\mathbf{g_I} - \mathbf{g_{II}}) \cdot \mathbf{v} = (T - S)(v_{DC} - v_{CD}) \geq 0$. $\square$

Fig. 2 shows the space of partner strategies (and the space of competitive strategies) as subsets of the nice memory-one strategies (cautious memory-one strategies), respectively. One can also define the dual properties, and derive the corresponding characterizations: a strategy for player *I* is said to be *submissive* if payoffs always satisfy $\pi_I \leq \pi_{II}$, irrespective of the strategy of player *II*; and a cautious strategy for player *I* is said to be *requiting* if $\pi_I > P$ implies $\pi_{II} > P$ (see Fig. 1 for a schematic representation of these strategy classes). The corresponding characterizations are:

Proposition 3. *Suppose player I applies the memory-one strategy* $\mathbf{p}$*. Then the following are equivalent:*

  (i) $\mathbf{p}$ *is submissive;*
 (ii) *If the co-player uses either AllC or the strategy* $(0, 1, 1, 1; 1)$*, then* $\pi_I \leq \pi_{II}$;
(iii) *The entries of* $\mathbf{p}$ *satisfy* $p_0 = p_{CC} = 1$ *and* $\delta(1 - p_{CD}) + \delta(1 - p_{DC}) \leq 1$.

Proposition 4. *Suppose the game payoffs satisfy* $2P < T + S$*. Then, for a player I with a cautious memory-one strategy* $\mathbf{p}$*, the following are equivalent:*

  (i) $\mathbf{p}$ *is requiting;*
 (ii) *If the co-player uses either AllC or the strategy* $(0, 0, 0, 1; 1)$*, then* $\pi_{II} > P$;
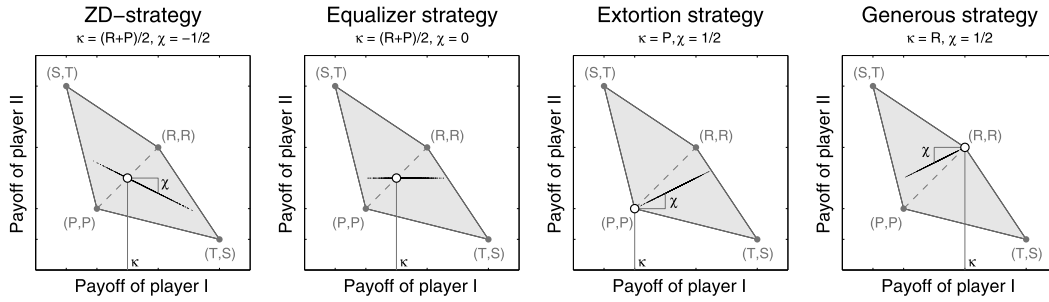
Fig. 3. Characteristic payoff relations for ZD strategies, equalizer strategies, extortion strategies and generous strategies. The grey-shaded area represents the set of feasible payoffs. In each graph, the strategy of player *I* was fixed, whereas for the strategy of player *II* we sampled 1000 random memory-one strategies. The resulting payoffs were drawn as black dots. For general ZD-strategies, these dots are on a line (intersecting the diagonal at κ, and having slope χ). Equalizer strategies have the additional property that the slope χ is zero, i.e. the payoff of co-player *II* is fixed to κ, independent of the co-player's strategy. Extortion strategies are ZD-strategies with $\kappa = P$ and $0 < \chi < 1$, and generous strategies fulfill $\kappa = R$ and $0 < \chi < 1$. For this figure, we have used the payoff values in Axelrod (1984), $T = 5$, $R = 3$, $P = 1$, $S = 0$, and continuation probability $\delta = 4/5$. For the strategy of player *I* we have used: (i) ZD-strategy $\mathbf{p} = (0.85, 0.725, 0.1, 0.35; 0.1)$; (ii) Equalizer strategy $\mathbf{p} = (0.875, 0.375, 0.375, 0.125; 0.5)$; (iii) Extortion strategy $\mathbf{p} = (1, 0.125, 0.75, 0; 0)$; (iv) Generous strategy $\mathbf{p} = (1, 0.125, 0.75, 0; 1)$.

