

Best response cycles in perfect information games

Citation for published version (APA):

Herings, P. J. J., & Predtetchinski, A. (2015). *Best response cycles in perfect information games*. Maastricht University, Graduate School of Business and Economics. GSBE Research Memoranda No. 017 <https://doi.org/10.26481/umagsb.2015017>

Document status and date:

Published: 01/01/2015

DOI:

[10.26481/umagsb.2015017](https://doi.org/10.26481/umagsb.2015017)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

P. Jean-Jacques Herings,
Arkadi Predtetchinski

**Best response cycles in
perfect information games**

RM/15/017

GSBE

Maastricht University School of Business and Economics
Graduate School of Business and Economics

P.O. Box 616
NL- 6200 MD Maastricht
The Netherlands

Best response cycles in perfect information games

P. Jean–Jacques Herings* and Arkadi Predtetchinski†

May 11, 2015

Abstract

We consider n –player perfect information games with payoff functions having a finite image. We do not make any further assumptions, so in particular we refrain from making assumptions on the cardinality or the topology of the set of actions and assumptions like continuity or measurability of payoff functions. We show that there exists a best response cycle of length four, that is, a sequence $(\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_0)$ of pure strategy profiles where every successive element is a best response to the previous one. This result implies the existence of point–rationalizable strategy profiles. When payoffs are only required to be bounded, we show the existence of an ϵ –best response cycle of length four for every $\epsilon > 0$.

Key words: Perfect information games, determinacy, best response cycles, rationalizability.

JEL codes: C72.

*Department of Economics, Maastricht University, P.O.Box 616, 6200 MD, The Netherlands. E-mail: p.herings@maastrichtuniversity.nl.

†Department of Economics, Maastricht University, P.O.Box 616, 6200 MD, The Netherlands. E-mail: a.predtetchinski@maastrichtuniversity.nl.

1 Introduction

We consider n -player perfect information games that are played over a tree of countably infinite length. At every node of the tree, called a history, a single player chooses an action. The resulting infinite sequence of actions is called a play and determines the payoff for every player.

We are very general with respect to the assumptions put on the set of available actions and the payoff functions. We make no assumptions on the set of available actions, so in particular we abstain from making assumptions on its cardinality or its topology. We assume payoff functions to have a finite image for some of our results and to have a bounded image for others, so we make no continuity or measurability assumptions with respect to payoff functions.

A special case of our set-up are win-lose games as studied by Gale and Stewart [1953]. Win-lose games are two-player zero-sum games where the payoff of a player is either zero or one. Another way to describe such games is to introduce a winning set of plays. Player 1 gets a payoff of one if and only if the play is an element of the winning set. Such a game is said to be determined if either player 1 has a winning strategy, i.e., a strategy that guarantees the play to be an element of the winning set irrespective of the strategy chosen by player 2, or player 2 has a winning strategy. It is easy to see that a win-lose game is determined if and only if it has a Nash equilibrium in pure strategies. Gale and Stewart [1953] show that there are win-lose games that are not determined.

We first consider n -player games under the assumption that the payoff functions have a finite image. Using elementary methods, we show that there exists a best response cycle of length four, that is, a sequence of pure strategy profiles $(\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_0)$ such that every successive element is a best response to its predecessor. The strategy profiles forming a best response cycle are point rationalizable (Bernheim [1984]) and they survive any procedure of iterated elimination of strictly dominated strategies (Dufwenberg and Stegeman [2002], Chen, Van Long, and Luo [2007]).

Since a Nash equilibrium corresponds to a best response cycle of length one, the results of Gale and Stewart [1953] demonstrate that even two-player zero-sum games may fail to have a best response cycle of length one. Our results imply that there always exists a best response cycle of length four.

We continue by weakening the assumption that payoff functions have a finite image to the assumption that payoff functions have a bounded image. If, moreover, the payoff functions were Borel measurable, then a result shown by Mertens and Neyman and reported in Mertens [1986] is that, for every positive ϵ , an ϵ -Nash equilibrium in pure strategies exists,

and therefore an ϵ -best response cycle of length one exists. We show that boundedness of the payoff functions is sufficient for the existence of an ϵ -best response cycle of length four.

