From Imitation Games to Kakutani*

Andrew McLennan[†] and Rabee Tourky[‡]

March 30, 2006

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. They give new insights into the "long" paths of the Lemke-Howson algorithm. They provide a rich class of games with "short" Lemke-Howson paths. They are useful for studying the complexity of a number of problems in computational economics.

Keywords: Computational economics, Lemke-Howson algorithm, Imitation games, Geometric imitation games, Kakutani's fixed point theorem, Approximate fixed point, Computational complexity, Long and short Lemke-Howson paths.

---

*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, the University of Minnesota Combinatorics Workshop, the University of Pensylvania Economic Theory Workshop, and the Penn State Micro Theory Seminar. We have benefitted from stimulating correspondence with Christos Papadimitriou and useful comments by Bernhard von Stengel.

[†]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.

[‡]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 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 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.

Very briefly, our proof and the associated computational procedure have the following description. Suppose that C is a nonempty compact convex subset of an inner product space and that $F\colon C \twoheadrightarrow C$ is an 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 $\{x_1, \dots, x_m\}$, the mover's payoff matrix is the $m \times m$ matrix

[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.

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}. The proof in Section 3 continues along the following lines: x_{m+1} is a convex combination of elements of the images $F(x)$ of nearby points x, because this is an imitation game the support of ρ^m is a subset of the support of ι^m; the diameters of the supports of ι^m and ρ^m decrease to zero as the number of iterations increases, because the sequence x_m is contained in a compact set; limit points of $\{x_m\}$ are exact fixed points, because F is upper semicontinuous and convex valued.

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. In Section 6 we show that the result of adopting this perspective is 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 consequence of this observation is that it is possible to derive a recent result

[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).

of Savani and von Stengel (2004, 2006) concerning long paths of the Lemke-Howson algorithm from earlier work by Morris (1994) concerning long paths of the Lemke paths algorithm. Morris (1994) uses dual cyclic polytopes to produce exponentially long Lemke paths. Savani and von Stengel (2004, 2006) observe that these paths correspond to a symmetric version of the Lemke-Howson algorithm and that the symmetric games obtained directly from Morris' construction have non-symmetric equilibria that are very quickly found by the Lemke-Howson algorithm. Savani and von Stengel (2004, 2006) therefore modify the construction and obtain games with unique equilibria and long Lemke-Howson paths. We observe that Morris' constructions correspond to imitation games with unique equilibria and long Lemke-Howson paths.

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 $x_1, \ldots, x_m, y_1, \ldots, y_m \in C$, 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, 2006) and Morris (1994).

We also give some additional results concerning imitation games. Gale and Tucker (1950) showed how to derive a symmetric game from a two player game in such a way that the Nash equilibria of the two player game are in one-to-one correspondence with the symmetric equilibria of the symmetric game. Consequently the problem of finding a symmetric equilibrium of a symmetric game 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 results in Gilboa and Zemel (1989) on the complexity of certain computational problems concerning Nash equilibria of two person games. The imitation games introduced in the present paper have also been recently used to study the complexity of various computational problems arising in economics see for instance Codenotti and Štefanovič (2005); Bonifaci, Di Iorio, and Laura (2005); Codenotti, Saberi, Varadarajanz, and Ye (2005).

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 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 problem 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 (1967a,b) 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—showing that an arbitrarily small neighborhood of the graph of the correspondence contain the graph of a continuous function, then applying 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. Very recently a striking result of Chen and Deng (2005b) (described in more detail in Section 5) has given a different, and quite profound, expression of this idea. We should also mention

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

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 (1971b), 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 (1967a) 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 (1967a) 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 (1971a) and refined by Merrill (1972b,a). 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 $(\sigma, \tau) \in \Delta^m \times \Delta^n$ such that $\sigma^T A\tau \geq \tilde{\sigma}^T A\tau$ for all $\tilde{\sigma} \in \Delta^m$ and $\sigma^T B\tau \geq \sigma^T B\tilde{\tau}$ for all $\tilde{\tau} \in \Delta^n$.

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*, respectively. Let $\mathcal{I} := \{1, \ldots, m\}$. 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:

Proposition 3.1. *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). Because 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 $S(X, Y) := \{\, \rho \in \Delta^m : \rho^\circ = X \text{ and } \overline{\rho} = Y \,\}$.

