



# A Semi-Potential for Finite and Infinite Games in Extensive Form

Stéphane Le Roux<sup>1</sup> · Arno Pauly<sup>2</sup>

© The Author(s) 2019

## Abstract

We consider a dynamic approach to games in extensive forms. By restricting the convertibility relation over strategy profiles, we obtain a semi-potential (in the sense of Kukushkin), and we show that in finite sequential games the corresponding restriction of better-response dynamics will converge to a Nash equilibrium in quadratic time. Convergence happens on a per-player basis, and even in the presence of players with cyclic preferences, the players with acyclic preferences will stabilize. Thus, we obtain a candidate notion for rationality in the presence of irrational agents. Moreover, the restriction of convertibility can be justified by a conservative updating of beliefs about the other players strategies. For infinite games in extensive form we can retain convergence to a Nash equilibrium (in some sense), if the preferences are given by continuous payoff functions; or obtain a transfinite convergence if the outcome sets of the game are  $\Delta_2^0$ -sets.

**Keywords** Sequential games · Convergence · Belief learning · Infinite games

## 1 Introduction

The Nash equilibria are the fixed points of the better (or best) response dynamics. In graph theory they would be called the sinks of these dynamics, and in computer science they may be called their terminal strategy profiles. In general these dynamics do not terminate, *i.e.* the corresponding binary relations over strategy profiles are not well founded. In finite games non-termination amounts to the existence of a cycle.

---

This work benefited from the Royal Society International Exchange Grant IE111233 (while Le Roux was at the TU Darmstadt and Pauly at the University of Cambridge). The authors were partially supported by the ERC inVEST (279499) project.

---

✉ Stéphane Le Roux  
leroux@lsv.fr

✉ Arno Pauly  
a.m.pauly@swansea.ac.uk

<sup>1</sup> LSV, CNRS, ENS Paris-Saclay, Université Paris-Saclay, Paris, France

<sup>2</sup> Department of Computer Science, Swansea University, Swansea, UK

A particular exception is found in potential games [28]: A potential is an acyclic relation over the profiles that includes all the players' individual better-response dynamics. In a potential game, better-response dynamics thus will always improve the potential, and hence terminates (at a Nash equilibrium) if the game is finite.

The notion of semi-potential was introduced by Kukushkin [15,16] in order to salvage some of the nice properties of potential games for a larger class of games. Here, each player's freedom to change strategies is restricted—however, only in such a way that if she could change the current outcome to a particular outcome in the absence of restriction, she can still do so in a way that is consistent with the restriction. In a *generic* normal form game this is equivalent to a potential, as there different strategies will induce different outcomes. Nevertheless, several classes of non-generic games have no potential but have a semi-potential: A study [15, Theorem 3] proved that it is the case for finite real-valued games in extensive form.

In this article we study Kukushkin's restriction of the convertibility relation (we call it *lazy convertibility*) as well as the resulting better-response dynamics (*lazy improvement*) in some more detail. Some relevant properties of the lazy improvement are:

- The dynamics are uncoupled: Each player bases her decisions on her own preference, but she does not need to know the other players' preferences.
- The dynamics are history independent: Unlike, e.g. fictitious play or typical regret-minimization approaches (e.g. [13]), the next step in the dynamics depends only on the current strategies of the players. In particular, players do not need memory for *learning*.
- We consider pure strategies, not stochastic ones. Thus, our approach has a very different flavour from the usual evolutionary game theory one (e.g. [8,9,35]).
- No restrictions akin to *generic payoffs* are required, and we merely need acyclic preferences to guarantee termination at a Nash equilibrium in finite games (and anyway this requirement cannot be avoided for existence of Nash equilibrium [18,19]).
- In a finite game with acyclic preferences, the dynamics stabilizes at a Nash equilibrium after a quadratic number of steps.
- The stabilization result for the rational players, *i.e.* with acyclic preferences, remains unaffected, if unpredictable players, *i.e.* with cyclic preferences, are added.
- Under some conditions, even in an infinite game in extensive form we can ensure stabilization at a Nash equilibrium after a transfinite number of steps.

## Our Contributions

We give two alternative proofs of the termination at a Nash equilibrium in finite games (originally shown by Kukushkin), one of which yields a tight quadratic bound on the number of steps required. Previously, no bounds on the rate of convergence had been given explicitly.

The main advantage of our two proofs of termination is that they work on a per-player basis: Each player with acyclic preference will terminate, even in the presence of players with cyclic preferences. This is far from obvious, as in most dynamics a single player who keeps altering their choices can induce the other players to keep changing, too. Having this individual stabilization result allows us to provide a micro-foundation of a notion of individual rationality derived from lazy improvement which is applicable even in the presence of irrational agents.

Finally, we consider the extension to infinite games in extensive form in several ways. For continuous payoff functions, we consider three variants of lazy improvement: The  $\varepsilon$ -variant, which will converge to an  $\varepsilon$ -Nash equilibrium of the game. The *deepening*-variant guarantees

that all accumulation points of the sequence are Nash equilibria, but the dynamics are history dependent. For the history-independent *fair* variant, we need (mild) extra assumptions to prove that all accumulation points are Nash equilibria. We then also consider having  $\Delta_2^0$ -outcome sets, and show that transfinite continuation of the dynamics still stabilizes.

## Related Work

Finite perfect information games where each player acts only once were studied by Kukushkin [15]. These games are potential games; and depending on the update rule, convergence to Nash equilibria or to subgame-perfect equilibria can be achieved. In the case where players may act more than once, [7] studied various update rules including subgame updates and updates by coalitions.

Some voting games that are neither generalized potential games nor perfect information games still have a semi-potential: See, for example, [16,26].

Apart from the notion of semi-potential, Kukushkin [16] also studied the more general notions of restricted acyclicity and weak acyclicity. Acyclicity was also studied in, e.g. [1].

The dynamics of a very specific infinite setting were explored in [5]. Kukushkin [17] considered the question of lazy improvement in infinite games. Under more general conditions than ours, he showed that there are always finite improvement paths getting  $\varepsilon$ -close to a Nash equilibrium. While our assumptions are more demanding, our results that all accumulation points are Nash equilibria is also a much stronger conclusion.

An extended abstract based on this work is [24]. Section 6.2 is based on [21, Section VI].

## Structure of the Article

The rest of the paper is organized as follows: Sect. 2 recalls the definitions of game in normal form, game in extensive form, and the better-response dynamics. Section 3 introduces the core concept of *lazy improvement*. Section 4 proves that in a finite game, lazy improvement terminates at a Nash equilibrium. Section 4.2 gives an alternative proof also showing that termination occurs after a quadratic number of improvement steps. Section 5 gives a basic epistemic justification for lazy convertibility. In Sect. 6 we discuss extensions to infinite games. Finally, Sect. 7 provides a number of (counter)examples showing that, to some extent, our definitions have to be the way they are.

## 2 Background and Notation

This section recalls the definitions of game in normal form, game in extensive form, and the better-response dynamics.

**Definition 1** A game in normal form is a tuple  $\langle A, (S_a)_{a \in A}, O, v, (\prec_a)_{a \in A} \rangle$  satisfying the following:

- $A$  is a non-empty set (of players, or agents),
- $\prod_{a \in A} S_a$  is a non-empty Cartesian product (whose elements are the *strategy profiles* and where  $S_a$  represents the strategies available to Player  $a$ ),
- $O$  is a non-empty set (of possible outcomes),
- $v : \prod_{a \in A} S_a \rightarrow O$  (the outcome function that values the strategy profiles),
- Each  $\prec_a$  is a binary relation over  $O$  (modelling the preference of Player  $a$ ).

**Definition 2** [*Nash equilibrium*] Let  $\langle A, (S_a)_{a \in A}, O, v, (\prec_a)_{a \in A} \rangle$  be a game in normal form. A strategy profile (profile for short)  $s \in S := \prod_{a \in A} S_a$  is a Nash equilibrium if it makes every Player  $a$  stable, i.e.  $v(s) \not\prec_a v(s')$  for all  $s' \in S$  that differ from  $s$  at most at the  $a$ -component.

$$NE(s) := \forall a \in A, \forall s' \in S, \neg(v(s) \prec_a v(s')) \wedge \forall b \in A - \{a\}, s_b = s'_b$$

Implicit in the concept of Nash equilibrium is the notion of *convertibility*: An agent can convert one strategy profile to another, if they differ only in her actions. As lazy improvement will be introduced in Sect. 3 by restricting the convertibility relation, we provide a formal definition:

**Definition 3** [*Convertibility, induced preference over profiles, and improvement*]

- Let  $\langle A, (S_a)_{a \in A}, O, v, (\prec_a)_{a \in A} \rangle$  be a game in normal form. For  $s, s' \in \prod_{a \in A} S_a$ , let  $s \xrightarrow{c}_a s'$  denote the ability of Player  $a$  to convert  $s$  to  $s'$  by changing her own strategy, formally  $s \xrightarrow{c}_a s' := \forall b \in A - \{a\}, s_b = s'_b$ .
- Given a game  $\langle A, (S_a)_{a \in A}, O, v, (\prec_a)_{a \in A} \rangle$ , let  $s \prec_a s'$  denote  $v(s) \prec_a v(s')$ . So in this article  $\prec_a$  may also refer to the induced preference over the profiles.
- Let  $\rightarrow_a := \prec_a \cap \xrightarrow{c}_a$  be the individual improvement relations of the players, and let  $\rightarrow := \cup_{a \in A} \rightarrow_a$  be the better-response dynamics.

Observation 4 is a direct consequence of Definitions 2 and 3.

**Observation 4** *The Nash equilibria of a game are exactly the sinks, i.e., the terminal profiles of the better-response dynamics  $\rightarrow$ .*

A (generalized)<sup>1</sup> *potential* is an acyclic relation containing  $\rightarrow$ . Clearly a game has a potential iff  $\rightarrow$  is acyclic. If  $\prod_{a \in A} S_a$  is finite, this is equivalent to the termination of the better-response dynamics. A less restrictive notion is a *semi-potential* (introduced in [15]). A semi-potential is an acyclic relation  $\leftrightarrow$  contained in  $\rightarrow$ , such that whenever  $s \rightarrow s'$  then there is some  $s''$  with  $s \leftrightarrow s''$  and  $v(s') = v(s'')$ . In words, if a strategy profile can be reached by an improvement step, there is an equivalent strategy profile (w.r.t the induced outcome) reachable via a step in the semi-potential. It follows that the sinks of a semi-potential are exactly the sinks of the better-response dynamics. Thus, in a finite setting, the existence of a semi-potential in particular implies the existence of sinks, i.e. Nash equilibria.

Our setting will be games in extensive form, rather than games in normal form. The idea here is that the players collectively choose a path through a tree, with each player deciding the direction at the vertices that she is controlling. The preferences refer only to the path created, and choices of the chosen path are irrelevant. Thus, the evaluation map  $v$  is highly non-injective, which in turn gives room for the notion of a semi-potential to be interesting. Formally, we define games in extensive form as follows:

**Definition 5** A game in extensive form is a tuple  $(A, T, O, d, v, (\prec_a)_{a \in A})$  where

- $A$  is the non-empty set of players,
- $T$  is a rooted tree (finite or infinite),
- $O$  is the non-empty set of outcomes,
- $d$  associates a player with each vertex in the tree,
- $v$  associates an outcome with each maximal path from the root through the tree,

<sup>1</sup> This article does not study more specific potentials, see [28] for the details.

- and for each Player  $a \in A$ ,  $\prec_a$  is a relation on  $O$  (the preference relation of  $a$ ).

The corresponding game in normal form is obtained as follows: Let a strategy of Player  $a$  associate an outgoing edge with each vertex controlled by  $a$ . If a strategy per player is given, the collective choices identify some maximal path  $p$  through the tree, called the *induced play*. Applying  $v$  to that path yields the outcome of the game, i.e. the valuation of the game in normal form is the composition of the map that identifies the induced play and the valuation of the game in extensive form.

In our concrete examples, the outcomes will be tuples of natural numbers, and the  $n$ th player will prefer a tuple  $(x_1, \dots, x_{|A|})$  to  $(y_1, \dots, y_{|A|})$  iff  $x_n > y_n$ .

