From Imitation Games to Kakutani

Andrew McLennan* and Rabee Tourky†

October 17, 2005

Abstract

We give a full proof of the Kakutani (1941) fixed point theorem that is brief, elementary, and based on game theoretic concepts. This proof points to a new family of algorithms for computing approximate fixed points that have advantages over simplicial subdivision methods. An *imitation game* is a finite two person normal form game in which the strategy spaces for the two agents are the same and the goal of the second player is to choose the same strategy as the first player. These appear in our proof, but are also interesting from other points of view.

Keywords: Imitation games, Lemke paths, Kakutani's fixed point theorem, Lemke-Howson algorithm, long and short paths, approximate fixed point, approximate eigenvectors, Krylov method.

1 Introduction

We give a new complete[1] proof of Kakutani's fixed point theorem. In comparison with earlier complete proofs of Brouwer's and Kakutani's fixed point theorems (to a greater or lesser extent depending on the proof) our argument has several advantages. It is elementary. It is direct, arriving at Kakutani's

*Department of Economics, The University of Minnesota, 1035 Management and Economics, 271 19th Ave. S. Minneapolis, MN 55455, and Discipline of Economics, H-04 Merewether Building, University of Sydney, NSW 2006 Australia. The hospitality of the Economic Theory Center of the University of Melbourne is gratefully acknowledged. We are grateful for comments by participants of the 2004 International Conference on Economic Theory at Keio University, the 2004 Canadian Economic Theory Conference, Games 2004, the International Conference in Honor of Jerzy Łoś, the 2004 UBC Summer Workshop in Economic Theory, and the University of Minnesota Combinatorics Workshop. We have benefitted from stimulating correspondence with Christos Papadimitriou and discussions with John Stachurski.

†Department of Economics, The University of Melbourne, Victoria 3010, Australia, and Purdue University Department of Economics, 403 W. State Street, West Lafayette, IN 47907-2056.

[1]The term 'complete' is intended to distinguish proofs "from scratch" from those (e.g., Shapley (1973a) and Shapley (1973b), Shapley and Vohra (1991), Herings (1997)) in which simple arguments are used to pass between members of the family of results that includes Kakutani's theorem, the KKMS theorem, the existence of a competitive equilibrium for an exchange economy, and the existence of a core point of a balanced NTU game.

theorem without an intermediate stop at Brouwer's theorem. It is based on game theoretic concepts, so it is complementary to the goals of instruction in theoretical economics.

There is a novel class of algorithms based on the proof that are simple to implement and flexible, and have various advantages in comparison with other algorithms. Numerical tests suggest that in practice the algorithms are quite fast and succeed on some types of problems for which other algorithms are unsuitable.

The proof also uses a new class of two person games in normal form, called *imitation games*, in which the two players' sets of pure strategies are "the same" and agent 2's goal is to choose the same pure strategy as player 1. We call agent 1 the *mover* and agent 2 the *imitator*. These games play a role in the proof and the algorithms. It turns out that they have several other interesting properties. They are useful in the analysis of the complexity of certain problems related to two person games and provide new insights into the Lemke-Howson algorithm (Lemke and Howson (1964)) for a computing Nash equilibrium of two person games.

Suppose that C is a nonempty compact convex subset of an inner product space and that $F\colon C \twoheadrightarrow C$ is a upper semicontinuous convex valued correspondence. Starting from any initial point $x_1 \in C$, we recursively define sequences $\{x_m\}$ and $\{y_m\}$ by choosing $y_m \in F(x_m)$ arbitrarily, then setting

$$x_{m+1} = \sum_{j=1}^{m} \rho_j^m y_j \, ,$$

where (ι^m, ρ^m) is a Nash equilibrium of the imitation game in which the common set of pure strategies is $\{x_1, \ldots, x_m\}$, the mover's payoff matrix is the $m \times m$ matrix A with entries

$$a_{ij} = -\|x_i - y_j\|^2,$$

and (due to the definition of an imitation game) the imitator's payoff matrix is the $m \times m$ identity matrix I. That is, the mover seeks to minimize the square of the distance between her choice x_i and the chosen image y_j of the choice x_j of the second player. A simple calculation ((2) in Section 3) shows that the support of ι^m is contained in the set of elements of $\{x_1, \ldots, x_m\}$ that are closest to x_{m+1}. Since this is an imitation game the support of ρ^m is a subset of the support of ι^m, so x_{m+1} is a convex combination of elements of the images $F(x)$ of nearby points x. Since the sequence x_m is contained in a compact set, the diameters of the supports of ι^m and ρ^m decrease to zero as the number of iterations increases. Since F is upper semicontinuous and convex valued, limit points of $\{x_m\}$ are exact fixed points of F.

If it is known that each imitation game has a Nash equilibrium, this argument proves Kakutani's fixed point theorem. The Lemke-Howson algorithm computes a Nash equilibrium of a two player normal form game, and the theorem stating that it is an algorithm (that is, it halts in finite time for any input) implies that every two person game has a Nash equilibrium. The Lemke-Howson algorithm is usually described as either a tableau method (e.g., Lemke and Howson (1964)) or as paths in the edges of a pair of polytopes (e.g., Shapley (1974)).

In an imitation game the imitator's best response correspondence has a simple description. This has the consequence that the "pivots" of the Lemke-Howson algorithm that change the mover's strategy have a predictable and trivial character, so that a simpler description involving only one simplex is obtained by ignoring these pivots. The result of adopting this perspective turns out to be precisely the "Lemke paths" algorithm[2] of Lemke (1965). In earlier literature the Lemke-Howson algorithm is treated as a specialization of the Lemke paths algorithm obtained by restricting the input of the latter to have a special form. The observation that the Lemke paths algorithm can be viewed as a projection of the Lemke-Howson algorithm, applied to an imitation game, is novel, at least so far as we know. One interesting consequence of this observation is that it is possible to derive a recent result of Savani and von Stengel (2004) concerning "long" paths of the Lemke-Howson algorithm from earlier work by Morris (1994) concerning long paths of the Lemke paths algorithm.

If one uses the Lemke paths algorithm to compute Nash equilibria of the imitation games that arise in our algorithm for finding a fixed point of F, one may be concerned that the computational burden will increase as m increases. General experience with the Lemke-Howson algorithm suggests that this is unlikely, and in practice our algorithm often finds a fixed point after a small number of iterations. When C is finite dimensional, there is also theoretical reassurance. A *geometric imitation game* is an imitation game derived from points $x_1, \ldots, x_m, y_1, \ldots, y_m$, as described above. Holding the dimension of C fixed, the length of the Lemke paths for a geometric imitation game is bounded by a polynomial function of m. This observation provides a rich class of examples that complement the works of Savani and von Stengel (2004) and Morris (1994).

We also give some additional results concerning imitation games. Gale and Tucker (1950) showed how to pass from a two player game to a symmetric game whose symmetric equilibria are in one-to-one correspondence with the Nash equilibria of the two player game. Consequently the problem of finding a symmetric equilibrium of a symmetric games is at least as hard (in a computational sense described precisely in Section 5) as the problem of finding a Nash equilibrium of a general two player game. We show how to pass from a symmetric game to an imitation game whose Nash equilibria are in one-to-one correspondence with the symmetric equilibria of the symmetric game, so the problem of finding a Nash equilibrium of an imitation game is at least as hard as the problem of finding a symmetric equilibrium of a symmetric game. Of course imitation games are a special type of two player games, so the three problems are equally hard. (Identical reasoning applies to related problems, such as finding all equilibria, or determining whether there is more than one equilibrium.) A related application of imitation games is our paper McLennan and Tourky (2005) that gives simple proofs of the results of Gilboa and Zemel

[2]There is a third algorithm (Lemke (1968)) that is usually called "the Lemke algorithm," or just "Lemke," so we have avoided that phrase in connection with the algorithm of Lemke (1965).

(1989) on the complexity of certain computational problems concerning Nash equilibria of two person games.

The remainder has the following organization. In the next section we survey the history of proofs of Brouwer's and Kakutani's fixed point theorems. Section 3 proves Kakutani's theorem using our recursive sequence and an elementary version of the Lemke paths algorithm. Section 4 is a discussion of the structure and properties of our algorithms for computing approximate fixed points. Section 5 discusses imitation games from the point of view of computational theory. Section 6 explains the derivation of the Lemke-Howson algorithm from the Lemke paths algorithm and shows how to derive the Lemke paths algorithm from the Lemke-Howson algorithm. Imitation games with a geometric derivation, as described above, are studied in Section 7, leading to a class of games with "short" Lemke-Howson paths. The last section gives some final remarks.

2 Historical Background

Brouwer's fixed point theorem is a celebrated achievement of early twentieth century mathematics. The literature now contains several proofs which typically involve advanced techniques or results. For instance the proof in Brouwer (1910) uses the ideas that were evolving into the field of mathematics now known as algebraic topology. An alternative proof due to Hirsch (cf. Milnor (1965)) is an application of the Morse–Sard theorem in differential analysis to prove the non-existence of a retraction from a disk to its boundary. An ingenious and elementary proof by Milnor (1978) uses a simple computation to reformulate the problems in terms of polynomials. The proof that is perhaps most popular in economics is similar to ours insofar as it has two phases. The first phase is a result in combinatoric geometry—Sperner's lemma in that proof and the existence of equilibrium for certain two person finite games via the Lemke-Howson algorithm here. The second phase uses approximations to pass to a topological conclusion. At this point we should call attention to Scarf (1967b,a) and Hansen and Scarf (1969) (see also Appendix C of Arrow and Hahn (1971)) which use methods related to those embodied in the Lemke-Howson algorithm to establish that the conclusion of Sperner's lemma holds in circumstances where Brouwer's theorem follows.

Motivated by a desire to provide a simple proof of von Neumann's minimax theorem, Kakutani (1941) extended Brouwer's fixed point theorem to convex valued correspondences. His method—show that arbitrarily small neighborhoods of the graph of the correspondence contain graphs of continuous functions, then apply Brouwer's fixed point theorem—is intuitive, but its implementation involves various details that are cumbersome and of slight interest to the other goals of instruction in theoretical economics. To the best of our knowledge all previous extensions of Brouwer's theorem to correspondences use either this method or (e.g., Eilenberg and Montgomery (1946)) algebraic topology.

The algorithm of Lemke and Howson (1964) is, in effect, an elementary proof of a result that had previously been proved by appealing to fixed point theorems. A sense of the thinking at the time is given by the following excerpt from the review of Lemke and Howson (1964) by R. J. Aumann[3].

> "The first algebraic proof of the theorem that every two-person non-constant-sum game with finitely many strategies for each player has a Nash equilibrium point. This important paper successfully culminates a series of developments beginning with Nash (1950, 1951). Nash's proofs are valid for n-person games with arbitrary n, but they use fixed point theorems ... [there has been] a long-standing suspicion that there must be an algebraic existence proof lurking in the background; the current paper provides it."

We regard the arguments given here as demonstrating that the Lemke-Howson algorithm embodies, in algebraic form, the fixed point principle itself, and not merely the existence theorem for finite two person games. In this respect our thinking was strongly influenced by the beautiful paper of Shapley (1974), which calls attention to the relation between the Lemke-Howson algorithm and the fixed point index. Another important predecessor is Eaves (1971a), which shows that a procedure for computing solutions to the linear complementarity problem can be used as the underlying engine of a procedure for computing approximate fixed points.

Scarf (1967b) introduced an algorithm, based on an extension of the argument used to prove Sperner's lemma, for computing an approximate fixed point. An extensive literature (e.g., Scarf (1973), Todd (1976), Doup (1988), Murty (1988)) elaborates on and refines this algorithm, and it continues to be an important approach to the computation of approximate fixed points. Scarf (1967b) also pointed out the possibility of using a linear complementarity problem to find approximate fixed points. This method was extended to Kakutani fixed points in Hansen and Scarf (1969), and versions involving a triangulation of the underlying space are given by Kuhn (1969) and Eaves (1971b) and refined by Merrill (1972a,b). Some aspects of these methods are discussed in Section 4.

3 From Imitation Games to Kakutani

A *two person game* is a pair (A, B) of $m \times n$ matrices of real numbers, where m and n are positive integers. For each integer $k \geq 1$ let Δ^k be the standard unit simplex in $\mathbb{R}^k$, i.e., the set of vectors whose components are nonnegative and sum to one. A *Nash equilibrium* of (A, B) is a pair $(\iota, \rho) \in \Delta^m \times \Delta^n$ such that $\iota^T A\rho \geq \tilde{\iota}^T A\rho$ for all $\tilde{\iota} \in \Delta^m$ and $\iota^T B\rho \geq \iota^T B\tilde{\rho}$ for all $\tilde{\rho} \in \Delta^n$.

