Economics Letters 250 (2025) 112300

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Can I afford to remember less than you? Best responses in repeated additive games

Franziska Lesigang [a,b,*], Christian Hilbe [b,c], Nikoleta E. Glynatsi [b,d]

[a] Institute of Discrete Mathematics and Geometry, TU Wien, Wiedner Hauptstrasse 8-10, Wien 1040, Austria
[b] Max Planck Research Group Dynamics of Social Behavior, Max Planck Institute for Evolutionary Biology, August-Thienemann-Straße 2, Plön, 24306, Germany
[c] IT:U Interdisciplinary Transformation University, Altenberger Straße 66c, Linz, 4020, Austria
[d] Discrete Event Simulation Research Team, RIKEN Center for Computational Science, Kobe, 650-0047, Japan

ARTICLE INFO

Keywords:
Repeated games
Best response
Additive games
Bounded rationality

ABSTRACT

In this paper, we study best responses in repeated additive games among two players. A stage game is additive if each player's payoff is the sum of two components, and each component only depends on the action of a single player. We suppose one player's strategy depends on the co-player's last n actions. Then we prove that the other player has a best response that only depends on their own $n-1$ actions. That is, for an important sub-class of games and strategies, players can achieve maximum payoffs even with less memory than their opponent.

1. Introduction

Repeated games are a foundational framework in game theory (Mailath and Samuelson, 2006). They allow the study of complex behaviors such as cooperation and reciprocity (Axelrod and Hamilton, 1981; García and van Veelen, 2016). When analyzing such behaviors, individuals are sometimes assumed to have bounded recall (e.g. Ueda, 2021; Mailath and Olszewski, 2011). For example, evolutionary studies often focus on players with *memory-n* strategies. Such players only take into account the outcome of the last n rounds (these strategies do not depend on calendar time). An important subset are *reactive-n* strategies; they only depend on the *co-player's* previous n actions (Glynatsi et al., 2024). Strategies with finite memory are relevant because they seem cognitively plausible, and can be studied analytically.

If one player has finite memory, an interesting question arises: Could the other player take advantage by adopting a strategy with more memory? As shown by Press and Dyson (2012), for memory-n strategies the answer is negative. They consider the repeated prisoner's dilemma with undiscounted payoffs. They show: Against a player with an arbitrary but fixed memory-n strategy, any payoff that can be realized in principle (possibly with a more complex strategy) can already be realized with some memory-n strategy. In particular, any memory-n strategy has a memory-n best response. This result has been considerably extended by Levínskỳ et al. (2020). They prove similar results for more general strategy spaces and game types (including stochastic games), and for games with discounting. Moreover, they show that the respective best responses can be taken to be *pure* (i.e., there is no better response in mixed or stochastic strategies).

Herein, we address an even stronger question. Suppose one player takes into account the last n rounds. Could the co-player afford to remember even less, without any harm to their payoff? We give a positive answer for reactive-n strategies when the game is additive. A two-player stage game is additive if each player's payoff is the sum of two components. One component only depends on the first player's action, and the other component only on the second player's action. Such games play a crucial role in evolutionary game theory, where the respective property has been termed 'equal gains from switching' (Nowak and Sigmund, 1990). We prove: Suppose the stage game is additive, and one player uses a reactive-n strategy. Then the co-player has a pure best response that only depends on their own last n–1 actions.

This result is remarkable for two reasons. First, it offers a sufficient condition for when individuals can afford to remember less than their opponent. Second, it allows researchers to identify equilibria among reactive-n strategies more efficiently, as it reduces the number of deviations that need to be checked.

