Zero-Sum Stochastic Games over the Field of Real Algebraic Numbers

Konstantin Avrachenkov, Vladimir Ejov, Jerzy Filar, Amir Moghaddam

HAL Id: hal-02413402

https://hal.inria.fr/hal-02413402

Submitted on 16 Dec 2019

ZERO-SUM STOCHASTIC GAMES OVER THE FIELD OF REAL ALGEBRAIC NUMBERS[*]

K. AVRACHENKOV[†], V. EJOV[‡], J. A. FILAR,[¶], AND A. MOGHADDAM,[‖][**]

Abstract. We consider a finite state, finite action, zero-sum stochastic games with data defining the game lying in the ordered field of real algebraic numbers. In both the discounted and the limiting average versions of these games we prove that the value vector also lies in the same field of real algebraic numbers. Our method supplies finite construction of univariate polynomials whose roots contain these value vectors. In the case where the data of the game are rational, the method also provides a way of checking whether the entries of the value vectors are also rational.

Key words. Stochastic games, ordered field property, algebraic numbers, algebraic variety, Gröbner basis, polynomial equations.

AMS subject classifications. 90D15

1. Introduction. Arguably, modern game theory was launched in 1928 by von Neumann (see [25]). His seminal paper proved that a finite, zero-sum, two person (matrix) game possesses a game-theoretic value and a pair of optimal strategies. Subsequently, Weyl (see [26]) supplied a simpler proof and, in addition, showed that if the entries of such a matrix game belong to an ordered field, then the value belongs to the same ordered field.

Stochastic games were introduced in 1953 by Shapley (see [20]). Shapley's formulation was analogous to what are now called discounted stochastic games. In these games the rewards at future stages are discounted by a factor $\beta \in [0, 1)$. However, in a remark, Shapley also observed that Weyl's ordered field property does not hold for stochastic games. In particular, when the data of these games lie in the field of rational numbers, the value vector need not be in the same field.

The latter remark, stimulated a whole line of research. Indeed, at least five, structured, classes of stochastic games have been identified and shown to possess the ordered field property, over the field $\mathbb{Q}$ of rational numbers. These classes include stochastic games of perfect information originally introduced by Gillette in 1957 [4], separable reward and state independent transition (SER-SIT) games [22], [16], single-controller games [15], switching-controller games [2] and additive reward and additive transition (ARAT) games [18]. However, a more comprehensive characterization of the ordered field property over $\mathbb{Q}$ proved to be challenging and was named as one of the open problems in the topic in a 1991 survey paper [17]. At that time the authors of [17] hoped to find an elegant set of necessary and sufficient conditions on the structure of the stochastic games defined over $\mathbb{Q}$ that will yield solutions in the same field. However, one of the consequences of the analysis presented below

[†]Inria Sophia Antipolis, `k.avrachenkov@inria.fr`
[‡]Collage of Science and Engineering Flinders University of South Australia Bedford Park SA5042 MSU, Faculty of Mechanics and Mathematics, Russia,119991, Moscow, GSP-1, 1 Leninskiye Gory
[¶]Centre for Applications in Natural Resource Mathematics, School of Mathematics and Physics, The University of Queensland, St Lucia, Queensland, 4072
[‖]`j.filar@uq.edu.au`
[**]`a.moghaddam1@uq.edu.au`

(see Section 5 and Example 3) is that this may occur only under rather exceptional data configurations and hence the restriction to the field of rationals is unnatural for stochastic games. This leads to the reformulation of the preceding problem to that of finding a natural, simple, ordered field containing $\mathbb{Q}$ over which the ordered field property will hold.

In this paper, we constructively solve the above problem over the ordered field $\mathbb{F}$ of real algebraic numbers. This is shown without any structural assumptions on the game and applies to both discounted and limiting average stochastic games that will be introduced later on. Arguably, $\mathbb{F}$ is the most natural ordered field to consider in view of the fact that the field of rational numbers $\mathbb{Q}$ was too small. This is because $\mathbb{F}$ is a finite degree field extension of $\mathbb{Q}$ is also countable and, in the sense of Rotman ([19], page 50), it is smallest field containing rationals. Our analysis exploits Gröbner basis methods as well as results reported in Szczechla et al. [23] and we also use similar notation to the latter paper. Perhaps, most importantly, our method supplies finite construction of univariate polynomials whose roots contain these value vectors. In the special case where the data of the game are rational, the method also provides a way of verifying whether the entries of the value vectors are also rational that is based on a classical divisor property of leading and constant term coefficients of polynomials.

We note that the asymptotic behavior of the value vector as $\beta \to 1$ has been studied by a number of authors. In an important, and related, contribution Bewley and Kohlberg [1] viewed Shapley's "optimality condition" as an elementary sentence in formal logic over the closed ordered field of real Puiseux series. These authors invoke a powerful theorem from mathematical logic, known as Tarski's principle, to conclude that in some neighborhood of $\beta = 1$, the value vector belongs to the field of real Puiseux series. The use of Tarski's principle, while ingenious, does not give insight into the manner in which fractional power series solutions arise naturally in this problem. After all, Puiseux series are objects of complex analysis encountered as branches of solutions of certain irreducible polynomials. Arguably, this has proved to be a difficulty for researchers in stochastic games, because Bewley and Kohlberg's result has become an important building block in subsequent developments. For instance, Mertens and Neyman [11] used it to prove their seminal theorem on the existence of the value vector in the limiting average stochastic games. This is why Szczechla et al. [23] supply a proof of Bewley and Kohlberg's Puiseux series expansion that is based on the Remmert-Stein Lemma of complex analytic varieties rather than on Tarski's principle.

We also note that our approach has similarities to the semialgebraic theory of stochastic games proposed in Milman [13] and continued by Neyman [14]. The latter permits analysis of sets defined by both polynomial equations and inequalities but it is not immediately clear how to apply Gröbner basis approach in that context. Since some works on complexity analysis of stochastic games (see e.g., [5], [3]) use the tools from algebraic and semi-algebraic geometries, we expect that the results of the present work may also be useful for this area of research.

Finally, we are grateful to two anonymous referees for pointing that the ordered field property for stochastic games over real algebraic numbers was already anticipated by Mertens and Zamir in [12] who allude to it in the process of demonstrating that in repeated games with incomplete information and data lying in the field of rationals, the entries of the value vector can be transcendental numbers. For two comprehensive recent surveys of stochastic games, the reader is referred to Jaśkiewicz and Nowak [6] and [7].

2. Definitions and preliminaries of matrix games. Any $m \times n$ real matrix $A = (a_{ij})_{i,j=1}^{m,n}$ can be regarded as a two-person, zero-sum matrix game with a_{ij} denoting the amount player II will pay player I if II chooses an action $j \in 1, 2, \cdots, n$ and I chooses an action $i \in 1, 2, \cdots, m$. A mixed (or randomized) strategy for player $I(II)$ in such a game is an m(n)-component probability vector x(y) whose $i^{th}(j^{th})$ entry $x_i(y_j)$ denotes the probability that player $I(II)$ will choose an action $i(j)$. A consequence of the celebrated "minimax theorem" for matrix games [25] is that there always exists a strategy pair (x^0, y^0) satisfying:

$$(2.1) \qquad x^T A y^0 \leq (x^0)^T A y^0 \leq (x^0)^T A y$$

for all mixed strategies $x(y)$ of player $I(II)$. The strategies (x^0, y^0) are then called *optimal strategies*, and the real number $val(A) := (x^0)^T A y^0$ is called the *value* of the matrix game A. It is well known that if $b_{ij} = ka_{ij} + c$, for all $i, j, k > 0$, and $B = (b_{ij})_{i,j=1}^{m,n}$, then $valB = kvalA + c$. Hence, there is no loss of generality in assuming that the value of a matrix game is either positive, or simply, not equal to zero.