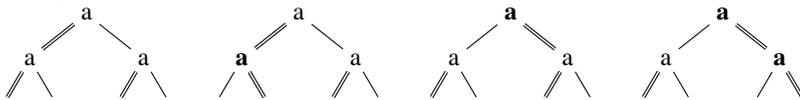
### 3 Defining Lazy Improvement

The idea underlying lazy improvement is that we do not let a player change their irrelevant choices, i.e. those choices not along the play induced after the improvement. Equivalently, we require a player to change choices at a minimal set (for the inclusion) of vertices when changing the induced play.

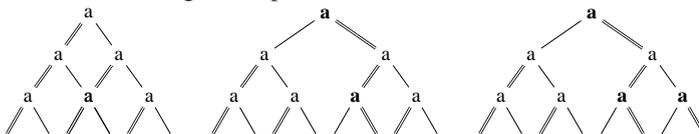
**Definition 6** [*Lazy convertibility and improvement*]

- For two strategy profiles  $s, s'$  in a game in extensive form let  $s \xrightarrow{c}_a s'$  (read: Player  $a$  can lazily convert  $s$  into  $s'$ ), if for every vertex  $t \in T$ , if  $s(t) \neq s'(t)$ , then  $d(t) = a$  and  $t$  lies along the play induced by  $s'$ . Let  $\xrightarrow{c} := \cup_{a \in A} \xrightarrow{c}_a$ .
- Let  $\rightarrow_a := \prec_a \cap \xrightarrow{c}_a$  be the lazy improvement of Player  $a$ , and let  $\rightarrow := \cup_{a \in A} \rightarrow_a$  be the lazy better-response dynamics, or lazy improvement.

Let us exemplify the notion of lazy convertibility, which has nothing to do with the preferences or the outcomes: Player  $a$  can lazily convert the leftmost strategy profile below into each of the three other profiles, but not into any other profile. Strategy choices are represented by double lines in the pictures, e.g. Player  $a$  chooses left instead of right at each node of the leftmost profile. Also for each other profile, Player  $a$  is written bold face at nodes where the profile differs from the leftmost one.



Contrary to the convertibility relations  $\xrightarrow{c}_a$  which are equivalence relations, the lazy convertibility relations  $\xrightarrow{c}_a$  are certainly reflexive but in general neither symmetric nor transitive. For instance, Player  $a$  cannot lazily convert the rightmost profile above back into the leftmost one. In the additional example below, Player  $a$  can convert the leftmost profile to the middle profile but not to the rightmost profile.



Since several forthcoming proofs invoke induction over the tree structure of the games, we note below that lazy convertibility could also be defined inductively.

**Observation 7** *The inductive definition below is equivalent to Definition 6.*

- If  $s$  is a leaf profile, let us define  $s \xrightarrow{c}_a s$  for all  $a \in A$ .
- Let two profiles  $s$  and  $s'$  have the profiles  $s_0, \dots, s_n$  and  $s'_0, \dots, s'_n$  as respective children with  $s_j = s'_j$  for  $j \neq k$ . Let Player  $a$  choose  $s_i$  at the root of  $s$  and  $s'_k$  at the root of  $s'$ . If  $s_k \xrightarrow{c}_b s'_k$  and if  $b = a$  or  $i = k$ , let us define  $s \xrightarrow{c}_b s'$ .

The lazy convertibility enjoys a useful property that the usual convertibility does not: If a player changes a play  $p$  into another play during a sequence of lazy convertibility, only the very same player might be later able to make the last step to induce  $p$  again, possibly induced by a different profile. This phenomenon is more formally stated by Lemma 1.

**Lemma 1** *If  $s \xrightarrow{c}_a s_0 \xrightarrow{c} \dots \xrightarrow{c} s_n \xrightarrow{c}_b s'$  where  $s$  and  $s'$  induce the same play, and if this play is different from the plays that are induced by  $s_0, \dots, s_n$ , then  $a = b$ .*

**Proof** Let us prove the claim by induction on the underlying game. Since the play induced by  $s_0$  is different from the play induced by  $s$ , these profiles are not just leaves, but proper trees instead. During the assumed  $\xrightarrow{c}$  reduction of  $s$ , its subprofile that is chosen by the root owner in  $s$  undergoes a  $\xrightarrow{c}$  reduction too, say  $t \xrightarrow{c}_a t_0 \xrightarrow{c} \dots \xrightarrow{c} t_n \xrightarrow{c}_b t'$ , where  $t$  and  $t'$  induce the same play (and the root owner chooses  $t'$  in  $s'$ ). If all these subprofiles are equal, Player  $a$  must be the root owner (of  $s$ ), since  $s$  and  $s_0$  induce different plays by assumption, and  $b$  is also the root owner since  $s_n$  and  $s'$  induce different plays, so  $a = b$ . Now let  $t_j$  be the first subprofile different from  $t$ , so  $t_j$  induces a play different from  $t$  and  $t'$ . For all  $k$  such that  $j \leq k < n$ , if  $t_k$  and  $t'$  induce different plays but  $t_{k+1}$  and  $t'$  induce the same play, then  $s_{k+1}$  and  $s'$  induce the same play by definition of  $\xrightarrow{c}$ , contradiction with the assumptions of the lemma, so all  $t_j, \dots, t_n$  induce plays different from that of  $t'$ . If  $t_1 \neq t$ , then  $a = b$  by the induction hypothesis, else  $a$  must be the root owner and does not choose  $t_1$  in  $s_1$ . The first time that  $a$  chooses some  $t_i$  again must be in  $s_j$ : Indeed if it were before,  $s$  and  $s_j$  would induce the same play, and if it were after,  $a$  could not change  $t_{j-1}$  into  $t_j$ . Therefore  $t_{j-1} \xrightarrow{c}_a t_j \xrightarrow{c} \dots \xrightarrow{c} t_n \xrightarrow{c}_b t'$  and  $a = b$  by the induction hypothesis.  $\square$

Observation 8 shows that despite the restrictive property from Lemma 1, the lazy convertibility is as effective as the usual convertibility, in the same sense as used in the definition of a semi-potential. (Thus, it will only remain to prove that lazy improvement is acyclic in order to establish lazy improvement as a semi-potential.)

**Observation 8** *If  $s \rightarrow s'$ , there is some strategy profile  $s''$  such that  $s \rightarrow s''$  and  $v(s') = v(s'')$ .*

**Proof** By definition, lazy convertibility does not restrict the choice of the new induced play, merely the ability to alter the strategy of the new induced play.  $\square$

**Corollary 1** *The Nash equilibria of a game are exactly the terminal profiles of the lazy improvement  $\rightarrow$ .*

## 4 Termination in Finite Games

This section presents two proofs. The first proof consists in showing acyclicity of the lazy improvement by contradiction, which carries over to infinite games. The second proof is closer to the original proof of [15], and it yields tight bounds on the number of lazy improvement steps occurring before termination.

### 4.1 First Proof, by Contradiction

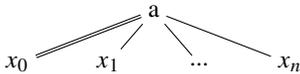
**Theorem 1** Consider a game in extensive form played on a finite tree, and some sequence  $(s_n)_{n \in \mathbb{N}}$  such that  $s_n \rightarrow_a s_{n+1}$  for all  $n \in \mathbb{N}$ . Assume that for a Player  $a$  there are infinitely many  $n$  with  $s_n \rightarrow_a s_{n+1}$ . Then  $a$  has a cyclic preference.

**Proof** Towards a contradiction let us assume that  $a$ 's preference is acyclic. Among the profiles  $s$  such that  $s = s_n \rightarrow_a s_{n+1}$  for infinitely many  $n$ , let  $s$  be minimal for  $a$ 's preference (i.e. least preferred), and let  $M$  be large enough such that every profile  $s_n$  with  $M < n$  occurs infinitely often in the sequence. Let  $s = s_n$  for some  $n > M$ , and let  $k > n$  be the least  $k$  such that  $s_n$  and  $s_k$  induce the same play. Lemma 1 implies that  $s_{k-1} \rightarrow_a s_k$ , so Player  $a$  prefers the outcome of  $s_n$  over that of  $s_{k-1}$ , contradiction.  $\square$

Together with Corollary 1 the following corollary shows the equivalence between all preferences being acyclic and universal existence of NE.

**Corollary 2** Consider outcomes  $O$ , players  $A$ , and their preferences  $(\prec_a)_{a \in A}$ : All  $\prec_a$  are acyclic iff for all finite games in extensive form built from  $O$ ,  $A$  and  $(\prec_a)_{a \in A}$  the lazy better-response dynamics terminates.

**Proof** The difficult implication of the equivalence is a corollary of Theorem 1. For the other implication, note that if  $x_0 \prec_a x_1 \prec_a \dots \prec_a x_n \prec_a x_0$ , then  $\rightarrow_a$  does not terminate on the profile below.



$\square$

**Corollary 3** [Kukushkin] In a finite game in extensive form where every player has acyclic preferences, lazy improvement is a semi-potential.

**Proof** Combine Corollaries 1 and 2.  $\square$

Kukushkin ([15, Theorem 3]) proved Corollary 3 in the case where the preferences are derived from payoffs. In this specific (yet usual) setting, it is not possible to consider players with cyclic preferences, so Theorem 1 or Corollary 2 cannot even be stated.

Based on Corollary 2 we obtain a reasonable candidate for rational behaviour in games in extensive form played with an unpredictable nature or erratic players: Perform lazy improvement until the players with acyclic preferences no longer change their strategies. It is always consistent with the observations to assume that the changes in another player's strategy are based on lazy convertibility. This argument is explored in more detail in Sect. 5. Nature can then be modelled as a player with the full relation as preferences, such that any convertible step for nature becomes an improvement step.

### 4.2 Second Proof, with Bounds

The proof of Theorem 1, by contradiction, gives a quick argument but no deep insight on how and how fast the relation terminates. A stronger statement can be proven by using the multiset of outcomes avoided by a Player  $a$  (i.e. the outcomes obtained in a subgame, where the decision not to play into that subgame was made by  $a$ , see Definition 9) to construct a measure that will decrease on any lazy improvement step by  $a$  (Lemma 3), and remain

unchanged by any lazy convertibility step by a different player (Lemma 2). Thus, we are in a situation very similar to potential games [28]—however, in a potential game a player *can* increase the potential (which is common to all the players) but does not *want* to, whereas here the players *cannot* impact the measure of another player as long as they are restricted to lazy convertibility.

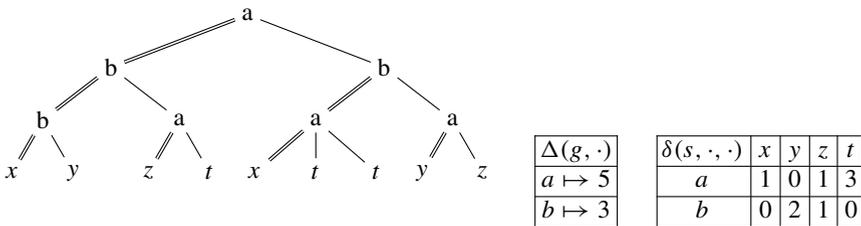
**Definition 9** [Avoided outcomes of a game and of a profile] The avoided outcomes of a game  $g$  are a function  $\Delta(g)$  of type  $A \rightarrow \mathbb{N}$ , and it is defined inductively below.

- $\Delta(g, a) := 0$  if  $g$  is a leaf game.
- If Player  $a$  owns the root of a game  $g$  whose children are  $g_0, \dots, g_n$  then
  - $\Delta(g, b) := \sum_{j=0}^n \Delta(g_j, b)$  for all  $b \neq a$ .
  - $\Delta(g, a) := (\sum_{j=0}^n \Delta(g_j, a)) + n$

The avoided outcomes of a profile  $s$  are a function  $\delta(s)$  of type  $A \rightarrow O \rightarrow \mathbb{N}$ , or equivalently in this case, of type  $A \times O \rightarrow \mathbb{N}$ , and it is defined inductively below.