(iii) *The two inequalities $B_1 > 0$ and $B_2 > 0$ hold, with*

$$B_1 = \delta(R - P)p_{DC} + \delta(P - S)p_{CC} - (P - S),$$
$$B_2 = \delta(T - P)p_{DC} + \delta(P - S)p_{CD} - (P - S). \tag{30}$$

4. ZD-strategies

The previous results have highlighted how Lemma 1 can be used to characterize several interesting strategy classes within the space of memory-one strategies (for example, the strategies that allow a player to outcompete the opponent, or the strategies that provide incentives to reach the social optimum). In the following, we present another application of Lemma 1: there are strategies with which a player can unilaterally enforce a linear relationship between the players' payoffs.

Definition 4. A memory-one strategy $\mathbf{p}$ is said to be a *ZD strategy* if there exist constants α, β, γ such that

$$\delta\tilde{\mathbf{p}} = \alpha\mathbf{g_I} + \beta\mathbf{g_{II}} + \big(\gamma - (1 - \delta)p_0\big)\mathbf{1} + \mathbf{g_0}. \tag{31}$$

Proposition 5. *Let $\delta < 1$, and suppose player I applies a memory-one strategy $\mathbf{p}$ satisfying Eq. (31). Then, irrespective of the strategy of the co-player,*

$$\alpha\pi_I + \beta\pi_{II} + \gamma = 0. \tag{32}$$

The same relation holds for $\delta = 1$, provided that the payoffs π_I and π_{II} exist.[6]

Proof. This follows directly from Lemma 1, using the identities $\pi_I = \mathbf{g_I} \cdot \mathbf{v}$, $\pi_{II} = \mathbf{g_{II}} \cdot \mathbf{v}$, and $1 = \mathbf{1} \cdot \mathbf{v}$. □

In the following, let $\delta < 1$. We proceed with a slightly different representation of ZD strategies, using the parameter transformation $\alpha = \phi\chi$, $\beta = -\phi$, and $\gamma = \phi\kappa(1 - \chi)$.[7] Under this transformation, ZD strategies take the form

$$\delta\tilde{\mathbf{p}} = \phi\Big[(1 - \chi)(\kappa\mathbf{1} - \mathbf{g_I}) + (\mathbf{g_I} - \mathbf{g_{II}})\Big] - (1 - \delta)p_0\mathbf{1} + \mathbf{g_0}, \tag{33}$$

and the enforced payoff relationship according to (32) becomes

$$\pi_{II} - \kappa = \chi(\pi_I - \kappa). \tag{34}$$

---

[6] If $\delta = 1$, and the payoffs π_I and π_{II} according to Eq. (6) do not exist, one can derive a slightly weaker result. In that case, it follows from Eq. (10) that

$$\lim_{\tau \to \infty} \frac{1}{\tau + 1} \sum_{t=0}^{\tau} \big(\alpha\pi_I(t) + \beta\pi_{II}(t) + \gamma\big) = 0.$$

[7] For $\delta < 1$, the proof of Proposition 6 shows that ZD strategies require $\phi > 0$ and $\chi < 1$ (and hence $\beta < 0$ and $\alpha + \beta < 0$). This allows us to conclude that the given parameter transformation is in fact bijective: the inverse is given by $\chi = -\alpha/\beta$, $\phi = -\beta$, and $\kappa = -\gamma/(\alpha + \beta)$.

Eq. (34) implies that the payoffs lie on a line segment intersecting the diagonal at some value κ (the payoff for the ZD-strategy against itself) and having a slope χ (see Fig. 3).

Players cannot use ZD strategies to enforce arbitrary payoff relationships of the form (34): since the entries of the continuation vector $\tilde{\mathbf{p}}$ correspond to conditional probabilities (and hence need to be in the unit interval), the parameters κ, χ and ϕ need to obey certain restrictions. This gives rise to the following definition.

