Fixed Point Solution to Stochastic Priority Games

Bruno Karelovic, Wieslaw Zielonka

Fixed Point Solution to Stochastic Priority Games

Bruno Karelović and Wiesław Zielonka
IRIF, Université Paris Diderot - Paris 7
case 7014, 75205 Paris Cedex 13, France
bruno.karelovic@gmail.com wieslaw.zielonka@irif.fr

March 1, 2017

Abstract

We define and examine a class of two-player stochastic games that we call priority games. The priority games contain as proper subclasses the parity games studied in computer science [4] and the games with the limsup and liminf payoff mappings. We show that the value of the priority game can be expressed as an appropriate nested fixed point of the value mapping of the one-day game. This extends the result of de Alfaro and Majumdar [4], where the authors proved that the value of the stochastic parity game can be expressed as the nested fixed point of the one-day value mapping.

The difference between our paper and [4] is two-fold.

The value of the parity game is obtained by applying the least and the greatest fixed points to the value mapping of the one-day game. However, in general, the greatest and the least fixed-points are not sufficient in order to obtain the value of the priority game.

To cope with this problem we introduce the notion of the nearest fixed point of a monotone bounded nonexpansive mapping. Our main result is that the value of the priority game can be obtained as the nested nearest fixed point of the value mapping of the one-day game.

The second point that makes our proof different from [4] is that our proof is inductive. We give a game interpretation for the nested fixed point formula where some variables are free (not bounded by the fixed point operator). Thus instead of proving the main result in one big step as in [4] we can limit ourselves to the case when just one fixed point is added to the nested fixed point formula.

1 Introduction

Stochastic two-player zero-sum games model the long-term interactions between two players that have strictly opposite objectives.

The study of stochastic games starts with the seminal paper of Shapley [10]. Since then stochastic games were intensively studied in game theory and, more recently, in computer science.

In stochastic games players' preferences are expressed by means of a payoff mapping. The payoff mapping maps infinite plays (infinite sequences of states and actions) to real numbers. The payoff mappings used in computer science tend to be different from the payoff mappings used in game theory. The payoffs prevalent in computer science are often expressed in some kind of logic which implies that they take only two values, 1 for the winning plays and 0 for the losing plays.

On the other hand, the payoff mappings used in game theory are rather real valued: mean-payoff, discounted payoff, lim sup and lim inf payoffs are among the most popular ones.

In this paper we define and examine the class of priority games. The priority games constitute a natural extension of parity games, this latter class is the class of games popular in computer science having applications in automata theory and verification.

However, the priority games are also relevant to the games traditionally studied in game theory. It turns out that the games with the limsup and liminf payoff [8] belong to the class of priority games.

To put the results of the paper in the context we recall below the relevant results concerning the stochastic parity games due to deAlfaro and Majumdar [4].

1.1 Parity games

A stochastic parity game is a zero-sum two-player game infinite game played by two players, player Max and player Min, on an arena with a finite set of states $\mathbf{S} = [n] = \{1, \ldots, n\}$.

For each state i, players Max and Min have nonempty sets of available actions, $\mathbf{A}(i)$ and $\mathbf{B}(i)$ respectively. At each stage, the players, knowing the current state and all the previous history, choose independently and simultaneously actions $a \in \mathbf{A}(i)$ and $b \in \mathbf{B}(i)$ respectively and the game moves to state j with probability $p(j|i, a, b)$. Immediately after each stage, and before the next one, both players are informed about the action played by the adversary player.

An infinite sequence of states and action occurring during the game is called a play.

The parity games are endowed with the reward vector $r = (r_1, \ldots, r_n)$, where $r_i \in \{0, 1\}$ is the reward of state i. The parity payoff $\varphi(h)$ of an infinite play h is defined to be equal[1] to the reward of the maximal state visited infinitely often in h, i.e. the payoff is equal to r_i if i was visited infinitely often in h and all states $j, j > i$, were visited only a finite number of times.

The set of all plays is endowed in the usual way with the Borel σ-algebra generated by the cylinders. Strategies σ, τ of players Max and Min and an initial state $i \in \mathbf{S}$ give rise to a probability measure $\mathbf{P}_i^{\sigma,\tau}$ over the Borel σ-algebra. The aim of player Max (respectively Min) is to maximize (respectively minimize) the expected payoff

$$\mathbf{E}_i^{\sigma,\tau}(\varphi) = \int \varphi(h)\mathbf{P}_i^{\sigma,\tau}(dh)$$

[1]The payoff of the parity game is usually formulated in a bit different way, however it is easy to see that the definition given here is equivalent to the usual one.

for each initial state i.

Since the parity payoff is Borel measurable, by the result of Martin [9], parity games have value v_i for each initial state i, i.e.

$$\sup_{\sigma}\inf_{\tau}\mathbf{E}_i^{\sigma,\tau}(\varphi) = v_i = \inf_{\tau}\sup_{\sigma}\mathbf{E}_i^{\sigma,\tau}(\varphi), \quad \forall i \in \mathbf{S}. \tag{1}$$

One of the techniques used to solve parity games relies on the μ-calculus. In this approach the point of departure is the one-day game[2] played at each state $i \in \mathbf{S}$. The one-day game has a value for each state $i \in \mathbf{S}$ and each reward vector $r = (r_1, \ldots, r_n)$. Let

$$f = (f_1, \ldots, f_n) \tag{2}$$

be the mapping that maps the reward vector $r \in \{0,1\}^n$ to the vector of values of the one-day game, i.e. for $r = (r_1, \ldots, r_n)$ and $i \in \mathbf{S}$, $f_i(r)$ is the value of the one-day game played at state i when the reward vector is r. We endow $[0,1]^n$ with the product order, $x = (x_1, \ldots, x_n) \leqslant (y_1, \ldots, y_n) = y$ if $x_i \leqslant y_i$ for all $i \in [n]$, which makes it a complete lattice. It is easy to see that

$$f : [0,1]^n \to [0,1]^n$$

is monotone under $\leqslant$, thus by Tarski's theorem [11], f has the least and the greatest fixed points.

Then one can define the nested fixed point

$$\mathbf{Fix}^n(f)(r) = \mu^{r_n}x_n.\mu^{r_{n-1}}x_{n-1}.\ldots.\mu^{r_2}x_2.\mu^{r_1}x_1.f(x_1, x_2, \ldots, x_{n-1}, x_n), \tag{3}$$

where $\mu^{r_i}x_i$ denotes either the greatest fixed point if $r_i = 1$, or the least fixed point if $r_i = 0$, and f the value function (2) of the one-day game. The main result obtained by de Alfaro and Majumdar [4] in the μ-calculus approach to parity games is that

$$v = (v_1, \ldots, v_n) = \mathbf{Fix}^n(f)(r),$$

where the left-hand side vector v is composed of the values v_i for the parity game starting at i, cf. (1). To summarize, the value vector of the parity game can be obtained by calculating the nested fixed point of the one-day value mapping[3].

The μ-calculus approach to parity games was first developed for deterministic parity games (perfect information games with deterministic transitions), see Walukiewicz [12]. The paper of de Alfaro and Majumdar [4] extended this approach to stochastic parity games.

[2] In computer science papers the one-day game is often not mentioned explicitly, but the value function f of the one-day game is used in the μ-calculus approach to parity games, where it is sometimes called the predecessor operator.

[3] The traditional presentation of this result is a bit different. Roughly speaking the variables are regrouped in blocks, each block consists of consecutive variables to which the same fixed point is applied. The each fixed point is applied to a group of variables rather than to separate variables. This allows to decreases the number of fixed points and the resulting formula alternates the least and the greatest fixed points. This is, however, only a technical detail which has no bearing on the result. For our purposes it is more convenient to apply fixed points to variables rather than to groups of variables.

1.2 From parity games to priority games

The parity games arose from the study of decidability questions in logic. In this framework the winning criteria are expressed in some kind of logic, where there is room for only two types of plays, the winning plays that satisfy a logical formula and the losing plays that do not satisfy the formula. For this reason the rewards in the parity games take only two values, 0 and 1, with the intuition that the reward 1 is favorable and the reward 0 unfavorable for our player (and the preferences are inverse for the adversary player).

However, the restriction to 0, 1 rewards does not allow to express finer player's preferences. This motivates the study of the games that allow any real rewards.

We define the priority game as the game where each state $i \in [n] = \mathbf{S}$ is equipped with a reward $r_i \in \mathbb{R}$. Like in the parity game the payoff $\varphi(h)$ of a play h is defined to be the reward r_i of the greatest state i that is visited infinitely often in h.

At first glance, the priority game is just a mild extension of the parity game. This impression is reinforced by the fact that deterministic priority games can be reduced to deterministic parity games. (However, we do not know if such reduction is possible for stochastic priority games.)

The interest in priority games is twofold. First, the priority games allow to quantify players' preferences in a more subtle way than it is possible in parity games. While in parity games there are only two classes of plays, the plays with the parity payoff 1 and the plays with the parity payoff 0, in priority games we can distinguish many levels of preferences. As a motivating simple example consider the priority game with three states $\mathbf{S} = \{1, 2, 3\}$ and rewards $r_1 = 0, r_2 = 1, r_3 = \frac{3}{4}$. This game gives rise to three distinct classes of infinite plays: player Max highest preference is for the plays such that the maximal state visited infinitely often is state 2 (these plays give him the payoff 1), his second preference is for the plays that visit state 3 infinitely often (such plays give him the payoff $\frac{3}{4}$), and his lowest preference is for the plays that from some moment onward stay forever in state 1 (this yields him the payoff 0). It is impossible to capture such a hierarchy of preferences when we limit ourselves to the parity payoff.

The second reason to be interested in the games with priority payoff stems from the fact that not only they generalize the parity games, but they contain as proper subclasses the games with the lim sup and lim inf payoffs [7]. Let $(i_k)_{k=1}^{\infty}$ be the infinite sequence of states visited during the play, where i_k is the state visited at stage k. Let $(r_{i_k})_{k=1}^{\infty}$ be the corresponding sequence of rewards. The limsup game (respectively liminf game) is the game with the payoff equal to $\limsup_k r_{i_k}$ (respectively $\liminf_k r_{i_k}$).

To see that a limsup game is a priority game let us take a finite state limsup game and rename the states in such a way that for any two states $i, j \in [n]$, if $i < j$ then $r_i \leqslant r_j$, i.e. the natural order of states reflects the reward order. Then the limsup payoff will be equal to the priority payoff.

For a liminf game we proceed in a similar way: we rename the states in such a way that, for any two states $i, j \in [n]$, $i < j$ implies that $r_j \leqslant r_i$. Under this condition the liminf payoff will be equal to the priority payoff.

Our approach to priority games is inspired by the μ-calculus approach to parity

games. There are two major differences however.

It is impossible to solve the priority games using only the least and the greatest fixed points, we need also other fixed points that we name "the nearest fixed points". To define this notion we use the well known fact that the one-day game value mapping (2) is not only monotone but it is also nonexpansive, which means that, for $x, y \in \mathbb{R}^n$, $\|f(x) - f(y)\|_\infty \leqslant \|x - y\|_\infty$, where $\|x\|_\infty = \sup_i |x_i|$ is the supremum norm.

In the study of parity games the fact that the one-day game value mapping f is nonexpansive is irrelevant, the monotonicity of f is all that we need in order to apply Tarski's fixed point theorem. When we study the priority games, other fixed points enter into consideration and the fact that f is nonexpansive becomes paramount.

It turns out that the priority games with rewards in $\mathbb{R}$ can be reduced through a linear transformation to the priority games with rewards in the interval $[0, 1]$. Therefore in the sequel we assume that the reward vector $r = (r_1, \ldots, r_n)$ belongs to $[0, 1]^n$. Under this condition value mapping f of the one-day game (2) is a monotone nonexpansive mapping from $[0, 1]^n$ to $[0, 1]^n$. Since our study of priority games is based on the analysis of the fixed points of f, in Section 3 we prepare the background and present basic facts concerning fixed points of monotone nonexpansive mappings from $[0, 1]^n$ to $[0, 1]^n$. All the results presented in Section 3 are either well known or are rather straightforward observations. The only purpose of Section 3 is to regroup in one place all the relevant facts and to introduce the notion of the nearest fixed point

$$\mu^r x.g(x)$$

of monotone nonexpansive mappings $g : [0, 1] \to [0, 1]$. Intuitively, $\mu^r x.g(x)$ is the fixed point of g which is nearest to $r \in [0, 1]$. Note that the least and the greatest fixed points of g are special cases of this notion, the greatest fixed point is the fixed point nearest to 1 and the least fixed point is the fixed point nearest to 0. We show that the notion of the nearest fixed point makes sense for monotone nonexpansive mappings from $[0, 1]$ to $[0, 1]$. In Section 3 we define also, for each vector $r = (r_1, \ldots, r_n) \in [0, 1]^n$ and a monotone nonexpansive mapping $f : [0, 1]^n \to [0, 1]^n$, the nested nearest fixed point

$$\mathbf{Fix}^n(f)(r) = \mu^{r_n} x_n.\mu^{r_{n-1}} x_{n-1}. \ldots \mu^{r_2} x_2.\mu^{r_1} x_1.f(x_1, x_2, \ldots, x_{n-1}, x_n), \tag{4}$$

which generalizes the nested least/greatest fixed point (3).

Section 4 introduces the one-day games.

Section 5 constitutes the core of the paper.

We prove that the value vector $v = (v_1, \ldots, v_n)$, where v_i is the value of state i in the priority games satisfies

$$v = (v_1, \ldots, v_n) = \mathbf{Fix}^n(f)(r),$$

where the right-hand side is the nested nearest fixed point (4) of the value mapping of the one-day game.