Lemma 3.2. *If $S(X, Y)$ is nonempty, then its dimension is at least $m - |X| - |Y|$.*

Proof. Let $A(Y, \mathcal{I} \setminus X)$ be the submatrix of A that is the intersection of the rows and columns indexed by the elements of Y and $\mathcal{I} \setminus X$ respectively. The claim follows from elementary linear algebra because $S(X, Y)$ is an open subset of the set of $z \in \mathbb{R}^{\mathcal{I} \setminus X}$ such that $\sum_j z_j = 1$ and $A(Y, \mathcal{I} \setminus X)z$ is an element of the diagonal in $\mathbb{R}^Y$. ■

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

Lemma 3.3. *The set of $m \times m$ matrices in general position is dense in the space of $m \times m$ matrices.*

Proof. Given an $m \times m$ matrix A_0 and $X, Y \subset \mathcal{I}$ with $|Y| - 1 \geq m - |X|$, we can find an A arbitrarily close to A_0 such that $\mathbb{R}^{\mathcal{I} \setminus X}$ is spanned by the vectors obtained by taking the difference between the first row of $A(Y, \mathcal{I} \setminus X)$ and each subsequent row. For such an A the origin is the only $z \in \mathbb{R}^{\mathcal{I} \setminus X}$ such that $A(Y, \mathcal{I} \setminus X)z$ is in the diagonal of $\mathbb{R}^Y$, so $S(X, Y) = \emptyset$. The set of such A is open, so we can repeat this maneuver, eventually obtaining an A arbitrarily near A_0 with $S(X, Y) = \emptyset$ for all $X, Y \subset \mathcal{I}$ with $|X| + |Y| > m$. ■

General position also has the following simple consequence.

Lemma 3.4. *If A is in general position and $S(X, Y)$ is nonempty, then $S(X, Y)$ is $(m - |X| - |Y|)$-dimensional.*

Proof. Suppose by way of contradiction that the dimension of $S(X,Y)$ is greater than $m-|X|-|Y|$. We may assume that (X,Y) is maximal in the sense that if $X' \supseteq X$ and $Y' \supseteq Y$ have this property, then $X'=X$ and $Y'=Y$. General position implies that $m-|X|-|Y| \geq 0$. Therefore, the closure of $S(X,Y)$ is a polytope of positive dimension, and any of its facets is the closure of $S(X',Y')$ for some (X',Y') with $X \subseteq X'$ and $Y \subseteq Y'$, where at least one inclusion is strict. Therefore, there exists such (X',Y') satisfying

$$\dim S(X',Y') = \dim S(X,Y) - 1 > m - |X| - |Y| - 1 \geq m - |X'| - |Y'|\,,$$

which contradicts maximality. ■

Lemma 3.5. *If A is in general position and $S(X,Y)$ is nonempty, then $S(X',Y')$ is not empty for any $X' \subseteq X$, and nonempty $Y' \subseteq Y$.*

Proof. Fix $i_0 \in Y'$. Then $S(X,Y)$ is the set of $\rho \in \mathbb{R}^m$ such that certain affine functionals have certain signs:

(i) $\rho_1 + \cdots + \rho_m - 1 = 0$;

(ii) $\rho_i = 0$ for all $i \in X$;

(iii) $\rho_i > 0$ for all $i \in \mathcal{I} \setminus X$;

(iv) $(A\rho)_i - (A\rho)_{i_0} = 0$ for all $i \in Y \setminus \{i_0\}$;

(v) $(A\rho)_i - (A\rho)_{i_0} < 0$ for all $i \in \mathcal{I} \setminus Y$.

Since $S(X,Y)$ is $m-|X|-|Y|$ dimensional, the affine functionals that vanish on $S(X,Y)$ have linearly independent linear parts. (That is, we think of each affine functional as a linear functional plus a constant.) Therefore any neighborhood of a $\rho \in S(X,Y)$ contains points at which these functionals take on any combination of signs. In particular, there is a ρ' arbitrarily near ρ such that (i) holds, $\rho'_i = 0$ for all $i \in X'$, $\rho'_i > 0$ for all $i \in X \setminus X'$, $(A\rho')_i - (A\rho')_{i_0} = 0$ for all $i \in Y' \setminus \{i_0\}$, and $(A\rho')_i - (A\rho')_{i_0} < 0$ for all $i \in Y \setminus Y'$. For ρ' sufficiently close to ρ it will also be the case that (iii) and (v) hold, in which case $\rho' \in S(X',Y')$. ■

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. If $V(X, Y)$ is a vertex and $i \in X$, then the last result implies that $S(X \setminus \{i\}, Y)$ is nonempty, so $E(X \setminus \{i\}, Y)$ is an edge that has $V(X, Y)$ as one of its endpoints. Similarly, if $i \in Y$, *and i is not the unique element of Y*, then $E(X, Y \setminus \{i\})$ is an edge that has $V(X, Y)$ as an endpoint.

The Lemke paths algorithm follows a path of edges described in the next proof.

Proposition 3.6. *The imitation game (A, I) has an I-equilibrium.*

Proof. 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, by Lemma 3.3, we may assume that A is in general position[4].