2. Model

Game setup. We consider a two-player iterated game without discounting. We assume the game is symmetric, although that assumption is not essential for our results. Each round, players choose one of m

* Corresponding author at: Institute of Discrete Mathematics and Geometry, TU Wien, Wiedner Hauptstrasse 8-10, Wien 1040, Austria.
E-mail addresses: franziska.lesigang@tuwien.ac.at (F. Lesigang), christian-hilbe@it-u.at (C. Hilbe), nikoleta.glynatsi@riken.jp (N.E. Glynatsi).

https://doi.org/10.1016/j.econlet.2025.112300
Received 14 February 2025; Received in revised form 20 March 2025; Accepted 20 March 2025
Available online 30 March 2025

actions. We refer to the action set as $\mathcal{A} := \{A_1, \dots, A_m\}$. The resulting payoffs are assumed to be additive (Maciejewski et al., 2014; McAvoy et al., 2021). This means one can find vectors $\mathbf{a} := (a_1, \dots, a_m)^T$ and $\mathbf{b} := (b_1, \dots, b_m)^T$ such that the payoff matrix $G = (g_{ij})$ of the stage game can be written as

$$G = \begin{pmatrix} a_1+b_1 & a_1+b_2 & \dots & a_1+b_m \\ a_2+b_1 & a_2+b_2 & \dots & a_2+b_m \\ \vdots & \vdots & \ddots & \vdots \\ a_m+b_1 & a_m+b_2 & \dots & a_m+b_m \end{pmatrix} \in \mathbb{R}^{m\times m}. \tag{1}$$

That is, a player's payoff can be decomposed into two independent components. Each component only depends on one player's action. An important example of an additive game is the donation game, a special variant of the prisoner's dilemma. Here, players choose among two actions, cooperation and defection. A cooperator pays a cost $c > 0$ to give a benefit of $b > c$ to the co-player. A defector pays nothing and gives nothing. This corresponds to an additive game with $\mathbf{a} = (-c, 0)$ and $\mathbf{b} = (b, 0)$.

We assume players use strategies with finite memory. To introduce such strategies formally, fix n. We use $\mathbf{h^i}$ to refer to player i's past n actions, $\mathbf{h^i} \in H_n^i := \mathcal{A}^n$. The elements of $\mathbf{h^i} = (h_1^i, \dots, h_n^i)$ are ordered such that the player's most recent actions come last. We call a tuple $\mathbf{h} = (\mathbf{h^1}, \mathbf{h^2})$ an n-history, or briefly history. We refer to $H_n := H_n^i \times H_n^{-i}$ as the set of all histories. In an m-action game, there are m^{2n} possible histories. We interpret the first player as the player whose payoff we wish to calculate. Accordingly, we speak of the *focal player*, and we call the co-player the *opponent*.

A memory-n strategy only depends on the past n rounds. Formally, it is a function that returns for each possible n-history a probability distribution over the possible actions. We denote this set of functions as

$$\mathcal{M}_n := (\Delta^{m-1})^{H_n}. \tag{2}$$

Here, $\Delta^{m-1} := \left\{ \mathbf{x} \in \mathbb{R}^m_{\geq 0} \;\middle|\; \sum_{i=1}^m x_i = 1 \right\}$. For given history $\mathbf{h}$, the ith entry of the vector $x_{\mathbf{h}} \in \Delta^{m-1}$ specifies the probability of taking action A_i. We note that for $n' < n$, any memory-n' strategy can be represented as memory-n strategy. We also note that the above definition does not determine the players' actions during the first n rounds, where no full n-history is yet available. However, because we consider games without discounting, we see further below that these initial actions are irrelevant for our result.

Two subsets of memory-n strategies are particularly important. First, *reactive-n strategies* are those that only depend on the opponent's last n actions. The respective set can be identified with

$$\mathcal{R}_n := (\Delta^{m-1})^{H_n^{-i}}. \tag{3}$$

Similarly, *self-reactive-n strategies* depend only on a player's own last n actions,

$$\mathcal{S}_n := (\Delta^{m-1})^{H_n^{i}}. \tag{4}$$

A strategy is called *pure* if the range of the function only contains unit vectors. Such strategies give a deterministic response to each history. For the set of pure self-reactive strategies, we write $\mathcal{S}_n^{pure}$.