Although the result of Section 5 can be seen as an extension of the μ-calculus characterization known for parity games [4], there is one point that distinguish our approach from the traditional μ-calculus approach to parity games. In the case of parity games[4],

[4]This remark concerns also deterministic parity games [12].

to the best of our knowledge, the μ-calculus proofs presented previously were not inductive, rather a formula similar to (3) was presented and it was shown, in one big step, that it represents the value of the parity game. The fact that the nested fixed point formula (3) is in some sense recursive, was not exploited to the full extent in the proof.

The novelty of the proof presented in Section 5 lies in the fact that it is genuinely inductive. We provide a clear game theoretic interpretation of the partial fixed point formula

$$\mathbf{Fix}^k(f)(r) = \mu^{r_k} x_k. \ldots \mu^{r_1} x_1. f(x_1, \ldots, x_k, r_{k+1}, \ldots, r_n), \tag{5}$$

where the fixed points are applied only to the low priority variables $x_1, \ldots, x_k$, while the free variables $x_{k+1}, \ldots, x_n$ take values $r_{k+1}, \ldots, r_n$ respectively.

Let $G(r)$ be the priority game endowed with the reward vector r. Let $G_k(r)$ be the priority game obtained from $G(r)$ by transforming all states $i, i > k$, into absorbing states[5], while the states j with $j \leqslant k$ have the same transitions in $G_h(r)$ as in $G(r)$. Both games have the same reward vector r.

It turns out that the partial nested fixed point (5) is equal to the value vector $v = (v_1, \ldots, v_n)$ of the priority game $G_k(r)$. We prove this fact by induction starting with the trivial priority game $G_0(r)$, where all states are absorbing. The inductive step consist in showing that, if (5) is the value of the game $G_k(r)$, then adding the new fixed point $\mu^{r_{k+1}} x_{k+1}$ we obtain the value vector of the game $G_{k+1}(r)$. In other words, adding one fixed point corresponds to the transformation of an absorbing state into a nonabsorbing one. Note that in priority games the absorbing states are trivial, if a state m is absorbing then $v_m = r_m$, i.e. the value of m is equal to the reward r_m. Thus transforming an absorbing state into a nonabsorbing we convert a trivial state into a nontrivial one. The crucial point is that in the inductive proof given in the paper we apply this transformation to just one state. And it is much easier to comprehend what happens if one state changes its quality from absorbing to nonabsorbing than when all states are nonabsorbing from the outset.

The preliminary version of this paper appeared in [5].

2 Stochastic priority games

An arena for a two-player stochastic priority game is composed of a finite set of states $\mathbf{S} = [n] = \{1, 2, \ldots, n\} \subset \mathbb{N}$ (we assume without loss of generality that $\mathbf{S}$ is a subset of positive integers) and finite sets $\mathbf{A}$ and $\mathbf{B}$ of actions of players Max and Min. For each state i, $\mathbf{A}(i) \subseteq \mathbf{A}$ and $\mathbf{B}(i) \subseteq \mathbf{B}$ are finite nonempty sets of actions that players Max and Min can play at i. We assume that $\mathbf{A}$ and $\mathbf{B}$ are disjoint and $(\mathbf{A}(i))_{i \in \mathbf{S}}$, $(\mathbf{B}(i))_{i \in \mathbf{S}}$ are partitions of $\mathbf{A}$ and $\mathbf{B}$.

For $i, j \in \mathbf{S}, a \in \mathbf{A}(i), b \in \mathbf{B}(i)$, $p(j|i, a, b)$ is the probability to move to j if players Max and Min execute respectively actions a and b at i.

[5] Recall that a state i is absorbing if it is impossible to leave i, i.e. for all possible actions executed in i with probability 1 the game remains in i.

An infinite game is played by players Max and Min. At each stage, given the current state i, the players choose simultaneously and independently actions $a \in \mathbf{A}(i)$ and $b \in \mathbf{B}(i)$ and the game moves to a new state j with probability $p(j|i,a,b)$. The couple (a,b) is called the *joint action*.

A finite history is a sequence $h = s_1, a_1, b_1, s_2, a_2, b_2, s_3 \ldots, a_{t-1}, b_{t-1}, s_t$ alternating states s_i and joint actions (a_i, b_i) and beginning and ending with a state. The length of h is the number of joint actions in h, in particular a history of length 0 consists of just one state and no actions. The set of finite histories is denoted H.

A strategy of player Max is a mapping $\sigma : H \to \Delta(\mathbf{A})$, where $\Delta(\mathbf{A})$ denotes the set of probability distributions over $\mathbf{A}$. We require that $\mathrm{supp}(\sigma(h)) \subseteq \mathbf{A}(i)$, where i is the last state of h and $\mathrm{supp}(\sigma(h)) := \{a \in \mathbf{A} \mid \sigma(h)(a) > 0\}$ is the support of the measure $\sigma(h)$.

A strategy σ is *memoryless* if $\sigma(h)$ depends only on the last state of h. Thus memoryless strategies of player Max can be identified with mappings from $\mathbf{S}$ to $\Delta(\mathbf{A})$ such that $\mathrm{supp}(\sigma(i)) \subseteq \mathbf{A}(i)$ for each $i \in \mathbf{S}$.

Strategies for player Min are defined in a similar way.

We use σ and τ (with subscripts or superscripts) to denote strategies of Max and Min.

Σ and $\mathcal{T}$ will stand for the sets of all strategies for players Max and Max respectively.

An infinite history or a play is an infinite sequence
$h = s_1, a_1, b_1, s_2, a_2, b_2, s_3, a_3, b_3, \ldots$ alternating states s_i and joint actions (a_i, b_i). The set of infinite histories is denoted H^∞. For a finite history h, by h^+ we denote the cylinder generated by h consisting of all infinite histories with prefix h. We assume that H^∞ is endowed with the σ-algebra $\mathcal{B}(H^\infty)$ generated by the set of cylinders.

Strategies σ, τ of players Max and Min and the initial state i determine a probability measure $\mathbf{P}_i^{\sigma,\tau}$ on $(H^\infty, \mathcal{B}(H^\infty))$.

We define inductively $\mathbf{P}_i^{\sigma,\tau}$ for cylinders in the following way.

Let $h_0 = s_1$ be a finite history of length 0. Then

$$\mathbf{P}_i^{\sigma,\tau}(h_0^+) = \begin{cases} 0 & \text{if } i \neq s_1, \\ 1 & \text{if } i = s_1. \end{cases}$$

Let $h_{t-1} = s_1, a_1, b_1, \ldots, s_{t-1}, a_{t-1}, b_{t-1}, s_t$ and $h_t = h_{t-1}, a_t, b_t, s_{t+1}$. Then

$$\mathbf{P}_i^{\sigma,\tau}(h_t^+) = \mathbf{P}_i^{\sigma,\tau}(h_{t-1}^+) \cdot \sigma(h_{t-1})(a_t) \cdot \tau(h_{t-1})(b_t) \cdot p(s_{t+1}|s_t, a_t, b_t).$$

Note that the set of cylinders is π-system (i.e. a family of sets closed under intersection) thus a probability defined on cylinders extends in a unique way to all sets of $\mathcal{B}(H^\infty)$.

To define the stochastic priority game we endow the arena with a reward vector

$$r = (r_1, \ldots, r_n)$$

associating with each state i a reward $r_i \in \mathbb{R}$.

Given the reward vector r, the *priority payoff* is a mapping

$$\varphi_r : H^\infty \to \mathbb{R}$$

such that for an infinite history $h = s_1, (a_1, b_1), s_2, (a_2, b_2), s_3, (a_3, b_3), \ldots$

$$\varphi_r(h) = r_\ell, \quad \text{where } \ell = \limsup_t s_t. \tag{6}$$

Thus the priority payoff is equal to the reward of the greatest (in the usual integer order) state visited infinitely often.

The aim of player Max (player Min) is to maximize (resp. minimize) the expected payoff

$$\mathbf{E}_i^{\sigma,\tau}[\varphi_r] = \int_{H^\infty} \varphi_r(h) \mathbf{P}_i^{\sigma,\tau}(dh).$$

The priority game has value v_i for a starting state i if

$$\inf_{\tau\in\mathcal{T}} \sup_{\sigma\in\Sigma} \mathbf{E}_i^{\sigma,\tau}[\varphi] = v_i = \sup_{\sigma\in\Sigma} \inf_{\tau\in\mathcal{T}} \mathbf{E}_i^{\sigma,\tau}[\varphi].$$

From the determinacy of Blackwell's games proved by Martin [9] it follows that the priority game has value for each initial state. (The Blackwell games do not have states but the result of Martin extends to the games with states as shown by Maitra and Sudderth [7].)

A strategy τ of player Min is ε-optimal, $\varepsilon \geqslant 0$, if for each state i and each strategy σ of player Max,

$$\sup_{\sigma\in\Sigma} \mathbf{E}_i^{\sigma,\tau}[\varphi] \leqslant v_i + \varepsilon.$$

Symmetrically, a strategy σ of player Max is ε-optimal if for each state i and each strategy τ of player Min,

$$\inf_{\tau\in\mathcal{T}} \mathbf{E}_i^{\sigma,\tau}[\varphi] \geqslant v_i - \varepsilon.$$

An ε-optimal strategy with $\varepsilon = 0$ is called *optimal*.

If the reward vector is such that $rew_i \in \{0, 1\}$ for each state i then we obtained the parity payoff. A proof of determinacy of stochastic parity games using fixed points was given by de Alfaro and Majumdar [4].

2.1 Normalizing the rewards

In the sequel it will be convenient to assume that all rewards belong to the interval $[0, 1]$ rather than to $\mathbb{R}$.

This can be achieved without loss of generality by a simple linear transformation. Let $a = \min_{i\in\mathbf{S}} r_i$, $b = \max_{i\in\mathbf{S}} r_i$ and $g(x) = \frac{1}{b-a}x - \frac{a}{b-a}$. Then $0 = g(a) \leqslant f(x) \leqslant g(b) = 1$ for $x \in \{r_1, \ldots, r_n\}$. Changing the reward vector from $r = (r_1, \ldots, r_n)$ to $g(r) = (g(r_1), \ldots, g(r_n))$ transforms linearly the priority payoff of all plays h since $\varphi_{g(r)}(h) = g(\varphi_r(h))$.

By the linearity of expectation, this implies that for all starting states i and all strategies σ and τ we have $g(\mathbf{E}_i^{\sigma,\tau}(\varphi_r)) = \mathbf{E}_i^{\sigma,\tau}(g(\varphi_r))$. This implies that v_i is the value of state i for the game with the priority payoff φ_r if and only if $g(v_i)$ is the value of i for the game with the priority payoff $\varphi_{g(r)}$. Similarly a strategy is ε-optimal for the priority payoff φ_r if and only if it is $\frac{\varepsilon}{b-a}$-optimal for the priority payoff $\varphi_{g(r)}$.

3 On fixed points of bounded monotone nonexpansive mappings

In this technical section we introduce monotone nonexpansive mappings, that play a crucial role in the study of stochastic priority games. The solution to stochastic priority games given in Section 5 relies heavily on fixed point properties of such mappings examined in Section 3.1. In Section 3.2 we define and examine the nested nearest fixed points of monotone nonexpansive mappings.

The duality of the nested nearest fixed points is studied in Section 3.3.

An element $x = (x_1, \dots, x_n)$ of $\mathbb{R}^n$ will be identified with the mapping x from $[n] = \{1, \dots, n\}$ to $\mathbb{R}$ and we can occasionally write $x(i)$ to denote x_i.

The set $\mathbb{R}^n$ is endowed with the natural componentwise order, for $x, y \in \mathbb{R}^n$, $x \leqslant y$ if $x_i \leqslant y_i$ for all $i \in [n]$.

A mapping $f : \mathbb{R}^n \to \mathbb{R}^k$ is *monotone* if for $x, y \in \mathbb{R}^n$, $x \leqslant y$ implies $f(x) \leqslant f(y)$ (we do not assume that $k = n$, thus $x \leqslant y$ and $f(x) \leqslant f(y)$ can relate to componentwise orders in two different spaces).

We assume that the Cartesian product $\mathbb{R}^n$ is endowed with the structure of a normed real vector space with the norm $\|\cdot\|_\infty$, for $x \in \mathbb{R}^n$, $\|x\|_\infty = \max_{i\in[n]}|x_i|$. Thus, for $x, y \in \mathbb{R}^n$, $\|x - y\|_\infty$ defines a distance between x and y.

We say that a mapping $f : \mathbb{R}^n \to \mathbb{R}^k$ is *nonexpansive* if, for all $x, y \in \mathbb{R}^n$, $\|f(x) - f(y)\|_\infty \leqslant \|x - y\|_\infty$.

Such a mapping f can be written as vector of k mappings $f = (f_1, \dots, f_k)$, where $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, k$. Clearly, f is monotone nonexpansive iff all f_i are monotone nonexpansive.

We say that a mapping $f : \mathbb{R}^n \to \mathbb{R}^k$ is *additive homogeneous* if for all $\lambda \in \mathbb{R}$ and $x \in \mathbb{R}^n$

$$f(x + \lambda e_n) = f(x) + \lambda e_k,$$

where e_n and e_k are the vectors $(1, \dots, 1)$ in $\mathbb{R}^n$ and $\mathbb{R}^k$ respectively having all components equal to 1.

Crandall and Tartar [2] proved the following result.

Lemma 1 (Crandall and Tartar [2])**.** *For additive homogeneous mappings $f : \mathbb{R}^n \to \mathbb{R}^k$ is the following conditions are equivalent:*

(i) f is monotone,

(ii) f is nonexpansive.