The paper is organized as follows. Section 2 introduces perfect information games and Section 3 presents our results on the existence of best response cycles. The proof of the main result, Theorem 3.1, is provided in Section 4. Section 5 concludes.

2 Perfect information games

Let \mathbb{N} be the set of natural numbers including zero. Given a set A , we let $A^{<\mathbb{N}}$ denote the set of finite sequences of elements of A , including the empty sequence \emptyset . A non-empty sequence $h \in A^{<\mathbb{N}}$ is written as $h = (a_0, \dots, a_t)$ for some $t \in \mathbb{N}$. We let $A^{\mathbb{N}}$ be the set of countably infinite sequences $p = (a_0, a_1, \dots)$ where $a_t \in A$ for every $t \in \mathbb{N}$.

A *perfect information game* consists of the following elements:

- A non-empty set of *actions* A .
- A set of *histories* $H \subseteq A^{<\mathbb{N}}$, which is required to be a pruned tree. That is: [1] $\emptyset \in H$, [2] if $h = (a_0, \dots, a_{t'}) \in H$ and $0 \leq t \leq t'$, then $(a_0, \dots, a_t) \in H$, and [3] if $h \in H$, then there is $a \in A$ such that $(h, a) \in H$. A *play* is an element $p = (a_0, a_1, \dots)$ of $A^{\mathbb{N}}$ such that $(a_0, \dots, a_t) \in H$ for every $t \in \mathbb{N}$. The set of all plays is denoted by P .
- A finite set of *players* $I = \{1, \dots, n\}$.
- A function $\iota : H \rightarrow I$ that assigns an *active player* to every history. We let $H^i = \{h \in H : \iota(h) = i\}$ be the set of histories where player i is active.
- For every $i \in I$, a *payoff function* $u^i : P \rightarrow \mathbb{R}$.

The reader will notice that we exclude “terminal histories” as we assume that some action can be played after every history $h \in H$. Nevertheless, games with terminal histories are encompassed as a special case of our set-up. For example, games of length two are those where the payoff $u^i(p)$ depends only on the first two coordinates of p .

For two histories $h, h' \in H$, we say that h' *follows* h , denoted $h \leq h'$, if either $h = \emptyset$ or $h = (a_0, \dots, a_t)$ for some $t \in \mathbb{N}$ and $h' = (a_0, \dots, a_{t'})$ for some $t' \in \mathbb{N}$ such that $t \leq t'$. A history $h \in H$ is said to be a *prefix* of the play $p = (a_0, a_1, \dots)$ if $h = \emptyset$ or $h = (a_0, \dots, a_t)$ for some $t \in \mathbb{N}$. Given $h \in H$, we let $P(h)$ denote the set of all plays $p \in P$ such that h is a prefix of p . Let $A(h) = \{a \in A : (h, a) \in H\}$ be the set of actions available at h .

A pure strategy for player i is a function s^i with domain H^i such that $s^i(h) \in A(h)$ for every $h \in H^i$. We only consider pure strategies. Let S^i be the set of player i 's pure strategies and let $S = \prod_{i \in I} S^i$ be the set of pure strategy profiles. Given $s \in S$ and $h \in H$, we write $s(h)$ to mean $s^i(h)$ for $i = \iota(h)$. The play induced by s is denoted by $\pi(s)$. We write $U^i(s) = u^i(\pi(s))$. Given $s \in S$ and $\sigma^i \in S^i$, we let s/σ^i denote the strategy profile obtained from s by replacing player i 's strategy s^i by σ^i .

We work in Zermelo–Fraenkel set theory with the axiom of choice.

3 Best response cycles

We first state our definition and results for the case where for every player $i \in I$ the payoff functions u^i have a finite image. A strategy $\sigma^i \in S^i$ is *player i 's best response* to the strategy profile $s \in S$ if for every $\tau^i \in S^i$ it holds that $U^i(s/\tau^i) \leq U^i(s/\sigma^i)$. Notice that if the image of the payoff function u^i is finite, player i always has a best response to any strategy profile. A strategy profile $\sigma \in S$ is said to be a *best response* to $s \in S$ if for every $i \in I$, σ^i is player i 's best response to s .