A matrix game A is called *completely mixed* if all of its optimal strategies are strictly positive in every component. Extending earlier results of Kaplansky [8], Shapley and Snow [21] demonstrated the following result.

PROPOSITION 2.1. *If A is a matrix game and $valA \neq 0$, then A has a square invertible submatrix $\bar{A}$, called a* Shapley–Snow kernel, *such that*

(K1) $val\bar{A} = valA = \frac{\det(\bar{A})}{\sum_{ij} \bar{A}^{ij}} = \left(\mathbf{1}^T[\bar{A}^{-1}]\mathbf{1}\right)^{-1}$ *where $\bar{A}^{ij}$ denotes the $(i,j)^{th}$ cofactor of $\bar{A}$.*

(K2) *There is a pair $\left(x^0, y^0\right)$ of strategies for $\bar{A}$ which is optimal for $\bar{A}$ (after inserting zeroes) and satisfies $\left(x^0\right)^T = (valA)\,\mathbf{1}^T[\bar{A}^{-1}]$ and $y^0 = (valA)\,[\bar{A}^{-1}]\mathbf{1}$.*

REMARK 2.2. *If $valA \neq 0$, Lemma 1.2 of [23] shows that it is always possible to find a Shapley-Snow kernel $\bar{A}$ that is a completely mixed matrix game and, of course, still satisfies $(K1) - (K2)$, above. Following [23] we shall refer to such a kernel as* cmv-Shapley-Snow kernel, *or simply* cmv-kernel. *Their usefulness stems from the fact that the algebraic formula for their value $val\bar{A} = \left(\mathbf{1}^T[\bar{A}^{-1}]\mathbf{1}\right)^{-1}$ is invariant under small perturbations.*

3. Discounted stochastic games.

3.1. Definitions and preliminaries of stochastic games. A *stochastic game* as formulated by Shapley [20] is played in stages. At each stage, the game is in one of finitely many states, $s = 1, 2, \ldots, N$, in which players I and II are obliged to play a matrix game $R(s) = (r(s, a, b))_{a,b=1}^{m_s, n_s}$, once. The "law of motion" is defined by $p(s'|s, a, b)$, where the event $\{s'|s, a, b\}$ is the event that the game will enter state s' at the next stage given that at the current stage the state of the game is s, and that players I and II choose the ath row and the bth column of $R(s)$, respectively.

In general, players' strategies will depend on complete past histories. In this paper, however, we shall only be concerned with *stationary strategies*. We may represent a typical stationary strategy μ for player I by a "composite" vector, $\mu = (\mu(1), \mu(2), \ldots, \mu(N))$, where each $\mu(s)$ is a probability distribution on $\{1, 2, \ldots, m_s\}$. Player II's stationary strategies ν are similarly defined.

It should be clear that once we specify the initial state s and a strategy pair (μ, ν) for players I and II, we implicitly define a probability distribution over all sequences of states and actions which can occur during the game and consequently over all sequences of payoffs to player I. In particular, if the random variable $\mathcal{R}_t$ denotes the payoff to player I at stage t, then the expected value of $\mathcal{R}_t$ given s and (μ, ν)

$$(3.1) \qquad\qquad E_{\mu\nu s}(\mathcal{R}_t) := E\{\mathcal{R}_t|\mu, \nu, s\}$$

is well defined. The *β-discounted stochastic game* Γ_β is then the game in which the overall payoff, normalized by a factor of $1 - \beta$, resulting from the strategy pair (μ, ν) and a starting state s is evaluated according to

$$v_\beta(\mu, \nu, s) := \sum_{t=1}^{\infty} \beta^{t-1}(1 - \beta)E_{\mu\nu s}(\mathcal{R}_t),$$

where $\beta \in (0, 1)$ is called the discount factor. The number $v_s(\beta)$ is called the *value* of the game Γ_β starting in state s if $v_s(\beta) = \sup_\mu \inf_\nu v_\beta(\mu, \nu, s) = \inf_\nu \sup_\mu v_\beta(\mu, \nu, s)$. The vector $\mathbf{v}(\beta) = (v_1(\beta), v_2(\beta), \ldots, v_N(\beta))$ is called the *value vector*. Furthermore, the pair (μ^0, ν^0) is called an optimal strategy pair for players I and II if, for each starting state s,

$$v_s(\beta) = v_\beta(\mu^0, \nu^0, s).$$

The existence of the value vector and of a pair of *optimal stationary strategies* was proved in 1953 in Shapley's seminal paper on the subject [20]. A key element in Shapley's proof was the construction of N auxiliary matrix games $R_\beta(s, \mathbf{u})$ that depend on an arbitrary vector $\mathbf{u} = (u_1, u_2, \ldots, u_N) \in \mathbb{R}^\mathbf{N}$ according to

$$(3.2) \qquad R_\beta(s, \mathbf{u}) = \left[(1 - \beta)r(s, a, b) + \beta \sum_{s'=1}^{N} p(s'|s, a, b)u_{s'}\right]_{a,b=1}^{m_s, n_s}.$$

In view of the fact that the value of a matrix game always exists, it is possible to define, for each $\beta \in (0, 1)$, an operator $T_\beta : \mathbb{R}^\mathbf{N} \longrightarrow \mathbb{R}^\mathbf{N}$, the sth component of which is given by

$$(3.3) \qquad\qquad [T_\beta(\mathbf{u})]_s := val[R_\beta(s, \mathbf{u})].$$

This operator is a contraction operator in the sup-norm with contraction constant $\leq \beta$; see [20]. It therefore follows from Banach's fixed-point theorem that there exists a unique fixed point $\mathbf{v}(\beta)$ of T_β; that is,

$$(3.4) \qquad \mathbf{v}(\beta) = T_\beta(\mathbf{v}(\beta)).$$

Also, any set of optimal strategy pairs for $R_\beta(s, \mathbf{v}(\beta))$ $(s = 1, 2, \ldots, N)$ can be shown to form an optimal strategy pair for Γ_β.

Assume $v_s(\beta) \neq 0$ for all β and s. In view of Proposition 2.1 we know that for each fixed $\beta \in (0, 1)$ and each $\mathbf{u}$ close enough to $\mathbf{v}(\beta)$ there exist cmv-kernels $\bar{R}_\beta(s, \mathbf{u})$ such that the fixed-point equation above reduces to

$$(3.5) \qquad v_s(\beta) = \frac{\det(\bar{R}_\beta(s, \mathbf{v}(\beta)))}{\sum_i \sum_j [\bar{R}_\beta(s, \mathbf{v}(\beta))]^{ij}} \quad \text{for each } s = 1, 2, \ldots, N,$$

where $[V]^{ij}$ denotes the $(i, j)^{th}$ cofactor of a matrix V.