We will need only the implication (i)→(ii) that we prove below for the reader's convenience. Moreover, if the result holds for mappings from $\mathbb{R}^n$ to $\mathbb{R}$ then it holds for mappings from $\mathbb{R}^n$ to $\mathbb{R}^k$. Thus we assume in the proof that that $f : \mathbb{R}^n \to \mathbb{R}$.

Proof. For $x, y \in \mathbb{R}^n$, $e_n = (1, 1, \ldots, 1) \in \mathbb{R}^n$ and $\lambda = \|x - y\|_\infty$ we have $y - \lambda e_n \leqslant x \leqslant y + \lambda e_n$. Thus for $f : \mathbb{R}^n \to \mathbb{R}$ monotone and additive homogeneous we obtain

$$f(y) - \lambda \leqslant f(x) \leqslant f(y) + \lambda.$$

Thus $|f(x) - f(y)| \leqslant \lambda = \|x - y\|_\infty$. □

3.1 Fixed points of monotone nonexpansive mappings

We say that a monotone mapping $f : \mathbb{R}^n \to \mathbb{R}^k$ is *bounded* if $f([0,1]^n) \subseteq [0,1]^k$.

The set of bounded monotone nonexpansive mappings will be denoted by $M_{n,k}[0,1]$. Moreover BMN will stand for the abbreviation for "bonded monotone nonexpansive".

In this section we introduce the notion of the nearest fixed point of BMN mappings generalizing the least and greatest fixed points.

In the following lemma states basic properties of fixed points of BMN mappings.

Lemma 2. *Let $f \in M_{1,1}[0,1]$. Define by induction, $f^{(0)}(x) = x$, $f^{(1)}(x) = f(x)$, $f^{(i+1)}(x) = f(f^{(i)}(x))$, for $x \in [0,1]$.*

Then

(i) for each $x \in [0,1]$ the sequence $(f^{(i)}(x)), i = 0, 1, \ldots,$ is monotone and converges to some $x^\infty \in [0,1]$. The limit x^∞ is a fixed point of f, $f(x^\infty) = x^\infty$,

(ii) if $x \leqslant y$ are fixed points of f, $f(x) = x$ and $f(y) = y$, then for each z such that $x \leqslant z \leqslant y$, $f(z) = z$,

(iii) the sequence $(f^{(i)}(0)), i = 0, 1, 2, \ldots,$ converges to the least fixed point $\bot_f$ of f while the sequence $(f^{(i)}(1)), i = 0, 1, 2, \ldots,$ converges to the greatest fixed point $\top_f$ of f. The interval $[\bot_f, \top_f]$ is the set of all fixed points of f.

If $0 \leqslant x \leqslant \bot_f$ then the sequence $(f^{(i)}(x))$ converges to $\bot_f$.

If $\top_f \leqslant x \leqslant 1$ then the sequence $(f^{(i)}(x))$ converges to $\top_f$.

If $0 \leqslant x < \bot_f$ then $x < f(x)$.

If $\top_f < x \leqslant 1$ then $f(x) < x$.

Proof. (i) Suppose that $f(x) \leqslant x$. Then inductively, since f is non-increasing, $f^{(i+1)}(x) \leqslant f^{(i)}(x)$ for all i, i.e. the sequence $f^{(i)}(x)$ is non-increasing. Since this sequence is bounded from below by 0 it converges to some x^∞.

The case of $f(x) \geqslant x$ can be treated in a similar way.

Since f is nonexpansive $|f(x^\infty) - f^{(i+1)}(x)| \leqslant |x^\infty - f^{(i)}(x)|$. Because the right-hand side tends to 0 we can see that $f^{(i)}(x)$ converges to $f(x^\infty)$. On the other hand, $f^{(i)}(x)$ converges to x^∞. Therefore $f(x^\infty) = x^\infty$.

(ii) Let $0 \leqslant x \leqslant z \leqslant y \leqslant 1$ and $f(x) = x$, $f(y) = y$. Since f is monotone, $f(x) \leqslant f(z) \leqslant f(y)$. Thus, since f is nonexpansive, $0 \leqslant f(y) - f(z) \leqslant y - z$ and $0 \leqslant f(z) - f(x) \leqslant z - x$. This implies that $f(z) = z$.

(iii) is a direct consequence of (i) and (ii). □

Let $f \in M_{1,1}[0,1]$. For $a \in [0,1]$ we define *the nearest fixed point* $\mu_a x.f(x)$ of f to be

$$\mu_a x.f(x) := \lim_i f^{(i)}(a).$$

Lemma 2 shows that this is really a fixed point of f which is closest to a, i.e. $|a - \mu_a x.f(x)| = \min_{z\in[0,1]}\{|a - z| \mid f(z) = z\}$.

Moreover, the least and the greatest fixed points of $f \in M_{1,1}[0,1]$ are respectively equal to $\mu_0 x.f(x)$ and $\mu_1 x.f(x)$.

We can see also that

$$\mu_a x.f(x) = \begin{cases} \mu_0 x.f(x) & \text{if } a \leqslant \mu_0 x.f(x), \\ a & \text{if } \mu_0 x.f(x) \leqslant a \leqslant \mu_1 x.f(x), \\ \mu_1 x.f(x) & \text{if } \mu_1 x.f(x) \leqslant a, \end{cases} \tag{7}$$

i.e. the fixed point nearest to a is equal either to the least or to the greatest fixed point or is equal to a itself.

Let $f \in M_{n,1}[0,1]$. Fixing $(r_1, \ldots, r_{k-1}, r_{k+1}, \ldots, r_n) \in [0,1]^{n-1}$ we can consider the mapping

$$x_k \mapsto f(r_1, \ldots, r_{k-1}, x_k, r_{k+1}, \ldots, r_n).$$

from $[0,1]$ to $[0,1]$. This mapping belongs to $M_{1,1}[0,1]$ thus, given $r_k \in [0,1]$, we can calculate the nearest fixed point

$$\mu_{r_k} x_k.f(r_1, \ldots, r_{k-1}, x_k, r_{k+1}, \ldots, r_n).$$

This fixed point depends on $r = (r_1, \ldots, r_{k-1}, r_k, r_{k+1}, \ldots, r_n)$, thus we can define the mapping

$$[0,1]^n \ni (r_1, \ldots, r_{k-1}, r_k, r_{k+1}, \ldots, r_n) \mapsto \mu_{r_k} x_k.f(r_1, \ldots, r_{k-1}, x_k, r_{k+1}, \ldots, r_n) \in [0,1] \tag{8}$$

Lemma 3. *If $(x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n)$ is BMN then the mapping (8) is BMN.*

Proof. Let $r = (r_1, \ldots, r_n), w = (w_1, \ldots, w_n) \in [0,1]^n$. Define two sequences $(r_k^i), i = 1, 2, \ldots$ and $(w_k^i), i = 1, 2, \ldots$, such that

$$r_k^1 = r_k \quad \text{and} \quad r_k^{i+1} = f(r_1, \ldots, r_{k-1}, r_k^i, r_{k+1}, \ldots, r_n)$$

and

$$w_k^1 = w_k \quad \text{and} \quad w_k^{i+1} = f(w_1, \ldots, w_{k-1}, w_k^i, w_{k+1}, \ldots, w_n).$$

By Lemma 2 both sequences converge to some r_k^∞ and w_k^∞ respectively and

$$r_k^\infty = \mu_{r_k} x_k . f(r_1, \ldots, r_{k-1}, x_k, r_{k+1}, \ldots, r_n)$$

and

$$w_k^\infty = \mu_{w_k} x_k . f(w_1, \ldots, w_{k-1}, x_k, w_{k+1}, \ldots, w_n).$$

We shall prove by induction that for all i, $|r_k^i - w_k^i| \leqslant \|r - w\|_\infty$.

Clearly, $|r_k^1 - w_k^1| = |r_k - w_k| \leqslant \max_i |r_i - w_i| = \|r - w\|_\infty$. Suppose that

$$|r_k^i - w_k^i| \leqslant \|r - w\|_\infty.$$

Then

$$|r_k^{i+1} - w_k^{i+1}| = |f(r_1, \ldots, r_{k-1}, r_k^i, r_{k+1}, \ldots, r_n) - f(w_1, \ldots, w_{k-1}, w_k^i, w_{k+1}, \ldots, w_n)| \leqslant$$
$$\max\{\max_{j \neq k} |r_j - w_j|, |r_k^i - w_k^i|\} \leqslant$$
$$\max\{\max_{j \neq k} |r_j - w_j|, \|r - w\|_\infty\} = \|r - w\|_\infty.$$

Taking the limit $i \nearrow \infty$ we obtain $|r_k^\infty - w_k^\infty| \leqslant \|r - w\|_\infty$. This proves that (8) is nonexpansive.

That (8) is monotone is obvious and left to the reader.

□

Note that the usual point of view (at least when only the greatest and the least fixed points are applied) is that, for a mapping $f \in M_{n,1}[0,1]$ taking the fixed point $\mu_{r_k} x_k . f(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n)$ bounds the variable x_k, i.e. we consider this expression as the function of the variables $x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n$ while r_k is considered as a constant. In other words, for a given fixed r_k we can consider the mapping

$$(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \mapsto \mu_{r_k} x_k . f(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n).$$

From Lemma **??** it follows that this mapping belongs to $M_{n-1,1}$.

Clearly, Lemma **??** adopts a lager point of view where, although is some sense the variable x_k becomes bound by the fixed point $\mu_{r_k} x_k$, at the same time r_k becomes a "new" variable. This larger point of view is interesting since it allows to examine how the nearest fixed point changes in function of r_k. In the next section we will define the nested nearest fixed point $\mu_{r_n} x_n . \ldots \mu_{r_1} x_1 . f(x_1, \ldots, x_n)$ of a mapping $f \in M_{n,n}[0,1]$. From the traditional point view this expression defines some special fixed point of f, i.e. some special element $d \in [0,1]^n$ such that $f(d) = d$.

However d depends on or more precisely is a function of $r = (r_1, \ldots, r_n)$. And it is interesting and fruitful to examine the function

$$(r_1, \ldots, r_n) \mapsto \mu_{r_n} x_n . \ldots \mu_{r_1} x_1 . f(x_1, \ldots, x_n).$$

Lemma 4. *If $f \in M_{k,m}[0,1]$ and $g \in M_{m,n}[0,1]$ then $g \circ f \in M_{k,n}[0,1]$, i.e. the composition of BMN mappings is BMN.*

Proof. For $x, y \in [0,1]^k$, we have $\|g(f(x)) - g(f(y))\|_\infty \leqslant \|f(x) - f(y)\|_\infty \leqslant \|x - y\|_\infty$. Trivially, monotonicity is also preserved by composition. □

3.2 Nested fixed points of bounded monotone nonexpansive mappings

In this section we define the *nested nearest fixed point operators*

$$\mathbf{Fix}^k : M_{n,n}[0,1] \to M_{n,n}[0,1], \quad k = 0, 1, \ldots, n.$$

Each $\mathbf{Fix}^k$ can be decomposed into n operators $\mathbf{Fix}^k_i$,

$$\mathbf{Fix}^k_i : M_{n,n}[0,1] \to M_{n,1}[0,1], \quad i \in [n],$$

so that, for $f \in M_{n,n}$,

$$\mathbf{Fix}^k(f) = (\mathbf{Fix}^k_1(f), \ldots, \mathbf{Fix}^k_n(f)).$$

Let $f = (f_1, \ldots, f_n) \in M_{n,n}[0,1]$, where $f_i \in M_{n,1}[0,1]$, for $i \in [n]$.
We set $\mathbf{Fix}^0(f)$ to be such that

$$\mathbf{Fix}^0(f)(r) = r, \quad \text{for } r \in [0,1]^n.$$

Thus $\mathbf{Fix}^0(f)$ is the identity mapping and does not depend of f. Note that $\mathbf{Fix}^0_i(f)(r) = r_i$, i.e. $\mathbf{Fix}^0_i(f)$ is the projection on the ith coordinate.

In general we set

$$\mathbf{Fix}^k_i(f)(r) = r_i, \quad \text{for all } 0 \leqslant k < i \leqslant n.$$

It remains to define $\mathbf{Fix}^k_i(f)(r)$ for $i \leqslant k$.

The definition is by induction on k. Suppose that $\mathbf{Fix}^{k-1}(f)$ is defined.
For $r \in [0,1]^n$ and $\zeta \in [0,1]$ let us set

$$F^{k-1}_i(\zeta; r) := \mathbf{Fix}^{k-1}_i(f)(r_1, \ldots, r_{k-1}, \zeta, r_{k+1}, \ldots, r_n), \quad \text{for } i \in [k-1]. \tag{9}$$

Note that $F^{k-1}_i(\zeta; r)$ depends on ζ and on $(r_1, \ldots, r_{k-1}, r_{k+1}, \ldots, r_n)$ but does not depend on r_k. Thus F^{k-1}_i is in fact a mapping from $[0,1]^n$ to $[0,1]$.

Then we define

$$\mathbf{Fix}^k_k(f)(r) := \mu_{r_k}\zeta.f_k(F^{k-1}_1(\zeta; r), \ldots, F^{k-1}_{k-1}(\zeta; r), \zeta, r_{k+1}, \ldots, r_n), \tag{10}$$
$$\mathbf{Fix}^k_i(f)(r) := F^{k-1}_i(r_1, \ldots, r_{k-1}, \mathbf{Fix}^k_k(f)(r), r_{k+1}, \ldots, r_n), \quad \text{for } i \in [k-1],$$
$$\mathbf{Fix}^k_i(f)(r) := r_i, \quad \text{for } i \in \{k+1, \ldots, n\}.$$

Since the definition of the nested fixed point mappings uses only the composition and the nearest fixed point operators, Lemmas 4 and 3 imply that

Corollary 5. *If $f \in M_{n,n}[0,1]$ then, for all $k \in \{0\} \cup [n]$, $\mathbf{Fix}^k(f) \in M_{n,n}[0,1]$.*

Let us note finally that $\mathbf{Fix}^k(f)$ depends only on $f_1, \ldots, f_k$ but is independent of $f_{k+1}, \ldots, f_n$.