Computation of payoffs. Given two (arbitrary) strategies σ and $\tilde{\sigma}$, let $\pi_{\sigma,\tilde{\sigma}}(t)$ denote the focal player's expected payoff in round t. We define the repeated-game payoff as the limiting average,

$$\pi(\sigma, \tilde{\sigma}) := \lim_{\tau\to\infty} \frac{1}{\tau} \sum_{t=1}^{\tau} \pi_{\sigma,\tilde{\sigma}}(t).$$

If the two strategies $(\sigma, \tilde{\sigma})$ are memory-n, this limit always exists (but it may depend on the initial n actions). To compute this payoff, we represent the game as a Markov chain. The states of the Markov chain are the game's n-histories $\mathbf{h} \in H_n$. The probability of transitioning from state $\mathbf{h}$ to state $\tilde{\mathbf{h}}$ is then given by

$$M_{\mathbf{h},\tilde{\mathbf{h}}} = m_\sigma \cdot m_{\tilde{\sigma}}. \tag{5}$$

Here, m_σ is defined as

$$m_\sigma = \begin{cases} \sigma(\mathbf{h})_i & \text{if } h_k = \tilde{h}_{k-1} \ \forall\ k \in \{2, \dots, n\} \text{ and } \tilde{h}_n = A_i, \\ 0 & \text{otherwise,} \end{cases}$$

and $m_{\tilde{\sigma}}$ similarly. For $t \geq n$, let $\mathbf{v}(t) \in \Delta^{|H_n|-1}$ denote the probability distribution of observing each history in round t. For a given distribution $\mathbf{v}(t)$, the next round's distribution is computed as $\mathbf{v}(t+1) = \mathbf{v}(t)M$. Since M is a stochastic matrix, the Perron–Frobenius theorem ensures that $\mathbf{v}(t)$ converges to a limiting distribution $\mathbf{v} = (v_{\mathbf{h}})$, which is a left eigenvector of M with eigenvalue 1. If M is primitive, this limiting distribution is unique; otherwise it is uniquely determined by the outcome of the first n rounds.

We can use this insight to compute the players' payoffs, by noting that the focal player's expected payoff in round $t \geq n$ can be written as

$$\pi_{\sigma,\tilde{\sigma}}(t) = \mathbf{v}(t) \cdot \mathbf{u}.$$

Here, $\mathbf{u} = (u_{\mathbf{h}})$ is the vector that assigns to every n-history the focal player's latest stage payoff. That is, if the history $\mathbf{h} = (\mathbf{h^1}, \mathbf{h^2})$ is such that $h_n^1 = A_i$ and $h_n^2 = A_j$, then $u_{\mathbf{h}} = g_{ij}$. If $\mathbf{v}(t) \to \mathbf{v}$ for $t \to \infty$, then so does $\frac{1}{\tau} \sum_{t=n}^{\tau} \mathbf{v}(t)$. We obtain

$$\pi(\sigma, \tilde{\sigma}) = \lim_{\tau\to\infty} \frac{1}{\tau} \sum_{t=1}^{\tau} \pi_{\sigma,\tilde{\sigma}}(t) = \lim_{\tau\to\infty} \frac{1}{\tau} \sum_{t=n}^{\tau} (\mathbf{v}(t) \cdot \mathbf{u})$$
$$= \left(\lim_{\tau\to\infty} \frac{1}{\tau} \sum_{t=n}^{\tau} \mathbf{v}(t) \right) \cdot \mathbf{u} = \mathbf{v} \cdot \mathbf{u}.$$

In this study, we assume the opponent uses some reactive-n strategy. We ask what kind of memory the focal player needs to have to implement a best response. To this end, let $\Sigma := \cup_n \mathcal{M}_n$ denote the set of all finite-memory strategies. For a reactive-n strategy p, we say $q \in \Sigma$ is a best response if

$$\pi(q, p) \geq \pi(\sigma, p) \quad \text{for all } \sigma \in \Sigma. \tag{6}$$

The work of Levínskỳ et al. (2020) guarantees there is always a best response among the pure self-reactive-n strategies:

Theorem 2.1. *Let $p \in \mathcal{R}_n$ be a reactive-n strategy. Then there exists a $q \in \mathcal{S}_n^{pure}$ that is a best response.*

This result does not only simplify the search for best responses. It also reduces the complexity of certain payoff calculations. Suppose we are to compute the payoff $\pi(q, p)$ with $q \in \mathcal{S}_n$ and $p \in \mathcal{R}_n$. Instead of interpreting these strategies as memory-n strategies and considering the $m^{2n} \times m^{2n}$ transition matrix M according to Eq. (5), we can consider a simpler transition matrix. Since both strategies p and q only depend on the past n actions of the self-reactive player 1, it suffices to consider a $m^n \times m^n$ dimensional matrix $\tilde{M}$, defined by:

$$\tilde{M}_{\mathbf{h^1},\tilde{\mathbf{h}}^1} = \begin{cases} q(\mathbf{h^1})_j & h_k^1 = \tilde{h}_{k-1}^1 \ \forall k \in \{2, \dots, n\} \text{ and } \tilde{h}_n^1 = A_j, \\ 0 & \text{otherwise.} \end{cases}$$

Based on the limiting distribution $\tilde{\mathbf{v}} = (\tilde{v}_{\mathbf{h^1}})$ of $\tilde{M}$, we can compute the payoff of q against p as

$$\pi(q, p) = \sum_{\mathbf{h^1} \in H_n^1} \tilde{v}_{\mathbf{h^1}} \sum_{j,k=1}^{m} g_{jk} \cdot q(\mathbf{h^1})_j \cdot p(\mathbf{h^1})_k. \tag{7}$$

3. Results

Theorem 2.1 holds for any payoff matrix. In the following we prove a stronger result when the game is additive.

Theorem 3.1. *Let $p \in \mathcal{R}_n$ be a reactive-n strategy and the game be additive. Then there exists a $q \in \mathcal{S}_{n-1}^{pure}$ that is a best response.*

The following observation simplifies the proof: Based on Theorem 2.1, there is a best response among the pure self-reactive-n strategies. Because the set of n-histories $\mathbf{h^1} \in H^1$ for this self-reactive player