If we now transform the above equations to

$$(3.6) \qquad v_s(\beta) \left\{ \sum_i \sum_j [\bar{R}_\beta(s, \mathbf{v}(\beta))]_{ij} \right\} - \det(\bar{R}_\beta(s, \mathbf{v}(\beta))) = 0, \ \ s = 1, 2, \ldots, N,$$

then for each fixed combination of the locations of these kernels we can regard this system as being a system of polynomials in the variables $z_0 := \beta, z_1 := v_1(\beta), z_2 := v_2(\beta), \ldots, z_N := v_N(\beta)$ of the form

$$(3.7) \qquad \begin{cases} f_1(z_0, \ldots, z_N) = 0 \\ \qquad \vdots \\ f_N(z_0, \ldots, z_N) = 0. \end{cases}$$

Consider any fixed collection of non-empty subsets of actions $K_s \subseteq \{1, \ldots, m_s\}$ and $L_s \subseteq \{1, \ldots, n_s\}$ such that $card \ K_s = card \ L_s$ for all $s = 1, \ldots, N$. Thus we have a finite sequence of sets $\kappa = (K_1, L_1, \ldots, K_N, L_N)$. By restricting the players' actions in each state s to K_s and L_s, we obtain a discounted stochastic game, $\Gamma_{\beta,\kappa}$ with auxiliary Shapley games

$$(3.8) \qquad R_{\beta,\kappa}(s, \mathbf{u}) = [R_\beta(s, \mathbf{u})_{ij}]_{i \in K_s, j \in L_s}.$$

For fixed κ, let

$$(3.9) \qquad \bar{R}_\beta(s, \mathbf{u}) = R_{\beta,\kappa}(s, \mathbf{u}) = [R_\beta(s, \mathbf{u})_{ij}]_{i \in K_s, j \in L_s}.$$

be the sth cmv-kernel associated with κ. Clearly the system of polynomials (3.7) depends on κ but this dependence is suppressed in order to simplify already complicated notation. However, it is important to note that $K = \{\kappa | \text{defining cmv-kernels}\}$ is finite. Hence, there are at most $|K|$ systems of polynomials of the form (3.7) to consider, where $|K|$ denotes the cardinality of K.

3.2. Discounted stochastic games over the ordered field of real algebraic numbers. Suppose now that all the data of the discounted stochastic game Γ_β, namely, $r(s,a,b)$, $p(s'|s,a,b)$ (for all, s, s', a, b) and β lie in the field $\mathbb{F}$ of real algebraic numbers. Besides purely mathematical interest, it is easy to conceive of applications of stochastic games with geometric structure where, say, real algebraic rewards arise naturally. For instance, imagine a situation where, in state 1, players I and II have two actions corresponding to choices of angles θ_1 and θ_2 from the set $\{\frac{\pi}{6}, \frac{\pi}{4}\}$. If these actions also correspond to points on the unit circle, then their choices determine triangles Δ_i whose areas are given by $\frac{1}{2}\cos(\theta_i)\sin(\theta_i)$, for $i = 1, 2$. If we defined the rewards in the first state as the difference between areas of Δ_2 and Δ_1, these would include real algebraic numbers.

From the outset, we wish to outline a very elegant - but non-constructive - proof of the fact that entries of the value vector $\mathbf{v}(\beta)$ of Γ_β lie in $\mathbb{F}$ that is based on a deep result from the field of mathematical logic that is known as *Tarski's Principle* [24] which states that:

An elementary sentence which is valid over one real closed field is valid over every real closed field.

Now, Bewley and Kohlberg [1] showed that Shapley's theorem captured in (3.4) constitutes a valid elementary sentence over the reals. Since the field $\mathbb{F}$ is known to be real closed, the result follows immediately.

We do not go into all details of the above argument because the main objective of this study, for Γ_β, is to explicitly exhibit the construction of decoupled, bivariate, polynomials whose roots contain the entries of $\mathbf{v}(\beta)$ and to demonstrate the usefulness of Gröbner basis techniques in the process. Furthermore, our approach lends itself to extension to the limiting average stochastic games covered in the next section. Since there is no analogue for Shapley's theorem in that case, it is not clear whether Tarski's principle could be used there.

Returning to the main line of discussion, suppose now that the Shapley-Snow kernels K have been fixed and chosen correctly in the sense that $z_s = v_s(\beta)$; $s = 1, 2, \ldots, N$ satisfy both (3.5) and (3.6) for $z_0 = \beta$. That is, $(z_0, \mathbf{z})$, with $\mathbf{z} = [z_1, ..., z_N]$, is among the real-valued roots of the system of coupled polynomial equations (3.7) which we now re-write as

$$(3.10) \qquad \begin{cases} f_1(z_0, \mathbf{z}) = 0 \\ \quad\vdots \\ f_N(z_0, \mathbf{z}) = 0. \end{cases}$$

The zero-set of (3.10) constitutes an algebraic variety, W, which can be difficult to analyse. Fortunately, the Gröbner basis methods (e.g., see the excellent book of Adams and Loustaunau [9]) enable us to, instead, examine a simpler variety $W_d \supset W$ (of the same dimension as W) in a neighbourhood of the solution of the stochastic game that corresponds to the zero-set of suitably constructed decoupled polynomials

$$(3.11) \qquad \begin{cases} g_1(z_0, z_1) = 0 \\ \quad\vdots \\ g_N(z_0, z_N) = 0, \end{cases}$$

where each $g_s(z_0, z_s)$ is a bivariate polynomial in only z_0 and z_s. That is, the zeroes of (3.11) contain the zeroes of (3.10) and hence the solution of the stochastic game.

The above simplification follows from the next result which can also be seen as a constructive extension of Lemmata $4.1 - 4.2$ and Theorem 4.3 proved in Szczechla et al. [23]. Interestingly, perhaps, the latter results exploit the fact that the contraction property of the Shapley operator T_β ensures that the Jacobian of the map corresponding to the fixed point equation is nonsingular.

THEOREM 3.1. *Consider a discounted stochastic game Γ_β with all data lying in the ordered field of real algebraic numbers. There exist bivariate, decoupled, polynomials $g_s(z_0, z_s)$ in variables z_0, z_s, for $s = 1, 2, \ldots, N$, whose zeroes define the variety W_d containing the solution of the stochastic game, for each $z_0 = \beta \in (0, 1)$. Furthermore, the corresponding $z_s = v_s(\beta)$ is a real algebraic number for each $s = 1, 2, \ldots, N$.*

Proof. Let $H_\beta := \{(z_0, z_1, \ldots, z_N) \in \mathbb{C}^{N+1} | z_0 = \beta\}$ and consider the variety W that is the zero set of (3.10). From Lemmata $4.1 - 4.2$ and the proof of Theorem 4.3 in [23] it follows that W has constant dimension 1 and that it intersects with H_β at a discrete subset of $\mathbb{C}^{N+1}$, that is, for any $\beta \in (0, 1)$

$$(3.12) \qquad \dim(W \bigcap H_\beta) = 0.$$

Now, apply the Buchberger algorithm to find the Gröbner basis GB_1 of the polynomials

$$\{f_1(z_0, \mathbf{z}), \ldots, f_N(z_0, \mathbf{z})\},$$

with the lexicographic term order $z_1 \prec z_2 \prec \ldots \prec z_N$. The last polynomial $g^m(z_0, \mathbf{z}) \in GB_1$ contains the least number of the variables $z_1, \ldots, z_N$ and has coefficients that are polynomial functions of z_0. It is certainly possible to choose some $z_0 = \beta^* \in (0, 1)$ that is not a root of any of these coefficient polynomials. But by (3.12) we have that $W_{\beta^*} := W \bigcap H_{\beta^*}$ is a zero dimensional variety and hence by Corollary 2.2.11 in [9] the polynomial $g^m(\beta^*, \mathbf{z})$ is univariate in z_1. Now, returning to arbitrary z_0 in place of β^*, define $g_1(z_0, z_1) := g^m(z_0, \mathbf{z})$, a bivariate polynomial in z_0, z_1.