3.3 Duality for the bounded monotone nonexpansive mappings

In this section we define and examine the notion of duality for the BMN mappings.

For $r = (r_1, \ldots, r_n) \in [0,1]^n$ we set $1 - r := (1 - r_1, \ldots, 1 - r_n)$.

Given a BMN mapping $f : [0,1]^n \to [0,1]$ the *dual* of f is the mapping $\overline{f} : [0,1]^n \to [0,1]$ such that

$$\overline{f}(r_1, \ldots, r_n) = 1 - f(1 - r_1, \ldots, 1 - r_n).$$

The dual of $f = (f_1, \ldots, f_k) \in M_{n,k}[0,1]$ is defined as $\overline{f} = (\overline{f}_1, \ldots, \overline{f}_n)$.

We can write this in a more explicit way if for $f = (f_1, \ldots, f_k) \in M_{n,k}[0,1]$ we define $1 - f := (1 - f_1, \ldots, 1 - f_k)$.

Then using this notation, for $f \in M_{n,k}[0,1]$, we can write succinctly

$$\overline{f}(r) = 1 - f(1 - r).$$

Lemma 6. *If f is BMN then $\overline{f}$ is BMN.*

Proof. Let $(r_1, \ldots, r_n) \leqslant (w_1, \ldots, w_n)$.

Then $(1 - r_1, \ldots, 1 - r_n) \geqslant (1 - w_1, \ldots, 1 - w_n)$ and $f(1 - r_1, \ldots, 1 - r_n) \geqslant f(1 - w_1, \ldots, 1 - w_n)$.

Thus $\overline{f}(r_1, \ldots, r_n) = 1 - f(1 - r_1, \ldots, 1 - r_n) \leqslant 1 - f(1 - w_1, \ldots, 1 - w_n) \leqslant \overline{f}(w_1, \ldots, w_n)$, i.e. $\overline{f}$ is monotone.

Finally $\|\overline{f}(r) - \overline{f}(w)\|_\infty = \|(1 - f(1-r)) - (1 - f(1-w))\|_\infty \leqslant \|(1-r) - (1-w)\|_\infty = \|r - w\|_\infty$, i.e. $\overline{f}$ is nonexpansive. □

Lemma 7. *If $f \in M_{n,1}[0,1]$ then, for all $k \in [n]$ and $r = (r_1, \ldots, r_n) \in [0,1]^n$,*

$$\mu_{r_k} x_k . f(r_1, \ldots, r_{k-1}, x_k, r_{k+1}, \ldots, r_n) = \\ 1 - \mu_{1-r_k} x_k . \overline{f}(1 - r_1, \ldots, 1 - r_{k-1}, 1 - x_k, 1 - r_{k+1}, \ldots, 1 - r_n).$$

Proof. Let $\top_f$ and $\bot_f$ be respectively the greatest and the least fixed points of the mapping

$$x_k \mapsto f r_1, \ldots, r_{k-1}, x_k, r_{k+1}, \ldots, r_n.$$

Similarly let $\top_{\overline{f}}, \bot_{\overline{f}}$ the greatest and the least fixed points of the mapping

$$x_k \mapsto \overline{f}(1 - r_1, \ldots, 1 - r_{k-1}, 1 - x_k, 1 - r_{k+1}, \ldots, 1 - r_n).$$

Since $\overline{f}(1-r_1, \ldots, 1-r_{k-1}, x_k, 1-r_{k+1}, \ldots, r_n) = 1 - f(r_1, \ldots, r_{k-1}, 1-x_x, r_{k+1}, \ldots, r_n)$ we have $\bot_{\overline{f}} = 1 - \top_f$ and $\top_{\overline{f}} = 1 - \bot_f$.

There are three possibilities concerning the position of r_k relative to $\bot_f$ and $\top_f$.

If $\top_f \leqslant r_k$ then

$$\mu_{r_k} x_k . f r_1, \ldots, r_{k-1}, x_k, r_{k+1}, \ldots, r_n = \top_f.$$

However, in this case we have also $1 - r_k \leqslant 1 - \top_f = \bot_{\overline{f}}$ implying that

$$\mu_{1-r_k} x_k.\overline{f}(1 - r_1, \ldots, 1 - r_{k-1}, x_k, 1 - r_{k+1}, \ldots, r_n) = \bot_{\overline{f}}.$$

In a similar way if $r_k \leqslant \bot_f$ then

$$\mu_{r_k} x_k.f(r_1, \ldots, r_{k-1}, x_k, r_{k+1}, \ldots, r_n) = \bot_f$$

and

$$\mu_{1-r_k} x_k.\overline{f}(1 - r_1, \ldots, 1 - r_{k-1}, x_k, 1 - r_{k+1}, \ldots, r_n) = \top_{\overline{f}}.$$

The last case to examine is when $\bot_f \leqslant r_k \leqslant \top_f$. Then

$$\mu_{r_k} x_k.f(r_1, \ldots, r_{k-1}, x_k, r_{k+1}, \ldots, r_n) = r_k$$

and, on the other hand,

$$\bot_{\overline{f}} \leqslant 1 - r_k \leqslant \top_{\overline{f}},$$

implying

$$\mu_{1-r_k} x_k.\overline{f}(1 - r_1, \ldots, 1 - r_{k-1}, x_k, 1 - r_{k+1}, \ldots, r_n) = 1 - r_k.$$

□

Lemma 8. *Let $g \in M_{m,k}[0,1]$ and $f \in M_{k,n}[0,1]$. Then $\overline{f \circ g} = \overline{f} \circ \overline{g}$, i.e. the dual of the composition of BMN mappings is equal to the composition of duals.*

Proof. For $r \in [0,1]^n$ we have $\overline{(f \circ g)}(r) = 1 - (f \circ g)(1 - r) = 1 - f(g(1 - r)) = 1 - f(1 - (1 - g(1 - r))) = 1 - f(1 - \overline{g}(r)) = (\overline{f}(\overline{g}(r))$. □

The following lemma examines the duality for the nested nearest fixed points.

Lemma 9. *Let $f = (f_1, \ldots, f_n) \in M_{n,n}[0,1]$. Then for all k, $0 \leqslant k \leqslant n$, and $r \in [0,1]^n$*

$$\mathbf{Fix}^k(f)(r) = 1 - \mathbf{Fix}^k(\overline{f})(1 - r). \tag{11}$$

Proof. Induction on k.

$r \mapsto \mathbf{Fix}^0(f)(r) = r$ is the identity mapping independently of f. Thus the left-hand side of (11) is equal to r and the right-hand side is $1 - (1 - r) = r$ as well.

For each $0 \leqslant k \leqslant n$, let us set

$$\mathbf{Fix}^k(f)(r) = H^k(r) = (H_1^k(r), \ldots, H_n^k(r))$$

and

$$\mathbf{Fix}^k(\overline{f})(r) = \overline{H}^k(r) = (\overline{H}_1^k(r), \ldots, \overline{H}_n^k(r)).$$

Using this notation (11) can be written as

$$\overline{H}^k(r) = 1 - H^k(1 - r). \tag{12}$$

Our aim is to prove the last equality for k under the assumption that it holds for $k-1$.

By definition

$$\begin{aligned}\overline{H}_k^k(1-r) = \mu_{1-r_k} x_k.\overline{f}_k(&\overline{H}_1^{k-1}(1-r_1,\ldots,1-r_{k-1},x_k,1-r_{k+1},\ldots,r_n),\\ &\ldots,\\ &\overline{H}_{k-1}^{k-1}(1-r_1,\ldots,1-r_{k-1},x_k,1-r_{k+1},\ldots,r_n),\\ &x_k,1-r_{k+1},\ldots,1-r_n).\end{aligned}$$

Let us define a mapping $G^k \in M_{n,n}[0,1]$:

$$G^k := (H_1^{k-1},\ldots H_{k-1}^{k-1},\pi_k,\pi_{k+1},\ldots,\pi_n),$$

where $\pi_i(x_1,\ldots,x_n) = x_i$, $i = k, k+1,\ldots,n$, is the projection on the i-th coordinate. Since $\overline{\pi}_i = \pi_i$, i.e. the dual of the projection is equal the same projection mapping we can see that the dual to G^k is

$$\overline{G}^k = (\overline{H}_1^{k-1},\ldots \overline{H}_{k-1}^{k-1},\pi_k,\pi_{k+1},\ldots,\pi_n).$$

Therefore, by Lemmas 8 and 7,

$$\begin{aligned}\overline{H}_k^k(1-r) =&\mu_{1-r_k} x_k.\overline{f}_k \circ \overline{G}^k(1-r_1,\ldots,1-r_{k-1},x_k,1-r_{k+1}\ldots,1-r_n) =\\ &\mu_{1-r_k} x_k.\overline{f_k \circ G^k}(1-r_1,\ldots,1-r_{k-1},x_k,1-r_{k+1}\ldots,1-r_n) =\\ &1-\mu_{r_k} x_k.f_k \circ G^k r_1,\ldots,r_{k-1},x_k,r_{k+1},\ldots,r_n = 1-H_k^k(r)\end{aligned}$$

For $m \in [k-1]$,

$$\begin{aligned}\overline{H}_m^k(1-r) &= \overline{H}_m^{k-1}(1-r_1,\ldots,1-r_{k-1},\overline{H}_k^k(1-r),1-r_{k+1},\ldots,1-r_n)\\ &= \overline{H}_m^{k-1}(1-r_1,\ldots,1-r_{k-1},1-H_k^k(r),1-r_{k+1},\ldots,1-r_n)\\ &= 1-H_m^{k-1}(r_1,\ldots,r_{k-1},H_k^k(r),r_{k+1},\ldots,r_n)\\ &= 1-H_m^k(r).\end{aligned}$$

Finally, for $m > k$,

$$1-\overline{H}_m^k(1-r) = 1-(1-r_m) = r_m = H_m^k(r).$$

This terminates the proof of (12).

□

4 The one-day game

In this section we define an auxiliary one-day game. This simple game constitutes an essential ingredient in our solution to the general priority games.

Let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ be a reward vector assigning to each state i the reward x_i.

A one-day game $\boldsymbol{M}(x)$ is the game played in the following way. If the game starts at a state k then players Max and Min choose independently and simultaneously actions $a \in \mathbf{A}(k)$ and $b \in \mathbf{B}(k)$. Suppose that upon execution of (a, b) the game moves to the next state m. This ends the game and player Max receives from player Min the payoff x_m. A one-day game played at state k given the reward mapping x will be denoted $\boldsymbol{M}_k(x)$.

Note that $\boldsymbol{M}_k(x)$ can be seen as a matrix game where

$$\boldsymbol{M}_k(x)[a, b] := \sum_{m \in \mathbf{S}} x_m \cdot p(m|k, a, b)$$

is the (expected) payoff obtained by player Max from player Min when the players play actions a and b respectively.

The *value mapping of the one-day game* is the mapping $f = (f_1, \ldots, f_n)$ from $\mathbb{R}^n$ to $\mathbb{R}^n$ such that, for each state $k \in [n]$,

$$f_k(x_1, \ldots, x_n) := \mathrm{val}(\boldsymbol{M}_k(x)), \tag{13}$$

where $\mathrm{val}(\boldsymbol{M}_k(x))$ is the value of the matrix game $\boldsymbol{M}_k(x)$, In other words, $f_k(x_1, \ldots, x_n)$ is the value of the one-day game played at state k seen as a function of the reward vector $x = (x_1, \ldots, x_n)$.

We will be interested in $f_k(x)$ seen as a function of the reward vector $x = (x_1, \ldots, x_n)$.

Since all entries in the matrix game $\boldsymbol{M}_k(x)$ belong to $\mathbb{R}$, $f_k(x) \in \mathbb{R}$, i.e. f_k is a mapping from $\mathbb{R}^n$ into $\mathbb{R}$.

Lemma 10. *The value mapping f of the one-day game defined in (13) is monotone and non-expansive.*

Proof. It is easy to see that f is monotone and it is also straightforward that f is additively homogeneous, i.e, for all $x \in \mathbb{R}^n$,

$$f(x + \lambda \cdot e_n) = f(x) + \lambda \cdot e_n,$$

where $e_n = (1, \ldots, 1) \in \mathbb{R}^n$ is the vector with 1 on all components. By Lemma 1 this implies that f is nonexpansive. □

5 Stopping priority games

Stopping priority games are a variant of priority games where some states are stopping or equivalently where some states are absorbing.

We solve the stopping priority games by induction on the number of non-stopping states and we show that the value function can be expressed as the nearest fixed point of the value function (13) of the one-day game.

Let $(S_t, t \geqslant 1)$ be the stochastic process such that S_t is the state visited at stage t.

For each state $k \in [n]$ we define the random variable

$$T_{>k} : H^\infty \to \mathbb{N} \cup \{\infty\}$$

such that

$$T_{>k} = \min\{t \mid S_t > k\}.$$

Thus $T_{>k}$ is the time of the first visit to a state greater than k.

We define a new stochastic process $S_t^{[k]}, t \in \mathbb{N}$, that we shall call the *stopped state process*:

$$S_t^{[k]} = \begin{cases} S_t & \text{if } T_{>k} \geqslant t, \\ S_q & \text{if } q = T_{>k} < t. \end{cases}$$

Thus if all previously visited states belong to $\{1, \ldots, k\}$ then $S_t^{[k]}$ is equal to the state visited at the current epoch t. However, if at some previous epoch a state $> k$ was visited then $S_t^{[k]}$ is the first such state. In other words, $S_t^{[k]}$ behaves as if the states $> k$ were absorbing, if $S_t^{[k]} > k$ then $S_q^{[k]} = S_t^{[k]}$ for all $q \geqslant t$.

