Parrondo's paradox: a priori and adaptive strategies

Ehrhard Behrends

Abstract. Several hundred pages in the WWW and a number of articles in (mostly) popular journals deal with Parrondo's paradox: it states that there are losing gambling games which, when being combined stochastically or in a suitable deterministic way, give rise to winning games.

Here we investigate the probabilistic background, methods from graph theory, (elementary) functional analysis and stochastic control theory come into play.

We show how the properties of the equilibrium distributions of the Markov chains under consideration give rise to the paradoxical behaviour, and we provide methods how to find the best a priori and adaptive strategies.

AMS-classification: 60J10, 60J20;

keywords: Parrondo, paradox, Markov chain, stochastic control theory.

1. Introduction

Let $\mathbf{P} = (p_{i,j})_{i,j=1,\ldots,s-1}$ be a stochastic $s \times s$-matrix and $\mathbf{x} = (x_i)_{i=0,\ldots,s-1}$ a real vector. $\mathbf{P}$ and $\mathbf{x}$ give rise to the following "game"$(\mathbf{P}, \mathbf{x})$:

> Consider the random walk on $\{0, \ldots, s-1\}$ which starts at 0 and which is driven by $\mathbf{P}$. At any stage of the game the player receives the "reward" x_i when being in state i.

Clearly the expected gain after m rounds of the game is the first component of the vector

$$(Id + \mathbf{P} + \mathbf{P}^2 + \cdots + \mathbf{P}^{m-1})\mathbf{x}.$$

If $\mathbf{P}$ is aperiodic and irreducible the $\mathbf{P}^\mu$ tend fast to a matrix $\mathbf{W}$ with identical rows $\pi = (\pi_0, \ldots, \pi_{s-1})$, the *equilibrium distribution of* $\mathbf{P}$, and thus the expected gain is – in the case of "large" m – roughly $m\langle \pi, \mathbf{x}\rangle$. (For the sake of readability we suppress the difference between row vectors and column vectors.) Thus it is natural to speak of a *winning* – resp. *losing* resp. *fair* – *game* if $\langle \pi, \mathbf{x}\rangle$ is positive (resp. negative resp. zero).

Now we consider r such games, $(\mathbf{P}_1, \mathbf{x}_1), \ldots, (\mathbf{P}_r, \mathbf{x}_r)$. Rather to deal with a fixed $(\mathbf{P}_\rho, \mathbf{x}_\rho)$ the player is now allowed to switch. More precisely, we consider the following *three variants*:

The a priori strategy: The number m of rounds of the game is fixed in advance, the player chooses probability measures $\mathbf{p}_1, \ldots, \mathbf{p}_m$ on $\{1, \ldots, r\}$. Then indices $\rho_1, \ldots, \rho_m$ are selected in $\{1, \ldots, r\}$, ρ_1 according to the law $\mathbf{p}_1$, ρ_2 according to $\mathbf{p}_2$ etc. (the choices are assumed to be independent).

The game is now played as in the $(\mathbf{P}, \mathbf{x})$-case, but in the μ'th round one "plays" with $(\mathbf{P}_{\rho_\mu}, \mathbf{x}_{\rho_\mu})$: the first reward is the first component of $\mathbf{x}_{\rho_1}$, and for the first step – from 0 to some i – one uses the transition probabilities from $\mathbf{P}_{\rho_1}$; the second reward is the i'th component of $\mathbf{x}_{\rho_2}$, the second step is according to $\mathbf{P}_{\rho_2}$, etc.

The constant a priori strategy: The same as before, but now all $\mathbf{p}_\mu$ are identical with a probability vector $\mathbf{p} = (p_1, \ldots, p_r)$.

The adaptive strategy: Here the player may choose any of the $(\mathbf{P}_\rho, \mathbf{x}_\rho)$, say $(\mathbf{P}_{\rho_1}, \mathbf{x}_{\rho_1})$, to start with: the first component of $\mathbf{x}_{\rho_1}$ will be his or her first reward, and the first step from 0 to a state i will be according to $\mathbf{P}_{\rho_1}$. Now a $(\mathbf{P}_{\rho_2}, \mathbf{x}_{\rho_2})$ may be chosen, and so on.

The Spanish physicist *Parrondo* has observed a strange phenomenon in connection with such games which is now called the *Parrondo paradox:* he has provided fair games in the case $s = 3$ and $r = 2$ such that there exist a priori strategies – even constant ones – where the expected gain tends to $+\infty$ with $m \to +\infty$ (for a survey of Parrondo's publications on his paradox cf. [10]).

In fact, in Parrondo's paper the approach is slightly different. He considers random walks on the integers driven by certain infinite stochastic matrices where the stepsize does depend only on the position modulo 3. If only expected total gains are of interest, one may instead pass to games as considered above, in Parrondo's particular case one deals with

$$\mathbf{P}_1 = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{pmatrix}, \ \mathbf{x}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{P}_2 = \begin{pmatrix} 0 & 0.1 & 0.9 \\ 0.25 & 0 & 0.75 \\ 0.75 & 0.25 & 0 \end{pmatrix}, \ \mathbf{x}_2 = \begin{pmatrix} -0.8 \\ 0.5 \\ 0.5 \end{pmatrix}.$$

Also, Parrondo treats a small perturbation of these games. In this way he can produce the even more spectacular paradox that there are losing games where a suitable a priori strategy leads to a winning game.

This paradox has given rise to a number of articles in mostly popular journals (see [4] and [5] and the literature cited there). Also, when searching for the key words "Parrondo" and "paradox" in the *WorldWideWeb*, one finds several hundred web pages. The fascination stems from the fact that the paradox seemingly "explains" certain aspects of the experience of everyone: in the game of chess, pieces can be sacrificed in order to win the overall game; a politician can win the votes although he was involved in a sex scandal and his party performs rather weakly, ...

Needless to say that these examples to "apply" the paradox – one finds them over and over again – have not the slightest justification in any provable result, be it by Parrondo or by someone else.

There are some natural problems associated with the paradox and with a priori and adaptive strategies which will be systematically studied in the sequel. Our paper is organized as follows:

Section 2 contains *preliminaries*. We introduce some definitions, prove results concerning the complexity of finding optimal walks in weighted graphs and study a certain finite dimensional normed quotient spaces. This space will be used later to apply Banach's fixed point theorem.

Section 3 is devoted to a *systematic study of a priori strategies*. It will easily turn out that suitable bang-bang strategies are optimal: all $\mathbf{p}_\mu$ can be chosen to be Dirac measures (propositon 3.2). There are, however, r^m such strategies, and one needs further results in order to find solutions in polynomial time which are at least "nearly optimal". The problem will be reduced to the study of finite graphs with "not too many" vertices, we show that nearly optimal strategies can be found in polynomial time (theorem 3.4).

We continue in *section 4* with the study of *constant a priori strategies*. Our main result, theorem 4.2, will relate properties of equilibrium distributions associated with the $\mathbf{P}_\rho$ with the problem whether or not a paradoxical behaviour is to be expected. In *section 5* we turn to *adaptive strategies*. The problems we are faced with are similar to those considered in stochastic control theory. One can use backward induction to find an optimal strategy explicitly (proposition 5.1), and the optimal strategies converge fast to a constant strategy. (theorem 5.3). Our short proof is self-contained, it uses Banach's fixed point theorem in the quotient space introduced in section 2.

For other approaches to explain the paradox and for its applications to understand certain phenomena in the physical and biological sciences we refer the reader to [4], [5], [6], [7], and [12]. Finally we note that there is a similar paradox in the theory of Markov processes in continuous time: in [11] it is shown that the random mixture of positive recurrent chains might be transient.

2. Some fundamental definitions and auxiliary results

A. Parrondo games

In the introduction we have described games which are defined by a stochastic $s \times s$-matrix $\mathbf{P}$ and a gain vector. As already noted this approach is slightly different form Parrondo's, a proper generalization of his examples could is given by the following definition.

Let $\tilde{\mathbf{P}} = (\tilde{p}_{i,j})_{i,j\in\mathbb{Z}}$ be an *infinite* stochastic matrix such that there is an $s \in \mathbb{N}$ for which $\tilde{p}_{i+s,j+s} = \tilde{p}_{i,j}$ for all i, j; we will call $\tilde{\mathbf{P}}$ *s-periodic* in this case. One can then define a stochastic $s \times s$-matrix $\mathbf{P} = (p_{i,j})$ by

$$p_{i,j} = \sum_{j'=j \bmod s} \tilde{p}_{i,j'}, \quad i, j = 0, \ldots, s-1,$$

and one can consider – provided these expressions exist – the numbers

$$x_i := \left(\sum_j j\tilde{p}_{i,j}\right) - i; \quad i = 0, \ldots, s-1.$$

Then, as is easy to see, the expected gain after m rounds played with the game defined by $\mathbf{P}$ and $\mathbf{x} := (x_i)$ is precisely the expectation of the position of the random walk on $\mathbb{Z}$ defined by $\tilde{\mathbf{P}}$. We note that every game $(\mathbf{P}, \mathbf{x})$ arises in this way by a suitable $\tilde{\mathbf{P}}$ (which is not unique).

