

# Best Response Cycles in Perfect Information Games

Citation for published version (APA):

Herings, P. J.-J., & Predtetchinski, A. (2017). Best Response Cycles in Perfect Information Games. *Mathematics of Operations Research*, 42(2), 427-433. <https://doi.org/10.1287/moor.2016.0808>

## Document status and date:

Published: 01/05/2017

## DOI:

[10.1287/moor.2016.0808](https://doi.org/10.1287/moor.2016.0808)

## Document Version:

Publisher's PDF, also known as Version of record

## Document license:

Taverne

## 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](http://www.umlib.nl/taverne-license)

## Take down policy

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

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.



## Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:  
<http://pubsonline.informs.org>

### Best-Response Cycles in Perfect Information Games

P. Jean-Jacques Herings, Arkadi Predtetchinski

To cite this article:

P. Jean-Jacques Herings, Arkadi Predtetchinski (2017) Best-Response Cycles in Perfect Information Games. *Mathematics of Operations Research* 42(2):427-433. <https://doi.org/10.1287/moor.2016.0808>

Full terms and conditions of use: <https://pubsonline.informs.org/page/terms-and-conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact [permissions@informs.org](mailto:permissions@informs.org).

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2016, INFORMS

Please scroll down for article—it is on subsequent pages

INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

# Best-Response Cycles in Perfect Information Games

P. Jean-Jacques Herings,<sup>a</sup> Arkadi Predtetchinski<sup>a</sup>

<sup>a</sup>Department of Economics, Maastricht University, 6200 MD Maastricht, Netherlands

Contact: [p.herings@maastrichtuniversity.nl](mailto:p.herings@maastrichtuniversity.nl) (PJ-JH); [a.predtetchinski@maastrichtuniversity.nl](mailto:a.predtetchinski@maastrichtuniversity.nl) (AP)

Received: July 3, 2015

Revised: April 13, 2016

Accepted: June 14, 2016

Published Online in Articles in Advance:  
November 4, 2016

MSC2010 Subject Classification: 91A06,  
91A10, 91A18, 91A26, 91A50

OR/MS Subject Classification: Primary:  
games/group decisions; secondary:  
noncooperative

<https://doi.org/10.1287/moor.2016.0808>

Copyright: © 2016 INFORMS

**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 of four 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  $\epsilon$ -best response cycle of length four for every  $\epsilon > 0$ .

**Keywords:** perfect information games • determinacy • best-response cycles • rationalizability

## 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 setup are win-lose games as studied by Gale and Stewart [4]. 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. A win-lose game is determined if and only if it has a Nash equilibrium in pure strategies. Gale and Stewart [4] show that there are win-lose games that are not determined.

We first consider  $n$ -player perfect information 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 [1]), and they survive any procedure of iterated elimination of strictly dominated strategies (Chen et al. [2], Dufwenberg and Stegeman [3]).

Since a Nash equilibrium corresponds to a best-response cycle of length one, the results of Gale and Stewart [4] 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 (reported in Mertens [7]) is that, for every positive  $\epsilon$ , an  $\epsilon$ -Nash equilibrium in pure strategies exists, and therefore an  $\epsilon$ -best-response cycle of length one exists. We show that boundedness of the payoff functions is sufficient for the existence of an  $\epsilon$ -best-response cycle of length four.

The rest of this 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 nonempty 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 nonempty 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 setup. 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 as  $h \leq h'$ , 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  $s/\sigma^i$  denote the strategy profile obtained from  $s$  by replacing player  $i$ 's strategy  $s^i$  by  $\sigma^i$ .

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  $i \in I$ ,  $\sigma^i$  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 four—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 nonempty subset  $E$  of  $S$ ,

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

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 [8]). Of course,  $\emptyset$  is a fixed point of  $B$ . The set  $F = \bigcup \{E \in 2^S: E \subseteq B(E)\}$  is the greatest fixed point of  $B$ . Theorem 3.1 implies that there is a nonempty 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 nonempty.*

### 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 [4]), Borel (Martin [5]), or quasi-Borel (Martin [6]). Under these conditions, a win-lose game has a best-response cycle of length one. Theorem 3.1 implies that any win-lose game has a best-response cycle of length four.

### Connection to Point Rationalizability

Bernheim ([1], p. 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  $\lambda$  is the set of point-rationalizable strategy profiles.