Definition 5. For a given δ, we call a payoff relationship $(\kappa, \chi) \in \mathbb{R}^2$ *enforceable* if there are $\phi \in \mathbb{R}$ and $p_0 \in [0, 1]$ such that each entry of the continuation vector $\tilde{\mathbf{p}}$ according to Eq. (33) is in [0,1]. We refer to the set of all enforceable payoff relationships as $\mathcal{E}_\delta$.

Proposition 6.

(i) *The set of enforceable payoff relationships is monotonically increasing in the discount factor: if $\delta' \leq \delta''$, then $\mathcal{E}_{\delta'} \subseteq \mathcal{E}_{\delta''}$.*
(ii) *There is a $\delta < 1$ such that $(\kappa, \chi) \in \mathcal{E}_\delta$ if and only if $-1 < \chi < 1$ and*

$$\max\left\{P, \frac{S - T\chi}{1 - \chi}\right\} \leq \kappa \leq \min\left\{R, \frac{T - S\chi}{1 - \chi}\right\}, \tag{35}$$

with at least one inequality in (35) being strict.

Proof.

(i) According to the definition, $(\kappa, \chi) \in \mathcal{E}_\delta$ if and only if one can find $\phi \in \mathbb{R}$ and $p_0 \in [0, 1]$, such that the corresponding continuation vector $\tilde{\mathbf{p}}$ according to Eq. (33) satisfies $\mathbf{0} \leq \delta\tilde{\mathbf{p}} \leq \delta\mathbf{1}$, or equivalently,

$$(1 - \delta)(1 - p_0) \leq \phi(1 - \chi)(R - \kappa) \leq 1 - (1 - \delta)p_0 \tag{36a}$$

$$(1 - \delta)(1 - p_0) \leq \phi\big[(1 - \chi)(S - \kappa) + T - S\big] \leq 1 - (1 - \delta)p_0 \tag{36b}$$

$$(1 - \delta)p_0 \leq \phi\big[(1 - \chi)(\kappa - T) + T - S\big] \leq \delta + (1 - \delta)p_0 \tag{36c}$$

$$(1 - \delta)p_0 \leq \phi(1 - \chi)(\kappa - P) \leq \delta + (1 - \delta)p_0 \tag{36d}$$

We note that in (36a)–(36d), the left hand side is monotonically decreasing in δ, whereas the right hand side is monotonically increasing in δ. In particular, if the conditions (36) are satisfied for some $\delta' \leq 1$ they are also satisfied for any $\delta'' \geq \delta'$.

(ii) ($\Rightarrow$) Suppose $(\kappa, \chi) \in \mathcal{E}_\delta$, and therefore the conditions (36) hold for appropriate parameters ϕ and p_0. Summing up the first inequality in (36a) and the first inequality in (36d) shows

$$1 - \delta \leq \phi(1 - \chi)(R - P). \tag{37}$$

Similarly, by taking the inequalities in (36b) and (36c), we get

$$1 - \delta \leq \phi(1 + \chi)(T - S). \tag{38}$$

In particular, $0 < \phi(1 - \chi)$ and $0 < \phi(1 + \chi)$, and therefore $\phi > 0$ and $-1 < \chi < 1$. Moreover, the conditions (36) imply

$$0 \leq \phi(1 - \chi)(R - \kappa)$$
$$0 \leq \phi\big[(1 - \chi)(S - \kappa) + T - S\big]$$
$$0 \leq \phi\big[(1 - \chi)(\kappa - T) + T - S\big]$$
$$0 \leq \phi(1 - \chi)(\kappa - P). \tag{39}$$

Since $\phi > 0$ and $\chi < 1$, these conditions are equivalent to condition (35). If none of the inequalities in (35) was strict, then (36a) or (36b) would require $p_0 = 1$, whereas (36c) or (36d) would require $p_0 = 0$.