- $\delta(s, a, o) := 0$  if  $s$  is a leaf profile.
- If Player  $a$  owns the root of a profile  $s$  and chooses the subprofile  $s_i$  among  $s_0, \dots, s_n$  then
  - $\delta(s, b, o) := \sum_{j=0}^n \delta(s_j, b, o)$  for all  $b \neq a$ .
  - $\delta(s, a, o) := (\sum_{j=0}^n \delta(s_j, a, o)) + |\{j \in \{0, \dots, n\} - \{i\} \mid v(s_j) = o\}|$

The smaller array below describes the function  $\Delta(g)$ , where  $g$  is the underlying game of the left-hand profile  $s$  below, and the right-hand array describes the function  $\delta(s)$ . For instance  $\delta(s, b, y) = 2$  because Player  $b$  avoids the outcome  $y$  twice: once at the leftmost internal node, after two leftward moves, when choosing outcome  $x$  rather than  $y$ , and also once after one rightward move, also when choosing  $x$  rather than  $y$ . Note that the only leaf that is not accounted for by the function of the avoided outcome of a profile/game is the leaf that is induced by the profile.



Observation 10 relates the two functions from Definition 9. It refers to  $s2g$ , a function that returns the underlying game of a given profile, see [18] or [19] for a proper definition.

**Observation 10** 1. Let  $s$  be a profile and  $a$  be a player, then  $\Delta(s2g(s), a) = \sum_{o \in O} \delta(s, a, o)$ .  
 2. Let  $g$  be a game, then  $1 + \sum_{a \in A} \Delta(g, a)$  equals the number of leaves of  $g$ .

**Proof** 1. By induction on  $s$ . If  $s$  is a leaf profile, the claim holds since  $\Delta(s2g(s), a) = 0 = \delta(s, a, o)$  by definition, so now let  $s$  be a profile where the root owner  $a$  chooses  $s_i$  among subprofiles  $s_0, \dots, s_n$ . For  $b \neq a$  Definition 9 and the induction hypothesis yield  $\Delta(s2g(s), b) = \sum_{j=0}^n \Delta(s2g(s_j), b) \stackrel{I.H.}{=} \sum_{j=0}^n \sum_{o \in O} \delta(s_j, b, o) = \sum_{o \in O} \sum_{j=0}^n \delta(s_j, b, o) = \sum_{o \in O} \delta(s, b, o)$ . Similarly we have  $\Delta(s2g(s), a) =$

$$\sum_{j=0}^n \Delta(s_2 g(s_j), a) + n \stackrel{I.H.}{=} \sum_{j=0}^n \sum_{o \in O} \delta(s_j, a, o) + |\{j \in \{0, \dots, n\} - \{i\} \mid v(s_j) \in O\}| = \sum_{o \in O} (\sum_{j=0}^n \delta(s_j, a, o) + |\{j \in \{0, \dots, n\} - \{i\} \mid v(s_j) = o\}|) = \sum_{o \in O} \delta(s, a, o).$$

2. By induction on  $g$ . This holds for every leaf game  $g$  since  $\Delta(g, a) = 0$  by definition. Let  $g$  be a game whose root is owned by Player  $a$  and whose subgames are  $g_0, \dots, g_n$ . The number of leaves in  $g$  is the sum of the numbers of leaves in the  $g_j$ , that is,  $\sum_{j=0}^n (1 + \sum_{b \in A} \Delta(g_j, b))$  by induction hypothesis. This, equals  $1 + \sum_{j=0}^n \sum_{b \in A - \{a\}} \Delta(g_j, b) + n + \sum_{j=0}^n \Delta(g_j, a)$ , which, in turn, equals  $1 + \sum_{b \in A - \{a\}} \Delta(g, b) + \Delta(g, a)$  by definition.  $\square$

Lemma 2 states conservation of the outcomes that are avoided by a player in a profile during a lazy convertibility step of another player. Intuitively, it is because a lazy convertibility step of a player cannot modify the subtrees that are avoided by the other players, even though she may own nodes therein. In the lemma and after  $\delta(s, b)$  denotes  $o \mapsto \delta(s, b, o)$

**Lemma 2**  $s \stackrel{c}{\leftarrow}_a s' \wedge b \neq a \Rightarrow \delta(s, b) = \delta(s', b)$

**Proof** By induction on the profile. It holds for leaves, so let  $s \stackrel{c}{\leftarrow}_a s'$  with subprofiles  $s_0, \dots, s_n$  and  $s'_0, \dots, s'_n$ , respectively. By definition of  $\stackrel{c}{\leftarrow}_a$  we have  $s_j \stackrel{c}{\leftarrow}_a s'_j$  for all  $j$ , and therefore  $\delta(s_j, b, o) = \delta(s'_j, b, o)$  by induction hypothesis. If the root owner is different from  $b$ , then  $\delta(s, b, o) = \sum_{j=0}^n \delta(s_j, b, o) = \sum_{j=0}^n \delta(s'_j, b, o) = \delta(s', b, o)$  by definition of  $\delta$ . If  $b$  is the root owner, she chooses the  $i$ th subprofile in both  $s$  and  $s'$  since  $b \neq a$ , and moreover  $s'_j = s_j$  for all  $j$  distinct from  $i$ . So  $\delta(s, b, o) = \sum_{j=0}^n \delta(s_j, b, o) + |\{j \in \{0, \dots, n\} - \{i\} \mid v(s_j) = o\}| = \sum_{j=0}^n \delta(s'_j, b, o) + |\{j \in \{0, \dots, n\} - \{i\} \mid v(s'_j) = o\}| = \delta(s', b, o)$ .  $\square$

However, the conservation does not fully hold for the player who converts the profile, unless the induced outcomes are the same for both profiles. The difference is little though, only depending on both induced outcomes. In Lemma 3, eq is just a boolean representation of equality:  $\text{eq}(x, x) := 1$  and  $\text{eq}(x, y) := 0$  for  $x \neq y$ .

**Lemma 3**  $s \stackrel{c}{\leftarrow}_a s' \Rightarrow \delta(s, a) + \text{eq}(v(s)) = \delta(s', a) + \text{eq}(v(s'))$

**Proof** By induction on the profile  $s$ . It holds for leaves, so let  $s \stackrel{c}{\leftarrow}_a s'$  with subprofiles  $s_0, \dots, s_n$  and  $s'_0, \dots, s'_n$ , respectively. If the root owner is distinct from  $a$ , she chooses the same  $i$ th subprofile in both  $s$  and  $s'$ , and therefore, for all outcomes  $o$  we have  $\delta(s, a, o) + \text{eq}(v(s), o) = \sum_{0 \leq j \leq n \wedge j \neq i} \delta(s_j, a, o) + \delta(s_i, a, o) + \text{eq}(v(s_i), o) = \sum_{0 \leq j \leq n \wedge j \neq i} \delta(s'_j, a, o) + \delta(s'_i, a, o) + \text{eq}(v(s'_i), o) = \delta(s', a, o) + \text{eq}(v(s'), o)$  by definition of  $\delta$ , since  $s_j = s'_j$  for  $j \neq i$ , and by induction hypothesis.

If  $a$  is the root owner, let  $a$  choose the  $i$ th and  $k$ th subprofiles in  $s$  and  $s'$ , respectively. Let  $N := \delta(s, a, o) + \text{eq}(v(s), o)$ , so  $N = \sum_{0 \leq j \leq n \wedge j \neq k} \delta(s'_j, a, o) + |\{j \in \{0, \dots, n\} - \{i\} \mid v(s_j) = o\}| + \delta(s_k, a, o) + \text{eq}(v(s_i), o)$  by unfolding Definition 9, since  $s'_j = s_j$  for all  $j \neq k$ , and since  $v(s) = v(s_i)$  by the choice at the root. Rewriting  $N$  twice with the easy-to-check equality  $|\{j \in \{0, \dots, n\} - \{x\} \mid v(s_j) = o\}| + \text{eq}(v(s_x), o) = |\{j \in \{0, \dots, n\} \mid v(s_j) = o\}|$ , first with  $x := i$  and then with  $x := k$  yields the equality  $N = \sum_{0 \leq j \leq n \wedge j \neq k} \delta(s'_j, a, o) + |\{j \in \{0, \dots, n\} - \{k\} \mid v(s_j) = o\}| + \delta(s_k, a, o) + \text{eq}(v(s_k), o)$ .

Since  $s_k \stackrel{c}{\leftarrow}_a s'_k$  by definition of lazy convertibility, and by the induction hypothesis, let us further rewrite  $\delta(s_k, a) + \text{eq}(v(s_k))$  with  $\delta(s'_k, a) + \text{eq}(v(s'_k))$  in  $N$ . Folding Definition 9 yields  $N = \delta(s', a, o) + \text{eq}(v(s'), o)$ .  $\square$

The invariant stated in Lemma 3 suggests that whenever a player will lazily convert a profile to obtain a better outcome, some measure will decrease a bit with respect to her preference. The invariant stated in Lemma 2 ensures that such a lazy conversion will leave the measure for the other players unchanged. The lazy improvement should therefore terminate, and even quite quickly, as proved below.

Recall that a finite preference relation  $\prec$  has height at most  $h$  if there is no chain  $s_1 \prec s_2 \prec \dots \prec s_{h+1}$ . Please bear in mind that the heights here refer to preferences, not trees. Also recall that  $\Delta(g, a)$  is the total number of choices available to Player  $a$ , minus the number of vertices where  $a$  is choosing.

**Theorem 2** [Strengthening Theorem 1 with bounds] *Consider a game  $g$  where Player  $a$  has an acyclic preference of height  $h$ . Then in any sequence (possibly infinite) of lazy improvement, the number of lazy improvement steps performed by Player  $a$  is bounded by  $(h - 1) \cdot \Delta(g, a)$ .*

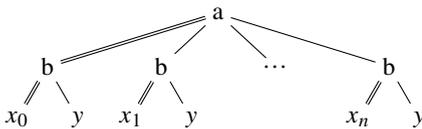
**Proof** For every outcome  $o$  let  $h(a, o)$  be the maximal cardinality of the  $\prec_a$ -chains whose  $\prec_a$ -maximum is  $o$ , and note that  $o \prec_a o'$  implies  $h(a, o) < h(a, o')$ . For every profile  $s$  let  $M(s, a) := \sum_{o \in O} (h(a, o) - 1) \cdot \delta(s, a, o)$  and note that  $0 \leq M(s, a) \leq (h - 1) \cdot \Delta(g, a)$  by Observation 10.1. Let  $s \xrightarrow{a} s'$  be a lazy improvement step, so  $s \xrightarrow{c} s'$  and  $v(s) \prec_a v(s')$  by definition, then  $M(s, a) - M(s', a) = \sum_{o \in O} (h(a, o) - 1) \cdot (\delta(s, a, o) - \delta(s', a, o)) = h(a, v(s')) - h(a, v(s)) > 0$  by Lemma 3. Let  $s \xrightarrow{c} s'$  be a lazy convertibility step where  $b \neq a$ , then  $M(s, a) = M(s', a)$  by Lemma 2. This shows that the  $\rightarrow_a$  steps are at most  $(h - 1) \cdot \Delta(g, a)$  in every sequence of  $\rightarrow$ .  $\square$

**Corollary 4** [Strengthen Corollary 2 with bounds] *The lazy improvement terminates for all games iff all preferences are acyclic, in which case the number of sequential lazy improvement steps is at most  $(h - 1) \cdot (l - 1)$  where  $h$  bounds the cardinality of the preference chains and  $l$  is the number of leaves.*

**Observation 11** 1. *The maximal length of a lazy improvement sequence is bounded in a quadratic manner in the size of the game in general and linearly when  $h$  from Corollary 4 is fixed.*

2. *The quadratic and linear bounds are tight.*

**Proof** (of 11.2) For the linear bound, let us consider the figure below and set  $x := x_0 = \dots = x_n$  and  $y \prec_a x$  and  $x \prec_b y$ . There is clearly a lazy improvement sequence starting from the figure and visiting each leaf exactly once.