The difference between the definition of  $B$  and that of  $\lambda$  could be understood as a difference in the order of the quantifiers: a strategy profile  $\sigma$  is an element of  $B(E)$  if there exists an  $s \in E$  such that for each  $i \in I$ , the strategy  $\sigma^i$  is a best response to  $s$ . On the other hand,  $\sigma$  is an element of  $\lambda(E)$  if for each  $i \in I$  there exists an  $s \in E$  such that  $\sigma^i$  is a best response to  $s$ . In essence, 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  $\lambda$ .

It holds that  $B(E) \subseteq \lambda(E)$  for each  $E \subseteq 2^S$ . Consequently, the greatest fixed point  $F$  of  $B$  is contained in the greatest fixed point of  $\lambda$ . 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 four, then the strategy profiles  $\sigma_0, \sigma_1, \sigma_2$ , and  $\sigma_3$  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 nonempty.*

### 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 [3], 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  $E = \bigcap_{m \in \mathbb{N}} E_m$ .

Let  $F$  be the greatest fixed point of the map  $B$ . If  $E \rightarrow D$  and  $F \subseteq E$ , then also  $F \subseteq D$ . Consequently, any reduction of  $S$  contains  $F$  as a subset and is nonempty. 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 et al. [2], rather than just countable sequences as in Dufwenberg and Stegeman [3].

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  $\epsilon$ -best response.

Let  $\epsilon$  be some positive real number. A strategy  $\sigma^i \in S^i$  is said to be player  $i$ 's  $\epsilon$ -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  $\epsilon$ -best response to  $s \in S$  if for every player  $i \in I$ ,  $\sigma^i$  is an  $\epsilon$ -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  $\epsilon$ -best-response cycle of length four—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  $\epsilon$ -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 four 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  $\epsilon$ -best-response cycle of length four.

We define the maps  $B_\epsilon^i$  and  $B_\epsilon$  analogously to  $B^i$  and  $B$ . The map  $B_\epsilon$  is called the  $\epsilon$ -best-response map. We let  $F_\epsilon$  denote the greatest fixed point of  $B_\epsilon$ . 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  $F_\epsilon$  is nonempty.*

Let  $\epsilon$  be a positive real number. A pure strategy profile  $s \in S$  such that  $s \in B_\epsilon(s)$  is called an  $\epsilon$ -Nash equilibrium. An  $\epsilon$ -Nash equilibrium corresponds to an  $\epsilon$ -best-response cycle of length one. A result due to Mertens and Neyman (reported by Mertens [7]) asserts that a perfect information game admits an  $\epsilon$ -Nash equilibrium for every  $\epsilon > 0$  provided that the payoff functions are bounded and Borel measurable. The proof of existence of  $\epsilon$ -equilibria in perfect information games by Mertens and Neyman relies on Borel determinacy as shown in Martin [5]. Therefore, if payoff functions are bounded and Borel measurable, then there exists an  $\epsilon$ -best-response cycle of length one. Theorem 3.4 implies that the boundedness of the payoff functions is sufficient to guarantee the existence of an  $\epsilon$ -best-response cycle of length four.

### 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  $\sigma$  is a strategy profile in  $G$  and  $\tau$  a strategy profile in  $G(h)$ , we say that  $\sigma$  agrees with  $\tau$  if  $\sigma(h, e) = \tau(e)$  for every history  $e$  in  $G(h)$ .

Lemmata 4.1 and 4.2 give some sufficient conditions for the game to have a best-response cycle of length four. Lemma 4.1 asserts that, if for every action  $a$  the subgame  $G(a)$  has a best-response cycle of length four, so does the entire game.

**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 four. Then the game  $G$  has a best-response cycle  $(\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_0)$  of length four such that for every  $a \in A(\emptyset)$  the strategy  $\sigma_k$  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  $\sigma_k$  by  $\sigma_k(\emptyset) = a_k$  and, for every  $a \in A(\emptyset)$ , by letting  $\sigma_k$  agree with  $\sigma_{a,k}$  in  $G(a)$ .