For the rest of the section we specialize to the case $m = n$ and $B = I$, where I is the $m \times m$ identity matrix. Such a game (A, I) is called an *imitation game*, and in such a game the two agents are called the *mover* and the *imitator*

[3] *Mathematical Reviews* MR0173556 (30 #3769).

respectively. Let $\mathcal{I} := \{1, \dots, m\}$. For any $\rho \in \Delta^m$, let

$$\rho^\circ := \{i \in \mathcal{I} \colon \rho_i = 0\} \quad \text{and} \quad \overline{\rho} := \operatorname*{argmax}_{i \in \mathcal{I}} (A\rho)_i.$$

An *I-equilibrium* of an imitation game (A, I) is a mixed strategy $\rho \in \Delta^m$ for the imitator such that the support of ρ is contained in $\overline{\rho}$: $\rho^\circ \cup \overline{\rho} = \mathcal{I}$. To prove that (A, I) has a Nash equilibrium it suffices to find an I-equilibrium:

Lemma 1. *A mixed strategy $\rho \in \Delta^m$ is an I-equilibrium of (A, I) if and only if there is $\iota \in \Delta^m$ such that (ι, ρ) is a Nash equilibrium of (A, I).*

Proof. If (ι, ρ) is a Nash equilibrium of (A, I), then the support of ρ is contained in the support of ι because ρ is a best response to ι for the imitator, and the support of ι is contained $\overline{\rho}$ because ι is a best response to ρ for the mover. Thus the support of ρ is contained in $\overline{\rho}$.

Now suppose ρ is an I-equilibrium of (A, I). Since the set of best responses to ρ contains the support of ρ, we may choose an $\iota \in \Delta^m$ that assigns all probability to best responses to ρ (so ι is a best response to ρ) and maximal probability to elements of the support of ρ (so ρ is a best response to ι). ■

For $X, Y \subset \mathcal{I}$ let $S(X, Y) := \{\, \rho \in \Delta^m \colon \rho^\circ = X \text{ and } \overline{\rho} = Y \,\}$.

Lemma 2. *If $S(X, Y)$ is nonempty, then it is convex, X is a proper subset of $\mathcal{I}$, Y is nonempty, and the closure of $S(X, Y)$ is*

$$S' := \bigcup_{X' \supset X, Y' \supset Y} S(X', Y').$$

Proof. Since $S(X, Y)$ is defined by a finite conjunction of inequalities, it is convex. If $\rho \in S(X, Y)$, then the support of ρ is nonempty, so X cannot be all of $\mathcal{I}$, and $\overline{\rho} = Y$ is nonempty. As a matter of continuity the closure of $S(X, Y)$ is contained in S', and the line segment between ρ and any point in S' is contained in $S(X, Y)$, so S' is contained in the closure of $S(X, Y)$. ■

We say that A is in *general position* if $|\rho^\circ| + |\overline{\rho}| \le m$ for all $\rho \in \Delta^m$.

Lemma 3. *If A is in general position and $S(X, Y)$ is nonempty, then $S(X, Y)$ is $(m - |X| - |Y|)$-dimensional and $S(X', Y') \neq \emptyset$ whenever $X' \subset X$ and $\emptyset \neq Y' \subset Y$.*

Proof. Elementary linear algebra implies that the dimension of $S(X, Y)$ is at least $m - |X| - |Y|$. Suppose that the dimension of $S(X, Y)$ is greater than this. Then the closure of $S(X, Y)$ is a polytope, and any of its facets is the closure of $S(X', Y')$ for some (X', Y') with $X \subset X'$ and $Y \subset Y'$, where at least one inclusion is strict, which implies that the dimension of $S(X', Y')$ is greater than $m - |X'| - |Y'|$. Proceeding to a facet of the closure of $S(X', Y')$, a facet of that facet, and so forth, we eventually arrive at a nonempty $S(X'', Y'')$ with $m - |X''| - |Y''| < 0$, contradicting general position.

For any $i_0 \in Y$ the quantities $\rho_1 + \cdots + \rho_m$, ρ_i ($i \in X$), and $(A\rho)_i - (A\rho)_{i_0}$ ($i \in Y \setminus \{i_0\}$) are affine functionals on $\mathbb{R}^m$, and since $S(X, Y)$ is $m - |X| - |Y|$ dimensional, they are affinely independent. Therefore any neighborhood of a point where they all vanish (e.g., an element of $S(X, Y)$) contains points attaining any combination of signs for these quantities, including the combination defining $S(X', Y')$ for any $X' \subset X$ and nonempty $Y' \subset Y$. ■

Suppose that A is in general position, and that $S(X, Y)$ is nonempty. If $|X| + |Y| = m$, then $S(X, Y)$ is a singleton whose unique element is denoted by $V(X, Y)$. Such points are called *vertices*. If $|X| + |Y| = m - 1$, then the closure of $S(X, Y)$, denoted by $E(X, Y)$, is a one dimensional line segment that is called an *edge*. Each edge has two endpoints, which are clearly vertices. The Lemke paths algorithm follows a path of edges described in the next proof.

Proposition 4. *The imitation game (A, I) has an I-equilibrium.*

Proof. Using elementary linear algebra, one may show that if A is not in general position, then its entries express a linear system of equations with more equations than unknowns that nonetheless has a solution. Therefore the matrices in general position are dense in the space of $m \times m$ matrices. If A is the limit of a sequence $\{A^r\}$ and, for each r, there is an I-equilibrium ρ^r of (A^r, I), then every accumulation point of the sequence $\{\rho^r\}$ is an I-equilibrium of (A, I). Thus we may assume that A is in general position[4].

Fix an arbitrary $s \in \mathcal{I}$. A vertex $V(X, Y)$ is said to be an *s-vertex* if $X \cap Y \subset \{s\}$. For an s-vertex $V(X, Y)$ the following are equivalent: (a) $V(X, Y)$ is an I-equilibrium; (b) $s \in X \cup Y$; (c) $X \cap Y = \emptyset$. If $V(X, Y)$ is an s-vertex that is not an I-equilibrium, then $X \cap Y = \{i\}$ for some $i \neq s$.

In general δ_j denotes the degenerate mixed strategy that assigns probability one to the pure strategy indexed by j. By general position, $\overline{\delta_s} = \{i^*\}$ for some i^*, so $\delta_s = V(\mathcal{I} \setminus \{s\}, \{i^*\})$ is an s-vertex. If $i^* = s$, then δ_s is an I-equilibrium and we are done, so we assume that $i^* \neq s$.

An edge $E(X, Y)$ is said to be an *s-edge* if $X \cup Y = \mathcal{I} \setminus \{s\}$. Observe that the endpoints of an s-edge are s-vertices.

Let $V(X, Y)$ be an s-vertex with $X \cap Y = \{i\}$. We now determine the number of s-edges that have $V(X, Y)$ as an endpoint. An edge having $V(X, Y)$ as an endpoint is necessarily $E(X \setminus \{j\}, Y)$ for some $j \in X$ or $E(X, Y \setminus \{j\})$ for some $j \in Y$, and these edges will not be s-edges unless $j = i$. On the other hand, Lemma 3 implies that $S(X \setminus \{i\}, Y) \neq \emptyset$, and that $S(X, Y \setminus \{i\}) \neq \emptyset$ if and only if $Y \setminus \{i\} \neq \emptyset$. Thus $V(X, Y)$ is an endpoint of two s-edges if Y has more than one element, and it is an endpoint of one s-edge if $Y = \{i\}$. In the latter case $X = \mathcal{I} \setminus \{s\}$ because $|X| = m - 1$ and $X \cup Y = \mathcal{I} \setminus \{s\}$, so $V(X, Y) = \delta_s$ and $i = i^*$.

[4] Alternatively we could describe a method of extending the the algorithm to games that are not in general position. There are standard degeneracy-resolution techniques for pivoting algorithms, including the simplex algorithm for linear programming, the Lemke-Howson algorithm, and the Lemke path algorithm.

Summarizing, if there is no I-equilibrium then δ_s is an end point of one s-edge, and every other s-vertex is an endpoint of two s-edges. Summing over s-vertices, we find that there is an odd number of pairs consisting of an s-edge and one of its endpoints. But of course this is impossible: the number of such pairs is even because each s-edge has two endpoints. ■

The Lemke path algorithm for computing an I-equilibrium begins at δ_s and proceeds from there along the path of s-edges. We denote this path $\mathcal{LP}$. Since each s-vertex that is not an I-equilibrium is an endpoint of precisely two I-edges, this path does not branch or return to any s-vertex that it visited earlier. Since it is finite, it must eventually arrive at an I-equilibrium.

A *correspondence* $F : C \twoheadrightarrow C$ assigns a nonempty $F(x) \subseteq C$ to each $x \in C$. It is *upper semicontinuous* if, for each $x \in C$: (i) $F(x)$ is closed; (ii) for each neighborhood $V \subseteq C$ of $F(x)$ there is a neighborhood U of x such that $F(x') \subseteq V$ for all $x' \in U$.

Theorem 5 (Kakutani (1941))**.** *Let C be a nonempty, compact, and convex subset of an inner product space. If $F : C \twoheadrightarrow C$ is an upper semicontinuous correspondence whose values are convex, then $x^* \in F(x^*)$ for some $x^* \in C$.*

Proof. Define sequences $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$ by choosing $x_1 \in C$ arbitrarily, then (recursively, given $x_1, \ldots, x_m$ and $y_1, \ldots, y_{m-1}$) letting y_m be any point of $F(x_m)$ and setting

$$x_{m+1} = \sum_{i=1}^{m} \rho_i^m y_i \,, \tag{1}$$

where ρ^m is an I-equilibrium of the imitation game (A, I) with m strategies for which the entries of A are $a_{ij} = -\|x_i - y_j\|^2$. Note that $\sum_{j=1}^m \rho_j^m (x_{m+1} - y_j) = 0$, so for each $i = 1, \ldots, m$ we have:

$$\begin{aligned}
(A\rho)_i &= -\sum_{j=1}^{m} \rho_j^m \|x_i - y_j\|^2 \\
&= -\sum_{j=1}^{m} \rho_j^m \|(x_i - x_{m+1}) + (x_{m+1} - y_j)\|^2 \\
&= -\sum_{j=1}^{m} \rho_j^m \langle x_i - x_{m+1}, x_i - x_{m+1} \rangle \\
&\qquad - 2\sum_{j=1}^{m} \langle x_i - x_{m+1}, \rho_j^m (x_{m+1} - y_j) \rangle \\
&\qquad\qquad - \sum_{j=1}^{m} \rho_j^m \langle x_{m+1} - y_j, x_{m+1} - y_j \rangle \\
&= -\|x_i - x_{m+1}\|^2 - \sum_{j=1}^{m} \rho_j^m \|x_{m+1} - y_j\|^2 .
\end{aligned} \tag{2}$$

Note that the second term does not depend on i. Since ρ is an I-equilibrium,

$$\operatorname{supp} \rho^m \subset \operatorname*{argmin}_{1 \leq i \leq m} \|x_i - x_{m+1}\| \,. \tag{3}$$

The sequence $\{x_m\}$ is contained in the compact set C, so it has an accumulation point x^*. Suppose by way of contradiction that $x^* \notin F(x^*)$. Since $F(x^*)$ is convex, it has a convex neighborhood V whose closure does not contain x^*. Applying upper semicontinuity, there is $\varepsilon > 0$ such that $\|x - x^*\| < 3\varepsilon$ implies both that $x \notin V$ and that $F(x) \subseteq V$. Choose m such that $\|x_{m+1} - x^*\| < \varepsilon$ and there is also some $\ell \leq m$ such that $\|x_\ell - x^*\| < \varepsilon$. By (3), if $\rho_i^m > 0$, then

$$\|x_i - x_{m+1}\| \leq \|x_\ell - x_{m+1}\| \leq \|x_\ell - x^*\| + \|x^* - x_{m+1}\| < 2\varepsilon$$

and

$$\|x_i - x^*\| \leq \|x_i - x_{m+1}\| + \|x_{m+1} - x^*\| < 3\varepsilon,$$

whence $y_i \in F(x_i) \in V$. But this implies that $x_{m+1} = \sum_i \rho_i^m y_i \in V$, which is the desired contradiction. ■

4 Approximate Fixed Points

Suppose C is a complete[5] convex subset of an inner product space, and let $F : C \twoheadrightarrow C$ be an upper semicontinuous and convex valued correspondence. Reviewing the argument above, we see that under these somewhat weaker hypotheses the recursive procedure (1) of the proof of Kakutani's fixed point theorem produces a sequence $x_1, x_2, \ldots$ whose accumulation points are fixed points of F. Such an accumulation point certainly exists when C is compact, and otherwise additional conditions on F may guarantee existence. Any algorithm that generates x_{m+1} is the engine of a computational procedure that will eventually produce an approximate fixed point up to any desired degree of accuracy. One such algorithm is to compute in each iteration a Lemke path for the imitation game (A, I) defined in connection with (1). For the non-degenerate case this Lemke path is described in Proposition 4. Thus, we have the raw materials for a class of algorithms that compute approximate fixed points.