By a *Parrondo game* we here mean a collection $(\mathbf{P}_1, \mathbf{x}_1), \ldots, (\mathbf{P}_r, \mathbf{x}_r)$ of r stochastic $s \times s$-matrices and gain vectors as in the introduction. We will assume that the $\mathbf{P}_1, \ldots, \mathbf{P}_r$ satisfy the following condition: there is an $L < 1$ such that the l^1-distance of the rows of the $\mathbf{P}_\rho$ is bounded by $2L$:

$$\sum_{j=0}^{s-1} |p^{\rho_1}_{i_1,j} - p^{\rho_2}_{i_2,j}| \leq 2L$$

for arbitrary $1 \leq \rho_1, \rho_2 \leq r$ and $0 \leq i_1, i_2 \leq s-1$ (the numbers $p^{\rho}_{i,j}$ are the entries of $\mathbf{P}_\rho$). The matrices are called *uniformly L-contractive* in this case.

This condition implies that the map $(y_0, \ldots, y_{s-1}) \mapsto (y_0, \ldots, y_{s-1})\mathbf{P}$ is a contraction with Lipschitz constant L on $\{(y_0, \ldots, y_{s-1}) \mid y_i \geq 0, \sum y_i = 1\}$ for every $\mathbf{P}$ in the convex hull of the $\mathbf{P}_\rho$ (see chapter 7 in [1]), in particular there exists a unique equilibrium distribution $\pi_{\mathbf{P}}$ for such $\mathbf{P}$. We note that the $\mathbf{P}_\rho$ are uniformly L-contractive for some $L < 1$ if in each row of any $\mathbf{P}_\rho$ there are more than $s/2$ strictly positive entries[1].

The proof of the following assertions are straightforward, they are left to the reader:

Lemma 2.1 *Let $(\mathbf{P}_1, \mathbf{x}_1), \ldots, (\mathbf{P}_r, \mathbf{x}_r)$ be a Parrondo game.*

(i) *If λ_ρ are nonnegative numbers for $\rho = 1, \ldots, r$ such that $\sum \lambda_\rho = 1$ then the constant a priori strategy associated with $\mathbf{p} := (\lambda_1, \ldots, \lambda_r)$ leads to the same expected gains as a play with the game defined by $\mathbf{P} := \sum_\rho \lambda_\rho \mathbf{P}_\rho$ and $\mathbf{x} := \sum \lambda_\rho \mathbf{x}_\rho$.*

(ii) *Let $\mathbf{P}_{\rho_1}, \ldots, \mathbf{P}_{\rho_m}$ be any matrices from $\{\mathbf{P}_1, \ldots, \mathbf{P}_r\}$, we consider the game which starts at zero and for which the first resp. second, ..., resp. m'th round is played with game $(\mathbf{P}_{\rho_1}, \mathbf{x}_{\rho_1})$ resp. $(\mathbf{P}_{\rho_2}, \mathbf{x}_{\rho_2})$, ..., resp. $(\mathbf{P}_{\rho_m}, \mathbf{x}_{\rho_m})$.*
Then the expected gain after m rounds is the first component of the vector

$$\mathbf{x}_{\rho_1} + \mathbf{P}_{\rho_1}\mathbf{x}_{\rho_2} + \mathbf{P}_{\rho_1}\mathbf{P}_{\rho_2}\mathbf{x}_{\rho_3} + \cdots + \mathbf{P}_{\rho_1} \cdots \mathbf{P}_{\rho_{m-1}}\mathbf{x}_{\rho_m}.$$

B. Nearly optimal walks in weighted graphs

Let us call an a priori strategy *deterministic* if all $\mathbf{p}_\mu$ are Dirac measures. We will prove later that the optimal possible expected gain among all a priori strategies is achieved at a deterministic one. There are, however r^m such strategies, and it is hopeless to find the best by pure inspection.

[1] This is satisfied in particular for Parrondo's matrices.

Here we prepare the results of section 3 which will provide approximate solutions in polynomial time. The idea is to transform the problem to a graph theoretical one.

Let $G = (V, E, f)$ be a finite connected directed graph with $n := |V|$ vertices and a weight function $f : E \to \mathbb{R}$. (Self-loops and parallel edges are admissible.) We fix an integer m, and we are interested in finding *a walk of length m such that the total weight is as large as possible.* (Recall that a *walk of length m* is a sequence $v_0, \ldots, v_m$ of vertices such that all $(v_{\mu-1}, v_\mu)$ lie in E, the *total weight* of such a walk is the number $\sum_\mu f(v_{\mu-1}, v_\mu)$.)

Let r be the smallest integer which bounds the number of edges leaving a vertex of G. One has to check about $n \cdot r^m$ possibilities to find a walk of length m with maximal weigth, even for moderate m this is a hopeless task. The situation is more favourable if one is interested in *approximate solutions* only. We need some further notation:

Definition 2.2 Let G be as above.

(i) A *cycle* γ of length l in G is a walk of length l for which the starting point and the endpoint coincide. The *average weight* $m(\gamma)$ of γ is the total weight of γ, devided by l.

(ii) A cycle γ is said to be *maximal* if $m(\gamma)$ is as large as possible.

(iii) Let η be a positive number. A walk of length m which starts at s will be called *η-optimal* if it is optimal up to a possible error of ηm (i.e., its total weight plus ηm dominates the total weight of all walks of length m).

If one doesn't insist in the best but is satisfied with an η-optimal walk it suffices to find optimal cycles, and these can be found in polynomial time:

Proposition 2.3 *With G as above the following hold:*

(i) *An optimal cycle can be found in $O(r \cdot n^2)$ steps.*

(ii) *For any $\eta > 0$ one can find an η-optimal walk of length m in $O(r \cdot n^2)$ steps. (More precisely there are numbers K and M such that $K \cdot r \cdot n^2$ steps suffice to find an η-optimal walk of length m provided that $m \geq M/\eta$.)*

Proof. (i) This is essentially due to Karp, cf. [9], it is only necessary to make a slight modification.

Let's repeat the essential steps of Karp's argument, we will use the notation of [9]. First one reduces the claim to the case of connected graphs where the maximal average weight is zero. (Karp considers cycles with *minimal* average weight, but this should not cause much confusion.) The main idea is to fix a vertex s and to define, for $k = 0, \ldots, n$ and any vertex v, the number $F_k(v)$ as the maximal weight of a walk of length k from s to v. One puts $F_k(v) := -\infty$ if no such walks exist[2]. It then can be shown that

$$\lambda^*(= 0) = \max_v \min_{0 \leq k < n} \frac{F_n(v) - F_k(v)}{n - k},$$

[2]Thus $F_0(v) = -\infty$ for all $v \neq s$; by definition, one puts $F_0(s) := 0$.

in particular the maximal average weight is known as soon as the $F_k(v)$ are determined[3].

It remains to show that this is possible with $O(r \cdot n^2)$ calculations. In Karp's paper this is done for the case of at most r *inbound* edges for any vertex whereas we need to consider graphs with at most r *outbound* ones.

We argue as follows. First we put $a_v^k := -\infty$ for $k = 1, \ldots, n$ and all vertices v.

Step 1: Consider, one after the other, all edges (s, v) which leave the distinguished vertex s and put $a_v^1 := \max\{a_v^1, f(s, v)\}$; it might happen that a_v^1 is updated several times during this procedure. After at most $O(r)$ steps, the a_v^1 coincide with the $F_1(v)$.

Step 2: For any v and any of the (at most r) v' such that there is an edge (v, v') put $a_{v'}^2 := \max\{a_{v'}^2, a_v^1 + f(v, v')\}$. After these $O(n \cdot r)$ calculations one has determined $F_2(v) = a_v^2$.

And so on.

If we continue this way for n steps all $F_k(v)$, $k = 1, \ldots, n$, are determined, and the total number of calculations is $O(r \cdot n^2)$ as claimed. Also, keeping track of the edges for which the optimal values are attained, we can construct from our data best possible walks of length k from s to v.

A cycle with maximal average weight can be found as a byproduct: choose a v such that $F_n(v)$ dominates all $(n-k)\lambda^* + F_k(v)$ and find an optimal walk of length n from s to v; this walk must contain a cycle which necessarily has maximal average weight (see [9]).

(ii) Let R be a positive constant such that the total weight of any walk which contains no cycles lies between $-R$ and R. (Such an R can easily be calculated from the $F_k(v)$.)