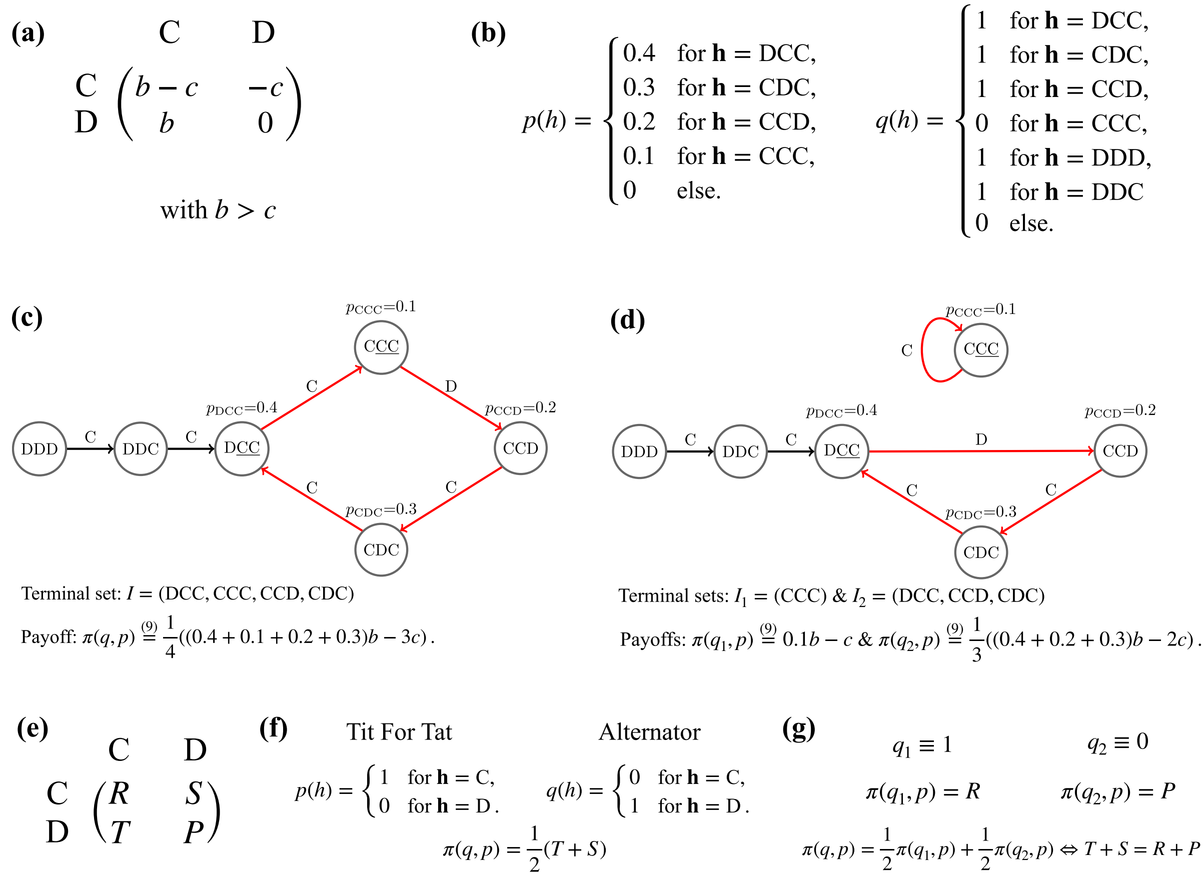

Fig. 1. (a) We illustrate the proof of Theorem 3.1 for the special case of a repeated donation game. **(b)** We suppose the co-player employs some reactive-3 strategy p, and the focal player adopts a pure self-reactive-3 strategy q. **(c)** Because the focal player's strategy is self-reactive, its actions are independent of the opponent. Starting from an initial sequence DDD, q's terminal set is $I=(\text{DCC},\text{CCC},\text{CCD},\text{CDC})$. Only the nodes in the terminal set are relevant for the long-run payoff. We observe that q behaves differently in the states CCC and DCC, even though these states coincide in the last 2 bits. This indicates that condition (8) is violated, and that q cannot be directly interpreted as a self-reactive-2 strategy. **(d)** Instead, suppose the player uses C after CCC and D after DCC. Then we can define appropriate self-reactive-2 strategies q_1 and q_2, with q_1 having the terminal set $I_1=(\text{CCC})$, and q_2 having the terminal set $I_2=(\text{DCC},\text{CCD},\text{CDC})$. In this example, q_2 gives a superior payoff compared to q, $\pi(q_2,p)\geq\pi(q,p)$. **(e)** We also illustrate why additivity of the payoff matrix is necessary. Let $\{\text{C},\text{D}\}$ define the action space, and consider a *non-additive* game defined by the given payoff matrix. **(f)** Let q be the pure self-reactive-1 strategy *alternator* and let q be the reactive-1 strategy *tit-for-tat*. In the game between p and q, players cooperate alternatingly, one in even rounds and the other in odd rounds. Thus, the average payoff of q is $\frac{1}{2}(T+S)$. **(h)** For the pure self-reactive-0 strategies $q_1\equiv 1$ and $q_2\equiv 0$ obtained in the proof, their payoffs are R and P respectively. The payoff of q is their average if and only if $T+S=R+P$, thus if and only if the game is additive.