Applying Buchberger algorithm again but with the term order $z_2 \prec z_3 \prec \ldots \prec z_N \prec z_1$ will, analogously, yield a new Gröbner basis GB_2 and a bivariate polynomial $g_2(z_0, z_2)$. Continuing in this fashion yields the polynomial system (3.11).

Note that if $(z_0, \mathbf{z}) = (z_0, z_1, \ldots, z_n)$ satisfies (3.10), then $g_s(z_0, z_s) = 0$, for every $s = 1, 2, \ldots, N$ since each g_s is part of some Gröbner basis of $\{f_1, \ldots, f_N\}$. Thus $W \subset W_d$. In particular, since $(\beta, \mathbf{v}(\beta)) \in W$ by Lemma 4.2 and Theorem 4.3 in [23], it follows that $(\beta, \mathbf{v}(\beta)) \in W_d$ and $g_s(\beta, v_s(\beta)) = 0$, for $s = 1, \ldots, N$.

Furthermore, since each g_s is a bivariate polynomial in z_0, z_s with coefficients that are algebraic numbers, it becomes a univariate polynomial in z_s with algebraic coefficients, when evaluated at $z_0 = \beta$ that is algebraic. Thus each $g_s(\beta, z_s)$ is a univariate polynomial with algebraic coefficients, hence its roots are also algebraic numbers, since the field of algebraic numbers is algebraically closed. Of course, $v_s(\beta)$ is real for every s by Shapley's fixed point equation (3.4). This completes the proof. $\square$

COROLLARY 3.2. *Consider a discounted stochastic game Γ_β with all data lying in the ordered field of real algebraic numbers $\mathbb{F}$. Then there exists a pair of optimal stationary strategies (μ^0, ν^0) with all entries lying in $\mathbb{F}$.*

Proof. Once we substitute $\mathbf{u} = \mathbf{v}(\beta)$ with entries lying in $\mathbb{F}$, the Shapley matrix games $R_\beta(s, \mathbf{v}(\beta)$ also have all entries that lie in $\mathbb{F}$. The statement of the corollary

now follows immediately from the contributions of Weyl [26] and Shapley and Snow [21], see also Proposition 2.1. ☐

REMARK 3.3. *It follows from the finiteness of the Buchberger algorithm (and the fact that there are only $|K|$ kernel selections κ), that the correct set of decoupled polynomials (3.11) can be found by a finite algorithm. Of course, the latter could still be of great complexity. Furthermore, the problem of finding the roots of these polynomials is, in general, not solvable by finite algorithms.*

4. Limiting average stochastic games over the ordered field of algebraic numbers. We now consider a stochastic game where the sequence of single stage expected rewards $\{E_{\mu\nu s}(R_t)\}_{t=0}^{\infty}$ is aggregated according to

$$v(\mu, \nu, s) = \lim_{\tau \to \infty} \inf \left(\frac{1}{\tau + 1} \right) \sum_{t=0}^{\tau} E_{\mu\nu s}(R_t).$$

Such a game is called a *limiting average (or Cesàro average) stochastic game* and will be denoted by Γ_{α}. It is still a zero-sum game to which the minimax solution concept applies. Indeed, the existence of the value vector $\mathbf{v} = (v_1, \ldots, v_N)^T$ of Γ_{α} was proved by Mertens and Neymann [11]. These authors showed that

(4.1) $$\mathbf{v} = \lim_{\beta \to 1^-} \mathbf{v}(\beta),$$

where $\mathbf{v}(\beta)$ is the value vector of the discounted game Γ_{β}. They exploited the following Puiseux series expansion of $\mathbf{v}(\beta)$ that is due to Bewley and Kohlberg [1]

(4.2) $$\mathbf{v}(\beta) = \sum_{r=0}^{\infty} c_r(s)(1-\beta)^{\frac{r}{M_s}} = c_0(s) + \delta(s, \beta),$$

for each $s = 1, 2, \ldots, N$. Here M_s is a natural number and $\delta(s, \beta) \to 0$ as $\beta \to 1$. Clearly, (4.1) and (4.2) imply that the value of the Γ_{α} game, starting at state s, is given by

$$\mathbf{v}_s = c_0(s), \ s = 1, 2, \ldots, N.$$

Unfortunately, we cannot immediately conclude from (4.1) that entries of $\mathbf{v}$ lie in the field $\mathbb{F}$ when the data of Γ_{α} lie in $\mathbb{F}$. This is because a limit of algebraic numbers need not be algebraic. Hence, we needed to establish the next result.

THEOREM 4.1. *Consider a limiting average stochastic game Γ_{α} with all data lying in the ordered field of real algebraic numbers. Then the value vector $\mathbf{v}$ of Γ_{α} has entries $\mathbf{v}_s$ that are also real algebraic numbers, for each $s = 1, 2, \ldots, N$.*

Proof. In view of (4.1) and (4.2), it is sufficient to prove that $c_0(s)$ is an algebraic number for each $s = 1, 2, \ldots, N$. Without loss of generality, we supply a proof for the case of $s = 1$ only.

The proof will be asymptotic, via a sequence of discounted games Γ_{β_k} such that $\beta_k \to 1^-$. In particular, $\{\beta_k\}_{k=0}^{\infty}$ is such that, for each k, there is a selection of kernels $\kappa_k \in K$ such, that the polynomials (3.10) define the variety W^k containing $(\beta_k, \mathbf{v}(\beta_k))$, where $\mathbf{v}(\beta_k)$ is the value vector of the game Γ_{β_k}. However, since $|K| < \infty$

at least one of these kernels, say $\hat{\kappa}$, must repeat itself infinitely often. Again, without loss of generality, we may assume that the entire sequence $\{\beta_k\}_{k=0}^{\infty}$ corresponds to the same kernel selection $\hat{\kappa}$ and hence the same variety $\hat{W}$. Now, for the term order $z_1 \prec z_2 \prec \ldots \prec z_N$ find the Gröbner basis of $\{f_1, \ldots, f_N\}$, and the decoupled polynomial $g_1(z_0, z_2)$, as in the proof of Theorem 3.1. We know that for each $k = 0, 1, 2, \ldots$

$$(4.3) \qquad g_1(\beta_k, v_1(\beta_k)) = 0.$$

Hence $g_1(\beta_k, z_1)$ can be factored as

$$(4.4) \qquad g_1(\beta_k, z_1) = (z_1 - v_1(\beta_k))q_1(\beta_k, z_1),$$

where $q_1(\beta_k, z_1)$ can be obtained by the long division of $g_1(\beta_k, z_1)$ by the linear factor $(z_1 - v_1(\beta_k))$. In case $v_1(\beta_k)$ is a root of multiplicity $t > 1$, the factor $(z_1 - v_1(\beta_k))^{t-1}$ would be included in a factorization of $g_1(\beta_k, z_1))$ into its irreducible factors. Substituting (4.2) into (4.4) we obtain for each $k = 0, 1, \ldots$

$$(4.5) \qquad g_1(\beta_k, z_1) = (z_1 - c_0(1))q_1(\beta_k, z_1) - \delta(1, \beta_k)q_1(\beta_k, z_1),$$

where $q_1(z_0, z_1)$ is a bivariate polynomial in z_0, z_1 that has coefficients that are algebraic numbers.

Next, with $z_0 = \beta$, we can rewrite the latter polynomial as $q_1(\beta, z_1) = (1 - \beta)^{\ell}\bar{q}_1(\beta, z_1)$, where $\bar{q}_1(\beta, z_1)$ does not have a root at $\beta = 1$. Hence, since $\beta_k \to 1$, we have that $\bar{q}_1(\beta_k, z_1) \to \bar{q}_1(1, z_1) = q_1^*(z_1)$, where the latter is a univariate polynomial in z_1 that is free of the parameter $z_0 = \beta$.