It is similar for the quadratic bound, but we need to be a bit more careful. For  $n \in \mathbb{N}$ , consider the game in the above figure, where  $y \prec_a x_0 \prec_a x_1 \dots \prec_a x_n$  and  $x_i \prec_b y$  for all  $i$ . Let us prove by induction on  $n$  the existence of a sequence of  $\frac{(n+2)(n+3)}{2} - 2$  lazy improvement steps when starting from the strategy profile above. For the base case  $n = 0$ , there are  $1 = \frac{(0+2)(0+3)}{2} - 2$  lazy improvement steps. For the inductive case, let Player  $a$  make  $n$  lazy improvements in a row, by choosing  $x_1$ , then  $x_2$ , and so on until  $x_n$ . At that point, let Player  $b$  improve from  $x_n$  to  $y$  and then let Player  $a$  come back to  $x_0$ . So far,  $n + 2$  lazy improvement steps have been performed. Now let us ignore the substrategy profile involving  $x_n$  (and  $y$ ). By induction hypothesis,  $\frac{(n+1)(n+2)}{2} - 2$  additional lazy improvement steps can be performed in a row. Since  $(n + 2) + \frac{(n+1)(n+2)}{2} - 2 = \frac{(n+2)(n+3)}{2} - 2$ , we are done.  $\square$

### 4.3 Lazy Non-worsening

In this section the outcomes are real-valued payoff tuples. In this case, Theorem 2 can be slightly generalized in a way that will prove useful for studying lazy improvement in infinite games in Sect. 6.3. Let a lazy non-worsening step be a lazy convertibility step that does not decrease the payoff of the converting player. Said otherwise, a lazy non-worsening step is either a lazy improvement step or a lazy convertibility step preserving the payoff of the converting player. As shown below, weakening the first (but not the second) “lazy improvement” in Theorem 2 into a “lazy non-worsening” still yields a correct statement.

**Definition 12** [*Non-worsening*] Let  $f_a$  be the payoff function of Player  $a$ , and let  $s \rightsquigarrow_a s'$  if  $f_a(s) = f_a(s') \wedge s \overset{c}{\prec}_a s'$ . Let  $\rightsquigarrow := \cup_{a \in A} \rightsquigarrow_a$  be the lazy preservation, and let  $\rightarrow \cup \rightsquigarrow$  be the lazy non-worsening.

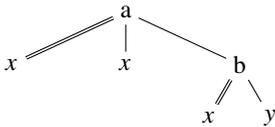
**Theorem 3** [Strengthen Theorem 1] *Consider a game  $g$  with real-valued payoffs, where Player  $a$  has at most  $h$  different payoffs. Then in every sequence (possibly infinite) of lazy non-worsening, the number of lazy improvement steps performed by Player  $a$  is bounded by  $(h - 1) \cdot \Delta(g, a)$ .*

**Proof** The proof is similar to that of Theorem 2. We modify  $\delta$  from Definition 9 to take payoffs into account instead of outcomes/payoff tuples. A modification of Lemma 3 is then easily obtained. Now, the old  $M(s, a)$  from the proof of Theorem 2 is redefined with the new  $\delta$ . We should additionally point out that a lazy preservation step preserves the new  $M(s, a)$ , by the new Lemma 3. Thus, lazy preservation steps preserve  $M$ , and have therefore no impact on the termination argument for the lazy improvement steps.  $\square$

**Corollary 5** *Let us restrict the lazy non-worsening such that preservation steps may only occur if the current profile is an NE. Every infinite sequence of such a restricted non-worsening is eventually made only of NE.*

Note that there is no bound on when at the latest the lazy improvement steps may occur in Corollary 5, as the following example shows:

**Example 1** We consider a game with two players,  $a$  and  $b$ , and two payoff tuples,  $x$  and  $y$  such that  $y \prec_a x$  and  $x \prec_b y$ . The game tree and initial strategy profile are as follows:



Player  $a$  can alternate between his leftmost and centre choice as often as he wishes using lazy equilibrium preservation. He can also at any time change to his rightmost choice. Then Player  $b$  has the opportunity to a lazy improvement step by changing to  $y$ , whereupon Player  $a$  can lazily improve by going back to the leftmost or centre choice to obtain the outcome  $x$ . After this happened, all remaining possible lazy non-worsening steps are Player  $a$  alternating between the leftmost and centre choice.

Lemma 4 will be used with Theorem 3 to deal with lazy improvement in infinite games in Sect. 6.

**Lemma 4** *Let  $(g_n)_{n \in \mathbb{N}}$  be a family of finite games in extensive form that differ only in payoffs and that converge towards some game  $g$  when  $n$  approaches infinity. Consider an infinite*

sequence of profiles  $(s_n)_{n \in \mathbb{N}}$  such that  $s_n (\rightarrow \cup \rightsquigarrow) s_{n+1}$  in  $g_n$ . Then there is some  $k \in \mathbb{N}$  such that for all  $n \geq k$  we have that  $s_n (\rightarrow \cup \rightsquigarrow) s_{n+1}$  in  $g$ .

**Proof** As lazy convertibility does not refer to the payoffs, we see that any lazy convertibility step in one of the  $g_n$  is also a lazy convertibility step in  $g$ . We only need to argue about the improvement and preservation aspects.

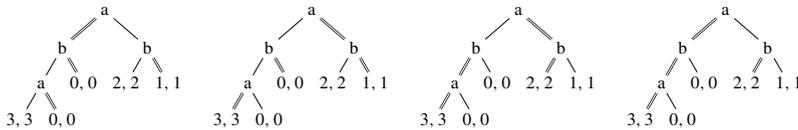
Let  $\delta$  be the minimum distance between different payoffs in  $g$ . We pick  $k \in \mathbb{N}$  such that for all  $n \geq k$  the difference between payoffs in  $g_n$  and  $g$  at the same leaf is less than  $\frac{\delta}{4}$  for all leaves. In particular, we find that in  $g_n$  for  $n \geq k$  payoffs differing by less than  $\frac{\delta}{2}$  correspond to identical payoffs in  $g$ , payoffs differing by at least  $\frac{\delta}{2}$  correspond to different payoffs in  $g$ .

Any lazy improvement step in  $g_n$  for  $n \geq k$  that improves the payoff of the acting player by at least  $\frac{\delta}{2}$  is a lazy improvement step in  $g$ . Any lazy improvement step in  $g_n$  for  $n \geq k$  that improves the payoff of the acting player by less than  $\frac{\delta}{2}$  corresponds to a preservation step in  $g$ , and so do preservation steps in  $g_n$ . □

### 5 Lazy Convertibility as Belief Updating

Let us discuss whether we should expect players to conform to lazy convertibility when playing a sequential game repeatedly. Observation 8 tells us that in the short term, a player has no incentive to deviate from lazy convertibility: If she desires some outcome she can reach by some deviation from her current strategy, she can obtain this outcome by converting a strategy in a lazy way. There is a caveat, though, in that restricting convertibility to lazy convertibility changes the overall reachability structure, as the following example shows.

**Example 2** The last profile of the three-step improvement relation below is a Nash equilibrium that cannot be reached from the first profile under lazy improvement.



From the perspective of any given player, it, however, makes a lot of sense to assume that all other players are updating their own strategies only in a lazy way—assuming that only relevant choices of the other players can be observed. The latter seems to be crucial in order to make the game truly sequential: If all players announced their entire strategy simultaneously, it would be a game in normal form after all.

To formalize this idea, let us fix a Player  $a$  and consider the game from her perspective. She may consider the game as a two-player game played by her against all other players aggregated into a single Player  $b$ . She starts with some initial strategy  $s_a^{(0)}$ , and some prior assumption  $s_b^{(0)}$  on the strategy of her opponent(s). She then updates her own strategy via lazy improvement to  $s_a^{(1)}$ . Then the game is actually played, and Player  $a$  observes the actual moves (but not the strategy) of her opponents. As she only observes the moves along the path actually taken, it is consistent with her observations to assume that the aggregated opponent player lazily converted  $s_b^{(0)}$  into some  $s_b^{(1)}$ . Then the Player  $a$  again performs a lazy improvement step to  $s_a^{(2)}$ , plays the game, etc. Provided that the Player  $a$  has acyclic preferences, Theorem 1 implies that her own strategy stabilizes to some strategy  $s_a$  eventually.

This learning procedure is the deterministic counterpart to the rational learning proposed by Kalai and Lehrer [10], and extended to define the *self-confirming* equilibria by Fudenberg

and Levine [10].<sup>2</sup> Wellman and Hu's conjectural equilibria [34] are based on the same intuition underlying the learning procedure—only actual actions are observed, not hypothetical ones, which are merely subject to conjecture.

Note that this procedure requires no assumptions on knowledge of rationality of other players or their payoff functions, not to speak of common knowledge. There is in general no reason to assume that the aggregated Player  $b$  acts according to some acyclic preference (given that the different players making up  $b$  may have partially antagonistic preferences). However, if each player has an acyclic preference and performs the same procedure as  $a$  above, then each player's actual strategy will stabilize. As any change in what a player assumes her aggregated opponents are playing has to be caused by either a change in her own, or someone else's strategy, this implies that also the believed strategy of the aggregated players  $s_b$  will stabilize. Furthermore, all the strategy profiles constructed in this way induce the same play, and combining them as follows yields a Nash equilibrium:

**Proposition 1** *Let a set of players play a finite sequential game by converting their own strategies lazily based on beliefs about the other players' strategies in order to maximize an acyclic preference relation. Then a Nash equilibrium can be obtained from the stable strategies they will settle to as follows: Along the common path chosen by their stable strategies, everyone follows their own strategy. In any subgame that is not reached, each player plays according to the beliefs held by the player controlling access to the subgame about their strategies.*

**Proof** At any vertex reached during the final play, the choice facing the current player is the same one she was anticipating due to her beliefs on her opponents' strategies. As her choice is consistent with the stable choice made during the dynamical updating, she has no incentive to change.  $\square$

In comparison, the investigation of the epistemic foundations of Nash equilibria by Aumann and Brandenburger [4] identified mutual knowledge of rationality, knowledge of the game and (in case of more than two players) a common prior as the prerequisite for playing a Nash equilibrium. A subgame-perfect equilibrium requires even stronger assumptions, namely well-aware players [3].

## 6 Lazy Improvement in Infinite Games

Infinite games in extensive form with win/lose preferences are generalizations of Gale–Stewart games [11], and are of great relevance for logic. That any two-player game in extensive form with antagonistic preferences and a Borel winning set actually has a Nash equilibrium is a highly non-trivial result by Martin [25]. It was used by Mertens and Neymann [27] to show that infinite games with finitely many players and bounded, Borel-measurable, non-necessarily antagonistic real-valued payoffs have  $\epsilon$ -Nash equilibria. It was generalized to infinitely many players and payoffs only bounded from above in [20]. Moreover, subgame-perfect equilibria do not always exist (cf. [21,32]).

The definition of lazy improvement applies to infinite games in extensive form as well, and we can adapt the results on finite games to see that it still constitutes a semi-potential:

**Proposition 2** *Consider an infinite game in extensive form where each player (there might be infinitely many) has acyclic preferences. Then lazy improvement is a semi-potential.*

<sup>2</sup> The notion has been corrected by Kamada later, but the difference is not present in the deterministic setting.

**Proof** As argued in Sect. 3, we only need to show that lazy improvement is acyclic. Assume the contrary, then there is some finite cycle  $s_1 \rightarrow s_2 \rightarrow \dots \rightarrow s_n \rightarrow s_1$ . Any subtree of the game tree not reached by any strategy profile  $s_i$  is irrelevant for the existence of the cycle, and could thus be pruned. Doing so yields a finitely branching game tree, with still a lazy improvement cycle.