For a given reward vector r and $k \in [n]$ we define the *stopping priority payoff* $\varphi_r^{[k]}$:

$$\varphi_r^{[k]} = r_\ell \quad \text{where } \ell = \limsup_t S_t^{[k]}.$$

The games with payoff $\varphi_r^{[k]}$ will be called *stopping priority games*. We will also speak about the $\varphi_r^{[k]}$-game to refer to the game with payoff $\varphi_r^{[k]}$. Similarly φ_r-game will stand for the usual priority game.

Note that once a state j greater than k is visited the game with payoff $\varphi_r^{[k]}$ is for all practical purposes over, independently of what can happen in the future the payoff is equal to the reward r_j of this state and the states visited after the moment $T_{>k}$ have no bearing on the payoff.

In the $\varphi_r^{[k]}$-game the states $[k]$ will be called *non-stopping* while the states $> k$, will be called *stopping*.

Note that since we have assumed that $\mathbf{S} = [n]$, i.e. n is the greatest state, we have $\varphi_r^{[n]} = \varphi_r$.

Note also that stopping states are trivial. If $i > k$ then for all plays h starting at i, $\varphi_r^{[k]}(h) = r_i$, thus $\mathbf{E}_i^{\sigma,\tau}(\varphi_r^{[k]}) = r_i$ for all strategies σ, τ, in particular the value of stopping state i, $i > k$, is r_i.

5.1 Dual game

We have constructed a ε-optimal strategy for Max and Min for the game starting at k but the strategy for Max was constructed under the condition $r_k < w_k$ while the strategy for Min was constructed under the condition $r_k \leqslant w_k$.

How to obtain ε-optimal strategies for both players for two remaining cases ($r_k \geqslant w_k$ for Max and $r_k > w_k$ for Min) we use the natural duality of the nested fixed points and the games.

Let G be a priority game. The dual game $\overline{G}$ is obtained in the following way:

(Di) $\overline{G}$ has the same states, actions and transition probabilities as G,

(Dii) if $r = (r_1, \dots, r_n)$ is the reward vector in G then $\overline{r} = (\overline{r}_1, \dots, \overline{r}_n)$ is the reward vector in $\overline{G}$, where for $z \in [0,1]$, $\overline{z} := 1 - z$,

(Diii) players Max and Min exchange the roles, in the dual game for each state $i \in \mathbf{S}$, $\mathbf{A}(i)$ are the actions of player Max while $\mathbf{B}(i)$ are the actions of player Min, moreover in the dual game player Max wants to minimize the priority payoff $\varphi_{\overline{r}}$ while Min wants to maximize the priority payoff $\varphi_{\overline{r}}$.

To avoid confusion, we write $\overline{\text{Max}}$ and $\overline{\text{Min}}$ to denote the players, respectively, maximizing and minimizing the priority payoff in the dual game.

A strategy σ is a strategy of player Max in G if and only if it is a strategy of player $\overline{\text{Min}}$ in the dual game $\overline{G}$. A symmetric property holds for strategies of player Min.

For each play h we have $\varphi_r(h) = 1 - \varphi_{\overline{r}}(h)$, thus $\mathbf{E}_i^{\sigma,\tau}(\varphi_r) = 1 - \mathbf{E}_i^{\tau,\sigma}(\varphi_{\overline{r}})$, where the left hand side is the expected payoff in G, while $\mathbf{E}_i^{\tau,\sigma}(\varphi_{\overline{r}})$ is the expected payoff in $\overline{G}$ when $\overline{\text{Max}}$ plays according to τ and $\overline{\text{Min}}$ plays according to σ.

This implies that $v_i = 1 - \overline{v}_i$, where v_i is the value of state i in G while $\overline{v}_i$ is the value of i in the $\overline{G}$. Moreover, a strategy is ε-optimal for player Max in G if and only if it is ε-optimal for player $\overline{\text{Min}}$ in $\overline{G}$. A symmetric property holds for strategies of player Min.

6 Constructing ε-optimal strategies

The rest of this section is devoted to the proof of the following main result characterizing the values of the stopping priority games by means of the nested nearest fixed points.

Theorem 11. *Let $f : [0,1]^n \to [0,1]^n$ be the value mapping of the one-day game defined in Section 4. For $0 \leqslant k \leqslant n$, let*

$$\mathbf{Fix}^k(f)$$

be the k-th nested fixed point of f, see Section 3.2. Then, for each reward vector r, for each initial state $i \in [n]$, the stopping priority $\varphi_r^{[k]}$-game starting at i has value equal to $\mathbf{Fix}_i^k(f)(r)$.

Proof. For each $\varepsilon > 0$ we construct ε-optimal strategies for both players.

The proof is carried out by induction on k.

The case $k = 0$ is trivial since when all states are stopping then the value of each state is equal to its reward, i.e. the value of state i is $\mathbf{Fix}_i^0(f)(r) = r_i$.

Under the assumption that the theorem holds for $k-1$, i.e. $\mathbf{Fix}_i^{k-1}(f)(r)$ is the value of the non-stopping state $i \in [k-1]$ in the $\varphi_r^{[k-1]}$-game, we shall prove that $\mathbf{Fix}_i^k(f)(r)$ is the value of the non-stopping state $i \in [k]$ in the $\varphi_r^{[k]}$-game.

We will use the following notation:

$$w_k := \mathbf{Fix}_k^k(f)(r) = \mu_{r_k} x_k . f_k(F_1^{k-1}(x_k; r), \ldots, F_{k-1}^{k-1}(x_k; r), x_k, r_{k+1}, \ldots, r_n) \tag{14}$$

and

$$w_i := \mathbf{Fix}_i^k(f)(r) = F_i^{k-1}(w_k; r), \quad i \in [k-1], \tag{15}$$

where F_i^{k-1} are defined as in (9). Thus our aim is to prove that $(w_1, \ldots, w_{k-1}, w_k)$ are the values of the states $\{1, \ldots, k-1, k\}$ in the $\varphi_r^{[k]}$-game.

Since w_k is a fixed point of (14) we have

$$w_k = f_k(w_1, \ldots, w_{k-1}, w_k, r_{k+1}, \ldots, r_n). \tag{16}$$

Let T_m be the random time of the $m-$th visit to state k of the stopping state process $(S_t^{[k]})_{t \geqslant 1}$, i.e.

$$\begin{aligned} T_1 &= \min\{t \mid S_t^{[k]} = k\}, \\ T_m &= \min\{t \mid t > T_{m-1} \text{ and } S_t^{[k]} = k\} \quad \text{for } m > 1. \end{aligned} \tag{17}$$

T_m can be infinite if the number of visits of the stopping state process $S_t^{[k]}$ to the state k is smaller than m and $T_1 = 1$ if the game starts at k. Since T_m is defined w.r.t. the stopping state process $S_t^{[k]}$, $T_m < \infty$ implies that all states visited prior to the moment T_m are $\leqslant k$.

Recall that $S_t, t \geqslant 1$, is the stochastic process that gives the state visited at stage t. $A_t, t \geqslant 1$ and $B_t, t \geqslant 1$ are the stochastic processes that give the actions played by players Max and Min respectively at stage t.

Let T be any random time, i.e. a mapping from plays to $\{1, 2, \ldots\} \cup \{\infty\}$ such that for each $m \in \{1, 2, \ldots\}$ the event $\{T = m\}$ belongs to the σ-algebra $\mathcal{F}_m = \sigma(S_1, A_1, B_1, S_2, \ldots, S_m)$. In other words, $\mathcal{F}_m$ is the σ algebra generated by the cylinders h_m^+, where h_m are histories of length m.

Intuitively that means that knowing the states and actions up to time m we can decide if $T = m$ or not.

Definition 12. *For a random time T, $\theta_T : H^\infty \to H^\infty$ will denote the shift mapping that maps plays to plays and is defined in the following way*

$$\theta_T(S_1, A_1, B_1, S_2, \ldots) = S_T, A_T, B_T, S_{T+1}, A_{T+1}, B_{T+1}, S_{T+2}, A_{T+2}, B_{T+2}, \ldots,$$

where S_t is the state process giving the state visited at stage t and A_t, B_t are action processes that give the actions played by players Max and Min at stage t.

Thus the shift θ_T "forgets" all history prior to time T. Of course, θ_T is well defined only on plays such that $T < \infty$.

Below we use the shift θ_{T_m+1}, where T_m is the time of the mth visit to state k. This shift will be applied only to the plays with $T_m < \infty$.

6.1 $\varepsilon/2$-optimal strategy $\sigma_\star$ for player Max when $r_k < w_k$ and k is the starting state.

We assume that

$$r_k < w_k \tag{18}$$

and the aim is to construct a strategy $\sigma_\star$ satisfying

$$\mathbf{E}_k^{\sigma_\star,\tau}(\varphi_r^{[k]}) \geqslant w_k - \varepsilon/2 \tag{19}$$

for each strategy τ of Min.

Let

$$\eta \in (w_k - \varepsilon/2, w_k)$$

and define

$$\xi_i = F_i^{k-1}(\eta; r), \quad \forall i \in [k-1]. \tag{20}$$

By the induction hypothesis, ξ_i is the value of the $\varphi^{[k-1]}_{(r_1,\dots,r_{k-1},\eta,r_{k+1},\dots,r_n)}$-game starting at the state i.

Let us consider the one-day game $\boldsymbol{M}_k(\xi_1,\dots,\xi_{k-1},\eta,r_{k+1},\dots,r_n)$ played at state k. Then

$$\eta_\star := f_k(\xi_1,\dots,\xi_{k-1},\eta,r_{k+1},\dots,r_n) \tag{21}$$

is the value of this game.

By the properties of monotone non-expansive mappings, (18) implies that w_k is in fact the least fixed point of the mapping

$$x_k \mapsto f_k(F_1^{k-1}(x_k; r),\dots,F_{k-1}^{k-1}(x_k; r), x_k, r_{k+1},\dots,r_n).$$

Thus $\eta < w_k$ implies that

$$\eta < f_k(\xi_1,\dots,\xi_{k-1},\eta,r_{k+1},\dots,r_n) = \eta_\star \leqslant w_k. \tag{22}$$

Fix δ such that

$$0 < \delta < \eta_\star - \eta. \tag{23}$$

We define the strategy $\sigma_\star$ of player Max in the following way:

- during the m-th visit to the state k, which takes place at time T_m, c.f. (17), player Max selects actions according to his optimal strategy in the one-day game $\boldsymbol{M}_k(\xi_1,\dots,\xi_{k-1},\eta,r_{k+1},\dots,r_n)$.

- during all stages j such that $T_m < j < T_{m+1}$, i.e. between the mth and $(m+1)$th visit to k, player Max plays according to his δ-optimal strategy for the $\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},\eta,r_{k+1},\ldots,r_n)}$-game.

  When he applies this strategy then we tacitly assume that after each visit to k player Max "forgets" all preceding history and he plays as if the game started afresh at the first state visited after the last visit to k.

From the optimality of $\sigma_\star$ in the one-day game $\boldsymbol{M}_k(\xi_1,\ldots,\xi_{k-1},\eta,r_{k+1},\ldots,r_n)$, we have

$$\begin{aligned}&\sum_{i<k}\xi_i\cdot\mathbf{P}_k^{\sigma_\star,\tau}(S_{T_m+1}=i\mid T_m<\infty)\\&+\eta\cdot\mathbf{P}_k^{\sigma_\star,\tau}(S_{T_m+1}=k\mid T_m<\infty)\\&+\sum_{i>k}r_i\cdot\mathbf{P}_k^{\sigma_\star,\tau}(S_{T_m+1}=i\mid T_m<\infty)\\&\geqslant\eta_\star.\end{aligned}\tag{24}$$

Indeed, when player Max plays according to the strategy $\sigma_\star$ at the moment T_m then the current state is k and he plays using his optimal strategy in the one-day game $\boldsymbol{M}_k(\xi_1,\ldots,\xi_{k-1},\eta,r_{k+1},\ldots,r_n)$. Now it suffices to notice that the left-hand side of (24) is nothing else but the payoff that player Max obtains in the one-day game $\boldsymbol{M}_k(\xi_1,\ldots,\xi_{k-1},\eta,r_{k+1},\ldots,r_n)$ (because S_{T_m+1} is the state visited at the next time moment T_m+1). Since $\eta_\star$ is the value of this one-day game the inequality follows.

In the sequel we will note $\mathbb{1}_A$ the indicator of the event A, i.e. the mapping that is equal to 1 on A and to 0 on the complement of A.

Let us note the following equality:

$$\sum_{i>k}r_i\cdot\mathbf{P}_k^{\sigma_\star,\tau}(S_{T_m+1}=i\mid T_m<\infty)=\mathbf{E}_k^{\sigma_\star,\tau}(\varphi_r^{[k]}\cdot\mathbb{1}_{\{S_{T_m+1}>k\}}\mid T_m<\infty).\tag{25}$$

Indeed, if a play belongs to the event $\{S_{T_m+1}=i, T_m<\infty\}$ for $i>k$ then $T_m<\infty$ means that at the moment T_m this play visits k and prior to T_m it never visited states $>k$ cf. (17), and at the next time moment T_m+1 such a play visits the stopping state $i>k$. But for such plays the payoff $\varphi_r^{[k]}$ is equal to r_i.

Consider now the event $\{S_{T_m+1}=i, T_m<\infty\}$, for $i<k$, see Figure 1.
This event consists of the plays such that

- the stopping state process $S_i^{[k]}$ visits k for the mth time at time T_m (this is guaranteed by $T_m<\infty$, cf.(17)) and
- at the next time moment T_m+1 the play visits the state $i<k$.