In this section we study the properties of the sequence (1), concluding with a discussion of the relative merits of our algorithms and two numerical examples. The emphasis is on the flexibility of our method, highlighting the relative paucity of information needed to apply the method and guarantee eventual convergence to a fixed point.

4.1 Properties of the procedure

We begin with the continuity requirement that guarantees that every accumulation point of $\{x_m\}$ is a fixed point. The correspondence $F\colon C \twoheadrightarrow C$ is said

[5]That is, every Cauchy sequence in C converges to a point in C.

to be *c-continuous* if for every $x \in C$ that is not a fixed point of F there is a neighbourhood V of x such that the closed convex hull of $F(V)$ does not contain x. Of course, every convex valued upper semicontinuous correspondence is c-continuous and the selection $f(x) = \arg\min_{y \in F(x)} \|y - x\|$ is a c-continuous function that is generally not continuous. Further, f has the same fixed points as F.

Proposition 6. *If $\{x_m\}$ is derived from a c-continuous F as per* (1)*, then every accumulation point x^* of $\{x_m\}$ is a fixed point of F.*

Proof. If $x^* \notin F(x^*)$, then there is a neighbourhood V of x^* such that the closed convex hull K of $F(V)$ is disjoint from x^*. Therefore, there exists a continuous linear functional p and a real number r satisfying $\langle p, x^* \rangle < r$ and $\langle p, x \rangle > r$ for every $x \in K$. But we know that there is m large enough such that the support of ρ^m is in V and $\langle p, x_{m+1} \rangle < r$. This contradicts the fact that $x_{m+1} = \sum_{i=1}^m \rho_i y_i$ and $y_i \in K$ for each i. ■

We turn to a technical result, which states that if x_i in the sequence $x_1, x_2, \ldots$ is not a fixed point, then eventually it is not played with positive probability. This is obvious when C is compact but not so obvious when the sequence is not bounded. Recall that a *cone* $K \subseteq \mathbb{R}^n$ is a closed set not equal to zero satisfying $K + K \subseteq K$, $\alpha K \subseteq K$ for any $\alpha \geq 0$, and $K \cap -K = \{0\}$. The next result makes no assumptions about the continuity of the correspondence F.

Proposition 7. *If C is a closed convex subset of $\mathbb{R}^n$, $\{x_m\}$ is a sequence derived from F as per* (1)*, and x_i is a point in $\{x_m\}$ that is not an accumulation point, then there exists m^* such that $\rho_i^m = 0$ for all $m \geq m^*$.*

Proof. We can assume without loss of generality that $x_i = 0$. There exist a finite number of cones $K_1, K_2, \ldots, K_k$ whose union is $\mathbb{R}^n$ and each satisfying $\langle x, y \rangle > 0$ for all $x, y \in K_j \setminus \{0\}$. For each j for which x_m enters the cone K_j let m_j be the first index satisfying $x_{m_j} \in K_j$. Since $\left\langle x_{m_j}, y \right\rangle > 0$ for all non-zero $y \in K_j$, the vector x_{m_j} defines a strictly positive linear functional on the cone K_j, and the set

$$P_j = \left\{ x \in K_j \cap C \colon \left\langle x_{m_j}, x \right\rangle \leq \tfrac{1}{2} \left\langle x_{m_j}, x_{m_j} \right\rangle \right\} ,$$

is compact.

Suppose $\rho_i^m > 0$ for some $m \geq m_j$. Then (since $x_i = 0$) $\|x_{m+1}\|^2 \leq \|x_{m_j} - x_{m+1}\|^2$. Expanding $\|x_{m_j} - x_{m+1}\|^2$ as an inner product, distributing, subtracting $\|x_{m+1}\|^2$ from both sides, and rearranging, we find that $x_{m+1} \in P_j$.

Suppose by way of contradiction that there are infinitely many m' with $\rho_i^{m'-1} > 0$. There are infinitely many such m' in the union of all P_j, which is a compact set. Since x_i is not an accumulation point, there exist two such m' and m'', say with $m' > m''$ such that $x_{m'}, x_{m''}$ are closer to each other than to x_i, which implies that $\rho_i^{m'-1} = 0$. This contradiction completes the proof. ■

We now study some cases for which C is not compact and the sequence $\{x_m\}$ has accumulation points. Recall that for any number i and function $f \colon C \to C$

we write $f^0(x) = x$ and $f^i(x) = f(f^{i-1}(x))$ for $i \geq 1$. Given any x the *f-orbit of x* is the set $\{x, f(x), f^2(x), \ldots\}$.

Proposition 8. *Let C be a closed convex subset of $\mathbb{R}^n$ and x_m be the sequence* (1) *applied to a c-continuous function $f\colon C \to C$. If any of the following hold, then x_m has accumulation points, each of which is a fixed point of f:*

(a) There exists some i^* such that the range $f^{i^*+1}(C)$ of f^{i^*+1} is bounded and $f^i(C)$ is convex for all $0 \leq i \leq i^*$.

(b) f is a linear operator and the f-orbit of x_1 is bounded.

(c) f has a non-repulsive point in C. That is, there exists $x^* \in C$ such that $\|f(y) - x^*\| \leq \|x - x^*\|$ for any $x \in C$.

Proof. For (a) suppose by way of contradiction that $\{x_m\}$ does not have an accumulation point. We claim that for every $i \leq i^*$ there is m_i such that $x_m \in f^i(C)$ for all $m \geq m_i$. This implies that $\{x_m\}$ is eventually in $f^{i^*}(C)$, so $\{y_m\}$ is eventually in $f^{i^*+1}(C)$, and $\{x_m\}$ is eventually in the convex hull of $f^{i^*+1}(C)$, which is bounded. Therefore x_m has accumulation points.

The proof of the claim is inductive. In the case $i = 0$ we have that $x_m \in f^0(C) = C$ whenever $m \geq m_0 := 1$. Suppose that the claim has been established for i. We may assume that $\{x_m\}$ has no accumulation points, in which case Proposition 7 implies that $\rho_j^m = 0$ for sufficiently large m if $j \leq m_i$ with $x_j \notin f^i(C)$. That is, there exists m_{i+1} such that for $m \geq m_{i+1}$ we have $\rho_j^m = 0$ for any $j \leq m_i$ with $x_j \notin f^i(C)$. Thus, for $m \geq m_{i+1}$ we have $y_j \in f(f^i(C)) = f^{i+1}(C)$ for all j such that $\rho_j^m > 0$. Since $f^{i+1}(C)$ is convex, it follows that $x_{m+1} \in f^{i+1}(C)$.

When (b) holds the existence of an accumulation point is a consequence of the fact that for a linear operator the sequences $\{x_m\}$ and $\{f(x_m)\}$ are always in the convex hull of the f-orbit of x_1.

In the case of (c) the sequence $\{x_m\}$ is contained in the ball with center x^* and radius $\|x^* - x_1\|$. ■

The procedure can be used to approximate fixed points of functions with non-convex domains. We study an example that can be extended in various ways. If C is not convex, then running a variant of our method needs at the very least an oracle that for any input x outputs whether $x \in C$ and if so outputs $f(x)$. For functions that map the boundary of C inside a star shaped set the extra information needed is a center of the star shaped set.[6] We don't need to *a priori* know anything extra about the geometry of C.

Proposition 9. *Let C be a closed subset of $\mathbb{R}^n$, and let $f\colon C \to C$ be a continuous function with bounded range that maps the boundary points of C into the*

[6]A *star shaped* set $K \subseteq \mathbb{R}^n$ is a set with an interior point x^* such that for any x in the interior of K the point $\alpha x + (1-\alpha)x^*$ is also in the interior of K for every $0 \leq \alpha \leq 1$. The point x^* is called a *center* of K.

interior of a star shaped set $K \subseteq C$ with center x^*. The sequence $\{x_m\}$ of (1) *applied to the function*

$$h \colon \mathbb{R}^n \to \mathbb{R}^n, \quad h(x) = \begin{cases} f(x) & \text{if } x \in C\,, \\ x^* & \text{otherwize}\,, \end{cases}$$

has accumulation points, which are fixed points of f.

Proof. We show that h is c-continuous. Let x be a point in $\mathbb{R}^n$. The required condition holds true if x is not in C or x is in the interior of C. So let x be on the boundary of C and assume that $f(x) \neq x$. We can pick a small enough neighbourhood V of x such that the closed convex hull Y of $f(V)$ is in the interior of K. Notice that the closed convex hull of $\{x^*, Y\}$ is also in the interior of K and is disjoint from x. Any accumulation point x of x_m is a fixed point of h and f. ■

An important problem associated with a linear operator $T \colon \mathbb{R}^n \to \mathbb{R}^n$ is that of approximating a real *eigenvector*, which is a non-zero vector $x \in \mathbb{R}^n$ satisfying $Tx = \lambda x$ for some $\lambda \in \mathbb{R}$. Here λ is called the *eigenvalue* of x. This problem is generally a fixed point problem in the projective space of lines passing through the origin.[7] Assume that T has n independent eigenvectors $z_1, z_2, \ldots, z_n \in \mathbb{R}^n$ with corresponding eigenvalues $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n| > 0$. Any eigenvector with an eigenvalue λ satisfying $|\lambda| = |\lambda_1|$ is called a *dominant* eigenvector. Let $H \subseteq \mathbb{R}^n$ be the hyperplane containing $z_2, z_3, \ldots, z_n$. Let Ω be $\mathbb{R}^n \setminus H$. Of course, we don't know the hyperplane H or its complement Ω, but we do know that Ω is open and dense in $\mathbb{R}^n$ and that H has positive codimension.

Proposition 10. *If $x_1 \in \Omega$, then $\frac{x_m}{\|x_m\|}$ accumulates at dominant eigenvectors of T when either:*

a. $\lambda_1 > |\lambda_2| \geq \cdots \geq |\lambda_n|$ *and $\{x_m\}$ is derived from* (1) *applied to T.*

b. $\lambda_1 \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$ *and $\{x_m\}$ is derived from* (1) *applied to $x \mapsto \frac{T(x)}{\|T(x)\|}$.*

If T has k distinct eigenvalues, then in both cases any x_m, y_m derived from (1) *remain in the linear span of $\{x_1, T^1 x_1, \ldots, T^k x_1\}$.*

Proof. Let M be the matrix whose columns are the eigenvectors $z_1, z_2, \ldots, z_n$. Let Λ be the the diagonal matrix with diagonal vector $(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Notice that $T = M\Lambda M^{-1}$ and $\Lambda = M^{-1}TM$. The vector $e = (1, 0, \ldots, 0)$ is a dominant eigenvector of Λ. Further, if α is a dominant eigenvector of Λ, then $M\alpha$ is a dominant eigenvector of T.

Define the inner product $\langle \alpha, \beta \rangle_M = \langle M\alpha, M\beta \rangle$ and norm $\|\alpha\|_M = \sqrt{\langle \alpha, \alpha \rangle_M}$ on $\mathbb{R}^n$. Pick $x_1 \in \Omega$ and let $\alpha_1 = M^{-1}x_1$. We know that the first coordinate α_1^1 of α_1 is not zero. Define the cone $C \subseteq \mathbb{R}^n$

$$C = \left\{ \beta \in \mathbb{R}^n \colon \beta^1 \alpha_1^1 > 0 \text{ and } \left| \frac{\beta^i}{\beta^1} \right| \leq \left| \frac{\alpha_1^i}{\alpha_1^1} \right| \right\} \cup \{0\}\,.$$

[7] The idea of of showing the existence of a real eigenvalue using Brouwer's fixed point theorem is due to Debreu and Herstein (1953), who give one of the alternative proofs of the Perron–Frobenius theorem on positive matrices.

Clearly, $\alpha_1 \in C$, $\beta \in C \setminus \{0\}$ implies $\beta^1 \neq 0$, and $\Lambda(C) \subseteq C$.