($\Leftarrow$) Conversely, let $-1 < \chi < 1$, and suppose $\max\left\{P, \frac{S - T\chi}{1 - \chi}\right\} \leq \kappa < \min\left\{R, \frac{T - S\chi}{1 - \chi}\right\}$. Then the inequalities (39) hold for any choice of $\phi > 0$, with the first two inequalities being strict. In particular, we can choose a ϕ sufficiently small such that each term on the right hand's side of (39) is bounded from above by 1/2. By setting $p_0 = 0$ and choosing a δ sufficiently close to one, it thus follows that all inequalities in (36) can be satisfied. An analogous argument holds when $\kappa = \min\left\{R, \frac{T - S\chi}{1 - \chi}\right\}$, in which case one needs to set $p_0 = 1$. □
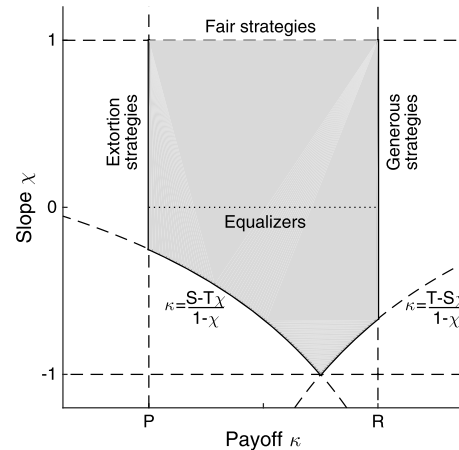
Fig. 4. Enforceable payoff relationships for players with a ZD strategy. The grey area depicts all pairs (κ, χ) that are enforceable when the discount factor δ is sufficiently close to one, as characterized in Proposition 6. The graph also depicts some particular subclasses of ZD strategies: equalizer strategies ($\chi = 0$), extortion strategies ($\kappa = P$, $\chi > 0$), and generous strategies ($\kappa = R$, $\chi > 0$). The so-called fair strategies (with $\chi = 1$) do only exist in the limit of no discounting, $\delta = 1$. For the illustration, we have taken the payoff values in Axelrod (1984), i.e. $T = 5$, $R = 3$, $P = 1$, $S = 0$.

The first part of Proposition 6 shows that a given linear payoff relationship of the form (34) is easier to enforce when players are sufficiently patient. As $\delta \to 1$, the limiting set of enforceable payoff relationships (κ, χ) is characterized by Proposition 6(ii); Fig. 4 provides an illustration.

There are various remarkable subclasses of ZD strategies (as depicted in Fig. 3 and Fig. 4). For $\chi = 0$, we encounter so-called *equalizer* strategies (see Boerlijst et al., 1997; Press and Dyson, 2012). By Eq. (34), player *I* can make use of such strategies to prescribe κ as payoff for player *II*. A player can thus determine the opponent's payoff (however, player *I* cannot fix the own payoff, since this would require χ to be unbounded, which is ruled out by Proposition 6). Press and Dyson (2012) also highlighted the class of *extortion* strategies (with $\kappa = P$ and $0 < \chi < 1$). Extortion strategies guarantee that the own 'surplus' over the minimax payoff P exceeds the opponent's surplus by a factor of χ^{-1}. Moreover, since $\chi > 0$, the payoffs of the two players are positively related. Hence, to maximize the own payoff, player *II* needs to maximize player *I*'s payoff: the best response against an extortioner is to cooperate unconditionally. As a counterpart to extortioners, Stewart and Plotkin (2012) defined the class of *generous* strategies, which satisfy Eq. (34) with $\kappa = R$ and $0 < \chi < 1$. Players using a generous strategy shoulder a larger burden of the loss (with respect to the social optimum R) than their co-player. Since $\chi > 0$, they also ensure that the payoffs of the two players are aligned, thereby motivating the co-player to cooperate. Finally, for games without discounting it was noted that strategies with $\chi = 1$ enforce $\pi_I = \pi_{II}$ (for $\delta = 1$, TFT is an example of such a *fair* strategy, see Press and Dyson, 2012; Hilbe et al., 2014b). However, as Proposition 6 shows, fair strategies cease to exist when future payoffs are discounted, and only approximately fair strategies (with χ close to one) may be feasible.