By the choice of the action  $a_k$  and since  $\sigma_{a_{k-1},k}^i$  is a best response to  $\sigma_{a_{k-1},k-1}$  in the subgame  $G(a_{k-1})$ , the strategy profile  $\sigma_k$  is a best response to  $\sigma_{k-1}$ . Thus the sequence  $(\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_0)$  is a best-response cycle of length four 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  $s(\emptyset) = a_0$  and  $s(a_0, \dots, a_t) = a_{t+1}$  for each  $t \in \mathbb{N}$ . 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

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

**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 four.*

**Proof.** Let  $\sigma_0$  be a strategy profile such that all players follow  $p$ . Clearly,  $\sigma_0^i$  is player  $i$ 's best response to  $\sigma_0$ . 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  $\tau_0$  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

$$\begin{aligned} s_0^i &= \sigma_0^i & \text{and} & & s_0^j(h) &= \begin{cases} \sigma_0^j(h) & \text{if } p(0) < h, \\ \tau_2^j(h) & \text{if } q(0) < h; \end{cases} \\ s_1^i &= \sigma_1^i & \text{and} & & s_1^j(h) &= \begin{cases} \sigma_1^j(h) & \text{if } p(0) < h, \\ \tau_0^j(h) & \text{if } q(0) < h; \end{cases} \\ s_2^i &= \tau_0^i & \text{and} & & s_2^j(h) &= \begin{cases} \sigma_2^j(h) & \text{if } p(0) < h, \\ \tau_0^j(h) & \text{if } q(0) < h; \end{cases} \\ s_3^i &= \tau_1^i & \text{and} & & s_3^j(h) &= \begin{cases} \sigma_0^j(h) & \text{if } p(0) < h, \\ \tau_1^j(h) & \text{if } q(0) < h; \end{cases} \end{aligned}$$



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  $\sigma_0$ . In the subgame  $G(p(0))$ , the strategy profile  $s_0$  coincides with  $\sigma_0$  while  $s_1^j$  coincides with  $\sigma_1^j$ . Since  $\sigma_1^j$  is a best response to  $\sigma_0$ , 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  $\sigma_1$ . In the subgame  $G(p(0))$ , the strategy profile  $s_1$  coincides with  $\sigma_1$  while  $s_2^j$  coincides with  $\sigma_2^j$ . Since  $\sigma_2^j$  is a best response to  $\sigma_1$ , 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  $\tau_0$ . In the subgame  $G(q(0))$ , the strategy profile  $s_2$  coincides with  $\tau_0$  while  $s_3^j$  coincides with  $\tau_1^j$ . Since  $\tau_1^j$  is a best response to  $\tau_0$ , 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  $\tau_1$ . In the subgame  $G(q(0))$ , the strategy profile  $s_3$  coincides with  $\tau_1$  while  $s_0^j$  coincides with  $\tau_2^j$ . Since  $\tau_2^j$  is a best response to  $\tau_1$ , 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  $l > n$  and assume that every  $n$ -player game  $G$  with  $c(G) < l$  has a best-response cycle of length four. By way of contradiction, suppose the  $n$ -player game  $G$  is such that  $c(G) = l$ , whereas  $G$  has no best-response cycle of length four.

Using recursion, we define a sequence  $h_0, h_1, \dots$  of histories as follows: Let  $h_0 = \emptyset$ . By our supposition,  $G(h_0) = G$  has no best-response cycle of length four. Let  $h_t$  be a history of length  $t \in \mathbb{N}$  such that  $G(h_t)$  has no best-response cycle of length four, and define  $i_t = i(h_t)$ . It holds that  $c(G(h_t)) = l$ . Lemma 4.2 now implies that there is a unique action  $a_t \in A(h_t)$  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,  $c(G(h_t, a')) < l$ , and the induction hypothesis implies that the subgame  $G(h_t, a')$  has a best-response cycle of length four. Since  $G(h_t)$  has no best-response cycle of length four, Lemma 4.1 implies that  $G(h_t, a_t)$  has no best-response cycle of length four. We define  $h_{t+1} = (h_t, a_t)$ , which completes the recursion.

The sequences  $h_t$  and  $a_t$  thus defined have the following properties: for every  $t \in \mathbb{N}$ ,

1.  $h_0 = \emptyset$  and  $h_{t+1} = (h_t, a_t)$ ,
2. the subgame  $G(h_t)$  has no best-response cycle of length four, and
3. for each  $a' \in A'(h_t)$ , the subgame  $G(h_t, a')$  has a best-response cycle of length four.

To gain some intuition into the remainder of the proof, we resort to an analogy with quitting games. We say that player  $i_t$  continues at period  $t$  if he plays action  $a_t$  at history  $h_t$ , and we say that  $i_t$  quits at period  $t$  if he plays an action  $a' \in A'(h_t)$  at  $h_t$ . By property 3 above, as soon as a player quits at some period, we enter a subgame that has a best-response cycle of length four.