For (a) suppose that $\{\alpha_m\}$ is derived from (1) applied to Λ using the inner product $\langle\cdot,\cdot\rangle_M$. Each α_m and $\Lambda(\alpha_m)$ are non-zero vectors in C. Let $C^i = \Lambda^i(C)$ for each $i = 1, 2, \ldots$. We have $C^i \subseteq C^{i+1}$ and $\cap_{i=1}^{\infty} C^i$ is the half line generated by either the dominant eigenvectors e or $-e$ of Λ. Notice that if $\alpha \notin C^i$, then α is not a fixed point of Λ. Thus, as in the proof of (a) of Proposition 8 we can apply Proposition 7 to show that for each i there exists m_i satisfying $\alpha_m \in C^i$ for all $m \geq m_i$. Therefore, $\frac{\alpha_m}{\|\alpha_m\|_M}$ converges to a dominant eigenvector of Λ. Now if $\{x_m\}$ is derived from (1) applied to T using the cannonical inner product $\langle\cdot,\cdot\rangle$, then $\{\alpha_m = M^{-1}x_m\}$ is derived from (1) applied to Λ using the inner product $\langle\cdot,\cdot\rangle_M$. This is because for each i, j we have

$$\|x_i - Tx_j\|^2 = \|\alpha_i - \Lambda\alpha_j\|_M^2$$

and in each iteration we define the same imitation game. We have therefore shown that $\frac{x_m}{\|x_m\|}$ converges to a dominant eigenvector of T.

For (b) let S be the unit sphere of the norm $\|\cdot\|_M$. Suppose that $\{\alpha_m\}$ is derived from (1) applied to $\alpha \mapsto \frac{\Lambda\alpha}{\|\Lambda\alpha\|_M}$ using the inner product $\langle\cdot,\cdot\rangle_M$. The sequence $\alpha_{m>1}$ remains in the convex hull of $S \cap C$, which is compact and disjoint from zero. Therefore, α_m accumulates at fixed points of $\alpha \mapsto \frac{\Lambda(\alpha)}{\|\Lambda(\alpha)\|}$ and $\frac{\alpha_m}{\|\alpha_m\|_M}$ accumulates at eigenvectors of Λ. Once again if $\{x_m\}$ is derived from (1) applied to $x \mapsto \frac{Tx}{\|Tx\|}$ using the cannonical inner product $\langle\cdot,\cdot\rangle$, then $\{\alpha_m = M^{-1}x_m\}$ is derived from (1) applied to $\alpha \mapsto \frac{\Lambda\alpha}{\|\Lambda\alpha\|_M}$ using the inner product $\langle\cdot,\cdot\rangle_M$. This is because in each i, j we have

$$\|x_i - \tfrac{Tx_j}{\|Tx_j\|}\|^2 = \|\alpha_i - \tfrac{\Lambda\alpha_j}{\|\Lambda\alpha_j\|_M}\|_M^2 .$$

Any eigenvector in C is a dominant eigenvector of Λ, proving (b).

For the last part, notice the Λ-orbit of α_1 is contained in the span of the k points $\alpha_1, \Lambda^1\alpha_1, \Lambda^2\alpha_1, \ldots, \Lambda^k\alpha_1$. ■

In the preceding result, if the number of distinct eigenvalues k is small, then we need to evaluate T at only k points to generate the whole sequence x_m. This and the other results indicate that the sequence x_m of (1) amplifies some of the useful properties of iterating a function, which is a method that is frequently used, both in economics and other disciplines, even when convergence is not guaranteed. Numerical experience also suggests that the sequence x_m of (1) converges to a fixed point x^* quite fast when x^* is locally contractive. Our understanding of this phenomenon is limited to one case which is rather interesting.

For the next result suppose that in each iteration m if the pure strategy m is an I-equilibrium, then $\rho_m^m = 1$ is selected. We assume that C is a convex complete subset of an inner product space and f is a function. We will say that x_{m+1} of (1) *enters a contractive neighbourhood with factor* $0 \leq q < 1$ if there is $n < m+1$ such that $\|f(x) - f(y)\| \leq q\|x - y\|$ for any x, y in the ball B with center x_{m+1} and radius $\eta = \|x_n - x_{m+1}\|$.

Proposition 11. *If x_{m+1} enters a contractive neighbourhood with factor $q \leq 1/2$, then for any $t \geq 1$ we have $x_{m+t+1} = f(x_{m+t})$ and there is a fixed point x^* satisfying $\|x_{m+1+t} - x^*\| \leq q^t \|x_{m+1} - x^*\|$.*

Proof. Let $\mu = \|x_i - x_{m+1}\|$ for some i in the support of ρ^m. By property (3) we have $\mu \leq \eta$. Thus, the support of ρ^m is in B. This implies that for any i in the support of ρ^m we have $\|f(x_{m+1}) - f(x_i)\| \leq \mu/2$. Since $x_{m+1} = \sum_{j=1}^m \rho_j^m f(x_j)$ we see that $\|f(x_{m+1}) - x_{m+1}\| \leq \mu/2$. Applying (3) again we see that for any $j \leq m$ we have $\|f(x_{m+1}) - x_j\| \geq 1/2\mu$. Thus, the pure strategy x^{m+1} is an I-equilibrium of the $(m+1)^{th}$ game. So $\rho_{m+1}^{m+1} = 1$ and $x_{m+2} = f(x_{m+1})$. Noting that the conditions of the Lemma hold for $m+2$, induction tells us that $x_{m+1+t} = f(x_{m+t}) \in B$ for all $t \geq 1$. Since C is complete x_{m+1+t} converges to a fixed point $x^* \in B$ in C and $\|x_{m+1+t} - x^*\| \leq q^t \|x_{m+1} - x^*\|$ for all t. ■

Finally, let us give a full description of the path of procedure in the one dimensional case with functions that have a single fixed point. Once again we assume that in each iteration m if the pure strategy m is an I-equilibrium, then $\rho_m^m = 1$ is selected.

Proposition 12. *Let $f\colon [a,b] \to [a,b]$ be a continuous function with a single fixed point $x^* \in (a,b)$. If $\{x_m\}$ is the result of applying (1) to f, then for each m the point x_m is uniquely chosen and*

$$x_m = \max_{\{x_i \leq x^* :\, i \leq m\}} x_i = \underline{x} \quad \text{or} \quad x_m = \min_{\{x_i \geq x^* :\, i \leq m\}} x_i = \overline{x}$$

with

$$x_{m+1} = \begin{cases} f(x_m) & \text{if } x_m = \overline{x} \text{ and } f(x_m) \geq 1/2(\overline{x} + \underline{x}), \\ f(x_m) & \text{if } x_m = \underline{x} \text{ and } f(x_m) \leq 1/2(\overline{x} + \underline{x}), \\ 1/2(\overline{x} + \underline{x}) & \text{otherwise}. \end{cases}$$

In particular, if there is $0 \leq q < 1$ such that $\|f(x) - x^*\| \leq q\|x - x^*\|$ for any x, then $\|x_{m+1} - x^*\| \leq q^m \|x_{m+1} - x^*\|$ for all m.

Proof. Choose any x_1. Because there is only one fixed point x^*, if $x_1 \geq x^*$, then $x_2 = f(x_1) \leq x_1$ and if $x_1 \leq x^*$, then $x_2 = f(x_1) \geq x_1$. Thus, the condition is satisfied. Suppose that the condition is satisfied for m. By assumption, x_m is equal to either $\overline{x}$ or $\underline{x}$. Suppose, without loss that $x^* \neq x_m = \overline{x} \neq \underline{x}$. This implies that $\underline{x}$ is not a fixed point, that $f(\underline{x}) \geq \underline{x}$, and that $f(\underline{x}) \geq 1/2(\underline{x} + \overline{x})$, because otherwise $f(\underline{x})$ could have been chosen in an earlier iteration but wasn't contradicting the uniqueness property.

We know that $f(x_m) < x_m$. The equilibrium ρ^m must satisfy $\rho_m^m > 0$ for otherwise there was a point in a previous iteration that could have been chosen but wasn't. If $f(x_m)$ is greater than $1/2(\underline{x} + \overline{x})$, then it is chosen since the only possible equilibrium is $\rho_m^m = 1$. If $f(x_m) \leq 1/2(\underline{x} + \overline{x})$, then $x_{m+1} = 1/2(\underline{x} + \overline{x})$. This is because it is chosen if $\rho_m^m = 1$ and if $0 < \rho_m^m < 1$, which are the only possible cases. ■

4.2 Discussion

Our algorithms have some similarities with simplicial continuation methods (e.g., Scarf (1973), Doup (1988), and references cited in those books) that compute approximate fixed points by finding completely labelled simplices in a simplicial subdivision of the space C. (Similar remarks pertain also to the other algorithms developed in Scarf (1967b,a) and Hansen and Scarf (1969).) In each case there is a combinatoric existence result, namely Sperner's lemma and the existence of Nash equilibrium for an imitation game respectively. In each case there is an algorithmic implementation of this result, namely Scarf's procedure for computing a completely labelled simplex and the Lemke paths algorithm.

Simplicial subdivision methods require a space C that can be subdivided into arbitrarily small simplices and an algorithm for doing so. For the main applications in economic theory, general equilibrium theory and game theory, such algorithms exist, but they are nontrivial. Each algorithm requires C to have a specific geometry or, at the very least, that C is contained in a known bounded set with a specific geometry. In contrast, our procedure requires an initial point x_1 and a computational procedure for passing from a point $x \in C$ to a point $y \in F(x)$, but it is not necessary to know C, and under various conditions our algorithm can be applied to correspondences with unbounded domains (see Proposition 8 for example).

Almost all procedures for computing approximate fixed points are iterative, using the output of one stage as the input to later iterations that refine the initial approximation. The first simplicial algorithms need to be started at one of the vertices of the simplex, and do not present an obvious method for taking advantage of the results of previous calculations. This problem was recognized in the earlier literature, and an important goal was to find some way to "restart" the algorithm at a point that was thought to be close to a fixed point. The idea was to first run the algorithm on a coarse subdivision of C, then, having found an initial approximate fixed point (i.e., a "completely labelled simplex") run the algorithm for a finer subdivision, starting near this point. Variants that achieve the desired effect by adding an additional dimension and, in effect, following a simplicial homotopy, were developed by Eaves (1971a) and Merrill (1972a,b), among others. The analogous problem for our procedure has a simple solution:the Lemke path algorithm can be started at any pure strategy of the imitation game, and the natural choice is the pure strategy representing the most recently computed x_{m+1}.

When applied to a linear operator $T\colon \mathbb{R}^n \to \mathbb{R}^n$ our algorithm belongs to the well studied class of Krylov iterative methods, which in the context of solutions to linear equations are discussed in Ipsen and Meyer (1998). These are methods that generate a sequence of approximations x_m such that to compute each new approximation x_{m+1} we need only to know the first m powers of the T-orbit of x_1. That is, $x_1, Tx_1, \ldots, T^m x_2$.[8] There always exists a $k \leq n$ for which $x_1, Tx_1, T^2x_1, \ldots, T^k x_1$ are linearly dependent and k could be much smaller than n, for instance when T has n independent eigenvectors with $k-1$ distinct

[8]For each m the vector x_{m+1} of (1) is a convex combination of $Tx_1, T^2x_1, \ldots, T^m x_1$.

eigenvalues. So as with any Krylov method we require at most k evaluations of T to compute any pair x_m, y_m of the sequence (1). This is significant when each evaluation Tx is costly. In contrast, any completely labeled simplex in a simplicial subdivision requires n evaluations of T and thus complete information about the operator.

Simplicial subdivision algorithms can spend a lot of time pivoting through parts of the space that are far from fixed points, because the speed of motion is bounded by the mesh[9] of the subdivision divided by the dimension. (Each pivot moves from a simplex to another simplex that shares all but one vertex, so one cannot move between vertices that have no common vertices in fewer pivots than the dimension plus one.) Our algorithms are potentially capable of quickly "zooming in" to the vicinity of a fixed point, thereby avoiding this sort of slow and steady march through C. Initial numerical experiments show this potential being realized.

Our discussion has focused on the specific procedure (1), but in fact there are many variants. The guarantee of eventual discovery of an approximate fixed point depends on the sequence of points eventually becoming dense in the convex hull of the sequence. It is possible to preserve this guarantee without keeping every element of the sequence $x_1, \ldots, x_m$ (see Proposition 7). For example, if several terms in this sequence are close to each other but distant from x_m, one might discard all but one representative of this cluster. In the case of a continuous function $f : C \to C$, when only x_m is retained our procedure amounts to iterative evaluation: $x_{m+1} = f(x_m)$. Although iterative evaluation is not guaranteed to converge unless f is a contraction mapping, it is nonetheless very popular in practice. Thus, apart from the local iterative behaviour described above our methods constitute a class of variations of iterative evaluation that may prove useful, either because they are faster or because they are applicable to a larger class of problems.