Theorem 3.1 *Suppose that for every $i \in I$ the image of the payoff function u^i is finite. Then there exists a best response cycle of length 4, that is a sequence $(\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_0)$ of pure strategy profiles such that every successive element is a best response to the previous one.*

This result has important implications for the behavior of the best response map. For $s \in S$ let $B^i(s)$ denote the set of player i 's best responses to s and let $B(s)$ denote the product set of best responses to s , thus

$$B(s) = \prod_{i \in I} B^i(s).$$

Define the *best response map* $B : 2^S \rightarrow 2^S$ by letting $B(\emptyset) = \emptyset$ and, for every non-empty subset E of S ,

$$B(E) = \bigcup_{s \in E} B(s).$$

By construction, the map B is monotone with respect to set inclusion, i.e. $B(E) \subseteq B(E')$ if $E \subseteq E'$. Consequently, the set of fixed points of B is a complete lattice (Tarski [1955]). Of course \emptyset is a fixed point of B . The set $F = \cup\{E \in 2^S : E \subseteq B(E)\}$ is the greatest fixed point of B . Now Theorem 3.1 implies that there is a non-empty set $E \in 2^S$ such that $E \subseteq B(E)$, for instance the set consisting of the four strategy profiles forming the best response cycle. We thus obtain the following corollary.

Corollary 3.2 *Suppose that for every $i \in I$ the image of the payoff function u^i is finite. Then the set F is non-empty.*

Connection to determinacy: A win–lose game is said to be *determined* if either player 1 has a winning strategy, i.e., a strategy that guarantees the play to be an element of the winning set irrespective of the strategy chosen by player 2, or player 2 has a winning strategy. If we endow the set of plays P with a topology having as base the collection of cylinder sets $P(h)$ for $h \in H$, then it has been shown that a win–lose game is determined if the winning set is closed (Gale and Stewart [1953]), Borel (Martin [1975]), or quasi-Borel (Martin [1990]). Under these conditions, a win–lose game has a best response cycle of length 1. Theorem 3.1 implies that any win–lose game has a best response cycle of length 4.

Connection to point rationalizability: Bernheim [1984, top of page 1015] defines the map $\lambda : 2^S \rightarrow 2^S$ by

$$\lambda(E) = \prod_{i \in I} \bigcup_{s \in E} B^i(s).$$

The greatest fixed point of λ is the set of point rationalizable strategy profiles.

In the 2–player case, it holds that $B(E) = \lambda(E)$ for every $E \in 2^S$. This follows from the fact that the set of best responses of player 1 to the strategy profile (s^1, s^2) depends only on s^2 . If $n > 2$, it holds that $B(E) \subseteq \lambda(E)$, while the converse inclusion is not true in general. Essentially, the definition of B embodies the restriction that any two players hold the same beliefs about a third player, a restriction that is not imposed under the definition of λ .

The greatest fixed point F of B is contained in the greatest fixed point of λ . In particular, if the sequence of strategy profiles $(\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_0)$ is a best response cycle of length 4, then the strategy profiles $\sigma_0, \sigma_1, \sigma_2$, and σ_3 are point rationalizable. Theorem 3.1 thus implies the existence of point rationalizable strategy profiles in perfect information games where the payoff functions have a finite image.

Corollary 3.3 *Suppose that for every $i \in I$ the image of the payoff function u^i is finite. Then the set of point rationalizable strategy profiles is non-empty.*

Connection to strict dominance: For every $i \in I$, consider sets $D^i \subseteq E^i \subseteq S^i$, and let $D = \prod_{i \in I} D^i$ and $E = \prod_{i \in I} E^i$. Following Dufwenberg and Stegeman [2002], we write $E \rightarrow D$ if D is obtained from E by eliminating strictly dominated strategies, that is, for every $i \in I$, for every $s^i \in E^i \setminus D^i$, there exists $\sigma^i \in E^i$ such that for every $s^{-i} \in E^{-i}$ it

holds that $U^i(s^i, s^{-i}) < U^i(\sigma^i, s^{-i})$. We say that $E \subseteq S$ is a reduction of S if there exists a sequence $S = E_0 \rightarrow E_1 \rightarrow \dots$ such that $E = \bigcap_{m \in \mathbb{N}} E_m$.