For each  $t \in \mathbb{N}$ , let  $Q_t$  be the set of histories of the form  $(h_t, a')$  for some  $a' \in A'(h_t)$ . We call any subgame  $G(q)$  of  $G$  where  $q \in Q_t$  for some  $t \in \mathbb{N}$  a *quitting subgame*. Using the axiom of choice, we can choose a best-response cycle  $(\sigma_{q,0}, \sigma_{q,1}, \sigma_{q,2}, \sigma_{q,3}, \sigma_{q,0})$  of length four in  $G(q)$  for every quitting subgame  $G(q)$ .

We next truncate the game  $G$  at some suitably chosen period  $t^*$  by deleting all quitting actions after period  $t^*$  and show that the game  $G^*$  thus obtained has a best-response cycle of length four. Furthermore, such a best-response cycle of length four, together with the best-response cycles of length four previously fixed in the quitting subgames, yields a best-response cycle of length four of the original game  $G$ . This last argument relies on the choice of  $t^*$ . Roughly speaking,  $t^*$  is to be chosen so that a player can do no worse by quitting at or before  $t^*$  than by quitting after  $t^*$ .

To choose the period  $t^*$ , suppose that in each quitting subgame  $G(q)$ , player  $i$ 's opponents follow the strategy profile  $\sigma_{q,k}$ . We can compute the highest payoff player  $i$  can get by quitting at any period  $t$ . The period  $t^*$  is chosen such that the highest payoff player  $i$  can get by quitting at or before  $t^*$  is at least as large as the highest payoff  $i$  can get by quitting after  $t^*$ .

We proceed to formal definitions of  $t^*$  and  $G^*$ . Consider a period  $t \in \mathbb{N}$ , a history  $q \in Q_t$ , a player  $i \in I$ , and an index  $k \in \{0, 1, 2, 3\}$ . Let

$$x_{q,k}^i = U_q^i(\sigma_{q,k}/\sigma_{q,k+1}^i).$$

This is the highest payoff that player  $i$  can get when playing against the strategy profile  $\sigma_{q,k}$  in the subgame  $G(q)$ . If  $i = i_t$  and  $Q_t \neq \emptyset$ , let

$$y_{t,k}^i = \max_{q \in Q_t} x_{q,k}^i.$$

In this case,  $y_{t,k}^i$  is the highest payoff player  $i = i_t$  can get by quitting at period  $t$ , assuming that in each quitting subgame  $G(q)$ , player  $i$ 's opponents follow the strategy profile  $\sigma_{q,k}$ . If  $i \neq i_t$  or  $Q_t = \emptyset$ , we define  $y_{t,k}^i$  to be a number that is smaller than any number in the image of  $u^i$ . Define

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

Since the player set  $I$  is finite, there is  $t^* \in \mathbb{N}$  such that for every  $i \in I$  and every  $k \in \{0, 1, 2, 3\}$ ,

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

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 there is no  $t > t^*$  and no  $q \in Q_t$  with  $q \leq h$ . Let  $G^*$  be the game with set of histories  $H^*$  and function  $u^*$  equal to the restriction of  $u$  to  $H^*$ . The set of plays in  $G^*$  is denoted by  $P^*$ . For  $i \in I$ , the payoff function in  $G^*$  is given by  $u^{*i}$ , the restriction of  $u^i$  to  $P^*$ , and the induced payoff function is denoted  $U^{*i}$ .

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 four. 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 of length four

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

such that, for every  $k \in \{0, 1, 2, 3\}$ , for every  $t \leq t^*$ , for every  $q \in Q_t$ , the strategy  $\tau_k^*$  agrees with  $\sigma_{q,k}$ . 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  $\tau_k^*$  to a strategy profile  $\tau_k$  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  $t > t^*$  and every  $q \in Q_t$ , we let  $\tau_k$  agree with the strategy profile  $\sigma_{q,k}$  in the subgame  $G(q)$ .

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 four.

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

**Proof.** Let  $i \in I$  and  $k \in \{0, 1, 2, 3\}$  be given. We show that  $\tau_{k+1}^i$  is a best response of player  $i$  to  $\tau_k$ . Let  $\alpha^i$  be any strategy of player  $i$ . We argue that  $U^i(\tau_k/\alpha^i) \leq U^i(\tau_k/\tau_{k+1}^i)$ .

If  $\pi(\tau_k/\alpha^i)$  is a play of the game  $G^*$ , the result follows by definition of  $\tau_k$  and  $\tau_{k+1}^i$ .