is finite, and because the self-reactive player's strategy is pure, the play of this player will eventually reach a finite terminal set. Once the self-reactive player reaches this set I, there is a fixed sequence of histories that will occur in an indefinite loop. As a result, the invariant distribution $\tilde{\mathbf{v}}$ puts full weight on a uniform distribution over all states in I. For the resulting payoffs, we obtain

$$\begin{aligned}
\pi(q,p) &\overset{(7)}{=} \sum_{\mathbf{h^1}\in H_n^1}\hat{v}_{\mathbf{h^1}}\sum_{j,k=1}^{m} g_{jk}\,q(\mathbf{h^1})_j\,p(\mathbf{h^1})_k\\
&= \sum_{\mathbf{h^1}\in I}\frac{1}{|I|}\sum_{j,k=1}^{m}(a_j+b_k)\,q(\mathbf{h^1})_j\,p(\mathbf{h^1})_k\\
&= \sum_{\mathbf{h^1}\in I}\frac{1}{|I|}\left(\sum_{j,k=1}^{m}a_j\,q(\mathbf{h^1})_j\,p(\mathbf{h^1})_k+\sum_{j,k=1}^{m}b_k\,q(\mathbf{h})_j\,p(\mathbf{h^1})_k\right)\\
&= \sum_{\mathbf{h^1}\in I}\frac{1}{|I|}\left(\sum_{j=1}^{m}a_j\,q(\mathbf{h})_j\underbrace{\sum_{k=1}^{m}p(\mathbf{h^1})_k}_{=1}+\sum_{k=1}^{m}b_k\,p(\mathbf{h^1})_k\underbrace{\sum_{j=1}^{m}q(\mathbf{h^1})_j}_{=1}\right)\\
&= \sum_{\mathbf{h^1}\in I}\frac{1}{|I|}\left(\sum_{j=1}^{m}a_j\,q(\mathbf{h^1})_j+\sum_{k=1}^{m}b_k\,p(\mathbf{h^1})_k\right)\\
&= \frac{1}{|I|}\sum_{\mathbf{h^1}\in I}\left(\mathbf{b}\cdot p(\mathbf{h^1})+\mathbf{a}\cdot q(\mathbf{h^1})\right).
\end{aligned}$$

We can now prove Theorem 3.1 (for an illustration of the main argument, see Fig. 1a–d).

Proof of Theorem 3.1. Let q be a self-reactive-n strategy that is a best response to p (which exists by Theorem 2.1). Let I denote q's terminal set, and let $N := |I|$.

If $N=1$, then after a finite number of actions, q plays a single action. Thus, there exists a pure self-reactive-0 strategy q_0 with the same terminal set and thus the same payoff against p. This concludes the proof for $N=1$.

Now let $N>1$. Since q is deterministic, for every state $\mathbf{h^1_k}\in I$, there is a unique $\mathbf{h^1_i}\in I\setminus\{\mathbf{h^1_k}\}$ that follows $\mathbf{h^1_k}$. We sort the elements of I in order of their occurrence $(\mathbf{h^1_1},\ldots,\mathbf{h^1_N})$. For $i\in\{1,\ldots,N\}$, let $\hat{\mathbf{h}}^1_i$ be the projection of $\mathbf{h^1_i}$ onto its last n–1 components. We distinguish two cases.

Case 1. Suppose all histories in I satisfy

$$\hat{\mathbf{h}}^1_k=\hat{\mathbf{h}}^1_j \Rightarrow \hat{\mathbf{h}}^1_{k+1}=\hat{\mathbf{h}}^1_{j+1}\ \text{for all } k,j\in\{1,\ldots,N\} \tag{8}$$

In that case, q's behavior on I depends only on its last $n-1$ actions. Then q is equivalent to a pure self-reactive-$(n-1)$ strategy q' with $\pi(q',p)=\pi(q,p)$. (During the initial n rounds, this q' could just play actions along the loop defined by the terminal set).