Put $W_m :=$ "the maximal weigth of a walk of length m" and consider any walk γ which achieves W_m. If γ contains a cycle of length k, then, by the definition of λ^*, we have $W_m \leq W_{m-k} + k\lambda^*$. If the optimal walk of length $m - k$ also has a cycle – of length k', say – we can continue with the estimate $\leq W_{m-k-k'} + (k + k')\lambda^*$. In this way we may get rid of cycles until we arrive at a $k_1 < n$ with

$$W_m \leq W_{k_1} + (m - k_1)\lambda^* \leq R + m\lambda^*.$$

Now let γ_0 be a cycle with optimal average weight, one had to perform $\leq K \cdot r \cdot n^2$ calculations to identify it. We consider a walk which leads from s to γ_0 as fast as possible, and then walks around γ_0 until m steps are completed. Suppose that there are d complete turns and $k_2 (< n)$ additional steps. The total weight of our walk is at least $-2R + dk\lambda^*$, where k denotes the length of γ_0 and $-2R$ bounds the weight of the initial and final segments of our walk of length k_1 and k_2. It then follows immediately from $dk = m - k_1 - k_2 > m - 2n$

[3] It should be noted that in this formula one might have expressions of the form $(-\infty) - (-\infty)$. One has to define them as $+\infty$ in order to arrive at a true result.

that $M := 3R + 2\lambda^*$ has the claimed properties. $\square$

C. A useful normed space

Let $\mathbf{1}$ be the vector in $\mathbb{R}^s$ for which all components are 1. Its linear span $\mathbb{R}\mathbf{1}$ in $\mathbb{R}^s$ consists of the constant vectors, and the quotient space $X := \mathbb{R}^s/\mathbb{R}\mathbf{1}$ can be considered as the space of all possible "forms" of maps from $\{0,\ldots,s-1\}$ to $\mathbb{R}$: graphs which only differ by a vertical translation are identified.

Now we provide $\mathbb{R}^s$ with the maximum norm $||\cdot||_\infty$ and X with the associated quotient norm:

$$||[\mathbf{x}]|| := \min_{c\in\mathbb{R}} \max_{i=0,\ldots,s-1} |x_i - c|.$$

We will need the following properties of this space:

Lemma 2.4

(i) $||[\mathbf{x}]|| = (\max_i x_i - \min_i x_i)/2.$

(ii) *Let $\mathbf{P} = (p_{i,j})_{i,j=0,\ldots,s-1}$ be a stochastic $s\times s$-matrix, L a number such that $\sum_j |p_{i,j} - p_{i',j}| \le 2L$ for all i, i'. Then $[\mathbf{x}] \mapsto [\mathbf{P}\mathbf{x}]$ is a well-defined Lipschitz map on X with Lipschitz constant L.*

Proof. The simple proof of (i) is left to the reader.
(ii) Since $\mathbf{x} \mapsto \mathbf{P}\mathbf{x}$ maps constants to constants the map under consideration is well-defined. Now let $[\mathbf{x}]$ in the unit ball of X be given. We choose a $c \in \mathbb{R}$ such that $\mathbf{y} := \mathbf{x} - c\mathbf{1}$ satisfies $\max_i y_i = 1$, $\min_i y_i = -1$. Therefore, for arbitrary i, i',

$$\begin{aligned}
|(\mathbf{P}\mathbf{y})_i - (\mathbf{P}\mathbf{y})_{i'}| &= |\sum_j (p_{i,j} - p_{i',j}) y_j| \\
&\le \sum_j |p_{i,j} - p_{i',j}| \\
&\le 2L.
\end{aligned}$$

Consequently, by (i), the norm of $[\mathbf{P}\mathbf{y}]$ and thus also that of $[\mathbf{P}\mathbf{x}]$ $(=[\mathbf{P}\mathbf{y}])$ is at most L. Since $[\mathbf{x}] \mapsto [\mathbf{P}\mathbf{x}]$ is linear this proves the claim. $\square$

3. A PRIORI STRATEGIES

What is, for a given Parrondo game, the best choice of the $\mathbf{p}_1,\ldots,\mathbf{p}_r$ in order to get a maximal expected gain? The following proposition shows that a suitable non-random choice will provide the best result. We need a

Lemma 3.1 *Let K be a compact convex set in $\mathbb{R}^r$ and K^m its m'th power, provided with the product topology. Then, for every function $h : K \to \mathbb{R}$ which is continuous and coordinate-wise convex[4], there are extreme points $\mathbf{e}_1,\ldots,\mathbf{e}_m$ of K such that $h(\mathbf{e}_1,\ldots,\mathbf{e}_m)$ is the maximum of h on K.*

[4] I.e., $\mathbf{x} \mapsto h(\mathbf{x}_1,\ldots,\mathbf{x}_{\mu-1},\mathbf{x},\mathbf{x}_{\mu+1},\ldots,\mathbf{x}_m)$ is convex on K for arbitrary μ and $\mathbf{x}_1,\ldots,\mathbf{x}_{\mu-1},\mathbf{x}_{\mu+1},\ldots,\mathbf{x}_m$.

Proof. Let $M := \max h$ and $\mathbf{x}_1, \ldots, \mathbf{x}_m$ be such that $M = h(\mathbf{x}_1, \ldots, \mathbf{x}_{m)})$. Consider the function $\mathbf{x} \mapsto h(\mathbf{x}, \mathbf{x}_2, \ldots, \mathbf{x}_m$. It is convex and continuous on K so that there is an extreme point $\mathbf{e}_1$ with $M = h(\mathbf{e}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$. It should be clear how to continue. □

Proposition 3.2 *Let a Parrondo game be given. Denote, for any a priori strategy $\mathbf{p}_1, \ldots, \mathbf{p}_m$, by $E_{\mathbf{p}_1,\ldots,\mathbf{p}_m}$ the expected gain when playing the game according to this strategy.*

Then the maximum of the $E_{\mathbf{p}_1,\ldots,\mathbf{p}_m}$ is attained at a strategy where all $\mathbf{p}_\mu$ are Dirac measures.

Proof. One only has to apply the preceding lemma with $K = \mathcal{P}_r$ (= the probability measures on $\{1, \ldots, r\}$) and

$$h : (\mathbf{p}_1, \ldots, \mathbf{p}_m) \mapsto E_{\mathbf{p}_1,\ldots,\mathbf{p}_m}.$$

This mapping is continuous and coordinate-wise affine, and it remains to note that the extreme points of K are the Dirac measures. □

With this proposition the problem how to find an optimal a priori strategy is solved: one has to check the r^m deterministic strategies, at one of these the optimum will be achieved.

However, r^m might be incredibly large, and therefore there it is desirable to find other possibilities. We are going to present a technique which provides an approximately optimal strategy in polynomial time (theorem 4.2).

The idea is to *combine the following facts:*

- Assertion (ii) of lemma 2.1: the expected total gain under the deterministic strategy $(\mathbf{P}_{\rho_1}, \mathbf{x}_{\rho_1}), \ldots, (\mathbf{P}_{\rho_m}, \mathbf{x}_{\rho_m})$ is the first component of

$$\mathbf{x}_{\rho_1} + \mathbf{P}_{\rho_1}\mathbf{x}_{\rho_2} + \mathbf{P}_{\rho_1}\mathbf{P}_{\rho_2}\mathbf{x}_{\rho_3} + \cdots + \mathbf{P}_{\rho_1} \cdots \mathbf{P}_{\rho_{m-1}}\mathbf{x}_{\rho_m}.$$

- Since the $\mathbf{P}_\rho$ are uniformly L-contractive, the matrix $\mathbf{P}_{\rho_1} \cdots \mathbf{P}_{\rho_{k-1}}$ in a typical summand of the preceding expression has nearly constant rows, say π, if k is not too small. Thus this summand contributes (roughly) with $\langle \pi, \mathbf{x}_{\rho_k} \rangle$ to the total gain.
- Therefore, if one is only interested in large m, it is natural to transform the problem to the search of an optimal walk in a directed graph: the vertices are constructed by means of suitable probability vectors and the weights of the edges are inner products.

Let $\mathbf{P}$ be a stochastic $s \times s$-matrix such that $L := \max_{i,i'} \sum_j |p_{i,j} - p_{i',j}|/2$ is smaller than one. Then $T_{\mathbf{P}} : \mathbf{x} \mapsto (\mathbf{x}^\top \mathbf{P})^\top$ is contraction on the set $\mathcal{P}^s$ of probability measures on $\{0, \ldots, s-1\}$, provided with the l^1-norm $|| \cdot ||_1$, with Lipschitz constant L (for notational convenience we will write in the sequel the vectors in $\mathcal{P}^s$ as row vectors; using this convention we can define $T_{\mathbf{P}}$ more simply as $T_{\mathbf{P}} : \mathbf{x} \mapsto \mathbf{xP}$). Thus Banach's fixed point theorem provides a $\pi_{\mathbf{P}}$ with $\pi_{\mathbf{P}}\mathbf{P} = \pi_{\mathbf{P}}$, and one knows that the $\mathbf{xP}^k$ converge geometrically fast with $k \to \infty$ to $\pi_{\mathbf{P}}$, the equilibrium of $\mathbf{P}$, for any $\mathbf{x}$ (see, e.g., lemma 10.6 in [1]).