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

A vertex $V(X, Y)$ is said to be an *s-vertex* if $\mathcal{I} \setminus \{s\} \subset X \cup Y$, in which case either $|X \cap Y| = 1$ or $|X \cap Y| = 0$. In the latter case $X \cup Y = \mathcal{I}$ and $V(X, Y)$ is an I-equilibrium. An edge $E(X, Y)$ is said to be an *s-edge* if $X \cup Y = \mathcal{I} \setminus \{s\}$. Of course the endpoints of an s-edge are s-vertices.

Let $V(X, Y)$ be an s-vertex with $X \cap Y = \{i\}$. The only s-edges that could have $V(X, Y)$ as an endpoint are $E(X \setminus \{i\}, Y)$ and $E(X, Y \setminus \{i\})$. The first of these is always defined, and the second is defined if and only if $Y \setminus \{i\} \neq \emptyset$. If $Y = \{i\}$, then $X = \mathcal{I} \setminus \{s\}$, so $V(X, Y) = \delta_s$, and $i = i^*$ because $S(X, Y)$ is nonempty.

Summarizing, if there is no I-equilibrium, then every s-vertex $V(X, Y)$ has $|X \cap Y| = 1$, the s-vertex $\delta_s = V(\mathcal{I} \setminus \{s\}, \{i^*\})$ is an end point of exactly one s-edge, and every other s-vertex is an endpoint of exactly 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. ■

[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.

The Lemke path algorithm for computing an I-equilibrium begins at δ_s and proceeds from there along the path of s-edges. We call this path the *Lemke path of A starting at δ_s*. This path does not branch or return to any s-vertex that it visited earlier, so it must eventually arrive at an I-equilibrium.

Let C be a closed convex subset of an inner product space. A *correspondence* $F : C \twoheadrightarrow C$ assigns a nonempty $F(x) \subseteq C$ to each $x \in C$. Given such an F, we may 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

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

where ρ^m is an I-equilibrium of the imitation game (A, I) with m strategies for which the entries of A are $a_{ij} = -\|x_i - y_j\|^2$. The following geometric fact underlies many of our results.

Proposition 3.7. *For each m we have*

$$\mathit{supp}\, \rho^m \subseteq \operatorname*{argmin}_{1 \le i \le m} \|x_i - x_{m+1}\| \,. \tag{2}$$

Proof. Note that $\sum_{j=1}^m \rho_j^m(x_{m+1} - y_j) = 0$, so for each $i = 1, \dots, m$ we have:

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

Since ρ^m is an I-equilibrium, its support is contained in the set of i that maximize this quantity, and the second term does not depend on i. ∎

For $x \in C$ let

$$G_F(x) := \bigcap_{\delta > 0} \overline{co\Big(\bigcup_{\|x' - x\| < \delta} F(x') \Big)}.$$

This correspondence majorizes F, i.e., $F(x) \subseteq G_F(x)$ for all x.

Proposition 3.8. *An accumulation point x^* of $\{x_m\}$ is a fixed point of G_F.*

Proof. Consider a particular $\delta > 0$. If the $\delta/3$-ball around x^* contains x_{m+1} and x_ℓ for some $\ell \leq m$, then x_ℓ is in the $2\delta/3$-ball around x_{m+1}, so the δ-ball around x^* contains all the points in $\{x_1, \dots, x_m\}$ closest to x_{m+1}. The last result implies that x_{m+1} is in the convex hull of $B_\delta := \bigcup_{\|x - x^*\| \leq \delta} F(x)$. There are arbitrarily large m such that x_{m+1} is arbitrarily close to x^*, so x^* is in the closure of the convex hull of B_δ. Since δ was arbitrary, the claim follows. ∎

The correspondence F is *upper semicontinuous* if, for each $x \in C$: (i) $F(x)$ is closed; (ii) for each open set V containing $F(x)$ there is a neighborhood U of x such that $F(x') \subseteq V$ for all $x' \in U$. Clearly, if F is upper semicontinuous with convex

values, then $F(x) = G_F(x)$ for all $x \in C$. Conversely, if $G_F = F$, then F is convex valued, and if the inner product space containing C is finite dimensional, then F is upper semi-continuous.

Theorem 3.9 (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. The sequence $\{x_m\}$ has an accumulation point x^* because C is compact, Proposition 3.8 gives $x^* \in G_F(x^*)$, and $G_F(x^*) = F(x^*)$ because F is upper semi-continuous with convex values. ■

4 Approximate Fixed Points

According to Proposition 3.8 any accumulation point of the sequence $\{x_m\}$ of (1) is a fixed point of G_F. If F and G_F have the same fixed points, then any algorithm that generates each 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 3.6. Thus, we have the raw materials for a class of algorithms that compute approximate fixed points.