Case 2. Otherwise, there exist $j,k\in\{1,\ldots,N\}$ for which the implication in (8) is false. We define $\pentagon := \hat{\mathbf{h}}^1_k=\hat{\mathbf{h}}^1_j$ and the distinct actions $X,Y\in\mathcal{A}$ such that $(X,\pentagon)=\mathbf{h^1_k}$ and $(Y,\pentagon)=\mathbf{h^1_j}$. Further, we denote their successors in I as $(\pentagon,Z)=\mathbf{h^1_{k+1}}$ and $(\pentagon,T)=\mathbf{h^1_{j+1}}$ where $T\neq Z$ and $T,Z\in\mathcal{A}$. Then, the Markov chain graph of the terminal set I takes the form

$$(X, \pentagon) \xrightarrow{Z} (\pentagon, Z) \overset{1}{\rightsquigarrow} (Y, \pentagon) \xrightarrow{T} (\pentagon, T) \overset{2}{\rightsquigarrow} (X, \pentagon),$$

where $\rightarrow$ denotes an edge, $\overset{1}{\rightsquigarrow}$ represents the path from $(\pentagon, Z)$ to $(Y, \pentagon)$, and $\overset{2}{\rightsquigarrow}$ represents the path from $(\pentagon, T)$ to $(X, \pentagon)$. Note that both paths can be empty. We observe that the two graphs

1. $(X, \pentagon) \xrightarrow{T} (\pentagon, T) \overset{2}{\rightsquigarrow} (X, \pentagon)$,
2. $(Y, \pentagon) \xrightarrow{Z} (\pentagon, Z) \overset{1}{\rightsquigarrow} (Y, \pentagon)$,

are well-defined deterministic Markov chains on two sets I_1 and I_2. Thus, there are pure self-reactive-n strategies q_1 and q_2 with I_1 and I_2 as their terminal sets. They are defined by

$$q_1(\mathbf{h^1}) = \begin{cases} q(\mathbf{h^1}) & \text{for } \mathbf{h^1} \in I_1 \setminus \{(X, \pentagon)\} \\ q\big((Y, \pentagon)\big) & \text{for } \mathbf{h^1} = (X, \pentagon) \end{cases}$$

and

$$q_2(\mathbf{h^1}) = \begin{cases} q(\mathbf{h^1}) & \text{for } \mathbf{h^1} \in I_2 \setminus \{(Y, \pentagon)\} \\ q\big((X, \pentagon)\big) & \text{for } \mathbf{h^1} = (Y, \pentagon). \end{cases}$$

(Again, to define the initial n moves of q_1 and q_2, one just requires them to play actions compatible with the loops defined by I_1 and I_2).

Note that $I_1 \cup I_2 = I$ and $I_1 \cap I_2 = \emptyset$. Therefore,

$$\begin{aligned} \pi(q,p) &= \frac{1}{|I|} \sum_{\mathbf{h^1} \in I} \big(\mathbf{b}p(\mathbf{h^1}) + \mathbf{a}q(\mathbf{h^1})\big) \\ &= \frac{1}{|I|}\left(\sum_{\mathbf{h^1} \in I_1} \big(\mathbf{b}p(\mathbf{h^1}) + \mathbf{a}q(\mathbf{h^1})\big) + \sum_{\mathbf{h^1} \in I_2} \big(\mathbf{b}p(\mathbf{h^1}) + \mathbf{a}q(\mathbf{h^1})\big) \right) \\ &= \frac{1}{|I|}\left(\sum_{\mathbf{h^1} \in I_1} \big(\mathbf{b}p(\mathbf{h^1}) + \mathbf{a}q_1(\mathbf{h^1})\big) + \sum_{\mathbf{h^1} \in I_2} \big(\mathbf{b}p(\mathbf{h^1}) + \mathbf{a}q_2(\mathbf{h^1})\big) \right) \\ &= \frac{1}{|I|}\left(\frac{|I_1|}{|I_1|} \sum_{\mathbf{h^1} \in I_1} \big(\mathbf{b}p(\mathbf{h^1}) + \mathbf{a}q_1(\mathbf{h^1})\big) + \frac{|I_2|}{|I_2|} \sum_{\mathbf{h^1} \in I_2} \big(\mathbf{b}p(\mathbf{h^1}) + \mathbf{a}q_2(\mathbf{h^1})\big) \right) \\ &= \frac{|I_1|}{|I|}\pi(q_1,p) + \frac{|I_2|}{|I|}\pi(q_2,p). \end{aligned}$$