Now, consider first, the more generic case of $\ell = 0$. Equation (4.5) becomes

$$(4.6) \qquad g_1(\beta_k, z_1) = (z_1 - c_0(1))\bar{q}_1(\beta_k, z_1) - \delta(1, \beta_k)\bar{q}_1(\beta_k, z_1),$$

for each $k = 0, 1, \ldots$. Passing to the limit with respect to k we obtain

$$(4.7) \qquad \lim_{k\to\infty} [g_1(\beta_k, z_1)] = g_1(1, z_1) = (z_1 - c_0(1))q_1^*(z_1),$$

since $\lim_{k\to\infty} [\delta(1, \beta_k)\bar{q}_1(\beta_k, z_1)] = 0$ and $q_1^*(z_1) \neq 0$. But by (4.3) at $z_1 = v_1(\beta_k)$, $g_1(\beta_k, v_1(\beta_k)) = 0$ for each k, so that the left side of (4.7) is zero. Thus, in the limit

$$(4.8) \qquad 0 = g_1(1, z_1) = (z_1 - c_0(1))q_1^*(z_1)$$

and hence $c_0(1) = v_1$ is an algebraic number.

To complete the proof, we next consider the case where ℓ is a positive integer. In this case we have

$$(4.9) \qquad g_1(\beta_k, z_1) = (z_1 - v_1(\beta_k))q_1(\beta_k, z_1) = (z_1 - v_1(\beta_k))(1 - \beta_k)^{\ell}\bar{q}_1(\beta_k, z_1),$$

for each $k = 0, 1, \ldots$. Set $\bar{g}_1(\beta, z_1) := (z_1 - v_1(\beta))\bar{q}_1(\beta, z_1)$, which is a bivariate polynomial in (β, z_1) with coefficients that are algebraic numbers. By continuity of polynomials, its limit $\bar{g}_1(1, z_1)$ as $\beta \to 1$ is a polynomial in z_1 with coefficients that are algebraic numbers. But, clearly, we still have that $0 = \bar{g}_1(\beta_k, v_1(\beta_k))$, for each $k = 0, 1, \ldots$, hence in the limit $0 = \bar{g}_1(1, v_1) = \bar{g}_1(1, c_0(1))$ by (4.1) and (4.2). Hence, $c_0(1)$ is a root of a polynomial with algebraic coefficients and is, therefore, an algebraic number. $\square$

5. Stochastic games over the rationals. In this section we comment briefly on the special case where both the data and the solution of the discounted stochastic game lie in the field of rational numbers $\mathbb{Q}$. Evidently, the system of decoupled polynomials (3.11) resulting from the repeated application of Gröbner basis technique can be used to search for possible rational solutions with the help of a classical result concerning rational roots of univariate polynomials with integer coefficients. In the terminology of Section 4, consider the sth polynomial equation $g_s(z_0, z_s) = 0$ of the system (3.11) with $z_0 = \beta \in \mathbb{Q}$. Without loss of generality assume that this equation has been normalized so that the polynomial $g_s(\beta, z_s)$ has integer coefficients. Namely, the latter is now a univariate polynomial of the form

$$(5.1) \qquad g_s(z_s) = c_{n_s}^s z_s^{n_s} + c_{n_s-1}^s z_s^{n_s-1} + \ldots + c_1^s z_s^1 + c_0^s,$$

where all the coefficients are integers. By the previously mentioned classical result (e.g., see [10] p. 171) if $z_s = \frac{p}{q} \in \mathbb{Q}$ is a root of $g_s(z_s) = 0$, then

$$(5.2) \qquad q \mid c_{n_s}^s \quad \& \quad p \mid c_0^s.$$

That is q, p must be divisors of $c_{n_s}^s$, c_0^s, respectively.

In view of the above there is a natural, finite, method of searching for rational roots of $g_s(z_s) = 0$, based on prime factorization of natural numbers. In particular, assume that we have a prime number representation of $c_{n_s}^s$, c_0^s as

$$(5.3) \qquad c_{n_s}^s = \prod_{j=1}^{m_s} (x_j)^{\alpha_j} \quad \& \quad c_0^s = \prod_{j=1}^{m_0} (y_j)^{\gamma_j}.$$

Now, the primes p, q defining the rational root $z_s = \frac{p}{q}$ must be of the form

$$(5.4) \qquad q = \prod_{j=1}^{m_s} (x_j)^{\theta_j}, \ \ \theta_j \le \alpha_j \quad \& \quad p = \prod_{j=1}^{m_0} (y_j)^{\tau_j}, \ \ \tau_j \le \gamma_j.$$

Clearly, an exhaustive search on $\theta_j \le \alpha_j$ and $\tau_j \le \gamma_j$ will supply all possible candidates for rational roots of $g_s(z_s) = 0$. Simple evaluation of $g_s(z_s)$ at these values will identify rational roots of that polynomial and hence potential rational candidates for the value of the stochastic game beginning in state s, for each $s = 1, \ldots, N$. Of course, the latter have to be substituted into Shapley's equations (3.4) to verify that they, indeed, form the value vector of the stochastic game. Example 3, in Section 6 illustrates that, generally, rational solutions arise only under very exceptional parameter configurations.

6. Conclusions and examples. As hinted in the introduction, we suggest that Theorems 4.1 and 5.1 provide a rather comprehensive characterization to the ordered field property problem posed in [17]. This is because the field $\mathbb{F}$ of algebraic numbers is still a countable ordered field that includes rationals $\mathbb{Q}$ and it is unlikely that any smaller extension of rationals would suffice. We have demonstrated that the Gröbner basis methods, in principle, allow us to identify exactly polynomials (with integer coefficients) whose roots contain the values of stochastic games Γ_β, for $s = 1, \ldots, N$. However, in general, these roots cannot be computed exactly in terms of radicals. This is illustrated in Example 2, below, where our decoupled polynomials $g_s(\beta, z_s)$

are quintic. Hence by the classical Abel-Ruffini theorem (e.g., see Theorem 75 and the following remark in [19], page 75), no radical solutions exist for these polynomials.

Let us now comment briefly on the structured classes of stochastic games that are known to possess the ordered field property, over the field $\mathbb{Q}$ of rational numbers. These classes include stochastic games of perfect information originally introduced by Gillette in 1957 [4], separable reward and state independent transition (SER-SIT) games [22], [16], single-controller games [15], switching-controller games [2] and additive reward and additive transition (ARAT) games [18]. For the sake of completeness, we briefly introduce these classes of games and supply an example of one of them which exposes their simple algebraic structure in the context of Theorem 4.1.

A *single controller* stochastic game is a game in which the transition probabilities satisfy the following:

$$p(s'|s, a, b_1) = p(s'|s, a, b_2) \ \forall \ s, s' \in \mathbf{S} \ \& \ \forall a \in \mathbf{A}(s), \ b_1, \ b_2 \in \mathbf{B}(s).$$

A *switching controller* stochastic game is a generalization of the above for which the state space S can be partitioned into disjoint subsets S^1 and S^2 such that $\forall \ s' \in S$:

$$p(s'|s, a, b_1) = p(s'|s, a, b_2) \ \forall \ s \in S^1 \ \& \ \forall \ a \in \mathbf{A}(s), \ b_1, b_2 \in \mathbf{B}(s),$$
$$p(s'|s, a_1, b) = p(s'|s, a_2, b) \ \forall \ s \in S^2 \ \& \ \forall \ a_1, a_2 \in \mathbf{A}(s), \ b \in \mathbf{B}(s).$$

Incidentally, the *perfect information* game is equivalent to a switching controller game in which one of the players has only a single action in each state.