In this section we study the properties of $\{x_m\}$, contrasting them with the properties of the simplex algorithm for finding approximate fixed points, presenting certain theoretical properties, and describing the results of numerical experiments. The emphasis is on the flexibility of our method and the relative paucity of information needed to apply it.

General discussion

The inputs for an algorithm for computing an approximate fixed point are a space C in which a fixed point is sought, an "oracle" that computes the value $f(x)$ of a function f at any point $x \in C$, and a bound $\varepsilon > 0$ on acceptable error. There is an a priori guarantee that f is continuous. (For particular classes of functions (e.g., affine, polynomial, smooth) there are, of course, many well studied procedures for solving systems of equations.) The goal is to find a point $x^* \in C$ such that $\|f(x^*) - x^*\| < \varepsilon$.

Note that there is no guarantee that x^* is near an actual fixed point of f. In fact such a guarantee cannot be attained without additional a priori restrictions: after the algorithm has run, The Devil could replace f with a continuous $f' : C \to C$ that has no fixed points near x^* even though it agrees with f at every point at which the value of f was computed. By virtue of similar reasoning, although the procedures we discuss here are algorithms—that is, they are guaranteed to halt in finite time—it is impossible to place an a priori upper bound on their worst-case running times. Consequently the complexity classes of theoretical computer science (polynomial time, **NP**, polynomial space, exponential time, etc.) cannot be applied to the comparison of algorithms for computing approximate fixed points. Since theoretical tools are lacking, such comparisons have an informal character, combining intuition, practical considerations, and actual computational experience.

Simplicial continuation methods (e.g., Scarf (1973), Doup (1988), and references cited in those books) are the best studied class of algorithms for computing an approximate fixed point. In these algorithms C (or sometimes a related space) is endowed with a simplicial subdivision. There is a rule for endowing each vertex of the subdivions with a "label." For example, the label of a vertex v may be defined to be the smallest index i such that $v_i \geq f_i(v)$. The procedure's objective is to find a "completely labelled simplex," which is one whose vertices have every label. If the values of f at the vertices are sufficiently close to each other, then any one of them is an approximate fixed point. A simplex is "almost completely labelled" if its vertices have all but one of the labels. An almost completely labelled simplex is either completely labelled or has two vertices with the same label, and the facet of the simplex opposite either of these vertices is either contained in the boundary of C or is also a facet of another almost completely labelled simplex. The main idea of the algorithm is to "pivot" from one almost completely labelled simplex to the next until a completely labelled simplex is reached. (There are rules for pivoting along simplices in the boundary that will not be described here.) The result asserting that this procedure eventually halts is a generalization of Sperner's lemma.

Our algorithm has certain broad similarities with these methods. 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. In each case continuity

implies that the results of these computations are eventually approximate fixed points.

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, as we will see below.

Many computational procedures (e.g., Newton's method for finding a zero of a differentiable univariate function) are iterative, using the output of one stage as the input to a subsequent calculation that refines the initial approximation. The first simplicial algorithms (Scarf (1967a,b)) require that C be a simplex, and they need to be started at one of its vertices. Having found an approximately labelled simplex for a coarse simplicial subdivision of C, we would like to find a completely labelled simplex in a finer subdivision, but the method does not present an obvious method for taking advantage of the results of the initial calculation. 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. Variants that achieve the desired effect by adding an additional dimension and, in effect, following a simplicial homotopy, were developed by Eaves (1971b) and Merrill (1972b,a), 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 candidate approximate fixed point x_m.

Each pivot of a simplicial algorithm moves from a simplex to another simplex that shares all but one vertex, so one cannot move between simplices that have no common vertices in fewer pivots than the dimension of the first simplex plus one. This implies that the speed of motion is bounded by the mesh[5] of the subdivision, and will also be slowed down by a factor proportional to the dimension of the problem unless the path of the algorithm is largely restricted to low dimensional

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

faces of C. In contrast, our algorithm is 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 $x_1, \ldots, x_m$ 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.

One might retain only the last n points, computing x_{m+1} from the imitation game derived from $x_{m-n+1}, \ldots, x_m$, which may be burdensome if m is large. There is no theoretical guarantee of convergence, but in the case of a continuous function $f : C \to C$, when only x_m is retained our procedure amounts to iterative evaluation—$x_{m+1} = f(x_m)$—which is very popular is practice, even though it is not guaranteed to converge unless f is a contraction mapping. Thus, our methods give a class of variations of iterative evaluation that may prove useful, either because they are faster or because they converge in 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. 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 for fixed d there is a polynomial time (in m) procedure that takes a $2m$-tuple $(x_1, \ldots, x_m, y_1, \ldots, y_m)$ of points in $\mathbb{R}^d$ as its input and outputs an I-equilibria of the imitation game (A, I) where A is the $m \times m$ matrix with entries $a_{ij} = -\|x_i - y_j\|^2$.