A potential disadvantage of our approach is that at the m^{th} stage we are computing an I-equilibrium for an $m \times m$ matrix. Roughly, one might expect the burden of this step to be proportional to m^2. Several factors mitigate this concern. First, in many of the examples we have examined to date, convergence occurs before m becomes large. Second, at every point ρ in the Lemke path, all but at most one of the elements of the support of ρ are best responses to ρ. Typically the number of best responses will be at most one more than the dimension of C, so the dimension of the space containing the Lemke path is in effect fixed, and does not grow indefinitely as m increases. There is also some assurance that the burden of computing an equilibrium of each of the imitation games need not explode when C is finite dimensional. Indeed, it can be shown that if C is d dimensional, then there is a polynomial time procedure that accepts a $2m$-tuples $(x_1, \ldots, x_m, y_1, \ldots, y_m)$ of points of (1) as an input and which outputs an I-equilibria of the imitation game (A, I) where A is the $m \times m$ matrix with entries $a_{ij} = -\|x_i - y_j\|^2$.

We conclude by noting that using the Lemke paths algorithm to compute

[9]The *mesh* of a simplicial subdivision is the maximum diameter of any simplex.

x_{m+1} is natural, and has certain advantages, but any other method of computing an I-equilibrium of the derived imitation game is acceptable. Although there is no obvious reason to think it would be practical, it is nonetheless interesting to note that recursive versions of the algorithms are possible because for large m one may use the algorithm itself to compute an I-equilibrium!

4.3 Examples

A very simple example highlights the "zooming in" property discussed above.

Example 4.1. Let $C = [-1, 1]$ and $f(x) = qx$ for some $-1 < q < 1$. The only fixed point of f is $x = 0$. Suppose, that we want to approximate this fixed point to a degree of accuracy ε. That is, we want the algorithms to find some $\|x\| < \varepsilon$. We begin the Scarf algorithm and the algorithm in (1) at the point $x_1 = 1$. The Scarf algorithm divides the line $\mathbb{R}$ into a mesh of size ε. It takes at least $1/\varepsilon$ steps to obtain a completely labeled order interval. If $q > 0$ or $-1/2 < q < 1/2$, then the sequence in (1) takes the form $x_{m+1} = f(x_m)$ because $f(x_m)$ is closer to x_m than to x_ℓ for any $1 \leq \ell < m$. For $-1 < q \leq -1/2$ the sequence takes the following form. For any odd m we have $x_{m+1} = f(x_m)$ and any even m we have $x_{m+1} = 1/2(x_m - x_{m-1})$. In both cases $\|x_{m+1}\| \leq q^m$ and the algorithm take at most $\ln(1/\varepsilon)/\ln(1/q)$ steps to obtain $\|x_{m+1}\| \leq \varepsilon$. Notice that for q close to -1 the algorithm is much faster than simply iterating the function with $x_{m+1} = f(x_m)$. ■

A numerical experiment shows how the predictions of Proposition 11 can occur in practice. In the experiment each randomly chosen function has expansionary regions but also has a locally contractive fixed point. For these functions we do not know how to implement the Scarf algorithm or of any other standard fixed point approximation method that is guaranteed to approximate a fixed point.

Example 4.2. For any $x \in \mathbb{R}^{500}$, let $x^3 := (x_1^3, x_2^3, \ldots, x_{500}^3)$ and

$$\arctan(x) := (\arctan(x_1), \arctan(x_2), \ldots, \arctan(x_{500}))\,.$$

We seek to solve the system of equations

$$f(x) := -\arctan(M(x-y)^3) + y = x\,,$$

where M is a randomly chosen 500×500 matrix with norm greater than 3 and y is a randomly chosen vector in $\mathbb{R}^{500}$.

The function f takes points from $C = \mathbb{R}^{500}$ to $[-\pi/2, \pi/2]^{500} + y$. Noting that the derivative of the one dimensional function arctan is $\frac{1}{1+x^2}$ we see that in a neighborhood close to y the function f approximates $-M(x-y)^3 + y$. Therefore, the one known fixed point of this function $x^* = y$ has a neighborhood N in which x^* is attractive. If we choose $M = 3I$, the function has a unique fixed point that is attractive in the neighbourhood (approximately) $[-0.62, 0.62]^{500} + y$, which

$f(x) := \arctan(M(x-y)^3) + y$ and $y_i \in \mathcal{N}(0,1)$.

	N	< 200	$x^* = y$	f^t	Mean	Med.	Max.	Min.
$M_{ij} \in -\lvert\mathcal{N}(0,1)\rvert$	1000	1000	1000	0	9.06	9	22	5
$M_{ij} \in \mathcal{N}(0, 1/13^2)$	1000	979	979	0	33.11	26	150	13

< 200: successful experiments stopped with $\|f(x^*) - x^*\|_\infty < 10^{-5}$.

$x^* = y$: successful experiment for which $\|x^* - y\| < 10^{-5}$.

f^t: experiments in which simple iteration converged.

Mean/Med.: the mean/median number of iterations of successful experiments.

Max./Min.: the maximum/minimum number of iterations amongst successful experiments.

Table 1: A numerical example with a locally attractive fixed point

has a much smaller volume than $[-\pi/2, \pi/2]^{500} + y$. Moreover, the function is neither a contraction or nonexpansionary.

We ran two sets of 1000 numerical experiments. In the first, for each ij a number m_{ij} is chosen randomly from a normal distribution with mean zero and variance one and we set $M_{ij} = |m_{ij}|$. In the second set the elements of M are randomly chosen from a normal distribution with zero mean and variance $1/13^2$. Numerical tests show that this ensures that the ℓ_2 norm of M remains around 3.4.

In each experiment the elements of y are chosen from a random distribution with zero mean and variance one. The algorithm is started at a random point whose elements are chosen from a normal distribution with mean zero and variance one. Each experiment stopped when

$$|(f(x) - x)_i| < 10^{-5} \quad i = 1, 2, \ldots, 500$$

or when the number of iterations exceeds 200. In each experiment we also simply iterated the function 1000 times and tested if these iterations converged to a fixed point. We did this using the same starting point as our algorithm and using a starting point whose elements are uniformly chosen from $[-\pi/2, \pi/2]^{500} + y$.

The results for this experiment are shown in Table 1. In the second set of experiments the algorithm is slower, since when M is not positive there could be other attractive fixed points that "draw" the sequence away from y. ■

5 Basic Properties of Imitation Games

This section discusses various issues related to imitation games. Collectively, these results show that imitation games constitute a simple subclass of the class of two person games that nonetheless embodies the complexity (in various senses) of general two person games. One manifestation of this, studied in this section, is a relationship between I-equilibria of imitation games and symmetric equilibria of symmetric games.

A *symmetric game* is a two person game (A, A^T) where A is an $m \times m$ matrix. A *symmetric equilibrium* of (A, A^T) is a $\rho \in \Delta^m$ such that (ρ, ρ) is a

Nash equilibrium of (A, A^T). The following result is essentially due to Gale and Tucker (1950).

Proposition 13. *Suppose B and C are $m \times n$ matrices whose entries are all positive, and let*

$$A = \begin{bmatrix} 0 & B \\ C^T & 0 \end{bmatrix}.$$

For $\rho \in \Delta^{m+n}$ the following are equivalent:

(a) ρ is a symmetric equilibrium of (A, A^T);

(b) there are $\sigma \in \Delta^m$, $\tau \in \Delta^n$, and $0 < \alpha < 1$ such that:

  (i) $\rho = ((1-\alpha)\sigma, \alpha\tau)$,

  (ii) (σ, τ) is a Nash equilibrium of (B, C), and

  (iii) $(1-\alpha)\sigma^T B\tau = \alpha\sigma^T C\tau$.

Proof. First suppose that ρ is a symmetric equilibrium of (A, A^T). Then $\rho = ((1-\alpha)\sigma, \alpha\tau)$ for some $\sigma \in \Delta^m$, $\tau \in \Delta^n$, and $0 \le \alpha \le 1$. Since $(\sigma, 0)^T A(\sigma, 0) = 0 = (0, \tau)^T A(0, \tau)$ and the entries of B and C are all positive, it cannot be the case that $\alpha = 0$ or $\alpha = 1$. Since $\alpha < 1$, in the game (B, C) the strategy σ is a best response for agent 1 to τ, and similarly τ is a best response for agent 2 to σ. In addition, $(\sigma, 0)$ and $(0, \tau)$ are both best responses to ρ, so $(1-\alpha)\sigma^T B\tau = \alpha\sigma^T C\tau$.

Now suppose that (b) holds. It is easily verified that $(\sigma, 0)$ and $(0, \tau)$ are best responses to $\rho := ((1-\alpha)\sigma, \alpha\tau)$ in (A, A^T), so any convex combination of $(\sigma, 0)$ and $(0, \tau)$, such as ρ, is also a best response to ρ. ■

This result implies that any computational problem related to Nash equilibrium of two player games, for instance finding a sample equilibrium or finding all equilibria, can be recast as a problem concerning symmetric equilibria of symmetric games. The symmetric games derived from two player games as above have a special structure, so it seems that the problems related to symmetric equilibria of symmetric games are at least as hard as those related to Nash equilibrium of two player games.

We now describe concepts from computer science that allow precise formal expression of this idea. An algorithm is *polynomial time* if its running time is bounded by a polynomial function of the size of the input. A computational task is *polynomial* if there is a polynomial time algorithm that accomplishes it, and the class of such tasks is denoted by **P**. Given two computational tasks P and Q, a *reduction* from Q to P is a pair of maps, one of which takes an input x for Q to an input $r(x)$ for P, and the other of which takes an output y of P to an output $s(y)$ of Q, such that s transforms the desired output of P for $r(x)$ to the desired output of Q for x. The reduction is a *polynomial time reduction*[10] if

[10] Other conditions on the reduction can also be considered (Papadimitriou, 1994, Section 8.1).

the size of the output of P for $r(x)$ is bounded by a polynomial function of the size of x and there are polynomial time algorithms that compute the values of r and s. The result above gives a polynomial time reduction passing from a two player game to a symmetric game whose set of symmetric equilibria mirrors the set of Nash equilibria of the given game. In this sense any computational task related to symmetric equilibria of symmetric games is at least as hard as the corresponding problem for Nash equilibria of two person games. For example, if the problem of finding a symmetric equilibrium of a symmetric game is in **P**, then so is the problem of finding a Nash equilibrium of a two person game.

We use imitation games to go in the other direction: problems associated with Nash equilibrium of two person games are at least as hard as the corresponding problems related to symmetric equilibria of symmetric games.

Proposition 14. *For an $m \times m$ matrix A and $\rho \in \Delta^m$ the following are equivalent:*

(a) ρ is a symmetric equilibrium of (A, A^T);

(b) ρ is an I-equilibrium of (A, I);

(c) there is $\iota \in \Delta^m$ such that (ι, ρ) is a Nash equilibrium of (A, I).

Proof. The equivalence of (a) and (b) is immediate. The equivalence of (b) and (c) is Lemma 1. ■

A fourth reformulation of the problem should also be mentioned. A *linear complementarity problem* is a problem of the form

$$z \geq 0,\ q + Az \leq 0,\ \langle z, q + Az \rangle = 0$$

where the $m \times m$ matrix A and the vector $q \in \mathbb{R}^m$ are given. The problem is said to be *monotone* if all the entries of A are positive. Suppose that this is the case, and that $q = (-1, \ldots, -1)$. If z is a solution, then $\rho := z / \sum_{i=1}^m z_i$ is an equilibrium of the imitation game (A, I), since the complementarity condition $\langle z, q + Az \rangle = 0$ means precisely that each pure strategy for the first agent is either unused (that is, $z_i = 0$) or gives the maximal expected payoff. Conversely, if ρ is an equilibrium of the imitation game (A, I), and $v := \rho^T A \rho$, then $z := \rho / v$ solves the linear complementarity problem above. The extensive literature on the linear complementarity problem is surveyed in Murty (1988) and Cottle et al. (1992).

We have shown that there is a polynomial time reduction passing between any two of the following problems.

(i) Find a Nash equilibrium of a two person game.

(ii) Find an I-equilibrium of an imitation game.

(iii) Find a symmetric equilibrium of a symmetric game.

(iv) Find a solution of a monotone LCP.