If both $T_{\mathbf{P}}$ and $T_{\mathbf{Q}} : \mathbf{x} \mapsto \mathbf{Qx}$ are contractions on $\mathcal{P}^s$ with Lipschitz constants L_1 and L_2, respectively, then $\mathbf{x} \mapsto \mathbf{xPQ}$ is Lipschitz with constant L_1L_2. From these elemenatary observations we may conclude that $\mathbf{x} \mapsto \mathbf{xP}_{\nu_1} \cdots \mathbf{P}_{\nu_k}$ is a contraction with Lipschitz constant L^k for arbitrary $\nu_1, \ldots, \nu_k$ in $\{1, \ldots, r\}$. Denote the associated fixed point by $\tilde{\pi}_{\nu_1,\ldots,\nu_k}$, the inner products with the $\mathbf{x}_\rho$ will be used to approximate the stepwise gains.

Let $\eta > 0$ be given. We choose an l such that $L^l \leq \eta$ and then a finite set $D \subset \mathcal{P}_s$ such that

- for $\nu_1, \ldots, \nu_l \in \{1, \ldots, r\}$ there is a $\pi \in D$ with $||\tilde{\pi}_{\nu_1,\ldots,\nu_l} - \pi||_1 \leq \eta$;
- for $\pi \in D$ there are $\nu_1, \ldots, \nu_l \in \{1, \ldots, r\}$ such that $||\tilde{\pi}_{\nu_1,\ldots,\nu_l} - \pi||_1 \leq \eta$.

(I.e., D and the collection of the $\nu_1, \ldots, \nu_l \in \{1, \ldots, r\}$ are mutually η-nets.)

Lemma 3.3

(i) *If $k \geq l$, then*

$$\pi\mathbf{P}_{\nu_1} \cdots \mathbf{P}_{\nu_k} =_{2\eta} \tilde{\pi}_{\nu_1,\ldots,\nu_k}$$

for arbitrary $\pi \in \mathcal{P}_s$ and $\nu_1, \ldots, \nu_k \in \{1, \ldots, r\}$; here "$a =_\tau b$" stands for "the l^1-distance between a and b is at most τ".

(ii) *For $\pi \in D$ and $\rho = 1, \ldots, r$ there is $\Pi_{\pi,\rho} \in D$ such that $\pi\mathbf{P}_\rho =_{4\eta} \Pi_{\pi,\rho}$.*

(iii) *Fix – for all η, ρ – such a $\Pi_{\pi,\rho}$. Define, for given $\rho_1, \ldots, \rho_{k_0}$ and $\pi_0 \in D$, a sequence (π_κ) in D by $\pi_{\kappa+1} := \Pi_{\pi_\kappa,\rho_{\kappa+1}}$ for $\kappa = 0, \ldots, k_0 - 1$. Then*

$$\pi_0\mathbf{P}_{\rho_1} \cdots \mathbf{P}_{\rho_k} =_{4k\eta} \pi_k$$

for every k.

(iv) *For each k with $l \leq k \leq k_0$ one has*

$$\pi_k =_{(4l+2)\eta} \tilde{\pi}_{\rho_1,\ldots,\rho_k}.$$

Proof. (i) One only has to apply the following elementary fact: if T is a Lipschitz map on a metric space (M, d) with Lipschitz constant η and fixed pont x_0, then for every x one has

$$d(Tx, x_0) = d(Tx, Tx_0) \leq \eta d(x, x_0) \leq \eta \mathrm{diam}(M).$$

(ii) Choose $\tilde{\pi}_{\nu_1,\ldots,\nu_l}$ with $\tilde{\pi}_{\nu_1,\ldots,\nu_l} =_\eta \pi$ and $\Pi_{\pi,\rho} \in D$ which is η-close to $\tilde{\pi}_{\nu_2,\ldots,\nu_{l-1},\rho}$. Then it follows that

$$\begin{aligned}
\pi\mathbf{P}_\rho &=_\eta \tilde{\pi}_{\nu_1,\ldots,\nu_l}\mathbf{P}_\rho \\
&= \tilde{\pi}_{\nu_1,\ldots,\nu_l}(\mathbf{P}_{\nu_1} \cdots \mathbf{P}_{\nu_l})\mathbf{P}_\rho \\
&= (\tilde{\pi}_{\nu_1,\ldots,\nu_l}\mathbf{P}_{\nu_1})(\mathbf{P}_{\nu_2} \cdots \mathbf{P}_{\nu_l}\mathbf{P}_\rho) \\
&=_{2\eta} \tilde{\pi}_{\nu_2,\ldots,\nu_{l-1},\rho} \\
&=_\eta \Pi_{\pi,\rho};
\end{aligned}$$

in the last but one estimate we have applied assertion (i).
(iii) This is a telescope sum argument:

$$\begin{aligned}\pi_0\mathbf{P}_{\rho_1}\cdots\mathbf{P}_{\rho_k}-\pi_k &= (\pi_0\mathbf{P}_{\rho_1}-\pi_1)(\mathbf{P}_{\rho_2}\cdots\mathbf{P}_{\rho_k})+\\ &+ (\pi_1\mathbf{P}_{\rho_2}-\pi_1)(\mathbf{P}_{\rho_3}\cdots\mathbf{P}_{\rho_k})+\\ &+ \cdots+\\ &+ \pi_{k-1}\mathbf{P}_{\rho_k}-\pi_k.\end{aligned}$$

Note that the norm of each of the summands is at most 4η and that the $T_{\mathbf{P}_\rho}$ are contractive.
(iv) If we apply (iii) with π_{k-l} instead of π_0 we get $\pi_k =_{4l\eta} \pi_{k-l}\mathbf{P}_{\rho_{k-l+1}}\cdots\mathbf{P}_{\rho_k}$, and the right hand side is 2η-close to the fixed point $\tilde{\pi}_{\rho_1,\ldots,\rho_k}$ of $\mathbf{x}\mapsto\mathbf{x}\mathbf{P}_{\rho_1}\cdots\mathbf{P}_{\rho_k}$. □

Let $\eta>0$ be given and D and the $\Pi_{\pi,\rho}$ as in the definition before lemma 3.3. As a last preparation of the main result of this section we introduce the following *directed weighted graph* $\mathcal{G}$:

- The vertices of $\mathcal{G}$ are the elements of $D\times\{1,\ldots,r\}$.
- With any vertex (π,i) there are associated r directed edges, namely to $(\Pi_{\pi,\rho},\rho)$ for $\rho=1,\ldots,r$; the ρ'th edge has weight $\langle\pi,\mathbf{x}_\rho\rangle$.

We note that, by proposition 2.3, the complexity to find an η-optimal walk of length m in $\mathcal{G}$ is $O(r(|D|r)^2)$.

Theorem 3.4 *Let $\eta_0>0$ be given. One can find a strategy $\rho_1,\ldots,\rho_m$ in $O(r^3\eta_0^{-2s-3})$ steps for which the expected total gain is $m\eta_0$-close to the optimal value.*

Proof. Let us assume for simplicity that all components of all $\mathbf{x}_\rho$ are bounded by 1 so that $\langle\pi,\mathbf{x}_\rho\rangle$ is τ-close to $\langle\pi',\mathbf{x}_\rho\rangle$ whenever $\pi=_\tau\pi'$.

Fix $\varepsilon>0$: we will show that the complexity is $O(r^3\eta_0^{-2s-2-2\varepsilon})$, where the constants depend on ε. To this end, we define $\eta:=\eta_0^{1+\varepsilon}/-2\log L$. Then it is true that

$$\frac{\log\eta}{\log L}\eta+\frac{\log\eta}{m\log L}\le\eta_0/3$$

provided that m is sufficiently large and η_0 is sufficiently small.

With this η we choose a set D and the $\Pi_{\pi,\rho}$ as in the preceding considerations. We note that we can bound the cardinality of D by $O(\eta^{-(s-1)})$ since $\mathcal{P}_s$ is a bounded $(s-1)$-dimensional set. Also, the number l above is $O(\log\eta/\log L)$.

Let $\rho_1,\ldots,\rho_m$ be a sequence in $\{1,\ldots,r\}$ for which $E:=E_{\rho_1,\ldots,\rho_m}$, the expected gain associated with the deterministic strategy $\rho_1,\ldots,\rho_m$, is maximal; recall that, by lemma 2.1(ii), this gain is the first component of

$$\mathbf{x}_{\rho_1}+\mathbf{P}_{\rho_1}\mathbf{x}_{\rho_2}+\mathbf{P}_{\rho_1}\mathbf{P}_{\rho_2}\mathbf{x}_{\rho_3}+\cdots+\mathbf{P}_{\rho_1}\cdots\mathbf{P}_{\rho_{m-1}}\mathbf{x}_{\rho_m}.$$

Claim 1: Choose an arbitrary vertex $(\pi_0, \rho_0) \in \mathcal{G}$ *and define a walk by*

$$(\pi_\mu, \rho_\mu) \mapsto (\pi_{\mu+1}, \rho_{\mu+1})$$

for $\mu = 0, \ldots, m-1$*; here* $\pi_{\mu+1} := \Pi_{\pi_{mu}, \rho_{\mu+1}}$.. *We claim that the weight of this walk is at least* $E - m\eta_0$.