Let  $p_1, \dots, p_n$  be the paths induced by the strategy profiles  $s_1, \dots, s_n$ , and choose  $k \in \mathbb{N}$  such that  $p_i|_{\leq k} = p_j|_{\leq k} \Leftrightarrow p_i = p_j$ . By choice of  $k$ , the path chosen inside any subgame rooted at depth  $k$  remains unchanged throughout the improvement cycle. Thus, replacing any such subgame with a leaf carrying the outcome induced by this path has no impact on the improvement cycle. We have obtained a finite game in extensive form with the same preferences and a cycle built from lazy improvement step, contradicting Theorem 1.  $\square$

Of course, in an infinite game acyclicity does not suffice to ensure termination or even convergence. In fact, [21, Example 26] (reproduced below as Example 6) shows that lazy improvement in infinite games will not always converge, and that even accumulation points do not have to be Nash equilibria. There are, however, several potential ways to extend the results on lazy improvement to infinite games in extensive form:

1. We can consider games where the preferences are expressed via continuous payoff functions. For some fixed  $\varepsilon > 0$ , we can then consider  $\varepsilon$ -lazy improvement (where only lazy convertibility is allowed, and improvement steps are only taken if the player can improve by more than  $\varepsilon$ ). Then Theorems 1 and 2 carry over, and as a counterpart to Corollary 1 we find that the terminal profiles of  $\varepsilon$ -lazy improvement are precisely the  $\varepsilon$ -Nash equilibria. See Sect. 6.1.
2. Again for continuous payoff functions, we can employ lazy improvement being done in a finitary way with increasing precision, and find that any accumulation point of particular subsequence is guaranteed to be a Nash equilibrium, see Sect. 6.2.
3. We define a notion of a *fair* lazy improvement sequence in Sect. 6.3. We then generalize the measure employed in the proof of Theorem 2 to infinite games with Lipschitz payoff functions for fair improvement sequences. Moreover, we prove that for continuous payoff functions, all accumulation points of a fair lazy improvement sequence with finitely many accumulation points are Nash equilibria.
4. Departing from the setting of continuous payoff functions, we can consider games where the players have win/lose objectives (i.e. their preference relations have height 2), and the winning sets are  $\Delta_2^0$ -sets. Then transfinite iteration of lazy improvement will reach a Nash equilibrium, see Sect. 6.4.

In the following we always assume that the game tree is the full infinite binary tree, and hence the set of resulting plays is  $\{0, 1\}^{\mathbb{N}}$ . This space carries a natural topology induced by the metric  $d(p, q) = 2^{-\min\{n | p(n) \neq q(n)\}}$  for  $p \neq q$ , and in particular is a compact zero-dimensional space. In the first three following subsections, we assume that the preferences of each player are given by payoff functions  $f_a : \{0, 1\}^{\mathbb{N}} \rightarrow \mathbb{R}$ , where  $p \prec_a q$  iff  $f_a(p) < f_a(q)$ . We can then speak about restrictions on the payoff functions such as being continuous or Lipschitz continuous.

### 6.1 $\varepsilon$ -Lazy Improvement

Consider preferences obtained from payoff functions. Then for every  $\varepsilon > 0$ , we can introduce  $\varepsilon$ -lazy improvement as the intersection of lazy convertibility and  $\varepsilon$ -improvement, where  $\varepsilon$ -improvement means considering only those improvement steps where the payoff for the

player increases by more than  $\varepsilon$ . Consequently, an  $\varepsilon$ -Nash equilibrium is a strategy profile where no player can improve by more than  $\varepsilon$ .

**Observation 13** *The sinks of  $\varepsilon$ -lazy improvement are precisely the  $\varepsilon$ -Nash equilibria.*

**Proposition 3** *Let  $(s_n)_{n \in \mathbb{N}}$  be a sequence of strategy profiles in an infinite binary game in extensive form with  $s_n \xrightarrow{c} s_{n+1}$ . Let Player  $a$  have a preference induced by a continuous payoff function  $f$ , and assume that for some  $\varepsilon > 0$ , whenever  $s_n \xrightarrow{c}_a s_{n+1}$ , with induced plays  $p_n, p_{n+1}$ , then  $f(p_{n+1}) > f(p_n) + \varepsilon$ . Then  $s_n \xrightarrow{c}_a s_{n+1}$  holds for only finitely many  $n$ .*

**Proof** We will essentially use a reduction to the case for finite trees, and invoke Theorem 1.

We consider the cover  $(A_k := ]\frac{\varepsilon(k-1)}{2}, \frac{\varepsilon(k+1)}{2}])_{k \in \mathbb{Z}}$  of  $\mathbb{R}$ . By continuity of  $f$  and compactness of  $\{0, 1\}^{\mathbb{N}}$ , there is a bar, i.e. a finite prefix-free family  $(w_i \in \{0, 1\}^*)_{i \leq N}$  such that  $\{0, 1\}^{\mathbb{N}} = \bigcup_{i \leq N} w_i \{0, 1\}^{\mathbb{N}}$  (where  $w_i \{0, 1\}^{\mathbb{N}}$  is the subset of the sequences of  $\{0, 1\}^{\mathbb{N}}$  having  $w_i$  as prefix), such that for every  $i$  there exists some  $k_i$  with  $f[w_i \{0, 1\}^{\mathbb{N}}] \subseteq A_{k_i}$ . Now consider the tree  $T$  with the  $w_i$  as the leaves. Clearly any strategy profile  $s_n$  restricts to some strategy profile  $s'_n$  on  $T$ , and moreover, if  $s_n \xrightarrow{c} s_{n+1}$ , then  $s'_n \xrightarrow{c} s'_{n+1}$ .

In the finite game played on  $T$ , let Player  $a$  have the preference  $w_i <_a w_j$  iff  $k_i < k_j$ . Clearly, this is an acyclic preference. For every other Player  $b$ , we just use the full preference  $w_i <_b w_j$  for every  $i, j$ . Whenever  $s_n \xrightarrow{c}_a s_{n+1}$ , then  $s'_n, s'_{n+1}$  must induce some  $w_i, w_j$  with  $k_i < k_j$ . Thus, we do not loose any convertibility steps performed by Player  $a$ , and have an instance of Theorem 1 which implies that  $a$  only converts finitely many times.  $\square$

**Corollary 6** *If there are finitely many players, each with continuous payoff function, performing  $\varepsilon$ -lazy improvement in an infinite binary game in extensive form, then the process terminates in finitely many steps.*

## 6.2 Deepening Lazy Improvement

Let us now assume that all (countably many) Player  $a$  have preferences derived from continuous payoff functions  $f_a$ . For each Player  $a$  we can use  $f_a$  to label every vertex  $v$  in the game with some rational interval  $I_v^a$  in a way<sup>3</sup> that the label of every vertex is a subset of its predecessor, and such that  $\bigcap_{n \in \mathbb{N}} I_{p_{\leq n}}^a = \{f_a(p)\}$ , i.e. the intersection of all labels along an infinite path is the singleton set containing the payoff for this path.

In the deepening lazy improvement dynamics, we start with some inspection depth  $d$ . The players consider the prefix of the game tree of depth  $d$ , where Player  $a$  prefers some vertex  $v$  (at depth  $d$ ) to some vertex  $u$  (also at depth  $d$ ) if all points in  $I_u^a$  are smaller than all points in  $I_v^a$ . Now any lazy improvement step in this finite game (on the tree cut at depth  $d$ ) induces an improvement step in the infinite tree game. By Theorem 1, improvement in every such finite game terminates.

Once all players are stable at the current inspection depth, the inspection depth is incremented by one. The incrementing shall be counted as an updating step, where the strategy profile is not modified. Thus, some infinite sequence of strategy profile always arises. We shall call the subsequence of the profiles right after the inspection depth is incremented, the *stable subsequence*.

<sup>3</sup> The idea behind this corresponds to the representation of real numbers in computable analysis [33].

Note that the choice of labelling system is not uniquely determined by the payoff function, and that the labelling in turn influences the lazy improvement dynamics. Moreover, note that while we are dealing with linear preferences only in the case of infinite games, we do make use of finite approximations that lack linear preferences—yet we are guaranteed that every preference occurring in our finite approximations is acyclic, which is sufficient for Theorem 1. Finally, the dynamics do depend on the history—however, only on the depth currently reached, not on any details.

**Observation 14** *The deepening lazy improvement dynamics are computable, i.e. given an infinite binary game and an initial strategy profile, we can compute a sequence of strategy profiles arising from deepening lazy improvement, as well as the indices of the stable subsequence.*

**Theorem 4** *The following properties are equivalent for a strategy profile  $s$ :*

1.  $s$  is a Nash equilibrium.
2.  $s$  is a fixed point<sup>4</sup> for deepening lazy improvement.
3.  $s$  is an accumulation point of the stable subsequence of some sequence obtained from deepening lazy improvement.

**Proof**

1.  $\Leftrightarrow$  2. By continuity of the preferences, a player prefers a strategy profile  $s$  to another profile  $s'$ , if and only if there is an inspection depth  $d$  such that he prefers the restriction of  $s$  to the restriction of  $s'$  in the corresponding finite approximation. This in turn implies that a strategy profile is a Nash equilibrium of the infinite game, if and only if all its finite prefixes are Nash equilibria in the corresponding finite games. The same holds for fixed points by construction of the lazy improvement steps for infinite games. Thus, the claim for infinite games follows from the result for finite games, i.e. Observation 4.
2.  $\Rightarrow$  3. If  $s$  is a fixed point, then the lazy improvement sequence with starting point  $s$  is constant, hence has  $s$  as accumulation point.
3.  $\Rightarrow$  2. Let the strategy profile  $s$  arise as an accumulation point of the stable subsequence of a sequence  $(s_n)_{n \in \mathbb{N}}$  obtained by deepening lazy improvement, and assume that  $s$  is not a fixed point. Then there is some minimal inspection depth  $d$  necessary to find a lazy improvement step in  $s$ , which is executed by some Player  $a$ . The detection at inspection depth  $d$  means that any strategy profile  $s'$  sharing a finite prefix of depth  $d$  with  $s$  will admit exactly the same lazy improvement step. The assumption that  $s$  is an accumulation point of the stable subsequence in particular implies that infinitely many strategy profiles occur that share a prefix of length  $d$  with  $s$ . In particular, there would have to be a strategy profile that shares a prefix of length  $d$  with  $s$ , and that is stable at inspection depth  $d' > d$ . But, as explained above, Player  $a$  would then wish to change his strategy, i.e. we have arrived at a contradiction. Hence,  $s$  has to be a fixed point.  $\square$

### 6.3 Fair Lazy Improvement

The third approach is based on what we call fair lazy improvement, and it is closely related to the outcomes' being real-valued payoff tuples. An infinite improvement sequence is fair

<sup>4</sup> Given the history dependence of the lazy improvement dynamics, a *fixed point* is understood to be any starting point of a lazy improvement sequence resulting in a constant sequence, i.e. no improvement step is found at any inspection depth.

if the following holds: Every player who could improve her payoff by more than some given value infinitely often also makes such an improvement infinitely often. This condition rules out two undesirable cases: First, a player keeps improving towards some lower payoff, while a larger payoff has been available all along; second, a player never gets the chance to improve at all, while she could improve significantly. The formal definition follows:

**Definition 15** [*Fair improvement*]<sup>5</sup> Consider a game with real-valued payoff functions  $(f_a)_{a \in A}$ . A lazy improvement sequence  $(s_n)_{n \in \mathbb{N}}$  is fair if the following holds: For all positive real numbers  $r$  and all players  $a \in A$ , if for all  $n$  there are  $m > n$  and a strategy profile  $s'$  such that  $s_m \rightarrow_a s'$  and  $f_a(s_m) + r < f_a(s')$ , then  $s_n \rightarrow_a s_{n+1}$  and  $f_a(s_n) + r < f_a(s_{n+1})$  for infinitely many  $n$ .

As of now we are unable to answer the following question.

**Open question 16** *In an infinite binary game with continuous real-valued payoff functions, are all accumulation points of a fair lazy improvement sequence Nash equilibria?*

We will show that the answer is positive in two special cases. First, if the sequence has only finitely many accumulation points, we can use the lazy non-worsening we introduced in Sect. 4.3 together with a limit argument to establish the following.

**Theorem 5** *If a fair lazy improvement sequence  $(s_n)_{n \in \mathbb{N}}$  in a binary game with continuous payoff functions has only finitely many accumulation points, then all of them are Nash equilibria.*

**Proof** As there are only finitely many accumulation points of  $(s_n)_{n \in \mathbb{N}}$ , there are only finitely many positions in the game tree where the current choice changes infinitely many times. Let  $d_0 \in \mathbb{N}$  be large enough that no such position occurs below depth  $d_0$  in the game tree. For any vertex  $v$  below depth  $d_0$ , the sequence of payoffs induced by  $s_n$  starting from  $v$  will converge.