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, $x_1, Tx_1, \ldots, T^m x_2$.[6] There always exists a $k \le n$ 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 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.

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

Univariate functions

The next result gives a full description of the path of the procedure in the one dimensional case for functions that have a single fixed point. Corresponding to the possibility of starting the Lemke path algorithm at the vertex corresponding to the most recently computed point, we assume that in the m^{th} iteration, if the pure strategy m is an I-equilibrium, then $\rho^m_m = 1$ is selected.

Proposition 4.1. *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*

$$x_m = \max_{\{x_i \le x^* :\, i \le m\}} x_i = \underline{x} \quad \textit{or} \quad x_m = \min_{\{x_i \ge x^* :\, i \le m\}} x_i = \overline{x}$$

with

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

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

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

Proof. Choose any x_1. Because there is only one fixed point x^*, if $x_1 \ge x^*$, then $x_2 = f(x_1) \le x_1$ and if $x_1 \le x^*$, then $x_2 = f(x_1) \ge x_1$. 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 $x^* \ne x_m = \overline{x} \ne \underline{x}$. This implies that $\underline{x}$ is not a fixed point, that $f(\underline{x}) \ge \underline{x}$, and that $f(\underline{x}) \ge 1/2(\underline{x} + \overline{x})$, 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 was not chosen even though it could have been. If $f(x_m)$ is greater than $1/2(\underline{x}+\overline{x})$, then it is chosen because the only possible equilibrium is $\rho^m_m = 1$. If $f(x_m) \le 1/2(\underline{x}+\overline{x})$, 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 possible cases. ■

A very simple example contrasts the the performance of the procedure with the performance of the simplex method. For clarity we choose the one dimensional case though similar examples can be constructed in higher dimensional spaces.

Example 4.2. 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. In contrast, for our algorithm if $-1/2 \le q < 1$, then $x_m = qx_{m-1}$ for all $m > 1$, so $\|x_m\| = |q|^m$. If $-1 \le q < -1/2$, then for every even m we have $x_m = qx_{m-1} < 0$ and for every odd $m > 1$ we have $x_m = 1/2(x_{m-1} - x_{m-2}) > 0$. So beginning with the point $x_1 = 1$ it is easy to check that we achieve $\|x_m\| \le \epsilon$ for

a. $m \ge \frac{\log_2(1/\varepsilon)}{\log_2(1/|q|)}$, if $0 \ne q \ge -1/2$.

b. $m \ge 2\frac{\log_2(1/\varepsilon)}{\log_2(2/(1+q))} + 1$, if $-1 \ne q < -1/2$.

c. $m = 3$, if $q = -1$, and $m = 2$ if $q = 0$.

We note that such fast rates of converges occur for a wide variety of univariate as well as multivariate functions. ■

Local contractions

The next result presents an important local property that indicates that the procedure can be very fast in sufficiently contractive neighborhoods of a fixed point. As above, we suppose that in each iteration m, if the pure strategy of the imitation game corresponding to the most recently computed point x_m is an I-equilibrium, then this I-equilibrium 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 4.3. *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 Proposition 3.7 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 $\|f(x_{m+1}) - f(x_i)\| \leq \mu/2$. Because $x_{m+1} = \sum_{j=1}^{m} \rho_j^m f(x_j)$ we see that $\|f(x_{m+1}) - x_{m+1}\| \leq \mu/2$. Applying Proposition 3.7 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 Proposition hold for $m+2$, induction tells us that $x_{m+1+t} = f(x_{m+t}) \in B$ for all $t \geq 1$. Because 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. ■

We note that if q in the above result is greater than $1/2$ but smaller than one, then we can achieve the same type of behavior by applying our algorithm to f^t with $q^t \leq 1/2$.[7]

[7] We write $f^1(x) := f(x)$ and $f^t(x) = f(f^{t-1}(x))$ for $t > 1$. To guarantee convergence to a fixed point of f we need the function f to be t-acyclic. That is, for any $x \in C$ satisfying $f(x) \neq x$ we have $f^t(x) \neq x$.