From the definition of $\sigma_\star$ it follows that starting from the time T_m+1 player Max plays using his δ-optimal strategy in the $\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},\eta,r_{k+1},\ldots,r_n)}$-game. Since, by the inductive hypothesis (20), the value of such a game for state i is ξ_i, we have

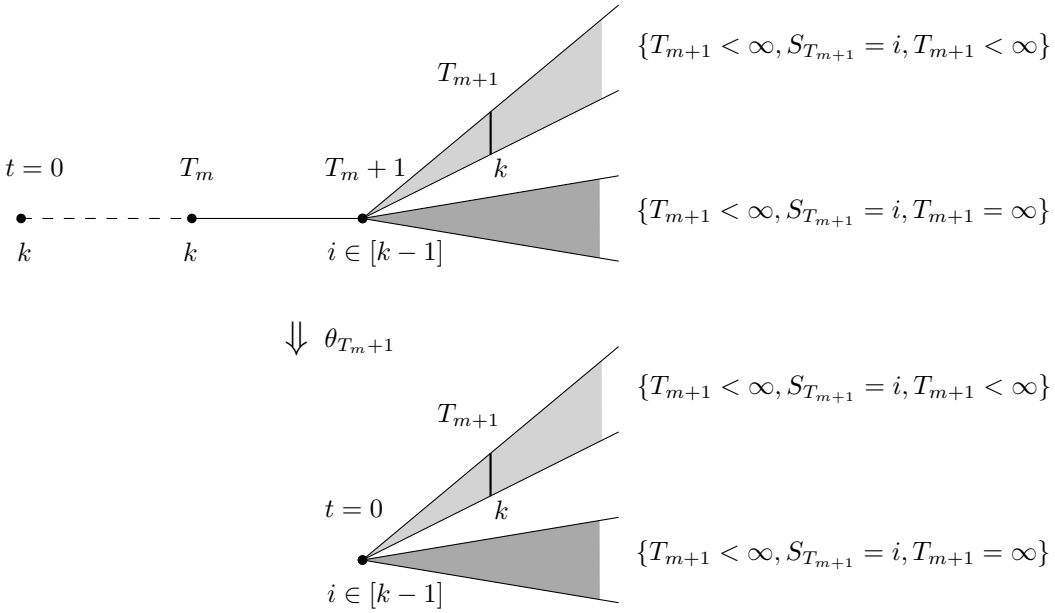

Figure 1: The upper figure : The event $\{S_{T_m+1}=i, T_m<\infty\}$ consists of the plays that at time T_m visit state k for the mth time without ever visiting the states $>k$ before, and at time T_m+1 they visit state i, where $i<k$. These plays are partitioned into two sets. The set $\{T_{m+1}<\infty, S_{T_m+1}=i, T_m<\infty\}$ of plays that will visit k for the $(m+1)$th time and the set $\{T_{m+1}=\infty, S_{T_m+1}=i, T_m<\infty\}$ of the plays for which the mth visit in k was the last one. The lower figure : The shift mapping θ_{T_m+1} "forgets" all the history prior to the time T_m+1.

$$\mathbf{E}_k^{\sigma_\star,\tau}(\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},\eta,r_{k+1},\ldots,r_n)}\circ\theta_{T_m+1} \mid S_{T_m+1}=i, T_m<\infty) \geqslant \xi_i-\delta, \quad \text{for all } i<k, \tag{26}$$

where θ_{T_m+1} is the shift mapping that deletes all history prior to the time T_m+1.

Using the fact that for all events A and B and each integrable mapping f we have $\mathbf{E}(f\mid A,B)\cdot P(A)=\mathbf{E}(f\cdot\mathbb{1}_{\{A\}}\mid B)$ we can rewrite (26) in the following form

$$\mathbf{E}_k^{\sigma_\star,\tau}(\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},\eta,r_{k+1},\ldots,r_n)}\circ\theta_{T_m+1}\cdot\mathbb{1}_{\{S_{T_m+1}=i\}} \mid T_m<\infty) \geqslant (\xi_i-\delta)\cdot\mathbf{P}_k^{\sigma_\star,\tau}(S_{T_m+1}=i \mid T_m<\infty), \quad \text{for } i<k. \tag{27}$$

We shall prove that for $i<k$,

$$\mathbf{E}_k^{\sigma_\star,\tau}(\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},\eta,r_{k+1},\ldots,r_n)}\circ\theta_{T_m+1}\cdot\mathbb{1}_{\{S_{T_m+1}=i\}} \mid T_m<\infty) = \eta\cdot\mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1}<\infty, S_{T_m+1}=i \mid T_m<\infty)+\mathbf{E}_k^{\sigma_\star,\tau}(\varphi_r^{[k]}\cdot\mathbb{1}_{\{T_{m+1}=\infty\}}\cdot\mathbb{1}_{\{S_{T_m+1}=i\}} \mid T_m<\infty). \tag{28}$$

Indeed the left-hand side of (28) is the sum of

$$\mathbf{E}_k^{\sigma_\star,\tau}(\varphi^{[k-1]}_{(r_1,\dots,r_{k-1},\eta,r_{k+1},\dots,r_n)} \circ \theta_{T_m+1} \cdot \mathbb{1}_{\{S_{T_m+1}=i\}} \cdot \mathbb{1}_{\{T_{m+1}=\infty\}} \mid T_m < \infty) \tag{29}$$

and

$$\mathbf{E}_k^{\sigma_\star,\tau}(\varphi^{[k-1]}_{(r_1,\dots,r_{k-1},\eta,r_{k+1},\dots,r_n)} \circ \theta_{T_m+1} \cdot \mathbb{1}_{\{S_{T_m+1}=i\}} \cdot \mathbb{1}_{\{T_{m+1}<\infty\}} \mid T_m < \infty). \tag{30}$$

Consider first (30). For plays h belonging to the event $\{T_{m+1} < \infty, S_{T_m+1} = i\}, i < k$, the shift θ_{T_m+1} removes all prefix history up to the time $T_m + 1$, see Figure 1. Since $T_{m+1} < \infty$ in the remaining suffix play $\theta_{T_m+1}(h)$ all visited states up to the next visit to k are $< k$. But for the plays that visit k at some moment and for which all states prior to this first visit to k are $< k$ the payoff $\varphi^{[k-1]}_{(r_1,\dots,r_{k-1},\eta,r_{k+1},\dots,r_n)}$ is constant and equal to the reward η associated with k. Thus (30) is equal to

$$\eta \cdot \mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} < \infty, S_{T_m+1} = i \mid T_m < \infty).$$

Let us examine now (29). The plays h belonging to the event $\{S_{T_m+1} = i, T_{m+1} = \infty, T_m < \infty\}$ have the following properties:

- at time T_m they visit k and all states visited prior to T_m are $\leqslant k$,
- at time $T_m + 1$, just after the mth visit to k, they visit the state i,
- since $T_{m+1} = \infty$ the suffix play $\theta_{T_m+1}(h)$ does not contain any occurrence of k (k is never visited for the $(m + 1)$th time).

These properties assure that for such plays $\varphi_r^{[k]}(h) = \varphi_r^{[k]}(\theta_{T_m+1}(h))$. However, $\theta_{T_m+1}(h)$ has no occurrence of k, which implies for the resulting payoff it is irrelevant if k is stopping or not and what is the reward of k. Thus $\varphi_r^{[k]}(\theta_{T_m+1}(h)) = \varphi^{[k-1]}_{(r_1,\dots,r_{k-1},\eta,r_{k+1},\dots,r_n)}(\theta_{T_m+1}(h))$. This terminates the proof that (29) is equal to

$$\mathbf{E}_k^{\sigma_\star,\tau}(\varphi_r^{[k]} \cdot \mathbb{1}_{\{T_{m+1}=\infty\}} \cdot \mathbb{1}_{\{S_{T_m+1}=i\}} \mid T_m < \infty).$$

This concludes also the proof of (28).

From (27) and (28) we obtain

$$\begin{aligned}\eta \cdot \mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} < \infty, S_{T_m+1} = i \mid T_m < \infty) + \mathbf{E}_k^{\sigma_\star,\tau}(\varphi_r^{[k]} \cdot \mathbb{1}_{\{T_{m+1}=\infty\}} \cdot \mathbb{1}_{\{S_{T_m+1}=i\}} \mid T_m < \infty) \\ \geqslant (\xi_i - \delta) \cdot \mathbf{P}_k^{\sigma_\star,\tau}(S_{T_m+1} = i \mid T_m < \infty).\end{aligned}$$

Summing both sides of this inequality for $i < k$ and rearranging the terms we obtain

$$\begin{aligned}\sum_{i<k} \xi_i \cdot \mathbf{P}_k^{\sigma_\star,\tau}(S_{T_m+1} = i \mid T_m < \infty) &\leqslant \eta \cdot \mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} < \infty, S_{T_m+1} < k \mid T_m < \infty) \\ &\quad + \mathbf{E}_k^{\sigma_\star,\tau}(\varphi_r^{[k]} \cdot \mathbb{1}_{\{T_{m+1}=\infty\}} \cdot \mathbb{1}_{\{S_{T_m+1}<k\}} \mid T_m < \infty) \\ &\quad + \delta \cdot \mathbf{P}_k^{\sigma_\star,\tau}(S_{T_m+1} < k \mid T_m < \infty) \\ &\leqslant \eta \cdot \mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} < \infty, S_{T_m+1} < k \mid T_m < \infty) \\ &\quad + \mathbf{E}_k^{\sigma_\star,\tau}(\varphi_r^{[k]} \cdot \mathbb{1}_{\{T_{m+1}=\infty\}} \cdot \mathbb{1}_{\{S_{T_m+1}<k\}} \mid T_m < \infty) \\ &\quad + \delta.\end{aligned}$$

The last inequality, (24) and (25) yield

$$\begin{aligned}\eta_\star \leqslant\ & \eta \cdot \mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} < \infty, S_{T_m+1} < k \mid T_m < \infty)\\ & + \mathbf{E}_k^{\sigma_\star,\tau}(\varphi_r^{[k]} \cdot \mathbb{1}_{\{T_{m+1}=\infty\}} \cdot \mathbb{1}_{\{S_{T_m+1}<k\}} \mid T_m < \infty)\\ & + \delta\\ & + \eta \cdot \mathbf{P}_k^{\sigma_\star,\tau}(S_{T_m+1} = k \mid T_m < \infty)\\ & + \mathbf{E}_k^{\sigma_\star,\tau}(\varphi_r^{[k]} \cdot \mathbb{1}_{\{S_{T_m+1}>k\}} \mid T_m < \infty).\end{aligned} \tag{31}$$

Notice that

$$\begin{aligned}\mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} < \infty, S_{T_m+1} < k \mid T_m < \infty) + \mathbf{P}_k^{\sigma_\star,\tau}(S_{T_m+1} = k \mid T_m < \infty)\\ = \mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} < \infty \mid T_m < \infty)\end{aligned} \tag{32}$$

which allows to regroup the first and the fourth summand of right-hand side of (31). Indeed, $\{T_{m+1} < \infty, T_m < \infty\}$ is the union of three disjoint events, depending on whether the state visited at the next time moment $T_m + 1$ is $< k$, $= k$, or $> k$. But for the second of these events we have $\{T_{m+1} < \infty, T_m < \infty, S^{[k]}_{T_m+1} = k\} = \{T_m < \infty, S^{[k]}_{T_m+1} = k\}$ since $S^{[k]}_{T_m+1} = k$ implies that $T_{m+1} = T_m + 1 < \infty$.

And finally the third event $\{T_{m+1} < \infty, T_m < \infty, S^{[k]}_{T_m+1} > k\}$ is empty since $S^{[k]}_{T_m+1} > k$ means that at time T_m+1 the game hits a stopping state thus the stopping state process will never return to k, therefore $T_{m+1} = \infty$. This terminates the proof of (32).

We can regroup also the second and the last summands of (31) since

$$\begin{aligned}\mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} = \infty, S_{T_m+1} < k \mid T_m < \infty) + \mathbf{P}_k^{\sigma_\star,\tau}(S_{T_m+1} > k \mid T_m < \infty)\\ = \mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} = \infty \mid T_m < \infty)\end{aligned}$$

We obtain this again by presenting the event $\{T_{m+1} = \infty, T_m < \infty\}$ as the union of three disjoint events depending on the value of S_{T_m+1}. However, $S_{T_m+1} = k$ contradicts $T_{m+1} = \infty$ and $S_{T_m+1} > k$ implies $T_{m+1} = \infty$.

Using these observations we deduce from (31) that

$$\begin{aligned}\eta_\star \leqslant\ & \eta \cdot \mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} < \infty \mid T_m < \infty)\\ & + \mathbf{E}_k^{\sigma_\star,\tau}(\varphi_r^{[k]} \cdot \mathbb{1}_{\{T_{m+1}=\infty\}} \mid T_m < \infty)\\ & + \delta.\end{aligned} \tag{33}$$

Since $\varphi_r^{[k]} \leqslant 1$, from (33) we obtain that

$$\eta \cdot \mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} < \infty \mid T_m < \infty) + \mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} = \infty \mid T_m < \infty) \geqslant \eta_\star - \delta.$$

But $\mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} = \infty \mid T_m < \infty) + \mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} < \infty \mid T_m < \infty) = 1$ thus the last inequality yields

$$\mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1} < \infty \mid T_m < \infty) \leqslant \frac{1 + \delta - \eta_\star}{1 - \eta} < \frac{1 + (\eta_\star - \eta) - \eta_\star}{1 - \eta} = 1.$$

Therefore

$$\begin{aligned}\mathbf{P}_k^{\sigma_\star,\tau}(\forall m, T_m<\infty) &= \lim_{m\to\infty}\mathbf{P}_k^{\sigma_\star,\tau}(\forall i\leqslant m, T_i<\infty)\\ &= \lim_{m\to\infty}\mathbf{P}_k^{\sigma_\star,\tau}(T_0<\infty)\cdot\prod_{q=0}^{m-1}\mathbf{P}_k^{\sigma_\star,\tau}(T_{q+1}<\infty\mid T_q<\infty)\\ &\leqslant \lim_{m\to\infty}\left(\frac{1-\eta_\star+\delta}{1-\eta}\right)^{m-1}\\ &= 0,\end{aligned}\tag{34}$$

i.e. if player Max uses the strategy $\sigma_\star$ then with probability 1 the state k is visited only finitely many times.