It is easy to see that if $E \rightarrow D$ and $F \subseteq E$, then also $F \subseteq D$, where F is the greatest fixed point of the map B . Consequently, any reduction of S contains F as a subset and is non-empty. We thus conclude that the strategy profiles in F survive any procedure of iterative elimination of strictly dominated strategies. This conclusion remains valid even under a more permissive definition of reductions that allows for arbitrary transfinite sequences of sets as in Chen, Van Long, and Luo [2007], rather than just countable sequences as in Dufwenberg and Stegeman [2002].

Now we turn to the case where the image of u^i is allowed to be any bounded set of real numbers. The generalization of Theorem 3.1 is obtained through the concept of ϵ -best response.

Let ϵ be some positive real number. A strategy $\sigma^i \in S^i$ is said to be player i 's ϵ -best response to the strategy profile $s \in S$ if for every $\tau^i \in S^i$ it holds that $U^i(s/\tau^i) < U^i(s/\sigma^i) + \epsilon$. A strategy profile $\sigma \in S$ is said to be an ϵ -best response to $s \in S$ if for every player $i \in I$, σ^i is an ϵ -best response to s .

Theorem 3.4 *Suppose that for every $i \in I$ the image of the payoff function u^i is bounded. Then for every $\epsilon > 0$ there exists an ϵ -best response cycle of length 4, that is a sequence $(\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_0)$ of pure strategy profiles such that every successive element is an ϵ -best response to the previous one.*

To deduce Theorem 3.4 from Theorem 3.1, for $i \in I$, let \bar{u}^i be defined by letting $\bar{u}^i(p) = \min\{m\epsilon : u^i(p) \leq m\epsilon\}$, where m takes integer values. Since u^i is bounded, \bar{u}^i has a finite image, and $u^i(p) \leq \bar{u}^i(p) < u^i(p) + \epsilon$. Let \bar{U}^i denote the induced payoff function over strategy profiles.

Let $(\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_0)$ be a best response cycle of length 4 as in Theorem 3.1 for the game with payoff functions $(\bar{u}^1, \dots, \bar{u}^n)$. For $k \in \{0, 1, 2, 3\}$, for every $i \in I$, for every $\tau^i \in S^i$, it holds that

$$U^i(\sigma_k/\tau^i) \leq \bar{U}^i(\sigma_k/\tau^i) \leq \bar{U}^i(\sigma_k/\sigma_{k+1}^i) < U^i(\sigma_k/\sigma_{k+1}^i) + \epsilon,$$

where $\sigma_4^i = \sigma_0^i$. It therefore follows that $(\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_0)$ is an ϵ -best response cycle of length 4.

We define the maps B_ϵ^i and B_ϵ analogously to B^i and B . The map B_ϵ is called the ϵ -best response map. We let F_ϵ denote the greatest fixed point of B_ϵ . The following corollary is immediate.

Corollary 3.5 *Suppose that for every $i \in I$ the image of the payoff function u^i is bounded. Then for each $\epsilon > 0$ the set F_ϵ is non-empty.*

Let ϵ be a positive real number. A pure strategy profile $s \in S$ such that $s \in B_\epsilon(s)$ is called an ϵ -Nash equilibrium. An ϵ -Nash equilibrium corresponds to a best response cycle of length 1. A result due to Mertens and Neyman (reported by Mertens [1986]) asserts that a perfect information game admits an ϵ -Nash equilibrium for every $\epsilon > 0$ provided that the payoff functions are bounded and Borel measurable. The proof of existence of ϵ -equilibria in perfect information games by Mertens and Neyman relies on Borel determinacy as shown in Martin [1975]. Therefore, if payoff functions are bounded and Borel measurable, then there exists an ϵ -best response cycle of length 1. Theorem 3.4 implies that boundedness of the payoff functions is sufficient to guarantee the existence of an ϵ -best response cycle of length 4.