Unbounded and nonconvex domains

Many fixed point problems arising from systems of equations have an unbounded domain. When, for example, the range $F(C)$ has compact closure the Bohnenblust and Karlin (1950) theorem guarantees the existence of a fixed point even in infinite dimensional spaces. Another example is the work of Brézis, Nirenberg, and Stampacchia (1972) on Ky Fan's minimax principle for functions with unbounded domains. We do not know of any standard fixed point approximation method that works for such functions.[8]

Our procedure works well when the image of some iterate of F is contained in a compact set. Let $S_C^0 := C$, let $S_C^1 := co\, F(C)$, and recursively let $S_C^i := co\, F(S_C^{i-1})$ for $i = 2, 3, \ldots$. A simple inductive argument shows that $S_C^{i+1} \subseteq S_C^i$ for all i. Let $\overline{S_C^\infty} := \bigcap_{i=1}^\infty \overline{S_C^i}$. In the finite dimensional case, $\overline{S_C^\infty}$ is non-empty and compact if and only if some $\overline{S_C^i}$ is compact.

Proposition 4.4. *Let C be a closed convex subset of $\mathbb{R}^n$ and let $F\colon C \twoheadrightarrow C$ be a correspondence with the same fixed points as G_F. If $\overline{S_C^\infty}$ is non-empty and compact, then $\{x_m\}$ is bounded.*

Proof. For some finite i the set S_C^i is bounded because $\overline{S_C^\infty}$ is compact. Thus, it suffices to show that for any i, eventually $x_m \in S_C^i$. Clearly, $x_m \in S_C^0 = C$ for all $m > 0$. Proceeding inductively, suppose that for some j there exists m_j such that $x_m \in S_C^j$ for all $m > m_j$.

Fix one of the finite number of points satisfying $x_i \notin S_C^j$, which is clearly not a fixed point. Assume by way of contradiction that there is a subsequence $\rho^{m'}$ of the I-equilibria satisfying $\rho_i^{m'} > 0$ for all m'. Proposition 3.8 implies that x_i cannot be a cluster point of $\{x_{m'+1}\}$, because x_i is not a fixed point, nor can the subsequence have any other cluster points, so $\|x_{m'+1}\| \to \infty$. We can assume that $x_i = 0$ because we are free to work in such a coordinate system. Let z be a limit point of the normalized sequence $\{x_{m'+1}/\|x_{m'+1}\|\}$. In general a point y' is as close to the origin as to another point y if and only if $\langle y', y\rangle \le \frac{1}{2}\langle y, y\rangle$, i.e., $\left\langle \frac{y'}{\|y'\|}, \frac{y}{\|y\|} \right\rangle \le \frac{1}{2}\frac{\|y\|}{\|y'\|}$. In particular, there must exist m', m'' with $m'' > m'$,

[8] From a computational point of view such problems can be unpleasant because, even though it is known that a suitable compact-convex restriction of F exists, it may be quite difficult to find one, or the ones that can be found may be very large, resulting in slower performance of subsequent steps in the procedure. In addition, computational methods for finding restrictions will typically pertain to very particular classes of problems.

$\|x_{m''+1}\| > \|x_{m'+1}\|$, and $x_{m'+1}/\|x_{m'+1}\|$ and $x_{m'+1}/\|x_{m'+1}\|$ arbitrarily close to z, but then the distance from $x_{m''+1}$ to $x_{m'+1}$ is less than $\|x_{m''+1}\|$. This is a contradiction because Proposition 3.7 implies that $x_i = 0$ is one of the points in $\{x_1, \dots, x_{m''}\}$ that are closest to $x_{m''+1}$. ∎

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. Let $C \subseteq \mathbb{R}^n$ be closed, but not necessarily convex, and let $f : C \to C$ be continuous with bounded range. The following 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 $f(x)$. If f maps the boundary of C inside a star shaped set the extra information needed is a center of the star shaped set.[9] We don't need to *a priori* know anything extra about the geometry of C.

Proposition 4.5. *Let C be an arbitrary 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 interior of a star shaped set $K \subseteq C$ with center x^*. The sequence $\{x_m\}$ of* (1) *applied to the function*

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

is bounded and its accumulation points are fixed points of f.

Proof. Each x_m for $m \geq 2$ is in the convex hull of $\{x_1\} \cup \{x^*\} \cup f(C)$, so $\{x_m\}$ is bounded. Proposition 3.8 tells us that any accumulation points x^* of $\{x_m\}$ is in $G_h(x^*)$. Any such x^* is in the interior of C and is thus a fixed point of f. ∎

Numerical experiments

A numerical experiment shows how the predictions of Proposition 4.3 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.

[9] 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.

Example 4.6. For any $x \in \mathbb{R}^{500}$, let $x^3 := (x_1^3, x_2^3, \dots, x_{500}^3)$ and

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

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 has a much smaller volume than $[-\pi/2, \pi/2]^{500} + y$. Moreover, the function is neither a contraction nor non-expansionary.

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

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

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

$f(x) := \arctan(M(x-y)^3) + y$ 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 among the successful experiments.