Multiplying both sides of (33) by $\mathbf{P}_k^{\sigma_\star,\tau}(T_m<\infty)$, taking into account that $0<\delta<\eta_\star-\eta$ and rearranging we get

$$\begin{aligned}\mathbf{E}_k^{\sigma_\star,\tau}(\varphi_r^{[k]}\cdot\mathbb{1}_{\{T_{m+1}=\infty\}}\cdot\mathbb{1}_{\{T_m<\infty\}}) &> \eta\cdot\mathbf{P}_k^{\sigma_\star,\tau}(T_m<\infty)\\ &\quad-\eta\cdot\mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1}<\infty,T_m<\infty)\\ &= \eta\cdot\mathbf{P}_k^{\sigma_\star,\tau}(T_{m+1}=\infty,T_m<\infty).\end{aligned}\tag{35}$$

Since the events $\{T_{m+1}=\infty,T_m<\infty\}_{m\geqslant 0}$ and $\{\forall m, T_m<\infty\}$ form a partition of the sets of plays but the last event has probability 0, summing up both sides of (35) for all $m\geqslant 1$ we obtain

$$\mathbf{E}_k^{\sigma_\star,\tau}(\varphi_r^{[k]})>\eta>w_k-\frac{\varepsilon}{2}$$

which terminates the proof of the right-hand side inequality in (**??**).

6.2 $\varepsilon/2$-optimal strategy $\tau_\star$ for player Min when $w_k\geqslant r_k$ and k is the starting state.

We assume that $w_k\geqslant r_k$ and $\varepsilon>0$. The aim of this section is to construct a strategy $\tau_\star$ for player Min such that

$$\mathbf{E}_k^{\sigma,\tau_\star}(\varphi_r^{[k]})\leqslant w_k+\varepsilon/2\tag{36}$$

for each strategy σ of Max.

The strategy $\tau_\star$ of player Min is constructed in the following way.

(i) If the current state is k then player Min selects actions with probability given by his optimal strategy in the one-day game $\boldsymbol{M}_k(w_1,\ldots,w_{k-1},w_k,r_{k+1},\ldots,r_n)$. Thus the strategy of player Min at k is "locally memoryless", the probability used to select actions to execute at k does not depend on the previous history.

(ii) During all stages j such that $T_m<j<T_{m+1}$ (between the mth and $(m+1)$th visit to state k) player Min plays using his $\varepsilon_m:=\varepsilon/2^{m+1}$-optimal strategy in the $\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},w_k,r_{k+1},\ldots,r_n)}$-game[6]. In general the strategy played by Min between two visits to state k is not memoryless because ε_m changes at each visit to k.

[6]This strategy exists by the induction hypothesis.

When player Min applies this strategy during all stages j, $T_m < j < T_{m+1}$, in the $\varphi_r^{[k]}$-game then we assume tacitly that starting from stage $T_m + 1$ player Min "forgets" all history preceding this stage and he plays this strategy as if the game started afresh at stage $T_m + 1$.

From the optimality of $\tau_\star$ in the one-day game $\boldsymbol{M}_k(w_1, \ldots, w_{k-1}, w_k, r_{k+1}, \ldots, r_n)$ we obtain

$$\begin{aligned}
&\sum_{j<k} w_j \cdot \mathbf{P}_k^{\sigma,\tau_\star}(S^{[k]}_{T_m+1} = j | T_m < \infty) \\
&+ w_k \cdot \mathbf{P}_k^{\sigma,\tau_\star}(S^{[k]}_{T_m+1} = k | T_m < \infty) \\
&+ \sum_{j>k} r_j \cdot \mathbf{P}_k^{\sigma,\tau_\star}(S^{[k]}_{T_m+1} = j | T_m < \infty) \\
&\leqslant w_k.
\end{aligned} \tag{37}$$

Indeed, at the time T_m the current visited state is k and player Min selects actions according to his optimal strategy in the one-day game $\boldsymbol{M}_k(w_1, \ldots, w_{k-1}, w_k, r_{k+1}, \ldots, r_n)$ and, by (16), the left-hand side of (37) gives the payoff in this one-day game while the right-hand side is the value of this game. Since he plays optimally the payoff cannot be greater than the value.

Let us consider the event

$$\{T_m < \infty, S_{T_m+1} = i\}, \quad \text{where } i < k. \tag{38}$$

This event, presented on the upper side of Figure 1, consists of plays h satisfying the following conditions:

(i) h visits k at least m times and prior to the m-th visit to k (which takes place at time T_m) the stopping states $\{k+1, \ldots, n\}$ were not visited, i.e. $S_t \in [k]$ for all $t < T_m$,

(ii) at time T_m the game moves from k to i, i.e. $S_{T_m+1} = i$.

The definition of $\tau_\star$ says that starting from time T_m+1, if the current state S_{T_m+1} is $< k$ and until the next visit to state k, player Min plays according to $\varepsilon/2^{m+1}$-optimal strategy in the $\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},w_k,r_{k+1},\ldots,r_n)}$-game. By (15), the value of the $\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},w_k,r_{k+1},\ldots,r_n)}$-game starting at state $i \in [k-1]$ is w_i.

Thus if we consider the game that, in some sense, restarts afresh at state i at time $T_m + 1$ and we apply to such residual game the payoff $\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},w_k,r_{k+1},\ldots,r_n)}$ and we assume that player Min plays $\tau_\star$ then the expected payoff will not be greater than $w_i + \varepsilon/2^{m+1}$, i.e.

$$\mathbf{E}_k^{\sigma,\tau_\star}(\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},w_k,r_{k+1},\ldots,r_n)} \circ \theta_{T_m+1} \mid S_{T_m+1} = i, T_m < \infty) \leqslant w_i + \varepsilon/2^{m+1}. \tag{39}$$

where $f \circ g$ denotes the composition of mapping f and g.

Now let us note that (37) closely resembles (24) while (39) resembles (26). What is different but symmetric is that the first two formulas concern strategies $(\sigma_\star, \tau)$ and the last two $(\sigma, \tau_\star)$. Moreover, the inequalities are reversed. The following table resumes the correspondence between constants appearing in the formulas:

Eq. (24), (26)	Eq. (37), (39)
η	w_k
$\eta_\star$	w_k
ξ_i	w_i
δ	$-\varepsilon_m$

Thus exactly in the same way as we deduced (33) from (26) and (24) we can deduce from (37) and (39) the following formula analogous to (33) (just reverse the inequality and replace the constants as indicated above):

$$\begin{aligned} & w_k \cdot \mathbf{P}_k^{\sigma,\tau_\star}(T_{m+1} < \infty \mid T_m < \infty) \\ +& \mathbf{E}_k^{\sigma,\tau_\star}(\varphi_r^{[k]} \cdot \mathbb{1}_{\{T_{m+1}=\infty\}} \mid T_m < \infty) \\ -& \varepsilon_m \leqslant w_k. \end{aligned}$$

Rearranging the terms and multiplying by $\mathbf{P}_k^{\sigma,\tau_\star}(T_m < \infty)$ we obtain from this inequality that

$$\begin{aligned} \mathbf{E}_k^{\sigma,\tau_\star}(\varphi_r^{[k]} \cdot \mathbb{1}_{\{T_{m+1}=\infty\}} \cdot \mathbb{1}_{\{T_m<\infty\}}) &\leqslant w_k \cdot \mathbf{P}_k^{\sigma,\tau_\star}(T_{m+1} = \infty, T_m < \infty) + \frac{\varepsilon}{2^{m+1}} \cdot \mathbf{P}_k^{\sigma,\tau_\star}(T_m < \infty) \\ &\leqslant w_k \cdot \mathbf{P}_k^{\sigma,\tau_\star}(T_{m+1} = \infty, T_m < \infty) + \frac{\varepsilon}{2^{m+1}}. \end{aligned}$$

The events $\{T_{m+1} = \infty, T_m < \infty\}$ are pairwise disjoint and their union is equal to $\{\exists m, T_m = \infty\}$ thus summing over $m \geqslant 1$ both sides of the inequality we obtain

$$\mathbf{E}_k^{\sigma,\tau_\star}(\varphi_r^{[k]} \cdot \mathbb{1}_{\{\exists m, T_m=\infty\}}) \leqslant w_k \cdot \mathbf{P}_k^{\sigma,\tau_\star}(\exists m, T_m = \infty) + \varepsilon/2.$$

On the other hand, for all plays in $\{\forall m, T_m < \infty\}$ the state k is visited infinitely often thus $\varphi_r^{[k]}$ is equal to r_k.

Thus

$$\begin{aligned} \mathbf{E}_k^{\sigma,\tau_\star}(\varphi_r^{[k]}) &= \mathbf{E}_k^{\sigma,\tau_\star}(\varphi_r^{[k]} \cdot \mathbb{1}_{\{\exists m, T_m=\infty\}}) + \mathbf{E}_k^{\sigma,\tau_\star}(\varphi_r^{[k]} \cdot \mathbb{1}_{\{\forall m, T_m=\infty\}}) \\ &= \mathbf{E}_k^{\sigma,\tau_\star}(\varphi_r^{[k]} \cdot \mathbb{1}_{\{\exists m, T_m=\infty\}}) + r_k \cdot \mathbf{P}_k^{\sigma,\tau_\star}(\forall m, T_m < \infty) \\ &\leqslant w_k \cdot \mathbf{P}_k^{\sigma,\tau_\star}(\exists m, T_m = \infty) + r_k \cdot \mathbf{P}_k^{\sigma,\tau_\star}(\forall m, T_m < \infty) + \varepsilon/2 \\ &\leqslant w_k + \varepsilon/2. \end{aligned}$$

6.3 $\varepsilon/2$-optimal strategies for the other cases when the starting state is k

In Sections 6.1 and 6.2 we have constructed $\varepsilon/2$-optimal strategies for player Max when $w_k > r_k$ and for player Min when $w_k \geqslant r_k$ under the condition that $\mathbf{Fix}^{k-1}(f)(r)$ is the value vector of the $\varphi_r^{[k-1]}$-game.

But passing to the dual game, the last condition implies that $\mathbf{Fix}^{k-1}(\overline{f})(\overline{r})$ is the value vector in the dual stopping game with payoff $\varphi_{\overline{r}}^{[k-1]}$.

Therefore, proceeding exactly as in Section 6.1, we can construct a strategy $\tau^\star$ for player $\overline{\text{Max}}$ in the dual game with payoff $\varphi_{\overline{r}}^{[k]}$ such that

$$\mathbf{E}_k^{\tau^\star,\sigma}(\varphi_{\overline{r}}^{[k]}) \geqslant \overline{w}_k - \varepsilon/2 \tag{40}$$

for all strategies σ of player $\overline{\text{Min}}$ if

$$\overline{w}_k > \overline{r}_k. \tag{41}$$

By duality of games and fixed points, $\mathbf{E}_k^{\tau^\star,\sigma}(\varphi_{\overline{r}}^{[k]}) = 1 - \mathbf{E}_k^{\sigma,\tau^\star}(\varphi_r^{[k]})$, $\overline{w}_k = 1 - w_k$ and $\overline{r}_k = 1 - r_k$. Thus (40) is equivalent to $\mathbf{E}_k^{\sigma,\tau^\star}(\varphi_r^{[k]}) \leqslant w_k + \varepsilon/2$ and (41) is equivalent to $w_k < r_k$, i.e. we get a $\varepsilon/2$-optimal strategy of player Min in the $\varphi_r^{[k]}$-game if $w_k < r_k$.

In the similar way, applying the construction of Section 6.2 to the dual game and coming back to the original game we get a strategy $\sigma^\star$ for player Max such that $\mathbf{E}_k^{\sigma^\star,\tau}(\varphi_r^{[k]}) \geqslant w_k - \varepsilon/2$ if $w_k \leqslant r_k$.

6.4 ε-optimal strategies for the $\varphi_r^{[k]}$-game starting at states $< k$.

It remains to prove that

$$\mathbf{Fix}_i^k(f)(r) := F_i^{k-1}(w_k; r)$$

is the value of the $\varphi_r^{[k]}$-game starting in the state $i < k$. To this end we must construct strategies $\sigma_\sharp$ and $\tau_\sharp$ for player Max and Min, respectively, such that

$$\mathbf{E}_i^{\sigma,\tau_\sharp}(\varphi_r^{[k]}) \leqslant \mathbf{Fix}_i^k(f)(r) + \varepsilon \quad \text{and} \quad \mathbf{E}_i^{\sigma_\sharp,\tau}(\varphi_r^{[k]}) \geqslant \mathbf{Fix}_i^k(f)(r) - \varepsilon \tag{42}$$

for all strategies σ, τ. We define only the strategy $\tau_\sharp$ for player Min and prove the first equation of (42). The definition of $\sigma_\sharp$ and the proof of the right-hand side of (42) are symmetrical and are left to the reader.

Recall that T_1 was defined as the (random) time of the first visit of the stopped state process $S_t^{[k]}$ to the state k, cf. (17). Let $\tau_\star$ be the strategy of player Min defined at page 26 that satisfies (36), i.e $\tau_\star$ is an $\varepsilon/2$-optimal for player Min in the $\varphi_r^{[k]}$-game starting at the state k.