4 Proof of Theorem 3.1

For a game G as in Section 2 and a history $h \in H$, we define the subgame $G(h)$ of G as the game with set of actions A , set of players I , set of histories $H_h = \{e \in A^{<\mathbb{N}} : (h, e) \in H\}$, an assignment of active players $\iota_h : H_h \rightarrow I$ given by $\iota_h(e) = \iota(h, e)$, and payoff functions given by $u_h^i(p) = u^i(h, p)$ for every play p of $G(h)$. Here (h, e) and (h, p) are concatenations defined in the obvious way. The induced payoff functions over strategy profiles are denoted by U_h^i .

If σ is a strategy profile in G and τ a strategy profile in $G(h)$, we say that σ agrees with τ if $\sigma(h, e) = \tau(e)$ for every history e in $G(h)$. We start out with a very easy lemma.

Lemma 4.1 *Suppose that for every $a \in A(\emptyset)$ the subgame $G(a)$ has a best response cycle $(\sigma_{a,0}, \sigma_{a,1}, \sigma_{a,2}, \sigma_{a,3}, \sigma_{a,0})$ of length 4. Then the game G has a best response cycle $(\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_0)$ of length 4 such that for every $a \in A(\emptyset)$ the strategy σ_k agrees with $\sigma_{a,k}$.*

Proof: Let $i = \iota(\emptyset)$. Let some $k \in \{0, 1, 2, 3\}$ be given and let $a_k \in A(\emptyset)$ be such that $U_a^i(\sigma_{a,k-1}/\sigma_{a,k}^i)$ is maximized over $a \in A(\emptyset)$. Define the strategy profile σ_k by $\sigma_k(\emptyset) = a_k$ and, for every $a \in A(\emptyset)$, by letting σ_k agree with $\sigma_{a,k}$ in $G(a)$. It is easy to see that $(\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_0)$ is a best response cycle of length 4 in G . \square

For a play $p = (a_0, a_1, \dots)$, we define $p(0) = a_0$. A strategy profile s is said to have the property that *all players follow p* if for every prefix h of p it holds that $s(h) = a_0$ if $h = \emptyset$

and $s(h) = a_{t+1}$ if $h = (a_0, \dots, a_t)$ for some $t \in \mathbb{N}$. Let w^i be the highest possible payoff for player i and let W^i be the set of plays where the maximum is attained, so

$$w^i = \max_{p \in P} u^i(p) \text{ and } W^i = \{p \in P : u^i(p) = w^i\}.$$

Lemma 4.2 *Let $i = \iota(\emptyset)$. Suppose that there exist two plays, say p and q , such that $u^i(p) = u^i(q) = w^i$ and $p(0) \neq q(0)$. Then the game G has a best response cycle of length 4.*

Proof: Let σ_0 be a strategy profile such that all players follow p . Clearly, σ_0^i is player i 's best response to σ_0 . Choose some $\sigma_1 \in B(\sigma_0)$ with $\sigma_1^i = \sigma_0^i$ and choose $\sigma_2 \in B(\sigma_1)$. Symmetrically, let τ_0 be a strategy profile such that all players follow q . Choose some $\tau_1 \in B(\tau_0)$ with $\tau_1^i = \tau_0^i$ and choose $\tau_2 \in B(\tau_1)$.

Let the strategy profiles s_0, s_1, s_2 , and s_3 be such that

$$s_0^i = \sigma_0^i \text{ and } s_0^j(h) = \begin{cases} \sigma_0^j(h) & \text{if } p(0) < h \\ \tau_2^j(h) & \text{if } q(0) < h, \end{cases}$$

$$s_1^i = \sigma_0^i \text{ and } s_1^j(h) = \begin{cases} \sigma_1^j(h) & \text{if } p(0) < h \\ \tau_0^j(h) & \text{if } q(0) < h, \end{cases}$$

$$s_2^i = \tau_0^i \text{ and } s_2^j(h) = \begin{cases} \sigma_2^j(h) & \text{if } p(0) < h \\ \tau_0^j(h) & \text{if } q(0) < h, \end{cases}$$

$$s_3^i = \tau_0^i \text{ and } s_3^j(h) = \begin{cases} \sigma_0^j(h) & \text{if } p(0) < h \\ \tau_1^j(h) & \text{if } q(0) < h, \end{cases}$$

where $j \neq i$. Notice that s_0 is a strategy profile such that all players follow p and s_2 is a strategy profile such that all players follow q .