Assume for the sake of contradiction that  $(s_n)_{n \in \mathbb{N}}$  has some accumulation point  $s$  which is not a Nash equilibrium. By continuity of the payoff functions, there is some  $d_1 \geq d_0$ , a Player  $a$  and some  $\delta > 0$  such that infinitely many  $s_n$  coincide with  $s$  at least up to depth  $d_1$ , that any two paths agreeing up to depth  $d_1$  grant Player  $a$  payoffs differing by less than  $\delta$ , and moreover, that in any strategy profile coinciding with  $s$  up to depth  $d_1$ , Player  $a$  has a lazy improvement step of at least  $\delta$  available.

Let  $g_n$  be the finite game of depth  $d_1$  where each leaf has the same payoff as the subgame starting at the corresponding vertex in the original game would yield using the strategy profile  $s_n$ , and let  $s'_n$  be the corresponding truncation of  $s_n$ . Let  $g$  be the limit of the  $g_n$ . We have either  $s'_n = s'_{n+1}$ , or  $s'_n \rightarrow s'_{n+1}$  in  $g_n$ . We can safely remove duplicates from the sequence. By Lemma 4 we find that  $s'_n \rightarrow \cup \rightsquigarrow s'_{n+1}$  in  $g$ , and then Theorem 3 implies that each player (in particular Player  $a$ ) makes only finitely many improvement steps in the sequence  $(s'_n)_{n \in \mathbb{N}}$ . Now any improvement step by Player  $a$  in  $(s_n)_{n \in \mathbb{N}}$  by more than  $\delta$  corresponds to an improvement step in the  $(s'_n)_{n \in \mathbb{N}}$ , and hence he makes only finitely many of those. This contradicts the fairness of  $(s_n)_{n \in \mathbb{N}}$ .  $\square$

Covering the case of only finitely many accumulation points does not suffice in general, as there can be uncountably many, as we shall proceed to show.

**Proposition 4** *Let  $A \subseteq \{0, 1\}^{\mathbb{N}}$  be a non-empty closed set with empty interior. Then there is a one player game with a continuous payoff function, and a fair lazy improvement sequence  $(s_n)_{n \in \mathbb{N}}$ , such that  $A$  is the set of runs induced by the accumulation points of  $(s_n)_{n \in \mathbb{N}}$ .*

<sup>5</sup> This is fair as in fair scheduler, not as in fair division of cake.

**Proof** We will use the payoff function  $p \mapsto (1 - d(p, A))$  for the player. As a closed subset of Cantor space,  $A$  can be represented as the set of infinite path through some pruned tree  $T_A \subseteq \{0, 1\}^{\mathbb{N}}$ . As  $A$  has empty interior, we know that for any  $v \in T_A$  there is some extension  $w \supseteq v$  with  $w \notin T_A$ . By iteratively applying this to children, we find that for all  $v \in T_A$  and  $k > |v|$ , there is some  $w_{v,k} \supseteq v$  with  $w_{v,k} \notin T_A$ ,  $|w_{v,k}| \geq k$ , and the longest proper prefix of  $w_{v,k}$  is in  $T_A$ .

Let  $T_A = \{v_n \mid n \in \mathbb{N}\}$ . We now construct a sequence of paths  $(p_n)_{n \in \mathbb{N}}$  iteratively, together with an auxiliary sequence  $(k_n)_{n \in \mathbb{N}}$  of integers. Let  $k_0 := 0$ , and  $p_0$  be some path extending  $w_{v_0,0}$ . Then let us always choose  $k_{n+1}$  such that  $d(p_n, A) > 2^{-k_{n+1}+1}$ , and  $p_{n+1}$  to be some path extending  $w_{v_{n+1},k_{n+1}}$ .

This construction ensures that  $d(p_m, A)$  converges monotonely to 0. We derive a sequence  $(s_m)_{m \in \mathbb{N}}$  of strategy profiles linked via lazy convertibility, such that  $s_m$  induces  $p_m$ . Then  $(s_m)_{m \in \mathbb{N}}$  is a fair lazy improvement sequence. It remains for us to argue that  $A$  is the set of accumulation points of  $(p_n)_{n \in \mathbb{N}}$ . Some open ball  $v\{0, 1\}^{\mathbb{N}}$  intersects  $A$  iff  $v \in T_A$ . Since any such  $v$  has infinitely many extensions  $v'$  also in  $T_A$ , we see that there are infinitely many  $p_n$  with prefix  $v$ . Thus, any  $p \in A$  is an accumulation point of  $(p_n)_{n \in \mathbb{N}}$ . Moreover, since  $\lim_{n \rightarrow \infty} d(p_n, A) = 0$ ,  $(p_n)_{n \in \mathbb{N}}$  cannot have any accumulation points outside of  $A$ .  $\square$

**Corollary 7** *There are fair lazy improvement sequences with uncountably many accumulation points.*

We can extend the argument based on a measure employed in the proof of Theorem 2 to the infinite case, provided that the payoff functions satisfy a rather strong Lipschitz condition. This conditions is used to ensure that (a modification of) the measure is a finite quantity.

**Proposition 5** *If the game tree is binary and if for each player there exists  $\eta > 2$  such that her payoff function is Lipschitz continuous for the distance  $d$  defined by  $d(h0\rho, h1\rho') = \frac{1}{\eta^{|h|}}$ , then all the accumulation points of a fair lazy improvement sequence are Nash equilibria.*

**Proof** To all strategy profiles  $s$  and all players  $a$  let us associate a real number:

$$M_a(s) := \sum_{h \in d^{-1}(a)} f_a(h \cdot (1 - s(h)) \cdot \rho(h \cdot (1 - s(h)), s)) - \min_{\rho \in \{0,1\}^\omega} (f_a(h\rho))$$

where  $d(h)$  is the player that plays at history  $h$ , and  $f_a(\rho)$  is the payoff for Player  $a$  and run  $\rho$ , and  $s(h)$  is the choice in  $\{0, 1\}$  that is prescribed by  $s$  at  $h$ , and  $\rho(h, s)$  is the run induced by strategy profile  $s$  from  $h$  on. Similarly to the finite case  $f_a(h \cdot (1 - s(h)) \cdot \rho(h \cdot (1 - s(h)), s))$  is the payoff that is avoided by  $a$  at history  $h$ . Note that the summands of  $M_a(s)$  are all non-negative by definition of the minimum, and that the sum converges absolutely: Indeed, by assumption  $|f_a(h0\rho) - f_a(h1\rho')| \leq \frac{L_a}{\eta^{|h|}}$  for some  $L_a > 0$  and for all  $h, \rho$ , and  $\rho'$ , so  $M_a(s) \leq \sum_{h \in \{0,1\}^*} \frac{L_a}{\eta^{|h|}} = L_a \sum_{l=0}^{+\infty} \left(\frac{2}{\eta}\right)^l = \frac{L_a}{1-\frac{2}{\eta}}$ . Also, each  $M_a$  is continuous.

Similarly to the finite case, it is easy to see that  $M_a$  is left unchanged when another player performs a lazy convertibility step. Also,  $M_a$  decreases by  $\delta$  when Player  $a$  performs a lazy convertibility step that improves her payoff by  $\delta \in \mathbb{R}$ : first prove the claim for convertibility step changing only one choice (at one node); then by induction the claim holds for finitely many changes; finally, the full claim holds by continuity of  $M_a$ . As  $M_a$  is non-negative, it follows that for all  $\delta > 0$ , in all lazy improvement sequences, no player can infinitely often improve by more than  $\delta$ .

Assume that  $(s_n)_{n \in \mathbb{N}}$  is a lazy improvements sequence, and let  $s$  be some accumulation point that is not an Nash equilibrium. So  $s \rightarrow_a t$  for some profile  $t$  and Player  $a$ . By continuity

of the payoffs there are  $\delta, \varepsilon > 0$  such that whenever  $d(s, s') < \delta$ , there is some  $t'$  such that  $s' \rightarrow_a t'$ , and the payoff for  $a$  in  $t'$  exceeds her payoff in  $s'$  by at least  $\delta$ .

Now if  $(s_n)_{n \in \mathbb{N}}$  were fair, then  $a$  would need to improve by at least  $\delta$  infinitely often, contradicting our observation above.  $\square$

### 6.4 Transfinite Lazy Improvement in the Difference Hierarchy

We start by formalizing what it means to do a transfinite number of improvement steps. The following definition generalizes the notion of finite sequence or  $\omega$ -sequence induced by a binary relation to  $\alpha$ -sequence for some ordinal  $\alpha$ : At limit ordinals, following a valid sequence amounts to picking an “accumulation point”.

**Definition 17** [*Ordinal sequence of a relation*] Let  $\rightarrow$  be a binary relation on some topological space  $S$ , and let  $\alpha$  be an ordinal number. An  $\alpha$ -sequence of  $\rightarrow$  is a family  $(s_\beta)_{\beta < \alpha}$  of elements in  $S$  such that for all  $\beta < \alpha$ , if  $\beta + 1 < \alpha$  then  $s_\beta \rightarrow s_{\beta+1}$ , and if  $\beta$  is a limit ordinal, then for every  $\beta' < \beta$  and every neighbourhood  $U$  of  $s_\beta$  there exists  $\gamma \in ]\beta', \beta[$  such that  $s_\gamma \in U$ .

Lemma 5 says that, given a binary relation over a compact set, the only reason why an ordinal sequence cannot be further extended is when a sink has been reached.

**Lemma 5** *Let  $(s_\beta)_{\beta < \alpha}$  be a countable ordinal sequence of  $\rightarrow$  over a compact set  $S$ . If  $(s_\beta)_{\beta \leq \alpha}$  is not a sequence of  $\rightarrow$  for any  $s_\alpha \in S$ , then  $\alpha = \alpha' + 1$  for some  $\alpha'$  and  $s_{\alpha'}$  is a sink of  $\rightarrow$ .*

**Proof** Let  $\alpha$  be a limit ordinal. Towards a contradiction let us assume that  $(s_\beta)_{\beta < \alpha}$  is not extendable. So for all  $s \in S$  there exist a neighbourhood  $U_s$  of  $s$  and an ordinal  $\beta_s < \alpha$  such that  $s_\gamma \notin U_s$  for all  $\gamma > \beta_s$ . The  $\{U_s\}_{s \in S}$  form an open cover of  $S$ , so by compactness of  $S$  there exists a finite subcover  $\{U_s\}_{s \in S'}$ . Let  $\gamma := \sup_{s \in S'} \beta_s + 1$ . So  $\gamma < \alpha$  by finiteness of  $S'$ . Moreover  $s_\gamma \notin U_s$  for all  $s \in S'$ , so  $s_\gamma \notin \cup_{s \in S'} U_s \supseteq S$ , contradiction.  $\square$

In this paper by countable we mean at most countable. We find that even in very simple games, we can have improvement sequences of any countable length.

**Proposition 6** *For every countable ordinal  $\alpha$  there exists a win–lose two-player game on a binary tree with open winning set for one player, and an  $\alpha$ -sequence of lazy improvement in the game.*

**Proof** By transfinite induction on  $\alpha$ . It holds for the case  $\alpha = 0$  (take the empty set as winning set). For the inductive case let us make a further case disjunction: First case,  $\alpha = \alpha' + 1$  is not a limit ordinal. Let  $(s_\beta)_{\beta < \alpha'}$  be an  $\alpha'$ -sequence on some game  $g$  with open winning set. Let  $X$  be the opponent of the player who wins according to  $s_{\alpha'}$ . Let us consider the supgame where  $X$  chooses between playing in  $g$  or winning directly. This leads to an  $(\alpha' + 1)$ -sequence. Second case,  $\alpha$  is a limit ordinal. Since it is countable, there exists a sequence  $(\beta_i)_{i \in \mathbb{N}}$  such that  $1 < \beta_i < \alpha$  for all  $i$  and  $\alpha = \sup_{i \in \mathbb{N}} \beta_i$ . Since  $\alpha$  is a limit ordinal,  $\beta_i + 1 < \alpha$  for all  $i$ , so by induction hypothesis let  $g_i$  be a game with open winning set for  $a$ , and that has a  $\beta_i + 1$ -sequence with starting profile  $s_i$ . Since ignoring the first profile of the sequence does not change its order type, we can further assume that  $s_i$  makes  $a$  lose. Now let us define a supgame by giving Player  $a$  the possibility to continue forever and lose, or stop at stage  $i$  and play in  $g_i$ . The winning set of  $a$  is a union of open sets and is therefore open. Let us build an  $\alpha$ -sequence as follows: Let us start with a profile where  $s^i$  is the subprofile in  $g_i$  for all  $i$ , and where  $a$  chooses to play  $g_0$  at the root of the supgame. Let the players change

strategies in  $g_0$  until  $b$  wins for the last time in the  $\beta_0 + 1$ -sequence, then let  $a$  change games to  $g_1$  and simultaneously perform the first change from  $s_1$  in  $g_1$ . Then let the players change strategies in  $g_1$  until  $b$  wins for the last time in the  $\beta_1 + 1$ -sequence, and so on.  $\square$

Lemma 6 uses the main proof technique in this section: From a putative uncountable ordinal sequence of lazy improvement, we can extract an uncountable factor (or substring) with more properties.