By the induction hypothesis, there exists an $\varepsilon/2$-optimal strategy α for player Min in the $\varphi_{(r_1,\dots,r_{k-1},w_k,r_{k+1},\dots,r_n)}^{[k-1]}$-game.

We define the strategy $\tau_\sharp$ for player Min by composing strategies α and $\tau_\star$ as follows:

$$\tau_\sharp(S_1, A_1, B_1, \cdots, S_m) = \begin{cases} \alpha(S_1, A_1, B_1, \cdots, S_m) & \text{if } T_1 > m, \\ \tau_\star(S_{T_1}, A_{T_1}, B_{T_1}, \cdots, S_m) & \text{if } T_1 \leqslant m. \end{cases}$$

Intuitively, $\tau_\sharp$ is the strategy such that player Min plays according to α until the first visit to k and starting from the moment of the first visit to k he switches to $\tau_\star$. Moreover, when he switches to $\tau_\star$ then he "forgets" all history prior to the moment T_1 and behaves as if the game have started afresh at k.

First we want to show that, for each strategy σ of player Max and for each state $i < k$,

$$\mathbf{E}_i^{\sigma,\tau_\sharp}(\varphi_r^{[k]} \mid T_1 < \infty) = \mathbf{E}_i^{\sigma,\tau_\sharp}(\varphi_r^{[k]} \circ \theta_{T_1} \mid T_1 < \infty) \leqslant w_k + \varepsilon/2$$

where θ_{T_1} is the shift operation, cf. Definition 12, and $w_k = \mathbf{Fix}_k^k(f)(r)$ is the value of k.

To justify the first equality let us notice that the plays with $T_1 < \infty$ do not visit the stopping states, i.e. the states $> k$, prior to T_1. Therefore the payoff $\varphi_r^{[k]}$ for such plays is not modified if we shift them by T_1.

The second inequality follows from the definition of $\tau_\sharp$. When the game hits state k at time T_1 player Min switches to strategy $\tau_\star$ and forgets the history prior to T_1. Since $\tau_\star$ is $\varepsilon/2$-optimal for player Min in the $\varphi_r^{[k]}$-game for plays starting at k, using this strategy limits the payoff to at most $w_k + \varepsilon/2$.

Now we examine the expected payoff for plays with $T_1 = \infty$. Such plays never visit k, therefore it is irrelevant for them if k is stopping or not like it is irrelevant what is the reward associated with k. Moreover, for such plays player Min plays according to strategy $\tau_\star$. For these reasons we have

$$\mathbf{E}_i^{\sigma,\tau_\sharp}(\varphi_r^{[k]} \mid T_1 = \infty) = \mathbf{E}_i^{\sigma,\tau_\star}(\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},w_k,r_{k+1},\ldots,r_n)} \mid T_1 = \infty). \tag{43}$$

From (43) we obtain

$$\begin{aligned} \mathbf{E}_i^{\sigma,\tau_\sharp}(\varphi_r^{[k]}) &= \mathbf{E}_i^{\sigma,\tau_\sharp}(\varphi_r^{[k]} \mid T_1 < \infty) \cdot \mathbf{P}_i^{\sigma,\tau_\sharp}(T_1 < \infty) \\ &\quad + \mathbf{E}_i^{\sigma,\tau_\sharp}(\varphi_r^{[k]} \mid T_1 = \infty) \cdot \mathbf{P}_i^{\sigma,\tau_\sharp}(T_1 = \infty) \\ &\leqslant (w_k + \varepsilon/2) \cdot \mathbf{P}_i^{\sigma,\tau_\sharp}(T_1 < \infty) \\ &\quad + \mathbf{E}_i^{\sigma,\tau_\star}(\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},w_k,r_{k+1},\ldots,r_n)} \mid T_1 = \infty) \cdot \mathbf{P}_i^{\sigma,\tau_\sharp}(T_1 = \infty). \end{aligned} \tag{44}$$

Since $\tau_\star$ is $\varepsilon/2$-optimal for player Min in the $\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},w_k,r_{k+1},\ldots,r_n)}$-game we have

$$\begin{aligned} F_i^{k-1}(w_k; r) + \varepsilon/2 &\geqslant \mathbf{E}_i^{\sigma,\tau_\star}(\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},w_k,r_{k+1},\ldots,r_n)}) \\ &= \mathbf{E}_i^{\sigma,\tau_\star}(\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},w_k,r_{k+1},\ldots,r_n)} \mid T_1 < \infty) \cdot \mathbf{P}_i^{\sigma,\tau_\star}(T_1 < \infty) \\ &\quad + \mathbf{E}_i^{\sigma,\tau_\star}(\varphi^{[k-1]}_{(r_1,\ldots,r_{k-1},w_k,r_{k+1},\ldots,r_n)} \mid T_1 = \infty) \cdot \mathbf{P}_i^{\sigma,\tau_\star}(T_1 = \infty). \end{aligned}$$

Notice that plays with $T_1 < \infty$ have payoff w_k in the $\varphi^{[k-1]}_{(r_1,\dots,r_{k-1},w_k,r_{k+1},\dots,r_n)}$-game because k is stopping in this game and the reward of k is equal to w_k. Hence we can rewrite (45) as

$$F_i^{k-1}(w_k;r) + \varepsilon/2 \geqslant w_k \cdot \mathbf{P}_i^{\sigma,\tau_\star}(T_1 < \infty) + \mathbf{E}_i^{\sigma,\tau_\star}(\varphi^{k-1}_{(r_1,\dots,r_{k-1},w_k,r_{k+1},\dots,r_n)} \mid T_1 = \infty) \cdot \mathbf{P}_i^{\sigma,\tau_\star}(T_1 = \infty).$$

Thus

$$\mathbf{E}_i^{\sigma,\tau_\star}(\varphi^{[k-1]}_{(r_1,\dots,r_{k-1},w_k,r_{k+1},\dots,r_n)} \mid T_1 = \infty) \cdot \mathbf{P}_i^{\sigma,\tau_\star}(T_1 = \infty) \leqslant F_i^{k-1}(w_k;r) + \varepsilon/2 - w_k \cdot \mathbf{P}_i^{\sigma,\tau_\star}(T_1 < \infty). \quad (45)$$

From (44) and (45) and since $\mathbf{P}_i^{\sigma,\tau_\sharp}(T_1 < \infty) = \mathbf{P}_i^{\sigma,\tau_\star}(T_1 < \infty)$ we get

$$\begin{aligned}
\mathbf{E}_i^{\sigma,\tau_\sharp}(\varphi_r^{[k]}) &\leqslant (w_k + \varepsilon/2) \cdot \mathbf{P}_i^{\sigma,\tau_\sharp}(T_1 < \infty) + F_i^{k-1}(w_k;r) + \varepsilon/2 - w_k \cdot \mathbf{P}_i^{\sigma,\tau_\star}(T_1 < \infty) \\
&= F_i^{k-1}(w_k;r) + \varepsilon/2 + (\varepsilon/2) \cdot \mathbf{P}_i^{\sigma,\tau_\sharp}(T_1 < \infty) \\
&\leqslant F_i^{k-1}(w_k;r) + \varepsilon \\
&= \mathbf{Fix}_i^k(f)(r) + \varepsilon
\end{aligned}$$

which terminates the proof of the ε-optimality of $\tau_\sharp$.

□

6.5 Discussion

Parity games form a special subclass of priority games where the winning regions (in the deterministic case [12]) or the values (for stochastic parity games [4]) can be expressed by means of μ-calculus formulas. The μ-calculus is a fixed point calculus over a complete lattice using the greatest and the least fixed points. From this point of view Theorem 11 is just an extension of the known result of de Alfaro and Majumdar [4] to a wider framework of priority games. However, there is one ingredient of Theorem 11 that seems to be new.

It is notoriously difficult to grasp the meaning of a μ-calculus formula alternating several greatest and least fixed point.

Theorem 11 provides a natural interpretation in the term of games of a formula where only some fixed points are applied and other variables remain free.

Let

$$(\xi_1,\dots,\xi_k) = (\mathbf{Fix}_1^{k-1}(r),\dots,\mathbf{Fix}_{k-1}^{k-1}(r))$$

be the values of the states $1,\dots,k-1$ in the $\varphi_r^{[k-1]}$-game. This game differs from the original priority game with the payoff φ_r in that the states $k, k+1, k+2, \dots, n$ are stopping.

Now when we add another fixed point to obtain

$$(\xi'_1, \ldots, \xi'_{k-1}, \xi'_k) = (\mathbf{Fix}^k_1(r), \ldots, \mathbf{Fix}^k_{k-1}(r), \mathbf{Fix}^k_k(r))$$

then this corresponds to the operation that transforms the state k from stopping in the $\varphi_r^{[k-1]}$-game into a non-stopping state in the $\varphi_r^{[k]}$-game.

In the game SKIRMISH, adapted by de Alfaro and Henzinger [3] from [6] (Figure 2) the players do not have optimal strategies and for one of the players a ε-optimal strategy cannot be memoryless. SKIRMISH has three states $\mathbf{S} = \{1, 2, 3\}$: state 1 is absorbing, state 3 has only one outgoing transition moving to state 2 independently of the actions played at 3, in state 2 each player has two actions: $\mathbf{A}(2) = \{\text{run}, \text{hide}\}$, $\mathbf{B}(2) = \{\text{fire}, \text{wait}\}$.

The reward vector is $r = (0, 0, 1)$. Thus player Max obtains payoff 1 if and only if the play visits infinitely often the state 3. Moreover since 1 is absorbing this state should never be visited.

The transitions are deterministic and given by $p(1|2, \text{run}, \text{fire}) = p(2|2, \text{hide}, \text{wait}) = p(3|2, \text{hide}, \text{fire}) = p(3|2, \text{run}, \text{wait}) = 1$.

Assume that the game starts at state 2.

If player Max plays a memoryless strategy σ_ε such that $\sigma_\varepsilon(2)(\text{hide}) = 1 - \varepsilon$ and $\sigma_\varepsilon(2)(\text{run}) = \varepsilon$, $\varepsilon > 0$, then player Min playing always action run at 2 will ensure that with probability 1 the game hits state 1 giving the payoff 0.

If player Max always plays action hide at 2 then player Min can play always action wait at 2 and the game will remain forever in 2 giving again payoff 0.

Nevertheless it turns out that the value of states 2 and 3 is 1.

The ε-optimal strategy of player Max decreases the probability to play action run after each visit to state 3 and is defined as follows:

$$\sigma(h)(\text{run}) = 1 - (1 - \varepsilon)^{1/2^{m+1}} \quad \text{and} \quad \sigma(h)(\text{hide}) = (1 - \varepsilon)^{1/2^{m+1}},$$

where h is a history ending at 2 and m is the number of occurrences of state 3 in h. Then for each strategy of Min the probability to visit state 3 infinitely often is at least $\prod_{m=0}^{\infty}(1 - \varepsilon)^{1/2^{m+1}} = 1 - \varepsilon$.

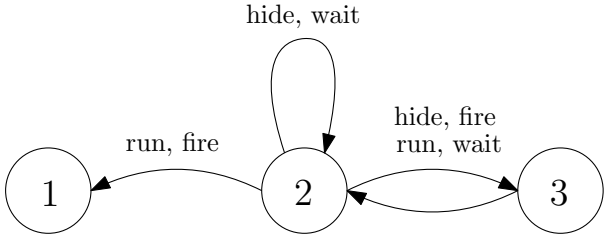

Figure 2: Game SKIRMISH [3].

References

[1] A. Arnold and D. Niwiński. *Rudiments of μ-calculus*, volume 146 of *Studies in Logic and the Foundations of Mathematics*. Elsevier, 2001.

[2] M. G. Crandall and L. Tartar. Some relations between nonexpansive and order preserving mappings. *Proceedings of the American Mathematical Society*, 78(3):385–390, 1980.

[3] L. de Alfaro and T.A. Henzinger. Concurrent omega-regular games. In *Proc. 15th Symp. Logic in Comp. Sci., LICS 2000*, pages 142–154. IEEE Computer Society Press, 2000.

[4] L. de Alfaro and R. Majumdar. Quantitative solution to omega-regular games. *Journal of Computer and System Sciences*, 68:374–397, 2004.

[5] B. Karelović and W. Zielonka. Nearest fixed points and concurrent priority games. In *International Symposium on Fundamentals of Computation Theory*, volume 9210 of *LNCS*, pages 381–393. Springer, 2015.

[6] P. R. Kumar and T. H. Shiau. Existence of value and randomized strategies in zero-sum discrete-time stochastic dynamic games. *SIAM Journal on Control and Optimization*, 19(5):617–634, 1981.

[7] A. Maitra and W. Sudderth. Stochastic games with Borel payoffs. In A. Neyman and S. Sorin, editors, *Stochastic Games and Applications*, volume 570 of *NATO Science Series C, Mathematical and Physical Sciences*, pages 367–373. Kluwer Academic Publishers, 2004.

[8] A. P. Maitra and W. D. Sudderth. *Discrete Gambling and Stochastic Games*. Springer, 1996.

[9] D.A. Martin. The determinacy of Blackwell games. *Journal of Symbolic Logic*, 63(4):1565–1581, 1998.

[10] L. S. Shapley. Stochastic games. *Proceedings Nat. Acad. of Science USA*, 39:1095–1100, 1953.

[11] A. Tarski. A lattice-theoretical fixpoint theoem and its aplications. *Pacific J. Math.*, 5:285–309, 1955.

[12] I. Walukiewicz. Monadic second-order logic on tree-like structures. *Theoretical Computer Science*, 275(1-2):311–346, 2002.