Proof of Claim 1:

For $k \geq l$ we have $\mathbf{P}_{\rho_1} \cdots \mathbf{P}_{\rho_k} \mathbf{x}_{\rho_{k+1}} =_{2\eta} \langle \tilde{\pi}_{\rho_1, \ldots, \rho_k}, \mathbf{x}_{\rho_{k+1}}$, and it follows that

$$\begin{aligned}
E \quad &= \quad \text{first component of} \\
& \qquad \mathbf{x}_{\rho_1} + \mathbf{P}_{\rho_1} \mathbf{x}_{\rho_2} + \mathbf{P}_{\rho_1} \mathbf{P}_{\rho_2} \mathbf{x}_{\rho_3} + \cdots + \mathbf{P}_{\rho_1} \cdots \mathbf{P}_{\rho_{m-1}} \mathbf{x}_{\rho_m} \\
&=_l \quad \mathbf{P}_{\rho_1} \cdots \mathbf{P}_l \mathbf{x}_{\rho_{l+1}} + \cdots + \mathbf{P}_{\rho_1} \cdots \mathbf{P}_{\rho_{m-1}} \mathbf{x}_{\rho_m} \\
&=_{2m\eta} \quad \sum_{\mu=l}^{m-1} \langle \tilde{\pi}_{\rho_1, \ldots, \rho_\mu}, \mathbf{x}_{\rho_{\mu+1}} \rangle \\
&=_{(4l+2)\eta m} \quad \sum_{\mu=l}^{m-1} \langle \pi_\mu, \mathbf{x}_{\rho_{\mu+1}} \rangle \\
&=_l \quad \sum_{\mu=1}^{m-1} \langle \pi_\mu, \mathbf{x}_{\rho_{\mu+1}} \rangle.
\end{aligned}$$

Thus the difference beween E and the weight $\sum_{\mu=1}^{m-1} \langle \pi_\mu, \mathbf{x}_{\rho_{\mu+1}} \rangle$ is at most $2l + 2m\eta + (4l+2)\eta m$, a number which is bounded by $m\eta_0$ by assumption.

Claim 2: Suppose that the maximal gain $G_{\mathcal{G}}$ *of a walk of length* m *in* $\mathcal{G}$ *is attained at a walk*

$$(\pi_0, \rho_0) \mapsto (\pi_1, \rho_1) \mapsto \cdots \mapsto (\pi_m, \rho_m).$$

Then $E_{\rho_1, \ldots, \rho_m} \geq E - m\eta_0$.

Proof of Claim 2: It suffices to read the preceding estimates from right to left.

Conclusio: Up to an error which is bounded by $m\eta_0$ one can find the optimal strategy by providing an optimal walk in $\mathcal{G}$. It only remains to apply proposition 2.3 to complete the proof of the theorem.

$\square$

4. CONSTANT A PRIORI STRATEGIES

Fix a constant a priori strategy $\mathbf{p} = (\lambda_1, \ldots, \lambda_r)$ and put

$$\mathbf{P} := \sum_\rho \lambda_\rho \mathbf{P}_\rho, \ \mathbf{x} := \sum_\rho \lambda_\rho \mathbf{x}_\rho.$$

Playing according to the constant strategy $\mathbf{p}, \mathbf{p}, \ldots$ amounts in playing with the game $(\mathbf{P}, \mathbf{x})$ (cf. lemma 2.1 (i)). Since $\mathbf{P}$ is L-contractive the game is easily analysed: The expected gain after m rounds is the first component of the vector

$$(Id + \mathbf{P} + \mathbf{P}^2 + \cdots + \mathbf{P}^{m-1})\mathbf{x}$$

and consequently, if m is large, roughly $m\langle\pi_{\mathbf{P}},\mathbf{x}\rangle$, where $\pi_{\mathbf{P}}$ stands for the equilibrium distribution of $\mathbf{P}$. Thus the number $\langle\pi_{\mathbf{P}},\mathbf{x}\rangle$ is of particular importance: if it is zero, positive or negative one plays a fair, winning or losing game, respectively.

In Parrondo's example the numbers $\langle\pi_{\mathbf{P}_1},\mathbf{x}_1\rangle$ and $\langle\pi_{\mathbf{P}_2},\mathbf{x}_2\rangle$ are zero, but $\langle\pi_{\mathbf{P}},\mathbf{x}\rangle$ is positive for $\mathbf{P}=\lambda\mathbf{P}_1+(1-\lambda)\mathbf{P}_2, \mathbf{x}=\lambda\mathbf{x}_1+(1-\lambda)\mathbf{x}_2$ with a suitable λ between 0 and 1, i.e., one may combine fair games (or even losing games) to arrive at a winning game.

Definition 4.1 Let $\mathbf{P}_1,\ldots,\mathbf{P}_r$ be stochastic $s\times s$-matrices which are uniformly contractive. We say that *a paradoxical behaviour for* $\mathbf{P}_1,\ldots,\mathbf{P}_r$ *is possible* provided there are vectors $\mathbf{x}_1,\ldots,\mathbf{x}_r$ with $\langle\pi_{\mathbf{P}_\rho},\mathbf{x}_\rho\rangle\le 0$ for every ρ and a probability vector $\mathbf{p}=(\lambda_1,\ldots,\lambda_r)$ such that, for the associated $(\mathbf{P},\mathbf{x})$, one has $\langle\pi_{\mathbf{P}},\mathbf{x}\rangle>0$.

Computer experiments reveal that, surprisingly, the possibility of paradoxical behaviour is rather the rule than the exception. (Clearly it is not expected always: e.g., if all $\mathbf{x}_\rho$ vanish, no interesting phenomena can be observed.)

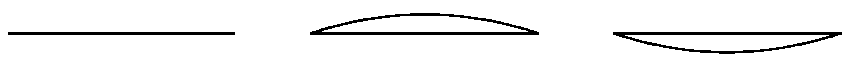
We have examined various examples in the case $r=2$, here the relevant $\mathbf{p}$ has the form $(\lambda,1-\lambda)$ for $0\le\lambda\le 1$. We have restricted ourselves to *fair games*, i.e., to situations where $\langle\pi_{\mathbf{P}_1},\mathbf{x}_1\rangle=\langle\pi_{\mathbf{P}_2},\mathbf{x}_2\rangle=0$.

It is natural to consider the function

$$\phi(\lambda):=\langle\pi_{\lambda\mathbf{P}_1+(1-\lambda)\mathbf{P}_2},\lambda\mathbf{x}_1+(1-\lambda)\mathbf{x}_2\rangle \text{ for } \lambda\in[0,1].$$

By our assumption $\phi(0)=\phi(1)=0$, and the $\phi(\lambda)$ with $0<\lambda<1$ are of interest if paradoxical behaviour has to be studied.

We have considered various $(\mathbf{P}_1,\mathbf{x}_1)$ and $(\mathbf{P}_2,\mathbf{x}_2)$ which have been defined by using a random generator. When plotting the associated ϕ-function it was very surprising to us to observe that only very few typical graphs occur. In the overwhelming majority of cases the graph of the ϕ-function is of one of the following three types:



This means that

- There are cases without paradoxical behaviour (not surprising).
- Very often *every* proper random mixture of fair games $(\mathbf{P}_1,\mathbf{x}_1)$ and $(\mathbf{P}_2,\mathbf{x}_2)$ gives rise to a winning game (or always to a losing game).

In the rest of this section we are going to try to explain these observations qualitatively. First we prove why paradoxical situation abound:

Theorem 4.2 *Let $\mathbf{P}_1, \ldots, \mathbf{P}_r$ be stochastic matrices which are uniformly L-contractive. The following are equivalent*

(i) *A paradoxical behaviour is possible.*

(ii) *At least two of the equilibrilum distributions $\pi_{\mathbf{P}_\rho}$ are different.*

Proof. Suppose first that all $\pi_{\mathbf{P}_\rho}$ coincide: $\pi_{\mathbf{P}_1} = \cdots = \pi_{\mathbf{P}_r} =: \pi$. Then π is also the equilibrium for any convex combination $\mathbf{P} = \sum_\rho \lambda_\rho \mathbf{P}_\rho$ so that, when $\langle \pi_{\mathbf{P}_\rho}, \mathbf{x}_\rho \rangle \leq 0$ for certain $\mathbf{x}_1, \ldots, \mathbf{x}_r$, one also has

$$\langle \pi_{\sum_\rho \lambda_\rho \mathbf{P}_\rho}, \sum_\rho \lambda_\rho \mathbf{x}_\rho \rangle = \sum \lambda_\rho \langle \pi, \mathbf{x}_\rho \rangle \leq 0.$$

Now suppose that at least two of the $\pi_{\mathbf{P}_\rho}$ are different, $\pi_{\mathbf{P}_1} \neq \pi_{\mathbf{P}_2}$ say. We will provide $\mathbf{x}_1, \mathbf{x}_2$ and $\lambda \in [0,1]$ such that $\langle \pi_{\mathbf{P}_1}, \mathbf{x}_1 \rangle = \langle \pi_{\mathbf{P}_2}, \mathbf{x}_2 \rangle = 0$ but $\langle \pi_{\lambda \mathbf{P}_1} + (1-\lambda)\mathbf{P}_2, \lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \rangle > 0$.