It is not known whether finding a Nash equilibrium of a two person game is in **P**. Papadimitriou (2001) has described this problem as (along with factoring) "the most important concrete open question on the boundary of **P**." The fact that (i) is in **P** if and only if (ii)-(iv) are each in **P** lends support to his view. The equivalence, up to polynomial time reduction, between these solution concepts holds also for other computational problems such as those shown by Gilboa and Zemel (1989) to be **NP**-complete (e.g., determining whether there is more than one solution) and finding all solutions.

6 Lemke Paths from Lemke-Howson

This section explains how the Lemke paths algorithm may be regarded as a projection of the Lemke-Howson algorithm applied to an imitation game. We begin by describing the Lemke-Howson algorithm for a general two player game that satisfies a general position condition. We then specialize to imitation games and relate what we obtain to the description of the Lemke paths algorithm given in Section 3.

Let (A, B) be a two person game, where A and B are $m \times n$ matrices. We index the rows and columns of A and B, and the components σ_i and τ_j of a mixed strategy profile $(\sigma, \tau) \in \Delta^m \times \Delta^n$, by the elements of

$$\mathcal{J}_1 := \{1, \dots, m\} \quad \text{and} \quad \mathcal{J}_2 := \{m+1, \dots, m+n\}$$

respectively. For $\sigma \in \Delta^m$ the indices of the first agent's unused strategies are the elements of $\sigma^\circ := \{\, i \in \mathcal{J}_1 \colon \sigma_i = 0 \,\}$, and the indices of the second agent's pure best responses are the elements of $\overline{\sigma} := \operatorname{argmax}_{j \in \mathcal{J}_2} (B^T \sigma)_j$. Similarly, for $\tau \in \Delta^n$ let $\tau^\circ := \{\, j \in \mathcal{J}_2 \colon \tau_j = 0 \,\}$ and $\overline{\tau} := \operatorname{argmax}_{i \in \mathcal{J}_1} (A\tau)_i$. Then (σ, τ) is a Nash equilibrium if and only if each pure strategy is either unused or a best response:

$$\sigma^\circ \cup \overline{\tau} = \mathcal{J}_1 \quad \text{and} \quad \tau^\circ \cup \overline{\sigma} = \mathcal{J}_2.$$

For $W_1, Z_1 \subset \mathcal{J}_1$ and $W_2, Z_2 \subset \mathcal{J}_2$ set $W := (W_1, W_2)$ and $Z := (Z_1, Z_2)$, and define

$$\tilde{S}(W, Z) := \{\, (\sigma, \tau) \in \Delta^m \times \Delta^n \colon \sigma^\circ = W_1,\, \overline{\tau} = Z_1,\, \tau^\circ = W_2,\, \overline{\sigma} = Z_2 \,\}.$$

The next result, and Lemma 16 below, are analogues of Lemmas 2 and 3, with proofs that are similar and consequently omitted.

Lemma 15. *If $\tilde{S}(W, Z)$ is nonempty, then it is convex and:*

(a) $\mathcal{J}_1 \setminus W_1$, Z_1, $\mathcal{J}_2 \setminus W_2$, *and* Z_2 *are all nonempty;*

(b) The closure of $\tilde{S}(W, Z)$ is $\bigcup \tilde{S}(W', Z')$ where the union is over all (W', Z') with $W_1' \supset W_1$, $Z_1' \supset Z_1$, $W_2' \supset W_2$, and $Z_2' \supset Z_2$.

We say that (A, B) is in *general position* if $|\sigma^\circ| + |\overline{\sigma}| \le m$ for all $\sigma \in \Delta^m$ and $|\tau^\circ| + |\overline{\tau}| \le n$ for all $\tau \in \Delta^n$. Throughout this section we assume that this is the case, so that if (σ, τ) is a Nash equilibrium, then these inequalities hold with equality and σ°, $\overline{\tau}$, τ°, and $\overline{\sigma}$ are pairwise disjoint. As was the case earlier, a failure of general position implies that the entries of A and B express a system of linear equations with more equations than unknowns that nonetheless has a solution, so the set of pairs (A, B) in general position is dense in the space of all such pairs.

Let

$$|(W, Z)| := |W_1| + |Z_1| + |W_2| + |Z_2|.$$

Lemma 16. *If (A, B) is in general position and $\tilde{S}(W, Z)$ is nonempty, then:*

(a) $\tilde{S}(W, Z)$ is $(m + n - |(W, Z)|)$-dimensional;

(b) $\tilde{S}(W', Z')$ is nonempty for all (W', Z') with $W'_1 \subset W_1$, $\emptyset \neq Z'_1 \subset Z_1$, $W'_2 \subset W_2$, and $\emptyset \neq Z'_2 \subset Z_2$.

Consider (W, Z) for which $\tilde{S}(W, Z) \neq \emptyset$. If $|(W, Z)| = m + n$, then $\tilde{S}(W, Z)$ is a singleton (by (a) above) whose unique element, denoted by $\tilde{V}(W, Z)$, is called a *vertex*. If $|(W, Z)| = m + n - 1$, then the closure of $\tilde{S}(W, Z)$, denoted by $\tilde{E}(W, Z)$, is a one dimensional (again by (a)) line segment that we call an *edge*.

We say that an edge $\tilde{E}(W, Z)$ is *horizontal* if $|W_1| + |Z_2| = m$ and $|W_2| + |Z_1| = n - 1$. If $|W_1| + |Z_2| = m - 1$ and $|W_2| + |Z_1| = n$, then we say that $\tilde{E}(W, Z)$ is *vertical*. A horizontal edge is a cartesian product of a singleton in Δ^m and a line segment in Δ^n, while a vertical edge is a cartesian product of a line segment in Δ^m and a singleton in Δ^n. As we will see shortly, the Lemke-Howson algorithm alternates between horizontal and vertical edges.

Each edge has two endpoints that are vertices. If $\tilde{V}(W, Z)$ is an endpoint of $\tilde{E}(W', Z')$, then $W'_1 \subset W_1$, $Z'_1 \subset Z_1$, $W'_2 \subset W_2$, $Z'_2 \subset Z_2$, and $|(W', Z')| = |(W, Z)| - 1$. If, for example $W'_1 = W_1 \setminus \{i\}$ while $Z'_1 = Z_1$, $W'_2 = W_2$, and $Z'_2 = Z_2$, then we say that (W', Z') *is obtained from* (W, Z) *by dropping* i *from* W_1. Suppose that $\tilde{V}(W, Z)$ is a vertex. If (W', Z') is obtained by dropping i from W_1 or W_2, then $\tilde{S}(W', Z')$ is nonempty by (b) above, so that $\tilde{E}(W', Z')$ is defined. Similarly, if (W', Z') is obtained by dropping i from Z_1 or Z_2, then $\tilde{E}(W', Z')$ is defined if and only if the resulting Z'_1 and Z'_2 are nonempty.

Fix an arbitrary $\tilde{s} \in \mathcal{J}_1 \cup \mathcal{J}_2$. A vertex $\tilde{V}(W, Z)$ is an $\tilde{s}$*-vertex* if

$$(\mathcal{J}_1 \cup \mathcal{J}_2) \setminus \{\tilde{s}\} \subset W_1 \cup Z_1 \cup W_2 \cup Z_2.$$

An edge $\tilde{E}(W, Z)$ is an $\tilde{s}$*-edge* if

$$W_1 \cup Z_1 \cup W_2 \cup Z_2 = (\mathcal{J}_1 \cup \mathcal{J}_2) \setminus \{\tilde{s}\}.$$

Clearly the endpoints of an $\tilde{s}$-edge are $\tilde{s}$-vertices; let $\tilde{e}(\tilde{E}(W, Z))$ be the two element set containing these endpoints. Let

$$\tilde{G}_{\tilde{s}} := (\tilde{V}_{\tilde{s}}, \tilde{E}_{\tilde{s}})$$

where $\tilde{V}_{\tilde{s}}$ is the set of $\tilde{s}$-vertices and $\tilde{E}_{\tilde{s}}$ is the set of $\tilde{s}$-edges. Then $\tilde{G}_{\tilde{s}}$ (with the relationship between edges and vertices given by $\tilde{e}$) is an undirected graph. The Lemke-Howson algorithm follows a path in $\tilde{G}_{\tilde{s}}$.

As in our analysis of the Lemke paths algorithm earlier, we give a taxonomy of pairs consisting of an $\tilde{s}$-vertex and an $\tilde{s}$-edge that has that vertex as an endpoint. For the sake of definiteness we assume that $\tilde{s} \in \mathcal{J}_1$ (the other case is similar) and we begin with those vertices in which the first agent's strategy is $\delta_{\tilde{s}}$. The general position assumption implies that there is a unique pure best response to $\delta_{\tilde{s}}$: $\overline{\delta_{\tilde{s}}} = \{\tilde{t}\}$ for some $\tilde{t} \in \mathcal{J}_2$. In turn general position implies that $\delta_{\tilde{t}}$ has a unique best response: $\overline{\delta_{\tilde{t}}} = \{i\}$ for some $i \in \mathcal{J}_1$. Then $(\delta_{\tilde{s}}, \delta_{\tilde{t}})$ is the $\tilde{s}$-vertex $\tilde{V}(W, Z)$ where

$$W_1 = \mathcal{I}_1 \setminus \{\tilde{s}\},\ Z_1 = \{i\},\ W_2 = \mathcal{I}_1 \setminus \{\tilde{t}\},\ Z_2 = \{\tilde{t}\}.$$

What $\tilde{s}$-edges have $(\delta_{\tilde{s}}, \delta_{\tilde{t}})$ as an endpoint? If $i = \tilde{s}$, so that $(\delta_{\tilde{s}}, \delta_{\tilde{t}})$ is a Nash equilibrium, then $(\delta_{\tilde{s}}, \delta_{\tilde{t}})$ is not an endpoint of any $\tilde{s}$-edge $\tilde{E}(W', Z')$. The only possibility is the (W', Z') obtained from (W, Z) by dropping $\tilde{s}$ from Z_1, but $\tilde{S}(W', Z') = \emptyset$ because $Z_1' = \emptyset$.

Next suppose that $i \neq \tilde{s}$, so that $(\delta_{\tilde{s}}, \delta_{\tilde{t}})$ is not a Nash equilibrium. If $\tilde{E}(W', Z')$ is an $\tilde{s}$-edge that has $(\delta_{\tilde{s}}, \delta_{\tilde{t}})$ as an endpoint, then (W', Z') is obtained from (W, Z) by dropping i from either W_1 or Z_1. If (W', Z') is obtained by dropping i from W_1, then Lemma 16 implies that $\tilde{E}(W', Z')$ is nonempty, but i cannot be dropped from Z_1 without making it empty, so there is exactly one $\tilde{s}$-edge having $(\delta_{\tilde{s}}, \delta_{\tilde{t}})$ as an endpoint.

Now consider an $\tilde{s}$-vertex $(\sigma, \tau) = \tilde{V}(W, Z)$ with $\sigma \neq \delta_{\tilde{s}}$. As above, the number of $\tilde{s}$-edges having (σ, τ) depends on whether this point is a Nash equilibrium.

If (σ, τ) is a Nash equilibrium, then the only possibility for an $\tilde{s}$-edge with (σ, τ) as an endpoint is $\tilde{E}(W', Z')$ where (W', Z') is obtained from (W, Z) by dropping $\tilde{s}$. There always is such an $\tilde{s}$-edge because $\sigma \neq \delta_{\tilde{s}}$ implies that $W_1 \setminus \{\tilde{s}\} \neq \emptyset$.

If (σ, τ) is not a Nash equilibrium, then one of $W_1 \cap Z_1$ and $W_2 \cap Z_2$ is a singleton while the other is empty. If $W_1 \cap Z_1 = \{j\}$, then there are two edges with (σ, τ) as an endpoint, namely $\tilde{E}(W', Z')$ and $\tilde{E}(W'', Z'')$ where (W', Z') and (W'', Z'') are obtained by dropping j from W_1 and Z_1 respectively. In particular, note that j cannot be the unique element of Z_1 because then $W_1 \subset \mathcal{J}_1 \setminus \{\tilde{s}\}$ would have $m - 1$ elements, and this can only happen if $\sigma = \delta_{\tilde{s}}$. Similarly, if $W_2 \cap Z_2 = \{j\}$, then there are two edges with (σ, τ) as an endpoint, namely $\tilde{E}(W', Z')$ and $\tilde{E}(W'', Z'')$ where (W', Z') and (W'', Z'') are obtained by dropping j from W_2 and Z_2 respectively. In this case j cannot be the unique element of Z_2 because W_2 has at most $n - 1$ elements, $W_2 \cup Z_2 = \mathcal{J}_2$, and $W_2 \cap Z_2 = \{i\}$. Thus, regardless of which of $W_1 \cap Z_1$ and $W_2 \cap Z_2$ is a singleton and which is empty, (σ, τ) is an endpoint of two $\tilde{s}$-edges. *Note that one of these edges is horizontal and the other is vertical.*

Summarizing, $(\delta_{\tilde{s}}, \delta_{\tilde{t}})$ is either a Nash equilibrium, in which case it is not an endpoint of any $\tilde{s}$-edge, or an endpoint of precisely one $\tilde{s}$-edge. Every other $\tilde{s}$-vertex is an endpoint of precisely one or two $\tilde{s}$-edges according to whether it is or is not a Nash equilibrium. The path of $\tilde{s}$-edges that begins at $(\delta_{\tilde{s}}, \delta_{\tilde{t}})$ is unbranching, alternates between horizontal and vertical edges, and cannot return to any $\tilde{s}$-vertex that it has already visited. Since the set of $\tilde{s}$-vertices is finite, it must eventually arrive at a Nash equilibrium. The *Lemke-Howson algorithm* follows this path. Note that there is one path of the algorithm for each $\tilde{s} = 1, \ldots, m+n$.