ZD strategies can also be connected to the strategy classes discussed in the previous section. Generous strategies, for example, are exactly the ZD strategies which are submissive partner strategies (in particular it follows that every generous strategy is a Nash equilibrium of the IPD). On the other hand, for stage games with $2P < T + S$ (which ensures $\pi_I + \pi_{II} \geq 2P$), extortion strategies are precisely those ZD strategies which are requiting and competitive.

We note that herein, we have entirely focused on the repeated prisoner's dilemma, due to the central role that this simple game situation takes in the literature on the evolution of cooperation (Rapoport and Chammah, 1965; Trivers, 1971; Sugden, 1986; Axelrod, 1984; Sigmund, 2010). However, the proofs of Lemma 1 and Proposition 5 did not require any assumptions on the payoff values (and in the proof of Proposition 6, we have only made use of the assumptions $R > P$ and $T > S$). Moreover, for $\delta = 1$, it was recently shown that similar results can also be obtained for stage games with 2 actions but $n \geq 2$ players (Hilbe et al., 2014b). Thus, while we believe that our results are most intuitive in the context of a prisoner's dilemma, the mathematics can be extended to more general strategic situations.

5. Discussion

The recent development began with the paper of Press and Dyson (2012) introducing ZD strategies for repeated games without discounting. In this context, Press and Dyson derived the linear relation (32). Their proof was based on a neat formula for the payoffs achieved if both players use memory-one strategies. This formula only involves vanishing determinants, which explains the name ZD. Press and Dyson highlighted those ZD strategies that fix the co-player's payoff to a given value between P and R, as well as the sinister properties of extortion strategies. They also stressed the fact that more complex strategies (based on larger memories, for instance) are not able to profit from their sophistication to gain the upper hand. The intriguing aspects of ZD strategies raised considerable attention (see, e.g., Ball, 2012). In the News section of the American Mathematical Society, it was stated that 'the world of game theory is currently on fire.' A more skeptical view could be found among economists. The well-known folk

theorem for repeated games states that trigger strategies can induce a rational co-player to agree to any feasible pay-off pair above the minimax level P, by threatening to switch to relentless defection otherwise (see Aumann, 1981; Kalai, 1990; Fudenberg and Maskin, 1986, 1990). Seen from this angle, the progress consisted merely in displaying *memory-one* strategies with a similar power to enforce specific payoff pairs. However, there is a subtle difference: whereas the Folk theorems are based on the assumption that players wish to maximize their payoffs, the results presented herein are independent of such an assumption. Interpreted in this way, we have explored how much control player I can exert on the resulting payoffs without being sure about the motives of player II.

Memory-one strategies able to fix the co-player's payoff had already been derived in Boerlijst et al. (1997) and Sigmund (2010), based on an approach different from that of Press and Dyson (2012). This method was used in Hilbe et al. (2013a) to provide another derivation of (32), not involving any determinants. It was substantially extended by Akin (2013) to yield a general equation for the mean distribution of memory-one strategies when $\delta = 1$. In this case, the mean distribution is understood in the sense of Cesaro, and need not always exist. In Lemma 1, we have extended this approach to cover the case $\delta < 1$, from which Akin's result for $\delta = 1$ immediately follows. Lemma 1 offers a geometric tool for the investigation of memory-one strategies. The vector $\delta\tilde{\mathbf{p}}$ consists of the conditional probabilities to play C in the next round (δ is the probability that there *is* a next round), whereas $\mathbf{g_0}$ can be viewed as 'conditional probability' to play C in the current round. In the limit of no discounting, Lemma 1 states that no matter which strategy player II is using, the limiting distribution $\mathbf{v}$ (if it exists) is on a hyperplane orthogonal to the difference of these two conditional probabilities. It was also Akin (2013) who extended the investigations beyond the case of ZD-strategies, to characterize partner strategies for $\delta = 1$ (calling them 'good' strategies, a term we feel is too general).