The *ARAT* stochastic game is one in which both the rewards and the transitions can be written as the sum of a term determined by player one and a term determined by player two. That is:

$$r(s, a, b) = r_1(s, a) + r_2(s, b) \ \forall \ s \in S \ \& \ \forall \ a \in \mathbf{A}(s), \ b \in \mathbf{B}(s),$$
$$p(s'|s, a, b) = p_1(s'|s, a) + p_2(s'|s, b) \ \forall \ s, s' \in S \ \& \ \forall \ a \in \mathbf{A}(s), \ b \in \mathbf{B}(s).$$

Finally, the *SER-SIT* stochastic game is one for which the rewards and transition probabilities have the following structure for some scalars $c(1), c(2), \ldots, c(N)$:

$$r(s, a, b) = c(s) + r(a, b) \ \forall \ s \in S \ \& \ \forall \ a \in \mathbf{A}(s), \ b \in \mathbf{B}(s)$$
$$p(s'|s_1, a, b) = p(s'|s_2, a, b) \ \forall \ s', s_1, s_2 \in S \ \& \ \forall \ a \in \mathbf{A}(s) \ b \in \mathbf{B}(s).$$

Next, we illustrate our approach with two simple examples: the first is a switching controller game and the second is a reward-diagonal game that does not belong to any of the above structured classes.

Example 1: We consider the switching controller stochastic game as illustrated in Figure 6.1. Its state space $S = S_1 \bigcup S_2 = \{1\} \bigcup \{2\}$. Each player has three actions in each state, that is, $\mathbf{A}(1) = \mathbf{A}(2) = \{a_1, a_2, a_3\}$ and $\mathbf{B}(1) = \mathbf{B}(2) = \{b_1, b_2, b_3\}$. For example, in Figure(6.1a) if players one and two choose the first column and row, respectively, the immediate reward will be -2 for player one and consequently $+2$ for player two. Furthermore, at the next stage the game stays in state 1 with probability $\frac{3}{10}$ and moves to state 2 with probability $\frac{7}{10}$.

We solved this game in MAPLE symbolic computing environment. In particular, postponing the normalization by the factor $(1 - \beta)$, Shapley's auxiliary matrix games

were evaluated and equations (3.5) were manipulated to arrive at the system of two polynomial equations of the form (3.6). The latter are (with $z_0 = \beta$):

$$f_1(\beta, z_1, z_2) = 32 - 12z_1 + \frac{22}{5}\beta z_1 + \frac{38}{5}\beta z_2$$

$$f_2(\beta, z_1, z_2) = 9z_2 - \frac{81}{20}\beta z_1 - \frac{99}{20}\beta z_2.$$

-2		4		6	
	$\left(\frac{3}{10}, \frac{7}{10}\right)$		$\left(\frac{3}{10}, \frac{7}{10}\right)$		$\left(\frac{3}{10}, \frac{7}{10}\right)$
6		2		0	
	$(0, 1)$		$(0, 1)$		$(0, 1)$
4		2		2	
	$\left(\frac{4}{5}, \frac{1}{5}\right)$		$\left(\frac{4}{5}, \frac{1}{5}\right)$		$\left(\frac{4}{5}, \frac{1}{5}\right)$

(a) State One

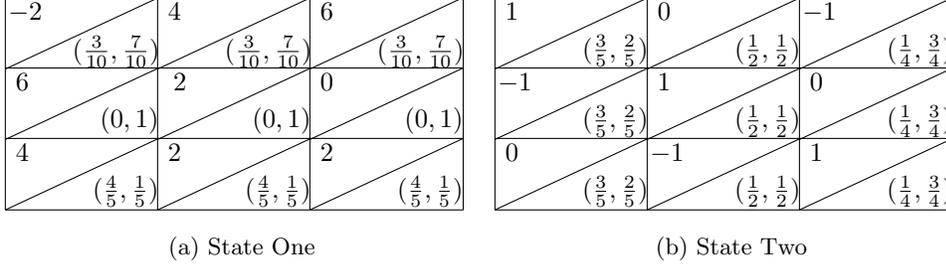| 1		0		-1	
	$\left(\frac{3}{5}, \frac{2}{5}\right)$		$\left(\frac{1}{2}, \frac{1}{2}\right)$		$\left(\frac{1}{4}, \frac{3}{4}\right)$
-1		1		0	
	$\left(\frac{3}{5}, \frac{2}{5}\right)$		$\left(\frac{1}{2}, \frac{1}{2}\right)$		$\left(\frac{1}{4}, \frac{3}{4}\right)$
0		-1		1	
	$\left(\frac{3}{5}, \frac{2}{5}\right)$		$\left(\frac{1}{2}, \frac{1}{2}\right)$		$\left(\frac{1}{4}, \frac{3}{4}\right)$

(b) State Two

Fig. 6.1: A switching controller stochastic game.

To compute the Gröbner bases, we first used the lexicographic term order $z_1 \prec z_2$ and then applied the same process again but with the term order $z_2 \prec z_1$. From the two Gröbner bases GB_1, GB_2 so obtained, we selected the two bivariate polynomials:

- For $z_1 \prec z_2$:

$$(6.1) \qquad g_1(\beta, z_1) = 5\beta^2 z_1 + 55\beta z_1 - 88\beta - 60z_1 + 160$$

- For $z_2 \prec z_1$:

$$(6.2) \qquad g_2(\beta, z_2) = 5\beta^2 z_2 + 55\beta z_2 + 72\beta - 60z_2.$$

It is not a coincidence, that the latter are linear functions in z_1 and z_2, respectively. Evidently, this is a common feature for all the five structured classes of stochastic games listed above. The zeroes of the above polynomials occur when

$$(6.3) \qquad z_1 = \frac{8}{5}\frac{11\beta - 20}{(\beta + 12)(\beta - 1)},$$

$$(6.4) \qquad z_2 = -\frac{72}{5}\frac{\beta}{(\beta + 12)(\beta - 1)}.$$

which are clearly positive for $\beta \in [0, 1)$. They are also rational for rational β and algebraic for algebraic β. Now, for $\beta \in (0, 1)$ the entries of the value vector $\mathbf{v}_\beta$ of Γ_β are obtained by normalization by the factor $(1 - \beta)$, namely

$$(6.5) \qquad v_1(\beta) = (1 - \beta)z_1 = \frac{160 - 88\beta}{5(\beta + 12)},$$

$$(6.6) \qquad v_2(\beta) = (1 - \beta)z_2 = \frac{72\beta}{5(\beta + 12)}.$$

State One:

18 $(0,1)$	0 $\left(\frac{1}{5}, \frac{4}{5}\right)$	0 $(1,0)$
0 $\left(\frac{2}{5}, \frac{3}{5}\right)$	12 $\left(\frac{1}{3}, \frac{2}{3}\right)$	0 $\left(\frac{3}{5}, \frac{2}{5}\right)$
0 $\left(\frac{3}{5}, \frac{2}{5}\right)$	0 $\left(\frac{2}{5}, \frac{3}{5}\right)$	6 $(0,1)$

(a) State One

State Two:

3 $\left(\frac{1}{5}, \frac{4}{5}\right)$	0 $\left(\frac{7}{10}, \frac{3}{10}\right)$	0 $\left(\frac{3}{10}, \frac{7}{10}\right)$
0 $\left(\frac{1}{2}, \frac{1}{2}\right)$	2 $\left(\frac{2}{5}, \frac{3}{5}\right)$	0 $\left(\frac{2}{5}, \frac{3}{5}\right)$
0 $\left(\frac{3}{10}, \frac{7}{10}\right)$	0 $\left(\frac{1}{5}, \frac{4}{5}\right)$	1 $(0,1)$

(b) State Two

Fig. 6.2: A reward-diagonal stochastic game.