We argue that $s_1 \in B(s_0)$. To see that s_1^i is a best response to s_0 for player i , notice that $\pi(s_0/s_1^i) = p$, which is a maximizer of u^i . For $j \neq i$, it also holds that s_1^j is a best response to s_0 . Indeed, under s_0 player i plays action $p(0)$ at history \emptyset , the same action as under σ_0 . In the subgame $G(p(0))$, the strategy profile s_0 coincides with σ_0 while s_1^j coincides with σ_1^j . Since σ_1^j is a best response to σ_0 , it follows that s_1^j is a best response to s_0 for player j .

We argue that $s_2 \in B(s_1)$. To see that s_2^i is a best response to s_1 for player i , notice that $\pi(s_1/s_2^i) = q$, which is a maximizer of u^i . For $j \neq i$, it also holds that s_2^j is a best response to s_1 . Indeed, under s_1 player i plays action $p(0)$ at history \emptyset , the same action

as under σ_1 . In the subgame $G(p(0))$, the strategy profile s_1 coincides with σ_1 while s_2^j coincides with σ_2^j . Since σ_2^j is a best response to σ_1 , it follows that s_2^j is a best response to s_1 for player j .

We argue that $s_3 \in B(s_2)$. To see that s_3^i is a best response to s_2 for player i , notice that $\pi(s_2/s_3^i) = q$, which is a maximizer of u^i . For $j \neq i$, it also holds that s_3^j is a best response to s_2 . Indeed, under s_2 player i plays action $q(0)$ at history \emptyset , the same action as under τ_0 . In the subgame $G(q(0))$, the strategy profile s_2 coincides with τ_0 while s_3^j coincides with τ_1^j . Since τ_1^j is a best response to τ_0 , it follows that s_3^j is a best response to s_2 for player j .

We argue that $s_0 \in B(s_3)$. To see that s_0^i is a best response to s_3 for player i , notice that $\pi(s_3/s_0^i) = p$, which is a maximizer of u^i . For $j \neq i$, it also holds that s_0^j is a best response to s_3 . Indeed, under s_3 player i plays action $q(0)$ at history \emptyset , the same action as under τ_1 . In the subgame $G(q(0))$, the strategy profile s_3 coincides with τ_1 while s_0^j coincides with τ_2^j . Since τ_2^j is a best response to τ_1 , it follows that s_0^j is a best response to s_3 for player j . \square

Let c_i denote the number of elements in the image of the function u^i and let $c(G) = c_1 + \dots + c_n$. We now prove Theorem 3.1 by induction on $c(G)$. Clearly, the conclusion of Theorem 3.1 holds for any n -player game G with $c(G) = n$. Consider some $\ell > n$ and assume that every n -player game G with $c(G) < \ell$ has a best response cycle of length 4. By way of contradiction, suppose the n -player game G is such that $c(G) = \ell$, whereas G has no best response cycle of length 4.

Using recursion, we define a sequence h_0, h_1, \dots of histories such that, for every $t \in \mathbb{N}$, history h_t is of length t , $h_{t'} \leq h_t$ if $t' \leq t$, and the subgame $G(h_t)$ has no best response cycle of length 4. Let $h_0 = \emptyset$. By our supposition, $G(h_0) = G$ has no best response cycle of length 4.

Let h_t be a history of length $t \in \mathbb{N}$ such that $G(h_t)$ has no best response cycle of length 4 and define $i_t = \iota(h_t)$. It holds that $c(G(h_t)) = \ell$. Lemma 4.2 now implies that there is a unique action $a_t \in A$ such that $W^{i_t} \cap P(h_t) \subseteq P(h_t, a_t)$. We define $A'(h_t) = A(h_t) \setminus \{a_t\}$. For every $a' \in A'(h_t)$, consider the subgame $G(h_t, a')$. Every play of the subgame $G(h_t, a')$ yields player i_t a payoff strictly below w^{i_t} . Consequently, $c(G(h_t, a')) < \ell$ and the induction hypothesis implies that the subgame $G(h_t, a')$ has a best response cycle of length 4. Since $G(h_t)$ has no best response cycle of length 4, Lemma 4.1 implies that $G(h_t, a_t)$ has no best response cycle of length 4. We define $h_{t+1} = (h_t, a_t)$, which completes the recursion.