In a comments article, Stewart and Plotkin (2012) introduced an example of a generous strategy, and showed that in a round robin tournament conducted after the fashion of Axelrod (1984), this generous strategy emerged as winner. Stewart and Plotkin also asked whether ZD strategies were relevant for evolutionary game theory. In this context, one considers a population of players, each equipped with a strategy. The players are then allowed to imitate other strategies, preferentially those with a higher payoff (see, e.g., Weibull, 1995; Samuelson, 1997; Hofbauer and Sigmund, 1998; Nowak, 2006; Sandholm, 2010).

It is obvious that extortion strategies cannot spread too much in such an evolutionary context; if they become too common, they are likely to encounter their own, which bides ill. If player I obtains twice the surplus of II, and II twice the surplus of I, each surplus is zero. However, Hilbe et al. (2013a) showed that extortion strategies can pave the way for the emergence of cooperative strategies, similar to *TFT* (which, for $\delta = 1$, can be regarded as a limiting case of an extortion strategy, Press and Dyson, 2012). This catalytic role of extortion strategies has also been confirmed for games on networks, in which players only interact within a small neighborhood (Szolnoki and Perc, 2014a, 2014b; Wu and Rong, 2014). Overall, these studies confirm that extortionate strategies have problems to succeed within a population. However, if the games are played between members of two distinct populations – for instance, between hosts organisms and their symbionts – then extortion strategies can emerge in whichever population is *slower* to adapt (Hilbe et al., 2013a). The slower rate of evolution acts as a commitment device. In effect, the slowly evolving organism becomes the Stackelberg leader in a sequential game, in which the slow player learns to adopt extortion strategies, whereas the faster evolving player learns to play the best response, and to cooperate unconditionally (Bergstrom and Lachmann, 2003; Damore and Gore, 2011; Gokhale and Traulsen, 2012).

But even in a one-population setup, certain ZD strategies prove successful: Stewart and Plotkin showed that evolutionary trajectories often visit the vicinity of generous strategies. The dynamics leads 'from extortion to generosity' (the title of Stewart and Plotkin, 2013). This is also confirmed, by analytical means based on adaptive dynamics, by Hilbe et al. (2013b). Remarkably, Stewart and Plotkin (2013, 2014) derived a characterization of all memory-one strategies which are *robust* in an evolutionary sense, for given population size N. This means that the replacement probability, as a resident strategy, is at most $1/N$ (which is the probability to be replaced if the mutant is neutral). In the limit of *weak selection*, which roughly means that the choice between two strategies is only marginally influenced by payoff (see Nowak et al., 2004), all robust ZD strategies need to be generous (Stewart and Plotkin, 2013). These predictions have also been tested in a recent behavior experiment, in which human subjects played against various computer opponents (Hilbe et al., 2014a). Although extortionate programs outcompeted their human co-players in every game, generous programs received, on average, higher payoffs against the human subjects than extortionate programs. Humans were hesitant to give in to extortion; although unconditional cooperation would have been their best response in all treatments, they only became more cooperative over time if their co-player was generous.

Intriguingly, if a player uses a generous strategy and the co-player does not go along, then the focal player will always shoulder a *larger* part of the loss (with respect to the mutual cooperation payoff R). Despite their forbearance, generous strategies do very well – which is not the least of the surprises offered by the Iterated Prisoner's Dilemma game.

References

Adami, Christoph, Hintze, Arend, 2013. Winning isn't everything: evolutionary stability of zero determinant strategies. Nat. Commun. 4, 2193.

Akin, Ethan, 2013. The iterated prisoner's dilemma: good strategies and their dynamics. Working paper, arXiv:1211.0969.

Aumann, Robert J., 1981. Survey of repeated games. In: Henn, R., Moeschlin, O. (Eds.), Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern. Wissenschaftsverlag, Mannheim.