Example 2: The next example is related to the same approach for a reward-diagonal stochastic game with data as given in Figure 6.2. By following the same method as in previous example we derived Shapley's polynomials (3.6). The latter are (with $z_0 = \beta$):

$$f_1(\beta, z_1, z_2) = -\frac{2}{75}\beta^2 z_1^3 + \frac{4}{75}\beta^2 z_1^2 z_2 - \frac{89}{5}\beta z_1^2 - \frac{2}{75}z_1\beta^2 z_2^2 + \frac{89}{5}z_1\beta z_2 + 396z_1 + \frac{1}{375}\beta^3 z_1^2 z_2 + 9\beta^2 z_1^2$$
$$- \frac{11}{375}\beta^3 z_1 z_2^2 - \frac{1}{5}\beta^2 z_1 z_2 - 144\beta z_1 + \frac{7}{375}\beta^3 z_2^3 - \frac{44}{5}\beta^2 z_2^2 - 252\beta z_2 - 1296 + \frac{1}{125}\beta^3 z_1^3$$

$$f_2(\beta, z_1, z_2) = -\frac{6}{25}z_1\beta^2 z_2^2 - 2z_1\beta z_2 + \frac{3}{25}\beta^2 z_2^3 + 2\beta z_2^2 + 11z_2 + \frac{3}{25}\beta^2 z_1^2 z_2 + \frac{33}{500}\beta^3 z_1^2 z_2 + \frac{69}{100}\beta^2 z_1^2$$
$$+ \frac{27}{500}\beta^3 z_1 z_2^2 + \frac{31}{50}\beta^2 z_1 z_2 - \frac{8}{5}\beta z_1 - \frac{31}{500}\beta^3 z_1^3 - \frac{29}{500}\beta^3 z_2^3 - \frac{131}{100}\beta^2 z_2^2 - \frac{47}{5}\beta z_2 - 6.$$

To compute the Gröbner bases, we first used the lexicographic term order $z_1 \prec z_2$ and then applied the same process again but with the term order $z_2 \prec z_1$. From the two Gröbner bases GB_1, GB_2 so obtained, we selected the two bivariate polynomials:

- For $z_1 \prec z_2$:

$$
\begin{aligned}
g_1(\beta, z_1) =\ & \big(573788072\beta^8 - 220687720\beta^7 - 7503382480\beta^6 + 20744645680\beta^5 \\
& -23613586040\beta^4 + 12667475128\beta^3 - 2648252640\beta^2\big) z_1^5 + \big(17339824845\beta^7 \\
& +777484096540\beta^6 - 5051043271810\beta^5 + 12082535624940\beta^4 \\
& -13976285582035\beta^3 + 7917677944720\beta^2 - 1767708637200\beta\big) z_1^4 \\
& + \big(-132203187917\beta^6 - 4061892203389\beta^5 + 41976312974469\beta^4 \\
& -139562105527903\beta^3 + 209430408589540\beta^2 - 146977072348800\beta \\
& +39326551704000\big) z_1^3 + \big(4540064732898\beta^5 - 32189023528776\beta^4 \\
& +111423218218458\beta^3 - 210428562654180\beta^2 + 192643241103600\beta \\
& -65988937872000\big) z_1^2 + \big(-36494410557024\beta^4 + 197044804007424\beta^3 \\
& -329307266563200\beta^2 + 153965516032800\beta + 14791357080000\big) z_1 \\
& + 89067118187808\beta^3 - 570684050648640\beta^2 + 1146714437289600\beta \\
& - 720144590112000.
\end{aligned}
$$

- For $z_2 \prec z_1$:

$$
\begin{aligned}
g_2(\beta, z_2) =& \left(573788072\beta^8 - 220687720\beta^7 - 7503382480\beta^6 + 20744645680\beta^5 - 23613586040\beta^4\right.\\
&\left. +12667475128\beta^3 - 2648252640\beta^2\right) z_2^5 + \left(-204848895\beta^7 - 858363627960\beta^6\right.\\
&\left. + 3391365456790\beta^5 - 4977728569660\beta^4 + 3171701981265\beta^3 - 682632847540\beta^2\right.\\
&\left. -44137544000\beta\right) z_2^4 + \left(-428935189842\beta^6 - 27966474140289\beta^5 + 100992652545319\beta^4\right.\\
&\left. -130647004403803\beta^3 + 72083867190015\beta^2 - 13791349509400\beta - 242756492000\right) z_2^3\\
&+ \left(-11840380819272\beta^5 - 294095993630166\beta^4 + 1011540632651148\beta^3\right.\\
&\left. -1175534529822210\beta^2 + 552033040308000\beta - 82102768687500\right) z_2^2\\
&+ \left(-150578129579184\beta^4 - 965930452772256\beta^3 + 3604616402786640\beta^2\right.\\
&\left. -3768621176239200\beta + 1280513355804000\right) z_2 - 737655098899392\beta^3\\
&+ 906896369212560\beta^2 + 449706809325600\beta - 673995164922000.
\end{aligned}
$$

For this case, the polynomials turn out to be quintic. As mentioned earlier, no radical solutions exist due to Abel-Ruffini's theorem. In order to find approximate real roots of these polynomials, we need to specify β. For instance, by setting $\beta = \frac{1}{2}$ the above polynomials simplify to:

$$
\begin{aligned}
g_1(\tfrac{1}{2}, z_1) =& -\frac{5359807932115}{128} z_1^4 - \frac{733786669}{32} z_1^5 + \frac{207671951406997}{64} z_1^3\\
&+ 31796998296714 z_1 - 278324994355884 - \frac{163463882465331}{16} z_1^2\\
g_2(\tfrac{1}{2}, z_2) =& -\frac{733786669}{32} z_2^5 + \frac{61773492718827}{8} z_2^2 - \frac{1895121071975}{128} z_2^4\\
&+ 167204428685829 z_2 - 314624555318484 - \frac{33907289629}{2} z_2^3.
\end{aligned}
$$

Approximating their real roots in MAPLE we obtain:

$$
z_1 = -1900.653702, 4.969443147, 71.41363761
$$

and

$$
z_2 = -643.7311436, 1.742768501, 28.64701004.
$$

Note that the unique solution of the stochastic game $\Gamma_{\frac{1}{2}}$ is obtained only with the help of one particular pair of the above real roots. Indeed, it follows that

(6.7) $$ v_1(\tfrac{1}{2}) = (1 - \tfrac{1}{2})z_1 = \tfrac{1}{2} 4.969443147 = 2.484721573 $$

(6.8) $$ v_2(\tfrac{1}{2}) = (1 - \tfrac{1}{2})z_2 = \tfrac{1}{2} 1.742768501 = 0.67138250. $$

Formal verification that the above are indeed values of $\Gamma_{\frac{1}{2}}$, involves substituting these into Shapley's auxilliary matrix games and finding their values.