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

experiments the algorithm is slower, because 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 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 5.1. *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). We have $\rho = ((1-\alpha)\sigma, \alpha\tau)$ for some $\sigma \in \Delta^m$, $\tau \in \Delta^n$, and $0 \leq \alpha \leq 1$. Because of the identity $(\sigma, 0)^T A(\sigma, 0) = 0 = (0, \tau)^T A(0, \tau)$ and the entries of B and C are all positive, it cannot be the case that $\alpha = 0$ or $\alpha = 1$. Because $\alpha < 1$, 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 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

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

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 5.2. *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 Proposition 3.1. ■

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 $q = (-1, \ldots, -1)$. (Since the rows of A can be rescaled, the essential point is that all entries of q are negative.) If z is a solution, then $\rho := z / \sum_{i=1}^m z_i$ is an equilibrium of the imitation game (A, I), because 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).

The computational problem of finding a Nash equilibrium of a finite two person normal form game (A, B) is called *2-Nash.* We have shown that there is a polynomial time reduction passing between any two of the following problems.

(i) 2-Nash.

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

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

(iv) Finding a solution of a monotone LCP.

Papadimitriou (2001) has described the problem of determining whether 2-Nash has a polynomial time algorithm as (along with factoring) "the most important concrete open question on the boundary of **P**." The fact that 2-Nash has such an algorithm if and only if (ii)-(iv) do 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.

Recently, there has been important progress on the computational complexity of 2-Nash. Culminating a rapid sequence of developments (Goldberg and Papadimitriou (2005); Daskalakis et al. (2005); Daskalakis and Papadimitriou (2005); Chen and Deng (2005a)) Chen and Deng (2005b) have shown that 2-Nash is complete for the computational class **PPAD** (Papadimitriou (1994b)) which includes many computational problems related to fixed points, Nash equilibrium, and competitive equilibrium in economies. In general a subclass $\mathbf{C}'$ of a class of computational problems $\mathbf{C}$ is *complete* for $\mathbf{C}$ if every problem in $\mathbf{C}$ has a polynomial time reduction to a problem in $\mathbf{C}'$. Concretely, Chen and Deng have shown that any problem in **PPAD** has a polynomial time reduction to an instance of 2-Nash.[11]

A problem in **PPAD** has two inputs. The first is a Turing machine $\mathcal{M}$ with the following properties. It accepts inputs from the space $\{0,1\}^r$ of bitstrings of length r some natural number r. Given such an input, $\mathcal{M}$ outputs a predecessor in $\{0,1\}^r$ (or an indication that the bitstring is a source) and a successor in $\{0,1\}^r$ (or an indication that the bitstring is a sink). The input cannot coincide with either the predecessor or the successor. The input is the successor of its predecessor when it is not a source, and it is the predecessor of its successor when it is not a sink. Thus the Turing machine computes a directed graph whose vertices are the bitstrings and whose maximum in-degree and maximum out-degree are both one. The second input is a source, which (by virtue of the possibility of "flipping" certain bits) may

[11] Note that although 2-Nash is conventionally described as *the* problem of finding a Nash equilibrium, in formal precision it is the subclass of **PPAD** consisting of all problems associated with particular games (A, B).

be taken to be $0^r = (0, \dots, 0)$. The desired output is a bitstring that is either a sink or a different source. The size of the problem $(\mathcal{M}, 0^r)$ is the maximum running time of $\mathcal{M}$ for any input bitstring.

Given a continuous function from the n-dimensional unit simplex to itself, the Scarf algorithm follows a path in a graph to an approximate fixed point. Papadimitriou (1994b) showed how to display the problem the Scarf algorithm solves as a member of **PPAD** if there is a Turing machine that performs the requisite function evaluations in an amount of time that is bounded by a polynomial function of np, where p is the desired number of bits of accuracy.

Hirsch et al. (1989) studied a discrete version of Brouwer's fixed point theorem that is based on function evaluation. Specifically, one is given an "oracle" or "black box" that computes the value of a function $f : [0,1]^n \to [0,l]^n$, and the goal is to find a point x satisfying $\|f(x) - x\| \leq 2^{-p}$. They show that if $n \geq 3$ and $x \mapsto f(x) - x$ is known to be Lipschitzian with Lipschitz constant L, then any algorithm must, in the worst case, evaluate the function at at least $(c(2^p - 10)L)^{n-2}$ points, where c is a positive constant. Specifically, they construct a collection of example functions such that, when np is sufficiently large, any smaller number of function evaluations will necessarily fail to distinguish between two examples that have different fixed points.