We now specialize to the case of an imitation game, so henceforth $m = n$ and $B = I$. Define $d : \mathcal{J}_1 \cup \mathcal{J}_2 \to \mathcal{I}$ by setting

$$d(j) = \begin{cases} j, & j \in \mathcal{J}_1, \\ j - m, & j \in \mathcal{J}_2. \end{cases}$$

Then $d(\sigma^\circ) \cap d(\overline{\sigma}) = \emptyset$ for all $\sigma \in \Delta^m$. If (σ, τ) is a vertex or an element of a horizontal edge, then $|\sigma^\circ| + |\overline{\sigma}| = m$, so $d(\sigma^\circ \cup \overline{\sigma}) = \mathcal{I}$ and σ is the uniform distribution on $d(\overline{\sigma})$. If (σ, τ) is an element of a vertical edge $\tilde{E}(W, Z)$, then σ is an element of the line segment between the uniform distribution on $d(Z_2)$ and the uniform distribution on $\mathcal{I} \setminus d(W_1)$.

The Lemke paths algorithm, as it was described in Section 3, has the following graph-theoretic expression. Let $s := \tilde{s}$. (Recall that we are assuming that $\tilde{s} \in \mathcal{J}_1$. If we were assuming that $\tilde{s} \in \mathcal{J}_2$ we would set $s := \tilde{s} - m$.) Let

$$G_s = (V_s, E_s)$$

where V_s is the set of s-vertices and E_s is the set of s-edges. For $E(X, Y) \in E_s$ let $e(E(X, Y))$ be the two element subset of V_s containing the endpoints of $E(X, Y)$. The Lemke path is the path of edges in G_s that starts at δ_s.

The next three results describe how $\tilde{G}_{\tilde{s}}$ and G_s are related.

Lemma 17. *If $(\sigma, \tau) = \tilde{V}(W, Z) \in \tilde{V}_{\tilde{s}}$, then $V(d(W_2), d(Z_1)) \in V_s$.*

Proof. Since $\tau^\circ = W_2$ and $\overline{\tau} = Z_1$, $\tau \in S(d(W_2), d(Z_1))$. General position implies $|W_2| + |Z_1| = m$, so (by another application of general position) $S(d(W_2), d(Z_1))$ is a singleton because $|d(W_2)| + |d(Z_1)| = m$. Thus $V(d(W_2), d(Z_1)) = \tau$. We claim that $\tau \in V_s$, i.e., $\mathcal{I} \setminus \{s\} \subset d(W_2 \cup Z_1)$. We have $|(W, Z)| = 2m$ and

$$(\mathcal{J}_1 \cup \mathcal{J}_2) \setminus \{\tilde{s}\} \subset W_1 \cup Z_1 \cup W_2 \cup Z_2,$$

so d maps at least two elements of the disjoint union of W_1, Z_1, W_2, and Z_2 to each element of $\mathcal{I} \setminus \{s\}$. In addition $|W_1| + |Z_2| = m$ and $d(W_1 \cup Z_2) = \mathcal{I}$. ■

This result implies that there is a function $\pi_V : \tilde{V}_{\tilde{s}} \to V_s$ given by

$$\pi_V(\tilde{V}(W, Z)) := V(d(W_2), d(Z_1)).$$

In this sense vertices in $\tilde{V}_{\tilde{s}}$ "project" onto vertices in V_s. The next result shows that a vertex $V(X, Y) \in V_s$ "lifts" to the pair of endpoints of a vertical edge unless it is an I-equilibrium and $s \in X$, in which case it lifts to a single vertex in $\tilde{V}_{\tilde{s}}$.

Lemma 18. *Suppose $V(X,Y) = \rho \in V_s$. If $X \cup Y = \mathcal{I}$ and $s \in X$, then there is a unique element of $\tilde{V}_{\tilde{s}}$ that is mapped to $V(X,Y)$ by π_V. Otherwise $V(X,Y)$ is the image of precisely two elements of $\tilde{V}_{\tilde{s}}$, and these are the endpoints of a vertical edge.*

Proof. Let τ be ρ reinterpreted, via the bijection $d|_{\mathcal{J}_2}$, as a probability measure on $\mathcal{J}_2$. Below we will construct various (W,Z) such that $\pi_V(\tilde{V}(W,Z)) = \rho$. Necessarily $d(W_2) = X$ and $d(Z_1) = Y$, so all of these will have

$$W_2 := d^{-1}(X) \cap \mathcal{J}_2 \quad \text{and} \quad Z_1 := d^{-1}(Y) \cap \mathcal{J}_1$$

in common. Let σ be the uniform distribution on Z_1.

First suppose that $X \cap Y = \{i\}$. Given W_2 and Z_1, if

$$W_1 \cup Z_1 \cup W_2 \cup Z_2 = (\mathcal{J}_1 \cup \mathcal{J}_2) \setminus \{\tilde{s}\},$$

then either

$$W_1 = d^{-1}(X) \cap \mathcal{J}_1 \quad \text{and} \quad Z_2 = d^{-1}((Y \setminus \{i\}) \cup \{d(\tilde{s})\}) \cap \mathcal{J}_2$$

or

$$W_1 = d^{-1}(X \setminus \{i\}) \cap \mathcal{J}_1 \quad \text{and} \quad Z_2 = d^{-1}(Y \cup \{d(\tilde{s})\})) \cap \mathcal{J}_2.$$

Both of these possibilities do in fact, have associated $\tilde{s}$-vertices (σ, τ) and (σ', τ), where σ' is the uniform distribution on $\mathcal{J}_1 \setminus d^{-1}(X \setminus \{i\})$, and these are clearly the two endpoints of a vertical edge.

Now suppose that $X \cup Y = \mathcal{I}$. One way to complete the definition of (W,Z) is by setting

$$W_1 := d^{-1}(X) \cap \mathcal{J}_1 = \mathcal{J}_1 \setminus Z_1 \quad \text{and} \quad Z_2 := d^{-1}(Y) \cap \mathcal{J}_2 = \mathcal{J}_2 \setminus W_2.$$

Clearly $\tilde{V}(W,Z) = (\sigma,\tau) \in \tilde{V}_{\tilde{s}}$ and $\pi_V(\tilde{V}(W,Z)) = V(X,Y)$.

If $\tilde{V}(W',Z')$ is another element of $\tilde{V}_{\tilde{s}}$ with $\pi_V(\tilde{V}(W',Z')) = V(X,Y)$, then $W_2' = W_2$ and $Z_1' = Z_1$, and of course

$$(\mathcal{J}_1 \cup \mathcal{J}_2) \setminus \{\tilde{s}\} \subset W_1' \cup Z_1' \cup W_2' \cup Z_2'.$$

If $s \in X$, then there is no such (W',Z') because (due to our maintained assumption that $\tilde{s} \in \mathcal{J}_1$) $\tilde{s} \in Z_1$, so that $W_1' \cup Z_1' \cup W_2' \cup Z_2' = \mathcal{J}_1 \cup \mathcal{J}_2$, whence $(W',Z') = (W,Z)$. If $\tilde{s} \in W_1$, then such a (W',Z') can be obtained by following the edge resulting from removing $d(\tilde{s})$ from W_1 to its other endpoint (σ',τ). Since we need $d(\sigma'^{\circ} \cup \overline{\sigma}') = \mathcal{I}$, necessarily Z_2' is obtained by adding $\tilde{s} + m$ to Z_2. All this is feasible: if $W_2' = W_2$, $Z_1' = Z_1$, $W_1' = W_1 \setminus \{\tilde{s}\}$, and $Z_2' := Z_2 \cup \{\tilde{s} + m\}$, then $\tilde{V}(W',Z') = (\sigma',\tau) \in \tilde{V}_{\tilde{s}}$ where σ' is the uniform distribution on $\mathcal{J}_1 \setminus W_1'$. Clearly $\tilde{V}(W,Z)$ and $\tilde{V}(W',Z)$ are the endpoints of the vertical edge $\tilde{E}((W_1',W_2),(Z_1,Z_2))$. ∎

There is a bijection between the horizontal edges in $\tilde{E}_{\tilde{s}}$ and the edges in E_s, with π_V projecting endpoints onto endpoints.

Lemma 19. *If $\tilde{E}(W,Z) \in \tilde{E}_{\tilde{s}}$ is a horizontal edge, then $E(d(W_2), d(Z_1)) \in E_s$. Let $\pi_E : \tilde{E}(W,Z) \mapsto E(d(W_2), d(Z_1))$ be the corresponding function from horizontal edges in $\tilde{E}_{\tilde{s}}$ to E_s. Then π_E is a bijection, and $\pi_V \circ \tilde{e} = e \circ \pi_E$.*

Proof. If $(\sigma, \tau) \in \tilde{S}(W,Z)$, then $\tau^\circ = W_2$ and $\overline{\tau} = Z_1$, so $S(d(W_2), d(Z_1))$ is nonempty because it contains τ. In addition, $|(W,Z)| = 2m - 1$,

$$(\mathcal{J}_1 \cup \mathcal{J}_2) \setminus \{\tilde{s}\} = W_1 \cup Z_1 \cup W_2 \cup Z_2,$$

$|W_1| + |Z_2| = m$, and $d(W_1 \cup Z_2) = \mathcal{I}$, so $d(W_2 \cup Z_1) = \mathcal{I} \setminus \{s\}$. Thus $E(d(W_2), d(Z_1)) \in E_s$ and π_E is defined.

Suppose we are given $E(X,Y) \in E_s$. Then $W_2 := d^{-1}(X) \cap \mathcal{J}_2$, $Z_1 := d{-}1(Y) \cap \mathcal{J}_1$, $W_1 := \mathcal{J}_1 \setminus (Z_1 \cup \{\tilde{s}\})$ and $Z_2 := \mathcal{J}_2 \setminus W_2$ are the unique sets with all the properties enumerated above. For (W,Z) defined in this way, $\tilde{S}(W,Z)$ is nonempty because it contains (σ,τ) whenever $\tau \in S(X,Y)$ and σ is the uniform distribution on $\mathcal{J}_1 \setminus W_1$, so $\tilde{E}(W,Z) \in \tilde{E}_{\tilde{s}}$. Thus π_E is bijective. The endpoints of $\tilde{E}(W,Z)$ are obtained by adding elements to W_2 and Z_1, so they must be mapped by π_V to endpoints of $E(d(W_2), d(Z_1))$, and it is easy to see how to obtain each endpoint of $E(d(W_2), d(Z_1))$ in this way. ■

Taken together, these results give the following picture. Any path in $\tilde{G}_{\tilde{s}}$ projects, via π_V and π_E, onto a path in G_s. Any path in G_s "lifts" to a path in $\tilde{G}_{\tilde{s}}$ in the sense that there is a path in $\tilde{G}_{\tilde{s}}$ that projects onto it, and the lifted path is unique up to vertical edges above the initial and final vertex of the given path in G_s.

An important consequence of this observation is that it provides a new proof of a recent result of Savani and von Stengel (2004) concerning "long" Lemke-Howson paths. Morris (1994) gives a sequence of examples of problems for which the length of the shortest Lemke path is an exponential function of the size of the problem. As Savani and Stengel point out, these paths can be interpreted as Lemke-Howson paths that compute symmetric equilibria of the derived symmetric game, but there are asymmetric equilibria that are reached very quickly by the Lemke-Howson algorithm. The construction above is a method of passing from an instance of the Lemke paths algorithm to an imitation game for which all Lemke-Howson paths project onto Lemke paths of the given problem, so it automatically produces a sequence of examples of two person games for which the length of the shortest Lemke-Howson path is an exponential function of the size of the game.