Aumann, Robert J., Shapley, Lloyd S., 1994. Long-term competition: a game-theoretic analysis. In: Meggido, N. (Ed.), Essays in Game Theory in Honor of Michael Maschler. Springer, New York, pp. 1–15.

Axelrod, Robert, 1984. The Evolution of Cooperation. Basic Books, New York.

Ball, Philip, 2012. Physicists suggest selfishness can pay. Nature. http://dx.doi.org/10.1038/nature.2012.11254.

Barlo, Mehmet, Carmona, Guilherme, Sabourian, Hamid, 2009. Repeated games with one-memory. J. Econ. Theory 144, 312–336.

Bergstrom, Carl T., Lachmann, Michael, 2003. The red king effect: when the slowest runner wins the coevolutionary race. Proc. Natl. Acad. Sci. USA 100, 593–598.

Boerlijst, Maarten C., Nowak, Martin A., Sigmund, Karl, 1997. Equal pay for all prisoners. Am. Math. Mon. 104, 303–307.

Camerer, Colin F., 2003. Behavioral Game Theory. Experiments in Strategic Interactions. Princeton University Press, Princeton.

Colman, Andrew M., 1995. Game Theory and Its Applications in the Social and Biological Sciences. Butterworth–Heinemann, Oxford.

Dal Bó, Pedro, Fréchette, Guillaume R., 2011. The evolution of cooperation in infinitely repeated games: experimental evidence. Amer. Econ. Rev. 101, 411–429.

Damore, James, Gore, Jeff, 2011. A slowly evolving host moves first in symbiotic interactions. Evolution 65, 2391–2398.

Duersch, Peter, Oechssler, Jörg, Schipper, Burkhard C., 2012. Unbeatable imitation. Games Econ. Behav. 76, 88–96.

Fehr, Ernst, Fischbacher, Urs, 2003. The nature of human altruism. Nature 425, 785–791.

Friedman, James W., 1971. A non-cooperative equilibrium for supergames. Rev. Econ. Stud. 38, 1–12.

Fudenberg, Drew, Maskin, Eric, 1986. The folk theorem in repeated games with discounting or with incomplete information. Econometrica 50, 533–554.

Fudenberg, Drew, Maskin, Eric, 1990. Evolution and cooperation in noisy repeated games. Amer. Econ. Rev. 80, 274–279.

Fudenberg, Drew, Dreber, Anna, Rand, David G., 2012. Slow to anger and fast to forgive: cooperation in an uncertain world. Amer. Econ. Rev. 102, 720–749.

Gokhale, Chaitanya, Traulsen, Arne, 2012. Mutualism and evolutionary multiplayer games: revisiting the Red King. Proc. - Royal Soc., Biol. Sci. 279, 4611–4616.

Hilbe, Christian, Nowak, Martin A., Sigmund, Karl, 2013a. The evolution of extortion in iterated prisoner's dilemma games. Proc. Natl. Acad. Sci. USA 110, 6913–6918.

Hilbe, Christian, Nowak, Martin A., Traulsen, Arne, 2013b. Adaptive dynamics of extortion and compliance. PLoS ONE 8, e77886.

Hilbe, Christian, Röhl, Torsten, Milinski, Manfred, 2014a. Extortion subdues human players but is finally punished in the prisoner's dilemma. Nat. Commun. 5, 3976.

Hilbe, Christian, Wu, Bin, Traulsen, Arne, Nowak, Martin A., 2014b. Cooperation and control in multiplayer social dilemmas. Proc. Natl. Acad. Sci. USA 111, 16425–16430.

Hofbauer, Josef, Sigmund, Karl, 1998. Evolutionary Games and Population Dynamics. Cambridge University Press, Cambridge.

Kagel, John H., Roth, Alvin E., 1997. The Handbook of Experimental Economics. Princeton University Press, Princeton.