Choose $\mathbf{y}_1, \mathbf{y}_2$ such that

$$\langle \pi_{\mathbf{P}_1}, \mathbf{y}_1 \rangle = 0 \neq \langle \pi_{\mathbf{P}_2}, \mathbf{y}_1 \rangle, \quad \langle \pi_{\mathbf{P}_2}, \mathbf{y}_2 \rangle = 0 \neq \langle \pi_{\mathbf{P}_1}, \mathbf{y}_2 \rangle;$$

this is possible since $\pi_{\mathbf{P}_1}, \pi_{\mathbf{P}_2}$, being different probability distributions, are linearly independent.

Consider, for $\eta_1, \eta_2 \in \{-1, +1\}$, the function

$$\phi_{\eta_1,\eta_2}(\lambda) := \langle \pi_{\lambda \mathbf{P}_1 + (1-\lambda)\mathbf{P}_2}, \lambda \eta_1 \mathbf{y}_1 + (1-\lambda)\eta_2 \mathbf{y}_1 \rangle$$

for $0 \leq \lambda \leq 1$. Fix any $\lambda_0 \in]0,1[$ and suppose that $\phi_{\eta_1,\eta_2}(\lambda_0) = 0$ for $\eta_1, \eta_2 = \pm 1$ would hold. Then it would follow that

$$\langle \pi_{\lambda_0 \mathbf{P}_1 + (1-\lambda_0)\mathbf{P}_2}, \mathbf{y}_1 \rangle = \langle \pi_{\lambda_0 \mathbf{P}_1 + (1-\lambda_0)\mathbf{P}_2}, \mathbf{y}_2 \rangle = 0.$$

But this ist not possible since, by the continuity of $\lambda \mapsto \pi_{\lambda \mathbf{P}_1 + (1-\lambda)\mathbf{P}_2}$, one has

$$\langle \pi_{\lambda_0 \mathbf{P}_1 + (1-\lambda_0)\mathbf{P}_2}, \mathbf{y}_1 \rangle \approx \langle \pi_{\mathbf{P}_2}, \mathbf{y}_1 \rangle \neq 0$$

for $\lambda_0 \approx 0$ and

$$\langle \pi_{\lambda_0 \mathbf{P}_1 + (1-\lambda_0)\mathbf{P}_2}, \mathbf{y}_2 \rangle \approx \langle \pi_{\mathbf{P}_1}, \mathbf{y}_2 \rangle \neq 0$$

for $\lambda_0 \approx 1$. It remains to fix $\eta_1, \eta_2, \lambda_0$ such that $\phi_{\eta_1,\eta_2}(\lambda_0) \neq 0$. If $\phi_{\eta_1,\eta_2}(\lambda_0) > 0$ we put $\mathbf{x}_1 := \eta_1 \mathbf{y}_1$, $\mathbf{x}_2 := \eta_2 \mathbf{y}_2$, otherwise we deal with $\mathbf{x}_1 := -\eta_1 \mathbf{y}_1$, $\mathbf{x}_2 := -\eta_2 \mathbf{y}_2$.

Next we try to explain – at least in a qualitative way – the phenomenon that there are few typical forms of the ϕ-function. The key result is the following

Proposition 4.3 *Fix $\mathbf{P}_1, \mathbf{P}_2$ which are uniformly L-contractive for some $L < 1$. Define, for λ between 0 and 1, $\gamma(\lambda)$ as the deviation of $\pi_{\lambda\mathbf{P}_1+(1-\lambda)\mathbf{P}_2}$ from the corresponding point of the line segment between $\pi_{\mathbf{P}_1}$ and $\pi_{\mathbf{P}_2}$:*

$$\gamma(\lambda) := \pi_{\lambda\mathbf{P}_1+(1-\lambda)\mathbf{P}_2} - (\lambda\pi_{\mathbf{P}_1} + (1-\lambda)\pi_{\mathbf{P}_2}).$$

Then $\gamma(\lambda)$ is the unique solution $\gamma \in \Delta := \{(a_0, \ldots, a_{s-1}) \mid \sum a_i = 0\}$ of the fixed point equation

$$\gamma = \gamma(\lambda\mathbf{P}_1 + (1-\lambda)\mathbf{P}_2) + \lambda(1-\lambda)(\pi_{\mathbf{P}_1} - \pi_{\mathbf{P}_2})(\mathbf{P}_1 - \mathbf{P}_2).$$

Proof. A straightforward calculation shows that $\gamma(\lambda)$ solves the equation. Uniqueness follows from the fact that

$$\gamma \mapsto \gamma(\lambda\mathbf{P}_1 + (1-\lambda)\mathbf{P}_2)$$

is a contraction with Lipschitz constant L on Δ; cf. [1], lemma 10.6.

With this proposition at hand we can try to make the observed facts plausible:

- If $\mathbf{P}_1, \mathbf{P}_2$ are random matrices chosen with respect ot the uniform distribution it is unlikely that $\pi_{\mathbf{P}_1}, \pi_{\mathbf{P}_2}$ have a particularly large mass at certain components. Thus $\pi_{\mathbf{P}_1} - \pi_{\mathbf{P}_2}$ "usually" will be a vector with "small" components. For the same reason $\mathbf{P}_1 - \mathbf{P}_2$ will have "small" entries so that it is to be expected that $(\pi_{\mathbf{P}_1} - \pi_{\mathbf{P}_2})(\mathbf{P}_1 - \mathbf{P}_2)$ is "small".
- If X is any Banach space and T a linear contraction, then the unique solution x of $x = Tx + y$ is $x = y + Ty + T^2y + \cdots$. Thus, if the contraction constant is small, one might approximate x better and better by y, $y + Ty$ etc.
- Conclusio: it is likely that – as a first approximation – one is justified to replace $\pi_{\lambda\mathbf{P}_1+(1-\lambda)\mathbf{P}_2}$ by $\lambda\pi_{\mathbf{P}_1} + (1-\lambda)\pi_{\mathbf{P}_2}$ or even better by

$$\lambda\pi_{\mathbf{P}_1} + (1-\lambda)\pi_{\mathbf{P}_2} + (\pi_{\mathbf{P}_1} - \pi_{\mathbf{P}_2})(\mathbf{P}_1 - \mathbf{P}_2).$$

- Having in mind that we have assumed that $\langle \pi_{\mathbf{P}_1}, \mathbf{x}_1 \rangle = \langle \pi_{\mathbf{P}_2}, \mathbf{x}_2 \rangle = 0$ it follows that a first approximation of the ϕ-function will be a multiple of the function $\lambda(1-\lambda)$ and that as a better approximation one will find an element in the linear span of $\lambda(1-\lambda), \lambda^2(1-\lambda), \lambda(1-\lambda)^2$.

It has to be emphasized that these considerations are of a very rough qualitative nature. Further research will be necessary until one is able to read off the structure of ϕ form $\mathbf{P}_1$ and $\mathbf{P}_2$ directly.

5. ADAPTIVE STRATEGIES

Again we consider a Parrondo game $(\mathbf{P}_1, \mathbf{x}_1), \ldots, (\mathbf{P}_r, \mathbf{x}_r)$, recall that we assume that there exists an $L < 1$ such that the $\mathbf{P}_\rho$ are uniformly L-contractive. The starting position of the game is zero, m rounds are to be played with an adaptive strategy, which $(\mathbf{P}_\rho, \mathbf{x}_\rho)$ should be chosen first ?

It is not obvious what has to be done since a big first reward ("nice" first component of $\mathbf{x}_\rho$) might be compensated by the fact that with the associated $\mathbf{P}_\rho$ one will have to move to states where the rewards are not favourable. Thus it is not likely that greedy strategies lead to good results, one has to be more careful.

There is, however, no need to re-invent the wheel, one finds already an elaborated theory which has been developed from a different starting point. What is necessary to know can be found in the literature concerned with *Random Decision Theory* or with *Random Control Processes*, we refer the reader to [3], [8] or [13]. Rather than adapting the general results of these theories to the present context we sketch a short self-contained approach.

E.g., our proposition 5.1 is a special case of the dynamic programming theorem 3.2.1 in [8], and theorem 5.3 can be found under another viewpoint as theorem 4.2.3 in [8] or in section 3.1 of [13]. We note, however, that our proof of theorem 5.3 might be new to the specialists.

By definition an adaptive strategy $\mathcal{S}$ is a rule to decide which $(\mathbf{P}_\rho, \mathbf{x}_\rho)$ should be used next given the present state of the game and the last moves. In a more formal approach one could define such a strategy as a sequence of maps which satisfy certain measurabilty assumptions with respect to an appropriate filtration, it should not be difficult to translate our more intuitive approach to this setting.

Suppose that the game starts at a certain $i \in \{0, \ldots, s-1\}$, that n rounds are to be played and that one uses a certain strategy $\mathcal{S}$. The expected total gain will be denoted by $g_n^{\mathcal{S}}(i)$, and $g_n(i)$ stands for the supremum over all $g_n^{\mathcal{S}}(i)$ where $\mathcal{S}$ runs through all adaptive strategies; additionally, we define $g_0(i) := 0$ for all i. By the following proposition this supremum is an fact a maximum and the best possible strategy can be described explicitly. Also it turns out that all $g_n(i)$ can easily be determined by induction, in particular one can then calculate $g_m(0)$: *this* is the value we are interested in when playing a Parrondo game where adaptive strategies are admitted.