**Lemma 6** *Let  $g$  be a game on a binary tree, where some open set  $X$  contains only worst runs for some Player  $a$ . If there exists an uncountable sequence of lazy improvement in  $g$ , it has an uncountable subsequence where improvements from Player  $a$  do not involve runs in  $X$ .*

**Proof** Since there is an uncountable sequence of lazy improvement in  $g$ , there is a  $\omega_1$ -sequence, where  $\omega_1$  is the first uncountable ordinal. Since there are only countably many vertices in the game, and since  $X$  is open and non-empty, it can be written  $\cup_{i \in \mathbb{N}} u_i \{0, 1\}^\omega$  where all  $u_i \in \{0, 1\}^*$ . If Player  $a$  avoids some  $u_i \{0, 1\}^\omega$  at some point in the  $\omega_1$ -sequence, it avoids it for ever, since it is open and since it contains only worst possible runs. So Player  $a$  escaping  $X$  by an improvement step only occurs countably many times in the  $\omega_1$ -sequence. Let  $\Gamma$  be the set of ordinals where such improvements occur. So, such improvements do not occur from  $(\sup \Gamma) + 1$  (a countable ordinal) to  $\omega_1$ . This truncated sequence witnesses the claim.  $\square$

Lemma 7 is the base case of the proof of Theorem 6, which is proved by transfinite induction.

**Lemma 7** *Let  $g$  be a game with finitely many players who have Boolean (i.e. win/lose) objectives. If every winning set is open or closed, every sequence of lazy improvement in  $g$  is countable.*

**Proof** By induction on the number of outcome tuples occurring in the game. The claim holds for one tuple, so let us assume that at least two tuples occur in the game. Towards a contradiction, let us consider an uncountable sequence of lazy improvement in  $g$ . Let us assume that the losing set of some Player  $a$  has non-empty interior  $X$ . By applying Lemma 6 there is an uncountable subsequence where improvements from Player  $a$  do not involve runs in  $X$ . So the above sequence is still valid in the game derived from  $g$  by moving  $X$  from the losing set of Player  $a$  to her winning set. Applying this to each player yields a game  $g'$  where the losing sets are all closed with empty interiors and where there is a  $\omega_1$ -sequence of lazy improvement.

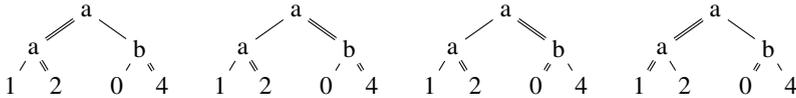
Let  $A$  be the set of the players occurring in the game and for all  $a \in A$  let  $W_a$  be the winning set of  $a$  in  $g'$ . Let  $A'$  have maximal cardinality under the constraint  $\cap_{a \in A'} W_a \neq \emptyset$ , so all runs in  $\cap_{a \in A'} W_a$  make all players in  $A \setminus A' \neq \emptyset$  lose. Since  $\cap_{a \in A'} W_a$  is open and non-empty,  $\{0, 1\}^\omega \setminus W_a$  has non-empty interior for all  $a \in A \setminus A'$ , which implies that  $A' = A$  to avoid a contradiction. So, the  $\omega_1$ -sequence of lazy improvement does not visit  $\cap_{a \in A'} W_a$  (because nobody would want to leave it), which induces an uncountable sequence with fewer tuples and allows us to conclude by IH.  $\square$

Lemma 8 will be useful during the transfinite induction step, when proving Theorem 6.

**Lemma 8** *Let  $g$  be a game, let  $a$  be a player with Boolean objectives, let  $u$  be a node of the game, let  $g_u$  be the subgame of  $g$  rooted at  $u$ , and for all profiles  $s$  in  $g$  let  $s_u$  be the corresponding profile in  $g_u$ . Consider a lazy improvement sequence in  $g$ . For all steps  $s \rightarrow a s'$  in the sequence (but possibly the first one entering  $g_u$ ), either  $s'_u = s_u$  or  $s_u \rightarrow a s'_u$  in  $g_u$ .*

**Proof** If the induced play does not reach  $u$  after the improvement step  $s \rightarrow_a s'$ , then  $s'_u = s_u$ . So let us assume it reaches  $u$  afterwards. If it also reaches  $u$  before,  $s_u \rightarrow_a s'_u$ , so let us assume it does not. Let us assume that some earlier profile induced a play that reached  $u$ . Since Player  $a$  is coming back to  $u$ , it must be her who left it. In particular, the outcome induced by  $s'_u$  makes player lose her objective, so  $s_u \rightarrow_a s'_u$ .  $\square$

**Example 3** In the following example, the numbers denote the payoffs for Player  $a$ . Player  $b$  may be assumed to be antagonistic to  $a$ . We depict a lazy improvement sequence such that its projection to the left subtree is not a lazy improvement sequence—in fact, the payoff is decreasing for the acting Player  $a$ . This shows that the restriction to boolean outcomes in Lemma 8 is not dispensable.



We recall from descriptive set theory (a standard reference is [14]) that a subset  $S$  of a metric space is called a  $\Delta^0_2$ -set, if it is expressible both as  $S = \bigcap_{i \in \mathbb{N}} U_i$  with open  $U_i$ , and as  $S = \bigcup_{i \in \mathbb{N}} A_i$  with closed  $A_i$ . By the Hausdorff–Kuratowski theorem, the  $\Delta^0_2$ -subsets of  $C^\omega$  are exactly those in the difference hierarchy. The difference hierarchy can be defined as follows:  $\mathcal{D}_0 = \{\emptyset\}$ . For some countable ordinal  $\alpha > 0$ ,  $\mathcal{D}_\alpha$  contains all sets of the form  $\bigcup_{i \in I} (u_i C^\omega \setminus A_i)$  where the  $u_i \in C^*$  are prefix independent, and each  $A_i$  appears in some  $\mathcal{D}_\beta$  with  $\beta < \alpha$ . That this indeed defines the difference hierarchy was observed by Motto-Ros [29, Section 7] extending previous work by Andretta and Martin [2]. A direct proof can be found in [23].

**Theorem 6** *Let  $g$  be a game with finitely many players who have Boolean objectives. If every winning set is  $\Delta^0_2$ , every sequence of lazy improvement in  $g$  is countable.*

**Proof** Let us proceed by transfinite induction on (the tuple of) the levels in the Hausdorff difference hierarchy of the winning sets  $W_a$  of the players  $a \in A$ . The claim holds when the  $W_a$  are open or closed by Lemma 7, so let us assume that some  $W_a$  is neither open nor closed.  $W_a$  can be written  $\bigcup_{i \in I} (u_i C^\omega \setminus A_i)$ , where the  $u_i$  are not prefixes of one another, where  $I$  is countable since there are countably many vertices, and where each  $A_i$  lies in some lower level of the difference hierarchy than  $W_a$ .

Towards a contradiction, let us assume that there is a  $\omega_1$ -sequence of lazy improvement in  $g$ . By Lemma 8 this induces sequences of equalities or lazy improvements in the subgame  $g_i$  rooted at  $u_i$  in  $g$ . Let  $\Gamma_i$  be the set of ordinals where improvement occurs in  $g_i$ . The induction hypothesis implies that  $\Gamma_i$  is countable. Let  $\Gamma'$  be the set of the ordinals where some  $g_i$  is reached for the first time. Then also  $\gamma := (\sup \Gamma' \cup \bigcup_{i \in \mathbb{N}} \Gamma_i) + 1$  is countable. In the truncated sequence from  $\gamma$  to  $\omega_1$ , the induced profiles in all  $g_i$  are constant. Let the  $t_i$  be the corresponding Boolean tuples, and let  $g'$  be derived from  $g$  by fixing the outcome tuple  $t_i$  all over  $g_i$ , for all  $i$ . In  $g'$  the winning set of  $a$  is open because it is a union of some of the  $u_i C^\omega$ , and the winning sets of the other players did not increase in complexity. So, by IH every sequence of lazy improvement in  $g'$  is countable, contradiction.  $\square$

**Corollary 8** *Let  $g$  be a game with finite branching and finitely many players who have Boolean objectives. If every winning set is  $\Delta^0_2$ , every sequence of lazy improvement in  $g$  is countable and ends at a Nash equilibrium.*

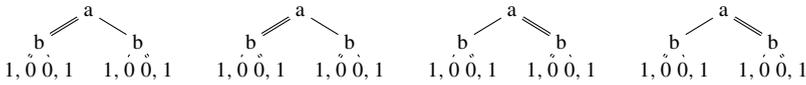
**Proof** By Theorem 6 and Lemma 5, since finite branching implies compactness.  $\square$

Regarding a potential extension of Corollary 8 to winning sets beyond  $\Delta_2^0$  we shall make a tangential remark: The computational task of finding a Nash equilibrium in a two-player game in extensive form with  $\Delta_2^0$  winning sets is just as hard as iterating the task of finding an accumulation point of a sequence over some countable ordinal. This follows from results in [6,22,23,30]. Finding a Nash equilibrium of a game with  $\Sigma_2^0$  winning sets is strictly more complicated. Thus,  $\Delta_2^0$  seems to be a natural boundary for results of the form of Corollary 8.

### 7 Some Counterexamples

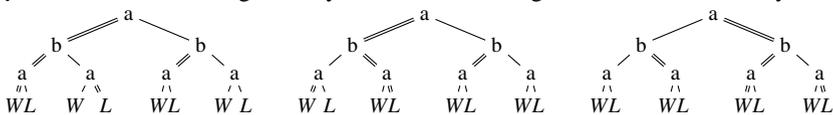
In order to obtain the termination result in the finite case (Theorem 1), some restriction on how players can improve is indeed necessary. We shall show below that the better-response dynamics  $\rightarrow$  may fail to terminate even for very simple games in extensive form:

**Example 4** An improvement cycle:



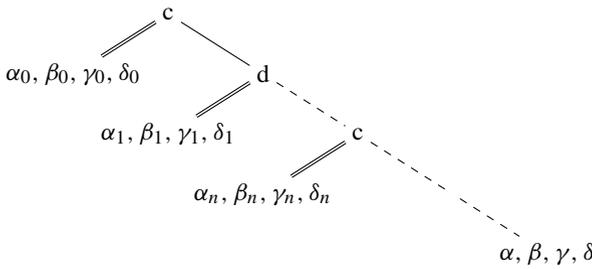
The technical notion of strategy that is used in this article to represent the intuitive concept of a strategy (in games in extensive form) is not the only possible notion. An alternative notion does not require choices from a player at every node that she owns, but only at nodes that are not ruled out by the strategy of the same player. The three objects in Example 5 are such minimalist, alternative strategy profiles, where double lines still represent choices. Up to symmetry, they constitute from left to right a cycle of improvements that could be intuitively described as lazy, so an actual cycle of length eight can easily be inferred from the short pseudocycle. This may happen because, although the improvements may look lazy, Player *a* forgets about her choices in a subgame (of the root) when leaving it, and may settle for different choices when coming back to the subgame. This suggests that even counter-factual choices are sometimes relevant. In particular, this means that lazy improvement is not a natural dynamics in the sense of Hart [12]; or a simple model in the sense of Roth and Erev [31].