For $i \in I$, we define $T^i = \{t \in \mathbb{N} : h_t \in H^i \text{ and } A'(h_t) \neq \emptyset\}$. Consider some $t \in T^i$. Using the Axiom of Choice we can choose, for every $a' \in A'(h_t)$, a best response cycle $(\sigma_{t,a',0}, \sigma_{t,a',1}, \sigma_{t,a',2}, \sigma_{t,a',3}, \sigma_{t,a',0})$ for the subgame $G(h_t, a')$. For $k \in \{0, 1, 2, 3\}$, let

$x_{t,k}^i(a') = u_{(h_t, a')}^i(\pi(\sigma_{t, a', k}/\sigma_{t, a', k+1}^i))$ be the payoff of player i in the subgame $G(h_t, a')$ resulting from the strategy profile $\sigma_{t, a', k}/\sigma_{t, a', k+1}^i$ and let

$$y_{t,k}^i = \max_{a' \in A'(h_t)} x_{t,k}^i(a').$$

For $i \in I$ such that $T^i \neq \emptyset$, for $k \in \{0, 1, 2, 3\}$, we define

$$y_k^i = \max_{t \in T^i} y_{t,k}^i.$$

Since the player set I is finite, there is $t^* \in T$ such that, for every $i \in I$ with $T^i \neq \emptyset$, for every $k \in \{0, 1, 2, 3\}$,

$$\max_{\{t \in T^i : t \leq t^*\}} y_{t,k}^i = y_k^i.$$

Consider the game G^* that is identical to the game G , except that for every $t > t^*$, the only action available at history h_t is a_t . Formally, let H^* be the set consisting of histories $h \in H$ such that if h follows a history of the form (h_t, a) where $t > t^*$, then $a = a_t$. Let G^* be a game with a set H^* of histories, a function ι^* equal to the restriction of ι to H^* , a set of plays P^* , payoff functions u^{*i} for $i \in I$ equal to the restriction of u^i to P^* , and induced payoff functions U^{*i} for $i \in I$.

The subgame $G^*(h_{t^*}, a_{t^*})$ has only one play, $(a_{t^*+1}, a_{t^*+2}, \dots)$, so trivially has a best response cycle of length 4. Applying Lemma 4.1 to the games $G^*(h_{t^*})$, $G^*(h_{t^*-1})$, \dots , $G^*(h_1)$, and $G^*(\emptyset) = G^*$ repeatedly, we find that the game G^* has a best response cycle

$$(\tau_0^*, \tau_1^*, \tau_2^*, \tau_3^*, \tau_0^*)$$

such that, for every $k \in \{0, 1, 2, 3\}$, for every $t \leq t^*$, for every $a' \in A'(h_t)$, the strategy τ_k^* agrees with $\sigma_{t, a', k}$. Notice that, for every $t > t^*$, $\tau_k^*(h_t) = a_t$ by definition of G^* .

Next, for every $k \in \{0, 1, 2, 3\}$, we extend τ_k^* to a strategy profile τ_k defined on the set H of histories in the game G . For every $h \in H^*$, we define $\tau_k(h) = \tau_k^*(h)$, and, for every $t > t^*$, for every $a' \in A'(h_t)$, we let τ_k agree with the strategy profile $\sigma_{t, a', k}$. Notice that $\pi(\tau_k)$ and $\pi(\tau_k/\tau_{k+1}^i)$, $i \in I$, are plays in P^* .

The following result completes the proof of the theorem as we obtain a contradiction to our supposition that G has no best response cycle of length 4.