Proposition 5.1

(i) *For* $n = 0, 1, \ldots$ *and* $i \in \{0, \ldots, s-1\}$ *one has*

$$g_{n+1}(i) = \max_{\rho=1,\ldots,r} x_i^\rho + \sum_j p_{i,j}^\rho g_n(j).$$

(ii) *The optimal strategy is defined as follows: when being at state* i *and* n *rounds remain to be played, choose* $(\mathbf{P}_{\rho_0}, \mathbf{x}_{\rho_0})$ *for the next game, where* ρ_0 *satisfies*

$$x_i^{\rho_0} + \sum_j p_{i,j}^{\rho_0} g_n(j) = \max_{\rho=1,\ldots,r} x_i^\rho + \sum_j p_{i,j}^\rho g_n(j).$$

Proof. Suppose first that $n = 1$ and that the present position is i. Then obviously the best strategy is to choose a $(\mathbf{P}_{\rho_0}, \mathbf{x}_{\rho_0})$ such that $x_i^{\rho_0} = \max_\rho x_i^\rho$. Since we have defined $g_n(j) = 0$ for all j this proves (i) and (ii) for the case $n = 1$.

Now assume that we are at i and $n+1$ rounds are still to be played. Suppose that we decide to choose a certain $(\mathbf{P}_\rho, \mathbf{x}_\rho)$. We will get x_i^ρ and we will have to move to another state j with probability $p_{i,j}^\rho$. If we play as clever as possible we have to expect $g_n(j)$ in the remaining n rounds so that our choice results in an expected gain of

$$x_i^\rho + \sum_j p_{i,j}^\rho g_n(j).$$

The maximum over all ρ is $g_{n+1}(i)$, and this proves the proposition.

(*Remark:* Note that the fact that the strategy has such a simple form is due to the Markov property of the underlying walk.) □

When one calculates the $g_n(i)$ for larger and larger n one can observe a *surprising phenomenon:* this sequence behaves rather regularly, there is a number γ such that all $(g_{n+1}(i) - g_n(i))_n$ tend with $n \to \infty$ to γ. Also the optimal strategy can be approximated quite simply: if n rounds are to be played and n is not too small, then the best choice, being at state i, is independent of n. To phrase it otherwise: there is a *selection function* $\sigma : \{0, \ldots, s-1\} \to \{1, \ldots, r\}$ such that a nearly optimal adaptive choice being at state i is $(\mathbf{P}_{\sigma(i)}, \mathbf{x}_{\sigma(i)})$.

The number γ has an obvious meaning, it stands for the expected gain in one step if one plays in an optimal way.

The following theorem explains the phenomenon, a characterization of γ is also given. We need a preparation, part (iii) will be crucial for the further investigations.

Lemma 5.2 *Define* $D' : \mathbb{R}^s \to \mathbb{R}^s$ *by*

$$(x_i)_{i=0,\ldots,s-1} \mapsto (\max_{\rho=1,\ldots,r} x_i^\rho + \sum_j p_{i,j}^\rho x_j)_{i=0,\ldots,s-1}.$$

(i) $D'(\mathbf{x} + a\mathbf{1}) = D'(\mathbf{x}) + a\mathbf{1}$ *for all* $a \in \mathbb{R}$*; recall that* $\mathbf{1}$ *stands for the constant vector* $(1, \ldots, 1)$.

(ii) *Let* X *be as in section 2C.* D' *induces a well-defined map* D *from* X *to* X *by* $[\mathbf{x}] \mapsto [D'\mathbf{x}]$.

(iii) D *is a contraction on* X *with Lipschitz constant* L.

Proof. (i) is obviously true, and (ii) follows from (i). For the proof of (iii) we first note that

$$\max a_\rho - \max b_\rho \le \max(a_\rho - b_\rho)$$

and consequently

$$|\max a_\rho - \max b_\rho| \le \max |a_\rho - b_\rho|$$

for real numbers $a_1, \ldots, a_r$ and $b_1, \ldots, b_r$.

Now let $[\mathbf{z}], [\mathbf{w}] \in X$ be given. With $a := ||[z-w]||$ we find $\mathbf{y} \in \mathbb{R}^s$ and a real α such that

$$\min y_i = -a, \ \max y_i = a, \ \mathbf{w} = \mathbf{z} + \alpha \mathbf{1} + \mathbf{y}.$$

Consider the numbers $\sum_j p^\rho_{i,j} y_j$ for $\rho = 1, \ldots, r$ and $i = 0, \ldots, s-1$. Since the $\mathbf{P}_\rho$ are uniformly L-contractive the distance between each two of them be estimated by $2La$. Therefore one may find a real c such that

$$\sum_j p^\rho_{i,j} y_j = c + c^\rho_i$$

with $|c^\rho_i| \leq La$.

From (i) we conclude that $D'(\mathbf{w}) = D'(\mathbf{z}+\mathbf{y}) + \alpha\mathbf{1}$ so that the distance between $D([\mathbf{z}])$ and $D([\mathbf{w}])$ is the distance between $D([\mathbf{z}+\mathbf{y}])$ and $D([\mathbf{z}])$, i.e., we have to estimate the diameter of the set of components of the vector $D([\mathbf{z}+\mathbf{y}]) - D([\mathbf{z}])$.

Let i be arbitrary. The i'th component of $D'(\mathbf{z}+\mathbf{y})$ (resp. of $D'(\mathbf{z})$) is

$$\max_\rho x^\rho_i + \sum_j p^\rho_{i,j} z_j + c + c^\rho_i = c + \max_\rho x^\rho_i + \sum_j p^\rho_{i,j} z_j + c^\rho_i$$

(resp. $\max_\rho x^\rho_i + \sum_j p^\rho_{i,j} z_j$). Thus the distance between $D([\mathbf{z}+\mathbf{y}])$ and $D([\mathbf{z}])$ can be estimated by the sup-norm distance of the vectors $(\max_\rho x^\rho_i + \sum_j p^\rho_{i,j} z_j + c^\rho_i)_i$ and $(\max_\rho x^\rho_i + \sum_j p^\rho_{i,j} z_j)_i$.

Put, for fixed i,

$$\begin{aligned} a_\rho &:= x^\rho_i + \sum_j p^\rho_{i,j} z_j + c^\rho_i, \\ b_\rho &:= x^\rho_i + \sum_j p^\rho_{i,j} z_j. \end{aligned}$$

By the preceding observation the absolute value of the difference can be estimated by $\max_\rho |a_\rho - b_\rho| = \max_\rho |c^\rho_i| \leq a \cdot L$, and this proves that

$$||D([\mathbf{z}+\mathbf{y}]) - D([\mathbf{z}])|| \leq a \cdot L.$$

□

Theorem 5.3 *Define, for $n = 0, 1, \ldots,$ the* vector of the adaptive gains $\mathbf{g}_n$ *by* $\mathbf{g}_n := (g_n(0), \ldots, g_n(s-1))$. *There are a vector $\mathbf{g} = (g_0, \ldots, g_{s-1})$ and a number γ such that the following hold:*

(i) $D'\mathbf{g} = \mathbf{g} + \gamma\mathbf{1}$, *i.e., $\mathbf{g}$ is a fixed point of D.*

(ii) *γ is uniquely determined, $\mathbf{g}$ is uniquely determined up to a vector $c\mathbf{1}$.*

(iii) *The $\mathbf{g}_{n+1} - \mathbf{g}_n$ tend to $\gamma\mathbf{1}$ geometrically fast.*

Proof. (i) Since D is a contraction on X there is a unique fixed point $[\mathbf{g}]$. Thus $D'(\mathbf{g})$ and $\mathbf{g}$ differ by a constant vector.
(ii) Let $\tilde{\mathbf{g}}$ and $\tilde{\gamma}$ be such that $D'(\tilde{\mathbf{g}}) = \tilde{\mathbf{g}} + \tilde{\gamma}\mathbf{1}$. Then $[\tilde{\mathbf{g}}]$ is also a fixed point of D, and by uniqueness we conclude that $\tilde{\mathbf{g}} = \mathbf{g} + c\mathbf{1}$ for a suitable c. With 5.2(ii) this implies that

$$\begin{aligned}\tilde{\mathbf{g}} + \tilde{\gamma}\mathbf{1} &= D'(\tilde{\mathbf{g}}) \\ &= D'(\mathbf{g}) + c\mathbf{1} \\ &= \mathbf{g} + (c+\gamma)\mathbf{1} \\ &= \tilde{\mathbf{g}} + \gamma\mathbf{1},\end{aligned}$$

i.e., $\gamma = \tilde{\gamma}$.
(iii) We have $\mathbf{g}_0 = \mathbf{0}$ and $[\mathbf{g}_{n+1}] = D([\mathbf{g}_n])$. Thus $[\mathbf{g}_n] \to [\mathbf{g}]$, and Banach's fixed point theorem even provides a c such that $||[\mathbf{g}_n] - [\mathbf{g}]|| \le cL^n$. Therefore we may choose $c_n \in \mathbb{R}$ such that $||\mathbf{g}_n - \mathbf{g} - c_n\mathbf{1}|| \le cL^n$.