Suppose that  $\pi(\tau_k/\alpha^i)$  is not a play of the game  $G^*$ . This implies that under the strategy profile  $\tau_k$  none of  $i$ 's opponents ever quits and that player  $i$ 's strategy  $\alpha^i$  requires player  $i$  to quit at a period  $t^+ > t^*$  using some action  $a^+ \in A'(h_{t^+})$ . Let  $q^+ = (h_{t^+}, a^+)$ . Since  $\tau_k$  agrees with the strategy profile  $\sigma_{q^+,k}$  in  $G(q^+)$ , it holds that

$$U^i(\tau_k/\alpha^i) \leq x_{q^+,k}^i \leq y_{t^+,k}^i.$$

By the choice of  $t^*$  there is some  $t^- \leq t^*$  such that

$$y_{t^+,k}^i \leq y_{t^-,k}^i.$$

It holds that  $i_{t^-} = i$  and  $Q_{t^-} \neq \emptyset$ . Take any history  $q^- = (h_{t^-}, a^-)$  in  $Q_{t^-}$  such that

$$y_{t^-,k}^i = x_{q^-,k}^i.$$

Let  $\beta^i$  be player  $i$ 's strategy that requires player  $i$  to continue at each period  $t < t^-$ , to quit at period  $t^-$  using the action  $a^-$ , and then to play according to the strategy  $\sigma_{q^-,k+1}^i$  in the quitting subgame  $G(q^-)$ . Then

$$x_{q^-,k}^i = U^i(\tau_k/\beta^i).$$

Combining the above inequalities, we see that  $U^i(\tau_k/\alpha^i) \leq U^i(\tau_k/\beta^i)$ . On the other hand, since the play induced by the strategy profile  $\tau_k/\beta^i$  is a play of the game  $G^*$ , we have  $U^i(\tau_k/\beta^i) \leq U^i(\tau_k/\tau_{k+1}^i)$ . This implies the desired inequality.  $\square$

By construction,  $\tau_{k+1}^i$  is the best strategy a player  $i \in I$  can use against the strategy profile  $\tau_k$ , if  $i$  is not allowed to quit after period  $t^*$ . To prove the lemma, we show that player  $i$  cannot do better by quitting later. To see this, suppose that under the strategy profile  $\tau_k$  none of the other players ever quits. Suppose that player  $i$  decides to quit at some period  $t^+ > t^*$ . The highest payoff player  $i$  can obtain in this way is  $y_{t^+,k}^i$ . The definition of  $t^*$  now guarantees that player  $i$  would not be worse off quitting at or before period  $t^*$  instead.



## 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  $\epsilon$ -best-response cycle of length four.

One might wonder whether there also exist best-response cycles of length two or three. In two-player games, the existence of a best-response cycle of length two is equivalent to the existence of Nash equilibrium. Indeed, suppose that the sequence  $(\sigma_0, \sigma_1, \sigma_0)$  of strategy profiles is a best-response cycle of length two, where  $\sigma_k = (\sigma_k^1, \sigma_k^2)$  for  $k = 0, 1$ . Then  $\sigma_1^1$  is player 1's best response to player 2's strategy  $\sigma_0^2$  while  $\sigma_0^2$  is player 2's best response to player 1's strategy  $\sigma_1^1$ . Consequently, the strategy profile  $(\sigma_1^1, \sigma_0^2)$  is a Nash equilibrium. We conclude that the existence of best-response cycles of length two cannot be guaranteed under the conditions of Theorem 3.1.

We do not know whether a best-response cycle of length three exists under these conditions. While parts of the proof of Theorem 3.1 could be generalized to deal with best-response cycles of any finite length, the construction of Lemma 4.2 crucially relies on best-response cycles of length four.

## References

- [1] Bernheim BD (1984) Rationalizable strategic behavior. *Econometrica* 52(4):1007–1028.
- [2] Chen Y-C, Van Long N, Luo X (2007) Iterated strict dominance in general games. *Games Econom. Behav.* 61(2):299–315.
- [3] Dufwenberg M, Stegeman M (2002) Existence and uniqueness of maximal reductions under iterated strict dominance. *Econometrica* 70(5):2007–2023.
- [4] Gale D, Stewart FM (1953) Infinite games with perfect information. Kuhn HW, Tucker AW, eds. *Contributions to the Theory of Games (AM-28)*, Annals of Mathematics Studies, Vol. 2 (Princeton University Press, Princeton, NJ), 245–266.
- [5] Martin DA (1975) Borel determinacy. *Ann. Math.* 102(2):363–371.
- [6] Martin DA (1990) An extension of Borel determinacy. *Ann. Pure Appl. Logic* 49(3):279–293.
- [7] Mertens JF (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 J. Math.* 5(2):285–309.