Note that for the third equality, additivity is crucial. We have now proven that the payoff $\pi(q,p)$ is a convex combination of $\pi(q_1,p)$ and $\pi(q_2,p)$. Therefore, it is dominated by at least one of the terms — either $\pi(q_1,p) \geq \pi(q,p)$ or $\pi(q_2,p) \geq \pi(q,p)$. We now check the two cases with the dominating strategy. After finitely many steps, we end up in Case 1. □

4. Discussion

Herein, we study optimal behavior in repeated games. We suppose the opponent uses a reactive-n strategy, taking into account the focal player's last n actions. We ask how many rounds the focal player needs to keep in memory to have a best response. For additive games, we give a surprising answer: it suffices to remember at most $n-1$ past actions.

Apart from being noteworthy on a conceptual level, this result has practical relevance for research. It simplifies the task of identifying Nash equilibria among reactive-n strategies. For example, suppose we ask whether a given reactive-3 strategy p is a Nash equilibrium of the repeated prisoner's dilemma. According to Theorem 2.1, this task requires comparing $\pi(p,p)$ to the payoff $\pi(q,p)$ of $2^{2^3} = 256$ self-reactive-3 strategies q. Instead, for additive games, we show it suffices to check $2^{2^2} = 16$ strategies only.

Interestingly, both requirements in Theorem 3.1 are essential. Neither is the result true if reactive-n strategies are replaced by the more general memory-n strategies. Nor does it hold if, say, the additive donation game is replaced by a non-additive prisoner's dilemma, see Fig. 1e–g. Nevertheless, our result is significant, as both reactive strategies and additive games have been key concepts in evolutionary game theory. There, they serve as major tools to describe the evolution of reciprocal cooperation in adaptive populations (e.g. Nowak and Sigmund, 1990; Imhof and Nowak, 2010; Glynatsi et al., 2024).

Acknowledgments

N.E.G. and C.H. acknowledge generous support from the European Research Council Starting Grant 850529: E-DIRECT. F.L., N.E.G., and C.H. also acknowledge support from the Max Planck Society. In addition, they acknowledge TU Wien Bibliothek for financial support through its Open Access Funding Program.

Data availability

No data was used for the research described in the article.

References

Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. Science 211, 1390–1396.
García, J., van Veelen, M., 2016. In and out of equilibrium I: Evolution of strategies in repeated games with discounting. J. Econom. Theory 161, 161–189.
Glynatsi, N.E., Akin, E., Nowak, M.A., Hilbe, C., 2024. Conditional cooperation with longer memory. In: Proceedings of the National Academy of Sciences.
Imhof, L.A., Nowak, M.A., 2010. Stochastic evolutionary dynamics of direct reciprocity. Proc. R. Soc. B 277, 463–468.
Levínskỳ, R., Neyman, A., Zelenỳ, M., 2020. Should I remember more than you? Best responses to factored strategies. Int. J. Game Theory 49, 1105–1124.
Maciejewski, W., Fu, F., Hauert, C., 2014. Evolutionary game dynamics in populations with heterogenous structures. PLoS Comput. Biol. 10, e1003567.
Mailath, G.J., Olszewski, W., 2011. Folk theorems with bounded recall under (almost) perfect monitoring. Games Econom. Behav. 71, 174–192.
Mailath, G.J., Samuelson, L., 2006. Repeated Games and Reputations. Oxford Univ. Press, Oxford, UK.
McAvoy, A., Rao, A., Hauert, C., 2021. Intriguing effects of selection intensity on the evolution of prosocial behaviors. PLoS Comput. Biol. 17, e1009611.
Nowak, M., Sigmund, K., 1990. The evolution of stochastic strategies in the prisoner's dilemma. Acta Appl. Math. 20, 247–265.
Press, W.H., Dyson, F.J., 2012. Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. Proc. Natl. Acad. Sci. 109, 10409–10413.
Ueda, M., 2021. Memory-two zero-determinant strategies in repeated games. R. Soc. Open Sci. 8, 202186.