With the notation of lemma 3.3 it follows that

$$\begin{aligned}\mathbf{g}_{n+1} - \mathbf{g}_n &= D'(\mathbf{g}_n) - \mathbf{g}_n \\ &= D'((\mathbf{g}_n - \mathbf{g} - c_n\mathbf{1}) + (\mathbf{g} + c_n\mathbf{1})) - \mathbf{g_n} \\ &=_{cL^n} D'(\mathbf{g} + c_n\mathbf{1}) - \mathbf{g}_n \\ &= \mathbf{g} + c_n\mathbf{1} + \gamma\mathbf{1} - \mathbf{g}_n \\ &=_{cL^n} \gamma\mathbf{1}\end{aligned}$$

so that $|g_{n+1}(i) - g_n(i)| \le 2cL^n$ for all n and i. □

With $\mathbf{g}$ and γ one can find a *constant adaptive strategy* which is close to optimal. More precisely, whenever $\sigma : \{0, \ldots, s-1\} \to \{1, \ldots, r\}$ is a map, the *constant adaptive strategy* $\mathcal{S}_\sigma$ *associated with* σ is defined by:

> "Suppose that you are now at state i. Then choose $(\mathbf{P}_{\sigma(i)}, \mathbf{x}_{\sigma(i)})$ for the next round."

Lemma 5.4 *Define a vector* $\mathbf{x}_\sigma$ *and a stochastic matrix* $\mathbf{P}_\sigma$ *as follows:*
the i'th component of $\mathbf{x}_\sigma :=$ *the i'th component of* $\mathbf{x}_{\sigma(i)}$.
the i'th row of $\mathbf{P}_\sigma :=$ *the i'th row of* $\mathbf{P}_{\sigma(i)}$.

(i) $\mathbf{P}_\sigma$ *is* L*-contractive. In particular there exists a unique equilibrium distribution* π_σ.
(ii) *The expected gain after* m *rounds of a game which uses* $(\mathbf{P}_{\sigma(i)}, \mathbf{x}_{\sigma(i)})$ *is precisely the gain which is to be expected when playing with the strategy* $\mathcal{S}_\sigma$*. There is a constant* c *such that this gain equals* $m\langle\pi_\sigma, \mathbf{x}_\sigma\rangle + c_m$*, where* $|c_m| \le c$ *for every* m.

Proof. Assertion (i) is an immediate consequence of the fact that the $\mathbf{P}_\rho$ are uniformly L-contractive. For the proof of (ii) we introduce the following notation:

$E_n^{(1)}(i)$:=the expected gain after n rounds when starting at i and playing according to $\mathcal{S}_\sigma$;
$E_n^{(2)}(i)$:=the same, but this time one plays with $(\mathbf{P}_{\sigma(i)}, \mathbf{x}_{\sigma(i)})$.

Then $E_1^{(1)}(i)$ and $E_1^{(2)}(i)$ coincide with the i'th component of $\mathbf{x}_{\sigma(i)}$, and

$$E_{n+1}^{(k)}(i) = (\mathbf{x}_{\sigma(i)})_i + \sum_j p_{i,j}^{\sigma(i)} E_n^{(k)}(j)$$

for $k = 1, 2$. Hence $E_n^{(1)}(i) = E_n^{(2)}(i)$ for all n and i.

It has already been noted in the introduction that the gain in the m'th round is $\langle \pi_\sigma, \mathbf{x}_\sigma \rangle$ up to an error which tends to zero with $m \to \infty$ geometrically fast. This proves the second part of the assertion. □

It follows that the "value" of a constant adaptive strategy σ is measured by $\langle \pi_\sigma, \mathbf{x}_\sigma \rangle$. The best possible among these r^s numbers can be found once one knows $\mathbf{g} = (g_0, \ldots, g_{s-1})$ and γ from theorem 5.3:

Theorem 5.5 *Let $\tilde{\sigma}$ be a map from $\{0, \ldots, s-1\}$ to $\{1, \ldots, r\}$ such that*

$$\max_\rho x_i^\rho + \sum_j p_{i,j}^\rho g_j = x_i^{\sigma(i)} + \sum_j p_{i,j}^{\sigma(i)} g_j$$

for every i. Then the constant adaptive strategy associated with $\tilde{\sigma}$ is optimal in the following sense:

(i) $\langle \pi_{\tilde{\sigma}}, \mathbf{x}_{\tilde{\sigma}} \rangle = \gamma$.

(ii) *There is a constant d such that the optimal gain $g_m(i)$ can be written as $m\gamma + d_m$ with $|d_m| \le d$ (so that, in view of the preceding lemma, the constant strategy $\mathcal{S}_{\tilde{\sigma}}$ is close to optimal).*

(iii) *For any σ one has $\langle \pi_\sigma, \mathbf{x}_\sigma \rangle \le \gamma$: no constant adaptive stategy gives better results than $\tilde{\sigma}$.*

Proof. (i) It follows from the definition of $\tilde{\sigma}$ that

$$\mathbf{x}_{\tilde{\sigma}} + \mathbf{P}_{\tilde{\sigma}} \mathbf{g} = \mathbf{g} + \gamma \mathbf{1}.$$

Consequently one has, for arbitrary n,

$$(\mathbf{P}_{\tilde{\sigma}})^n \mathbf{x}_{\tilde{\sigma}} = ((\mathbf{P}_{\tilde{\sigma}})^n - (\mathbf{P}_{\tilde{\sigma}})^{n+1}) \mathbf{g} + \gamma \mathbf{1}.$$

With $n \to \infty$ the left hand side tends to $\langle \pi_{\tilde{\sigma}}, \mathbf{x}_{\tilde{\sigma}} \rangle \mathbf{1}$ and the right hand side to $\gamma \mathbf{1}$. Therefore $\langle \pi_{\tilde{\sigma}}, \mathbf{x}_{\tilde{\sigma}} \rangle = \gamma$.
(ii) By theorem 5.3(ii) the $g_{n+1}(i) - g_n(i)$ tend to γ geometrically fast. It remains to write $g_m(i)$ as

$$g_m(i) = (g_m(i) - g_{m-1}(i)) + (g_{m-1}(i) - g_{m-2}(i)) + \cdots + (g_1(i) - g_0(i)).$$

(iii) For *any* adaptive strategy we know that the expected gain is bounded by $g_m(i)$. Thus

$$m \langle \pi_\sigma, \mathbf{x}_\sigma \rangle + c_n \le g_n(i) = m\gamma + d_m$$

with suitable bounded sequences (c_m) and (d_m). This implies $\langle \pi_\sigma, \mathbf{x}_\sigma \rangle \le \gamma$. □

REFERENCES

[1] E. BEHRENDS. *Introduction to Markov Chains – with Special Emphasis on Rapid Mixing.* Vieweg Verlag, Wiesbaden 2000.

[2] L. BREIMAN. *Optimal gambling systems for favorable games.* Fourth Berkeley Symp. on Math. Stat. and Prob. **1** (1961), 65–78.

[3] E.B. DYNKIN, A.A. YUSHKEVICH. *Controlled Markov Processes.* Springer, Berlin-Heidelberg-New York, 1979.

[4] G.P. HARMER, D. ABBOTT. *Parrondo's paradox.* Statistical Science **14** (1999), 206–213.

[5] G.P. HARMER, D. ABBOTT. *Losing strategies can win by Parrondo's paradox.* Nature **402** (1999), 264.

[6] G.P. HARMER, D. ABBOTT, P. TAYLOR. *The paradox of Parrondo's games.* Proc. of the Royal Society **456** (1994), 247–259.

[7] G.P. HARMER, D. ABBOTT, P. TAYLOR, J. PARRONDO. *Brownian ratchets and Parrondo's games.* Chaos **11** (2001), 705–714.

[8] O. HERNÁNDEZ-LERMA, J.B. LASSERRE. *Discrete-Time Markov Control Processes.* Springer, Applications of Mathematics 30, Berlin, Heidelberg, New York, 1996.

[9] R. KARP. *A characterization of the minimal cycle mean in a digraph.* Discrete mathematics **23** (1978), 309–311.

[10] J.M.R. PARRONDO. *Juegos de azar paradójicos.* La Gaceta de la Real Sociedad Matemática Española **4** (2001), 355–365.

[11] R. PINSKY, M. SCHEUTZOW. *Some remarks an dexamples concerning the transience and recurrence of random diffusions.* Ann. Inst. Henri Poincaré **28** (1992), 519–536.

[12] S. RAHMANN. *Optimal adaptive atrategies for games of the Parrondo type.* Preprint.

[13] D.J. WHITE. *Markov Decision Processes.* John Wiley and Sons, 1993.

Freie Universität Berlin,
Fachbereich Mathematik und Informatik,
Arnimallee 2–6,
D-14 195 Berlin, Germany;
e-mail: behrends@math.fu-berlin.de