7 Short Paths in Geometric Games

In the recursive sequence (1) of Kakutani's fixed point theorem we need to find an I-equilibrium of an imitation game with a particular geometric derivation. In this section we study such imitation games in more detail. Let L be an inner product space of fixed dimension d. The *geometric imitation game* (A, I) induced by $x_1, \ldots, x_m, y_1, \ldots, y_m \in L$ is the one in which the entries of the $m \times m$ matrix A are $a_{ij} = -\|x_i - y_j\|^2$.

How rich is the class of geometric imitation games? As we have defined the concept, if (A, I) is a geometric imitation game, then the entries of A are non-positive, so there are imitation games that are not geometric imitation games. However, there is a less restrictive sense in which every imitation game can be realized as an imitation game. We will say that $m \times m$ matrices A and A' are *equivalent* if it is possible to pass from A to A' by some finite sequence of the following transformations: (a) adding a constant to all entries in some column of the matrix; (b) multiplying all entries by a positive scalar. We say that imitation games (A, I) and (A', I) are *equivalent* if A and A' are equivalent in this sense. Equivalent imitation games have the same best response correspondence for the mover, and consequently they have the same Lemke paths.

The *dimension* of (A, I) is the smallest d such that (A, I) is equivalent to the imitation game induced by some $x_1, \ldots, x_m, y_1, \ldots, y_m \in \mathbb{R}^d$.

Proposition 20. *For any $m \times m$ matrix A the dimension of (A, I) is at most $m - 1$.*

Proof. Let $e_1, e_2, \ldots, e_m$ be the standard unit basis vectors of $\mathbb{R}^m$, and let $\mathbf{e} := (1, \ldots, 1) \in \mathbb{R}^m$. Set $x_1 := e_1, \ldots, x_m := e_m$. Let $p : \mathbb{R}^m \to \mathbb{R}^m$ be the function

$$p(y) = (-\|y - x_1\|^2, \ldots, -\|y - x_m\|^2).$$

Let π be the orthogonal projection of $\mathbb{R}^m$ onto the hyperplane

$$H := \{\, z \in \mathbb{R}^m : \mathbf{e}^T z = 0 \,\},$$

and let $q := \pi \circ p$. Obviously A is equivalent to a matrix A' whose columns are all contained in H. It is easy to compute that the matrix of $Dp(\frac{1}{m}\mathbf{e})$ is $2(I - \frac{1}{m}\mathbf{e}\mathbf{e}^T)$. In particular, if $v \in H$, then $Dq(\frac{1}{m}\mathbf{e})v = Dp(\frac{1}{m}\mathbf{e})v = 2v$. Applying the inverse function theorem, for a sufficiently small neighborhood U of $\mathbf{e}/m$ there exist $\varepsilon > 0$ and $y_1, \ldots, y_m \in U \cap (\frac{1}{m}\mathbf{e} + H)$ such that $q(y_1), \ldots, q(y_m)$ are the columns of $\varepsilon A'$. Let A'' be the matrix whose columns are $p(y_1), \ldots, p(y_m)$. Since $p(y) - q(y)$ is always a scalar multiple of $\mathbf{e}$, $\varepsilon A'$ and A'' are equivalent. Finally note that $x_1, \ldots, x_m, y_1, \ldots, y_m$ are contained in the $(m-1)$-dimensional hyperplane $\frac{1}{m}\mathbf{e} + H$. ■

We now show that the Lemke paths of a d-dimensional geometric imitation game are "short," relative to m, if d is fixed but m is allowed to grow. Given the work in the previous sections, we see that geometric imitation games in a fixed finite dimensional space are a class of two person games in which the paths of the Lemke-Howson algorithm are "short" and a class of symmetric games for which the computational problem of finding a symmetric Nash equilibrium is in **P**. In this sense the next result complements the work of of Savani and von Stengel (2004) and Morris (1994) concerning "long" Lemke-Howson and Lemke paths.

Lemma 21. *If L is a finite dimensional inner product space, then for an open dense subset in L^{2m} of geometric imitation games the length of the sequences $\mathcal{LP}$ for the induced imitation games are bounded by a polynomial function of m.*

Proof. We will use the terminology in the proof of Kakutani's fixed point theorem. By (3) if $\rho \in \Delta^m$, then $i \in \overline{\rho}$ if and only if x_i minimizes the distance from $x_1, x_2, \ldots, x_m$ to $\sum_{j=1}^m \rho_i y_i$. Therefore, for an open dense subset of geometric imitation games, $|\overline{\rho}| \leq d+1$ for all $\rho \in \Delta^m$, where d is the dimension of L. If ρ is an element of an edge $E(X,Y) \in E_s$, then $\overline{\rho}$ is either Y or $Y \cup \{s\}$, and $X \cup Y$ contains every index except s, so (X,Y) is completely determined by Y. Thus the number of nonempty edges in E_s is not greater than the number of $d+1$ element subsets of $\{1,2,\ldots,m\}$, which is bounded by a polynomial function of m. ■

8 Concluding Remarks

We have given a new proof of Kakutani's fixed point theorem that passes quickly from the existence of Nash equilibria in two person games to the desired conclusion. The two person games arising in this argument are imitation games, and the Lemke paths algorithm provides a simple proof of Nash equilibrium existence for these games.

The proof of Kakutani's theorem is based on a new algorithm for computing approximate fixed points and points to a number of variations on the algorithm. Such algorithms have attractive features that may be useful. There is certainly a great deal to do in the direction of understanding the dynamics of the algorithms. This seems likely to be a fruitful direction for further theoretical research.

The study of imitation games has led to other interesting findings. The Lemke paths algorithm has been displayed as resulting from applying the Lemke-Howson algorithm to a suitable imitation game. This shows that "long" Lemke paths are in fact "long" Lemke-Howson paths. The geometric version of imitation games have been shown to provide a class of games with "short" Lemke and Lemke-Howson paths. These two observations provide additional insights on the works of Savani and von Stengel (2004) and Morris (1994).

Several important equilibrium concepts have been shown to be of comparable computational complexity, insofar as there are polynomial time reductions pass between them. In McLennan and Tourky (2005) we use imitation games to give simple proofs of results of Gilboa and Zemel (1989) that have also recently been reproved by Conitzer and Sandholm (2003), Codenotti and Štefanovič (2005), and Blum and Toth (2004). In their work Codenotti and Štefanovič (2005) and Bonifaci et al. (2005) also use imitation games to prove new computational complexity results. One may hope that in other ways as well the imitation game concept will have a unifying and simplifying influence on the study of computational issues related to two person games and the linear complementarity problem.

References

K. J. Arrow and F. H. Hahn. *General Competitive Analysis*. Holden Day, San Francisco, 1971.

A. Blum and B. Toth. NP-hardness involving Nash equilibria in a non-degenerate setting. unpublished, 2004.

V. Bonifaci, U. Di Iorio, and L. Laura. On the complexity of uniformly mixed Nash equilibria and related regular subgraph problems. unpublished, 2005.

L.E.J. Brouwer. Uber Abbildung von Mannigfaltikeiten. *Mathematische Annalen*, 71:97–115, 1910.

B. Codenotti and D. Štefanovič. On the computational complexity of Nash equilibria for $(0,1)$ bimatrix games. forthcoming in *Information Processing Letters*, 2005.

V. Conitzer and T. Sandholm. Complexity results about Nash equilibria. *Proceedings of the 18th International Joint Conference on Artificial Intelligence*, pages 765–771, 2003.

R. W. Cottle, J. Pang, and R. E. Stone. *The linear complementarity problem*. Computer Science and Scientific Computing. Academic Press Inc., Boston, MA, 1992. ISBN 0-12-192350-9.

G. Debreu and I. N. Herstein. Nonnegative square matrices. *Econometrica*, 21: 597–607, 1953.

T. Doup. *Simplicial Algorithms on the Simplotope*. Springer-Verlag, Berlin, 1988.

B. Curtis Eaves. The linear complementarity problem. *Management Science*, 17:612–634, 1971a.

B. Curtis Eaves. Computing Kakutani fixed points. *SIAM J. Appl. Math.*, 21: 236–244, 1971b.

S. Eilenberg and D. Montgomery. Fixed-point theorems for multivalued transformations. *American Journal of Mathematics*, 68:214–222, 1946.

H. Gale, D. Kuhn and A. W. Tucker. On symmetric games. *Annals of Math. Studies*, 24:81–87, 1950.

I. Gilboa and E. Zemel. Nash and correlated equilibria: some complexity considerations. *Games Econom. Behav.*, 1(1):80–93, 1989.

T. Hansen and H. Scarf. On the applications of a recent combinatorial algorithm. Cowles Foundation Discussion Paper, No. 272, 1969.

P. J.-J. Herings. An extremely simple proof of the K-K-M-S theorem. *Economic Theory*, 10:361–367, 1997.

I. C. F. Ipsen and C. D. Meyer. The idea behind Krylov methods. *Amer. Math. Monthly*, 105(10):889–899, 1998.

S. Kakutani. A generalization of Brouwer's fixed point theorem. *Duke Math. J.*, 8:457–459, 1941.

H. W. Kuhn. Approximate search for fixed points. In *Computing Methods in Optimization Problems*. Springer-Verlag, Berlin, 1969. Proceedings of San Remo 1968 conference.

C. E. Lemke. Bimatrix equilibrium points and mathematical programming. *Management Sci.*, 11:681–689, 1965.

C. E. Lemke. On complementarity pivot theory. In *Mathematics of the Decision Sciences*, pages 95–114. American Mathematical Society, Providence, 1968.

C. E. Lemke and J. T. Howson. Equilibrium points of bimatrix games. *J. Soc. Indust. Appl. Math.*, 12:413–423, 1964.

A. McLennan and R. Tourky. Simple complexity from imitation games. unpublished, 2005.

O. H. Merrill. A summary of techniques for computing fixed points of continuous mappings. In Richard H. Day and Stephen M. Robinson, editors, *Mathematical Topics in Economic Theory and Computation*. Society for Industrial and Applied Mathematics, Philadelphia, 1972a.

O. H. Merrill. Application and extensions of an algorithm that computes fixed points of upper semicontinuous point to set mappings. Ph.D. Dissertation, University of Michigan, 1972b.

J. Milnor. *Topology from the Differentiable Viewpoint*. University Press of Virginia, Charlottesville, 1965.

J. Milnor. Analytic proofs of the "hairy ball theorem" and the Brouwer fixed-point theorem. *The American Mathematical Monthly*, 85(7):521–524, 1978.

W. D. Morris, Jr. Lemke paths on simple polytopes. *Math. Oper. Res.*, 19(4): 780–789, 1994.

K. G. Murty. *Linear complementarity, linear and nonlinear programming*. Heldermann, Berlin, 1988. Sigma serties in applied mathematics; 3.

J. Nash. Non-cooperative games. *Ann. of Math. (2)*, 54:286–295, 1951.

J. F. Nash. Equilibrium points in n-person games. *Proc. Nat. Acad. Sci. U. S. A.*, 36:48–49, 1950.

C. H. Papadimitriou. Algorithms, games and the internet. In *Annual ACM Symposium on the Theory of Computing*, pages 749–253, 2001.

C. H. Papadimitriou. *Computational Complexity*. Adison Wesley Longman, New York, 1994.

R. Savani and B. von Stengel. Exponentially many steps for finding a nash equilibrium in a bimatrix game. *Proc. 45th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2004)*, 2004.

H. Scarf. *The computation of economic equilibria*. Yale University Press, New Haven, Conn., 1973. With the collaboration of Terje Hansen, Cowles Foundation Monograph, No. 24.

H. E. Scarf. The core of an N person game. *Econometrica*, 35:50–69, 1967a.

H. E. Scarf. The approximation of fixed points of a continuous mapping. *SIAM J. Appl. Math.*, 15:1328–1343, 1967b.

L. Shapley and R. Vohra. On Kakutani's fixed point theorem, the K-K-M-S theorem, and the core of a balanced game. *Economic Theory*, 1:108–116, 1991.

L. S. Shapley. On balanced games without side payments. In T.C. Hu and Stephen Robinson, editors, *Mathematical Programming*, pages 261–290. Academic Press, New York, 1973a.

L. S. Shapley. On balanced games without side payments—a correction. The Rand Paper Series Report No. p-4910/1, 1973b.

L. S. Shapley. A note on the Lemke-Howson algorithm. *Math. Programming Stud.*, 1:175–189, 1974. Pivoting and extensions.

M. J. Todd. *The computation of fixed points and applications*. Springer-Verlag, Berlin, 1976. Lecture Notes in Economics and Mathematical Systems, Vol. 124.