Kalai, Ehud, 1990. Bounded rationality and strategic complexity in repeated games. In: Ichiishi, T., Neyman, A., Tauman, Y. (Eds.), Game Theory and Applications. Academic Press, San Diego, pp. 131–157.

Kalai, Adam T., Kalai, Ehud, Lehrer, Ehud, Samet, Dov, 2010. A commitment folk theorem. Games Econ. Behav. 69, 127–137.

Mailath, George J., Olszewski, Wojciech, 2011. Folk theorems with bounded recall under (almost) perfect monitoring. Games Econ. Behav. 71, 174–192.

Mailath, George J., Samuelson, Larry, 2006. Repeated Games and Reputations. Oxford University Press, Oxford.

Milinski, Manfred, Wedekind, Claus, 1998. Working memory constrains human cooperation in the prisoner's dilemma. Proc. Natl. Acad. Sci. USA 95, 13755–13758.

Myerson, Roger B., 1991. Game Theory: Analysis of Conflict. Harvard University Press, Cambridge, MA.

Nowak, Martin A., 2006. Evolutionary Dynamics. Harvard University Press, Cambridge, MA.

Nowak, Martin A., Sigmund, Karl, 1995. Invasion dynamics of the finitely repeated prisoner's dilemma. Games Econ. Behav. 11, 364–390.

Nowak, Martin, Sasaki, Akira, Taylor, Christine, Fudenberg, Drew, 2004. Emergence of cooperation and evolutionary stability in finite populations. Nature 428, 646–650.

Press, William H., Dyson, Freeman J., 2012. Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. Proc. Natl. Acad. Sci. USA 109, 10409–10413.

Rapoport, Anatol, Chammah, Albert M., 1965. The Prisoner's Dilemma. University Michigan Press, Ann Arbor.

Roth, Alvin E., Murnighan, J. Keith, 1978. Equilibrium behavior and repeated play of the prisoner's dilemma. J. Math. Psychol. 17, 189–198.

Samuelson, Larry, 1997. Evolutionary Games and Equilibrium Selection. MIT Press, Cambridge, MA.

Sandholm, William H., 2010. Population Games and Evolutionary Dynamics. MIT Press, Cambridge MA.

Sigmund, Karl, 2010. The Calculus of Selfishness. Princeton University Press, Princeton.

Stahl, Dale O., 1991. The graph of prisoners' dilemma supergame payoffs as a function of the discount factor. Games Econ. Behav. 3, 368–384.

Stewart, Alexander J., Plotkin, Joshua B., 2012. Extortion and cooperation in the prisoner's dilemma. Proc. Natl. Acad. Sci. USA 109, 10134–10135.

Stewart, Alexander J., Plotkin, Joshua B., 2013. From extortion to generosity, evolution in the iterated prisoner's dilemma. Proc. Natl. Acad. Sci. USA 110, 15348–15353.

Stewart, Alexander J., Plotkin, Joshua B., 2014. Collapse of cooperation in evolving games. Proc. Natl. Acad. Sci. USA 111, 17558–17563.

Sugden, Robert, 1986. The Economics of Rights, Co-Operation and Welfare. Blackwell, Oxford.

Szolnoki, Attila, Perc, Matjaz, 2014a. Evolution of extortion in structured populations. Phys. Rev. E 89, 022804.

Szolnoki, Attila, Perc, Matjaz, 2014b. Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations. Sci. Rep. 4, 5496.

Trivers, Robert L., 1971. The evolution of reciprocal altruism. Q. Rev. Biol. 46, 35–57.

Weibull, Jorgen W., 1995. Evolutionary Game Theory. MIT Press, Cambridge, MA.

Wu, Zhi-Xi, Rong, Zhihai, 2014. Boosting cooperation by involving extortion in spatial prisoner's dilemma games. Phys. Rev. E 90, 062102.

Yamagishi, Toshio, Kanazawa, Satoshi, Mashima, Rie, Terai, Shigeru, 2005. Separating trust from cooperation in a dynamic relationship: prisoner's dilemma with variable dependence. Ration. Soc. 17, 275–308.