Example 3: In this, final and simple, example we illustrate that the solutions of Shapley's equations lie in the field of rationals for only some exceptional values of rational data defining a stochastic game. In particular we considered a simply structured two state stochastic game where the second state is absorbing and has just one action, for each player. The rational data defining the game are as in Figure 6.3, where $a > 0$ and $\beta \in (0,1)$ are also rational numbers.
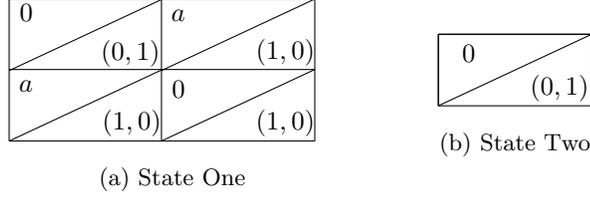

(a) State One

(b) State Two

Fig. 6.3: A game with one absorbing state.

As in previous examples, postponing the normalization by the factor $(1 - \beta)$, Shapley's fixed-point equations (3.4) now become

$$(6.9) \qquad z_1 = val \begin{bmatrix} \beta z_2 & a + \beta z_1 \\ a + \beta z_1 & \beta z_1 \end{bmatrix}, \qquad z_2 = val\,[\beta z_2]\,.$$

Note that a pure strategy saddle-point is impossible in the first Shapley auxiliary matrix. These fixed point equations were expressed in their kernel form (3.5) and then manipulated to obtain the system of two polynomial equations of the form (3.7)

$$f_1(\beta, z_1, z_2) = \beta z_1 z_2 - 2az_1 - \beta z_1^2 - \beta^2 z_1 z_2 + (a + \beta z_1)^2,$$
$$f_2(\beta, z_1, z_2) = z_2 - \beta z_2.$$

Since $z_2 = 0$ is the only root of $f_2 = 0$, f_1 immediately simplifies to an already decoupled quadratic in z_1 with coefficients that depend on a and β

$$f_1(\beta, z_1, 0) = -2az_1 - \beta z_1^2 + (a + \beta z_1)^2 = a^2 + 2a(\beta - 1)z_1 + \beta(\beta - 1)z_1^2,$$

whose discriminant is given by

$$D = 4a^2(\beta - 1)^2 - 4a^2\beta(\beta - 1) = 4a^2(1 - \beta).$$

Clearly, $f_1(\beta, z_1, 0) = 0$ has rational solutions only if $4a^2(1-\beta) = d^2$ for some rational d. This happens for some exceptional data configurations. For instance, if $\beta = \frac{8}{9}$, $1 - \beta = \frac{1}{9}$ and $D = \frac{4}{9}a^2$ then

$$z_1 = \frac{\frac{2a}{9} \pm \frac{2a}{3}}{-\frac{16}{81}} \Rightarrow z_1 = \frac{9}{4}a > 0,$$

since the other root is negative and hence cannot be the value of the stochastic game with non-negative rewards. Note that the value of the normalized stochastic game, starting in state 1, $v_1(\frac{8}{9}) = (1 - \frac{8}{9})z_1 = \frac{1}{4}a$.

Finally, we note that upon substituting, say, $a = 1$ into $f_1(\frac{8}{9}, z_1, 0)$ we obtain $f_1(\frac{8}{9}, z_1, 0) = 1 - \frac{2}{9}z_1 - \frac{8}{81}z_1^2$. Multiplying the preceding by the lowest common denominator of its coefficients we obtain the polynomial $g_1(z_1) = -8z_1^2 - 18z_1 + 81$. Now it is easy to verify that the denominator of the rational root $z_1 = \frac{9}{4}$ is a divisor of 8 and the numerator is a divisor of 81, thereby illustrating (5.2).

REFERENCES

[1] T. Bewley and E. Kohlberg, *The asymptotic theory of stochastic games*, Mathematics of Operations Research, 1 (1976), pp. 197–208.

[2] J. A. Filar, *Ordered field property for stochastic games when the player who controls transitions changes from state to state*, Journal of Optimization Theory and Applications, 34 (1981), pp. 503–515.

[3] S.K.S. Frederiksen, *Semi-algebraic Tools for Stochastic Games*, PhD thesis, Aarhus Universitet, Aarhus, Denmark, 7 2015.

[4] D. Gillette, *Stochastic games with zero stop probabilities*, Contributions to the Theory of Games, 3 (1957), pp. 179–187.

[5] K.A. Hansen, M. Koucky, N. Lauritzen, P.B. Miltersen, and E.P. Tsigaridas, *Exact algorithms for solving stochastic games: extended abstract*, STOC 11 Proceedings of the forty-third annual ACM symposium on theory of computing, (2011), pp. 205–214.

[6] A. Jaśkiewicz and A. S. Nowak, *Nonzero-sum stochastic games*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2018, pp. 281–344.

[7] ———, *Zero-sum stochastic games*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2018, pp. 215–280.

[8] I. Kaplansky, *A contribution to von neumann's theory of games*, Annals of Mathematics, 46 (1945), pp. 474–479.

[9] P. Loustaunau and W. W. Adams, *An introduction to gröbner bases*, American Mathematical Soc., 1994.

[10] S. MacLane and G. Birkhoff, *Algebra*, The Macmillan Company, 1967.

[11] J. F. Mertens and A. Neyman, *Stochastic games*, International Journal of Game Theory, 10 (1981), pp. 53–66.

[12] J. F. Mertens and S. Zamir, *Incomplete information games with transcendental values*, Mathematics of Operations Research, 6 (1981), pp. 313–318.

[13] E. Milman, *The semi-algebraic theory of stochastic games*, Mathematics of Operations Research, 27 (2002), pp. 401–418.

[14] A. Neyman, *Real algebraic tools in stochastic games*, Springer Netherlands, Dordrecht, 2003, pp. 57–75.

[15] T. Parthasarathy and T. E. S. Raghavan, *An orderfield property for stochastic games when one player controls transition probabilities*, Journal of Optimization Theory and Applications, 33 (1981), pp. 375–392.

[16] T. Parthasarathy, S. H. Tijs, and O. J. Vrieze, *Stochastic games with state independent transitions and separable rewards*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1984, pp. 262–271.

[17] T. E. S. Raghavan and J. A. Filar, *Algorithms for stochastic games — a survey*, Zeitschrift für Operations Research, 35 (1991), pp. 437–472.

[18] T. E. S. Raghavan, S. H. Tijs, and O. J. Vrieze, *On stochastic games with additive reward and transition structure*, Journal of Optimization Theory and Applications, 47 (1985), pp. 451–464.

[19] J. Rotman, *Galois theory*, Universitext (Berlin. Print), Springer New York, 1998.

[20] L. S. Shapley, *Stochastic games*, Proceedings of the National Academy of Sciences, 39 (1953), pp. 1095–1100.

[21] L. S. Shapley and R. N. Snow, *Basic solutions of discrete games*, Princeton University Press, 1952, pp. 27–36.

[22] M. J. Sobel, *Myopic solutions of markov decision processes and stochastic games*, Operations Research, 29 (1981), pp. 995–1009.

[23] W. W. Szczechla, S. A. Connell, J. A. Filar, and O. J. Vrieze, *On the puiseux series expansion of the limit discount equation of stochastic games*, SIAM Journal on Control and Optimization, 35 (1997), pp. 860–875.

[24] A. Tarski, *A decision method for elementary algebra and geometry*, University of California Press, Berkeley and Los Angeles, 1951.

[25] J. v. Neumann, *Zur theorie der gesellschaftsspiele*, Mathematische Annalen, 100 (1928), pp. 295–320.

[26] H. Weyl, *Elementary proof of a minimax theorem due to Von Neumann*, Princeton University Press, 1952, pp. 19–26.