**Example 5** Let *W* be winning for Player *a* and *L* be losing, and vice versa for Player *b*.



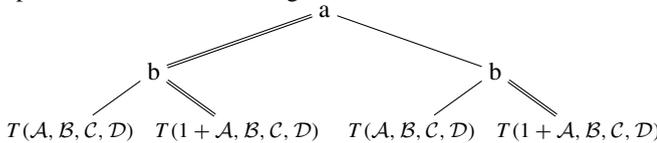
The example below shows that for infinite games, a sequence of lazy improvement steps may have multiple accumulation points even for continuous payoff functions; and moreover, that not all accumulation points have to be Nash equilibria.

**Example 6** ([21, Example 26])

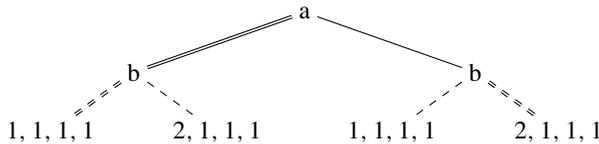


Let us consider games with four players  $a, b, c$ , and  $d$ . Given four real-valued sequences  $A = (\alpha_n)_{n \in \mathbb{N}}$ ,  $B = (\beta_n)_{n \in \mathbb{N}}$ ,  $C = (\gamma_n)_{n \in \mathbb{N}}$ , and  $D = (\delta_n)_{n \in \mathbb{N}}$  converging towards  $\alpha, \beta, \gamma$ , and  $\delta$ , let  $T(A, B, C, D)$  be the following game and strategy profile. Note that apart from the payoffs, the underlying game effectively involves players  $c$  and  $d$  only. If  $C$  and  $D$  are increasing, the lazy improvement dynamics sees players  $c$  and  $d$  alternating in switching their top left-move to a right-move.

Let  $A := B := (1 + \frac{1}{n+1})_{n \in \mathbb{N}}$  and let  $C := D := (1 - \frac{1}{n+1})_{n \in \mathbb{N}}$ . Starting from the profile below, players  $c$  and  $d$  will continue to unravel the subgame currently chosen jointly by  $a$  and  $b$ . Player  $b$  will keep alternating her choices to pick the least-unravalled subgame available to her. Player  $a$  will prefer to chose a subgame where Player  $b$  currently chooses right, and also prefers less-unravalled subgames.

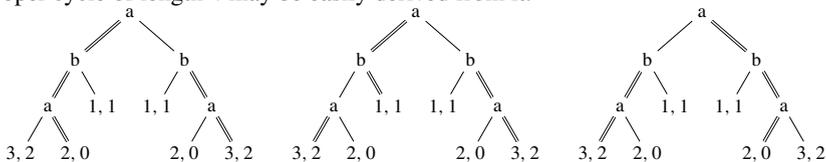


First of all, already the subgame where  $b$  moves first demonstrates that the lazy improvement dynamics will not always converge, hence we have to consider accumulation points rather than limit points. For the next feature, note that there is an infinite sequence of lazy improvement where players  $a$  and  $b$  (at both nodes that she owns) switch infinitely often, and where Player  $a$  switches only when Player  $b$  chooses the right subgame (on the induced play). Then the following strategy profile is an accumulation point, but it is clearly not a Nash equilibrium.



In our current model the players perform lazy improvement updates in a sequential manner. If simultaneity was allowed (yet not compulsory), cycles could occur, as shown in the example below.

**Example 7** The following profiles constitute the beginning of a cycle of synchronous updates. A proper cycle of length 4 may be easily derived from it.



This behaviour can be avoided by considering lazy best-response dynamics, rather than merely lazy better-response. In the sequential case, clearly the termination of the latter implies termination of the former. In the simultaneous case we find the following.

**Proposition 7** *The synchronous lazy best-response sequences in a game with  $n$  internal nodes have length at most  $2^n$ , provided that the players have acyclic preferences.*

**Proof** It suffices to prove the claim for preferences that are linear orders, which we prove by induction on the number of internal nodes of the game  $g$ . (It holds for zero.) Let  $v$  be an internal node in  $g$  whose children are all leaves, let  $a$  be the owner of  $v$ , and let us consider a sequence where  $a$  always chooses the same outcome  $x$  at  $v$ . Let  $g'$  be the game derived from  $g$  by replacing  $v$  with a leaf enclosing the outcome  $x$ . The synchronous lazy best-response sequence in  $g$  corresponds, by restriction of the profiles, to a sequence in  $g'$ , so it has length at most  $2^{n-1}$  by I. H. Now let us consider an arbitrary sequence, and note that  $a$  can change choices only once at  $v$ , from some non-preferred outcome to her preferred one (among the outcomes occurring below  $v$ ). So the length of a sequence in  $g$  is at most  $2^{n-1} + 2^{n-1} = 2^n$ .  $\square$

**Acknowledgements** We are grateful to Dietmar Berwanger, Victor Poupet, and Martin Ziegler for helpful discussions, and to an anonymous referee for his or her helpful comments on a previous version of this paper.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Andersson D, Gurvich V, Hansen TD (2010) On acyclicity of games with cycles. *Discrete Appl Math* 158(10):1049–1063. <https://doi.org/10.1016/j.dam.2010.02.006>
2. Andretta A, Martin DA (2003) Borel–Wadge degrees. *Fundam Math* 177(2): 175–192. <http://eudml.org/doc/283071>
3. Aumann RJ (1995) Backward induction and common knowledge of rationality. *Games Econ Behav* 8(1):6–19. [https://doi.org/10.1016/S0899-8256\(05\)80015-6](https://doi.org/10.1016/S0899-8256(05)80015-6)
4. Aumann RJ, Brandenburger A (1995) Epistemic conditions for Nash equilibrium. *Econometrica* 63(5):1161–1180. <https://doi.org/10.2307/2171725>
5. Boros E, Elbassioni K, Gurvich V, Makino K (2012) On Nash equilibria and improvement cycles in pure positional strategies for chess-like and backgammon-like  $n$ -person games. *Discrete Math* 312(4):772–788. <https://doi.org/10.1016/j.disc.2011.11.011>
6. Brattka V, Gherardi G, Marcone A (2012) The Bolzano–Weierstrass Theorem is the jump of Weak König’s Lemma. *Ann Pure Appl Logic* 163(6):623–625. <https://doi.org/10.1016/j.apal.2011.10.006>. [arXiv:1101.0792](https://arxiv.org/abs/1101.0792)
7. Brihaye T, Geeraerts G, Hallet M, Le Roux S (2017) Dynamics and coalitions in sequential games. In: *Proceedings eighth international symposium on games, automata, logics and formal verification, GandALF 2017*, pp 136–150. <https://doi.org/10.4204/EPTCS.256.10>
8. Cressman R (2003) *Evolutionary dynamics and extensive form games*. MIT Press, Cambridge
9. Cressman R, Schlag K (1998) The dynamic (in)stability of backwards induction. *J Econ Theory* 83(2):260–285. <https://doi.org/10.1006/jeth.1996.2465>
10. Fudenberg D, Levine DK (1993) Self-confirming equilibrium. *Econometrica* 61(3): 523–545 . <http://www.jstor.org/stable/2951716>
11. Gale D, Stewart F (1953) Infinite games with perfect information. In: Kuhn HW, Tucker AW (eds) *Contributions to the theory of games*, vol 28. *Annals of mathematical studies*. Princeton University Press, Princeton, pp 245–266. <https://doi.org/10.1515/9781400881970-014>
12. Hart S (2008) Dynamics and equilibria. In: *GAMES 2008*
13. Hart S, Mas-Colell A (2000) A simple adaptive procedure leading to correlated equilibrium. *Econometrica* 68:1127–1150. <https://doi.org/10.1111/1468-0262.00153>

14. Kechris A (1995) Classical descriptive set theory, vol 156. Graduate texts in mathematics. Springer, New York
15. Kukushkin NS (2002) Perfect information and potential games. *Games Econ Behav* 38(2):306–317. <https://doi.org/10.1006/game.2001.0859>
16. Kukushkin NS (2011) Acyclicity of improvements in finite game forms. *Int J Game Theory* 40(1):147–177. <https://doi.org/10.1007/s00182-010-0231-0>
17. Kukushkin NS (2011) Nash equilibrium in compact-continuous games with a potential. *Int J Game Theory* 40(2):387–392. <https://doi.org/10.1007/s00182-010-0261-7>
18. Le Roux S (2008) Generalisation and formalisation in game theory. Ph.D. thesis, Ecole Normale Supérieure de Lyon
19. Le Roux S (2009) Acyclic preferences and existence of sequential Nash equilibria: a formal and constructive equivalence. In: TPHOLS, international conference on theorem proving in higher order logics. Lecture notes in computer science. Springer, pp 293–309. [https://doi.org/10.1007/978-3-642-03359-9\\_21](https://doi.org/10.1007/978-3-642-03359-9_21)
20. Le Roux S (2013) Infinite sequential Nash equilibria. *Log Methods Comput Sci* 9(2). [https://doi.org/10.2168/LMCS-9\(2:3\)2013](https://doi.org/10.2168/LMCS-9(2:3)2013)
21. Le Roux S, Pauly A (2014) Infinite sequential games with real-valued payoffs. In: CSL-LICS '14. ACM, pp 62:1–62:10. <https://doi.org/10.1145/2603088.2603120>
22. Le Roux S, Pauly A (2014) Weihrauch degrees of finding equilibria in sequential games. [arXiv:1407.5587](https://arxiv.org/abs/1407.5587)
23. Le Roux S, Pauly A (2015) Weihrauch degrees of finding equilibria in sequential games. In: Beckmann A, Mitraná V, Soskova M (eds) *Evolving computability*, vol 9136. Lecture notes in computer science. Springer, pp 246–257. [https://doi.org/10.1007/978-3-319-20028-6\\_25](https://doi.org/10.1007/978-3-319-20028-6_25)
24. Le Roux S, Pauly A (2016) A semi-potential for finite and infinite sequential games (extended abstract). In: Cantone D, Delzanno G (eds) *Proceedings of GANDALF, EPTCS*, vol 226, pp 242–256. <https://doi.org/10.4204/EPTCS.226.17>
25. Martin DA (1975) Borel determinacy. *Ann Math* 102(2):363–371. <https://doi.org/10.2307/1971035>
26. Meir R, Polukarov M, Rosenschein JS, Jennings NR (2017) Iterative voting and acyclic games. *Artif Intell* 252:100–122. <https://doi.org/10.1016/j.artint.2017.08.002>
27. Mertens JF (1987) Repeated games. In: *Proceedings of the international congress of mathematicians*. American Mathematical Society, pp 1528–1577
28. Monderer D, Shapley L (1996) Potential games. *Games Econ Behav* 14(124):124–143. <https://doi.org/10.1006/game.1996.0044>
29. Motto-Ros L (2009) Borel-amenable reducibilities for sets of reals. *J Symb Logic* 74(1):27–49. <https://doi.org/10.2178/jsl/1231082301>
30. Pauly A (2015) Computability on the countable ordinals and the Hausdorff–Kuratowski theorem. [arXiv:1501.00386](https://arxiv.org/abs/1501.00386)
31. Roth AE, Erev I (1995) Learning in extensive form games: experimental data and simple dynamic models in the intermediate term. *Games Econ Behav* 8:164–212. [https://doi.org/10.1016/S0899-8256\(05\)80020-X](https://doi.org/10.1016/S0899-8256(05)80020-X)
32. Solan E, Vieille N (2003) Deterministic multi-player Dynkin games. *J Math Econ* 39(8):911–929. [https://doi.org/10.1016/S0304-4068\(03\)00021-1](https://doi.org/10.1016/S0304-4068(03)00021-1)
33. Weihrauch K (2000) *Computable Analysis*. Springer, Berlin
34. Wellman MP, Hu J (1998) Conjectural equilibrium in multiagent learning. *Mach Learn* 33(2):179–200. <https://doi.org/10.1023/A:1007514623589>
35. Xu Z (2013) Convergence of best-response dynamics in extensive-form games. SSE/EFI working paper series in economics and finance 745, Stockholm School of Economics