Lemma 4.3 *The sequence $(\tau_0, \tau_1, \tau_2, \tau_3, \tau_0)$ is a best response cycle of length 4 in G .*

Proof: Let $i \in I$ and $k \in \{0, 1, 2, 3\}$ be given. We show that τ_{k+1}^i is a best response of player i to τ_k . Let α^i be any strategy of player i . Consider the play $p = \pi(\tau_k/\alpha^i)$. If p belongs to P^* , then $u^i(p) \leq U^{*i}(\tau_k^*/\tau_{k+1}^{*i}) = U^i(\tau_k/\tau_{k+1}^i)$, where the inequality follows since

τ_{k+1}^{*i} is a best response of player i to τ_k^* in G^* and the equality holds because τ_k^*/τ_{k+1}^{*i} and τ_k/τ_{k+1}^i are identical on H^* .

Now consider the case where p does not belong to P^* . Then it has a prefix of the form (h_{t^+}, a^+) , where $t^+ > t^*$ and $a^+ \in A'(h_{t^+})$. Since $a^+ \neq a_{t^+}$, it follows that $i_{t^+} = i$. Since τ_k agrees with the strategy profile $\sigma_{t^+, a^+, k}$ in $G(h_{t^+}, a^+)$, it holds that $u^i(p) \leq x_{t^+, k}^i(a^+) \leq y_{t^+, k}^i$. By the choice of t^* there is some $t^- \in T^i$ such that $t^- \leq t^*$, $i_{t^-} = i$, and $y_{t^+, k}^i \leq y_{t^-, k}^i$. Take any action $a^- \in A'(h_{t^-})$ such that $x_{t^-, k}^i(a^-) = y_{t^-, k}^i$ and a strategy β^i of player i in the game G^* such that $\beta^i(h_t) = a_t$ for every $t < t^-$, $\beta^i(h_{t^-}) = a^-$, and β^i agrees with $\sigma_{t^-, a^-, k+1}^i$ in the subgame $G(h_{t^-}, a^-)$. Since $U^{*i}(\tau_k^*/\beta^i) = y_{t^-, k}^i$, we find that

$$U^i(\tau_k/\alpha^i) \leq U^{*i}(\tau_k^*/\beta^i) \leq U^{*i}(\tau_k^*/\tau_{k+1}^{*i}) = U^i(\tau_k/\tau_{k+1}^i).$$

□

5 Conclusion

Even win–lose games, two–player zero-sum perfect information games where the payoff of a player is either zero or one, may fail to have Nash equilibria or, equivalently, best–response cycles of length one. We consider general n -player perfect information games, where the payoff functions of the players have a finite image. We do not make any assumptions on the set of available actions and we do not make any further assumptions on the payoff functions. We show that in any such game there is a best–response cycle of length four. This result implies the existence of point–rationalizable strategy profiles. The strategies in a best response cycle cannot be eliminated by any procedure of iterated elimination of strictly dominated strategies. For the case of payoff functions with a bounded image, we find that there is always an ϵ -best response cycle of length four.

References

- [1] Bernheim, B.D. (1984): Rationalizable Strategic Behavior. *Econometrica* 52, 1007–1028.
- [2] Chen Y.-C., Van Long, N., and Luo X. (2007): Iterated strict dominance in general games. *Games and Economic Behavior* 61, 299–315.
- [3] Dufwenberg, M., and Stegeman, M. (2002): Existence and uniqueness of maximal reductions under iterated strict dominance. *Econometrica* 70, 2007–2023.

- [4] Gale, D., and Stewart, F. M. (1953): Infinite games with perfect information, Contributions to the theory of games, vol. 2, *Annals of Mathematics Studies*, no. 28, Princeton University Press, Princeton, New Jersey, pp. 245–266.
- [5] Martin, D.A. (1975): Borel determinacy. *Annals of Mathematics* 102, 363–371.
- [6] Martin, D.A. (1990): An extension of Borel determinacy. *Annals of Pure and Applied Logic* 49, 279–293.
- [7] Mertens, J. F. (1986): Repeated games, CORE Discussion Paper 8624, Center for Operations Research and Econometrics, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.
- [8] Tarski, A. (1955): A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics* 5, 285–309.