Chen and Deng (2005b) embed the basic logical operations used to describe $\mathcal{M}$ in the two player game they construct. Consider the application of this to Turing machines $\mathcal{M}$ that express instances of **PPAD** that seek approximate fixed points of continuous functions $f : [0,1]^n \to [0,1]^n$. A polynomial (or merely subexponential) time algorithm for 2-Nash would then yield a polynomial (subexponential) time algorithm for such fixed point problems. This might happen because the algorithm's embedded "intelligence" was able to exploit special properties of functions f that are described by small Turing machines, or because it embedded methods for generating deep analyses of the Turing machines themselves. In view of the state of our toolbox for analyzing Turing machines, this seems fantastical, but by the same token there is currently little hope for a proof that it is impossible. In this slightly tentative sense Chen and Deng's result can be regarded as showing that there is no polynomial (subexponential) time algorithm for 2-Nash.

6 Lemke Paths from Lemke-Howson

The Lemke-Howson algorithm (Lemke and Howson (1964)) finds an equilibrium of a two person game. It can be described geometrically as a path in the cartesian product of the two agents' simplices of mixed strategies consisting of pivots, where each pivot holds one of the two mixed strategies fixed and changes the other mixed strategy by passing from one endpoint of a line segment to the other. The path alternates between the two agents, first changing one mixed strategy, then the other. Thus, the odd numbered pivots trace a piecewise linear path in one simplex, while the even numbered pivots trace a path in the other. This section shows that when the Lemke-Howson algorithm is applied to the imitation game (A, I), the path in the imitator's simplex of mixed strategies is the path of the Lemke paths algorithm applied to A.

We begin by describing the Lemke-Howson algorithm for a two player game satisfying a general position condition. 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

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

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 $\overline{\tau} := \operatorname{argmax}_{i \in \mathcal{J}_1} (A\tau)_i$. The pair (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

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

Let

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

We say that (A, B) is in *general position* if $|\sigma^\circ| + |\overline{\sigma}| \leq m$ for all $\sigma \in \Delta^m$ and $|\tau^\circ| + |\overline{\tau}| \leq n$ 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.

The next result is analogous to Lemmas 3.4 and 3.5, with proofs that are similar and consequently omitted.

Proposition 6.1. *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*.

Each edge has two endpoints that are vertices. If $\tilde{V}(W, Z)$ is an endpoint of $\tilde{E}(W', Z')$, then it must be the case that $W_1' \subset W_1$, $Z_1' \subset Z_1$, $W_2' \subset W_2$, $Z_2' \subset Z_2$, 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. The set $\tilde{G}_{\tilde{s}}$ (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 $\tilde{s} \in \mathcal{J}_1$. By symmetry this is without loss of generality.

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 the strategy $\delta_{\tilde{s}}$: $\overline{\delta_{\tilde{s}}} = \{\tilde{t}\}$ for some $\tilde{t} \in \mathcal{J}_2$. If $(\delta_{\tilde{s}}, \tau)$ is an $\tilde{s}$-vertex, then $\overline{\delta_{\tilde{s}}} \cup \tau^\circ = \mathcal{J}_2$, and $\mathcal{J}_2 \setminus \{\tilde{t}\} \subset \tau^\circ$ implies that $\tau = \delta_{\tilde{t}}$. 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 the only $\tilde{s}$-edge that could have $(\delta_{\tilde{s}}, \delta_{\tilde{t}})$ as an endpoint is $\tilde{E}(W', Z')$ where (W', Z') is obtained from (W, Z) by dropping $\tilde{s}$ from Z_1. But $\tilde{S}(W', Z') = \emptyset$ because $Z_1' = \emptyset$, so $(\delta_{\tilde{s}}, \delta_{\tilde{t}})$ is not an endpoint of any $\tilde{s}$-edge.

Now suppose that $i \neq \tilde{s}$. 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 Proposition 6.1 implies that $\tilde{E}(W', Z')$ is nonempty, but i cannot be dropped from Z_1 without making it empty. Therefore, 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 $\sigma \neq \delta_{\tilde{s}}$. 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 $\tilde{S}(W', Z')$ could not possibly be empty unless $Z_1 = \{\tilde{s}\}$, and this would imply $\sigma = \delta_{\tilde{s}}$ because $W_1 = \mathcal{J}_1 \setminus \{\tilde{s}\}$, contrary to assumption.

If (σ, τ) is not a Nash equilibrium, then the conditions

$$|W_1| + |Z_2| = m,\ |Z_1| + |W_2| = n, \text{ and } W_1 \cup Z_1 \cup W_2 \cup Z_2 = (\mathcal{J}_1 \cup \mathcal{J}_2) \setminus \{\tilde{s}\}$$

imply that $W_1 \cup Z_1 = \mathcal{J}_1 \setminus \{\tilde{s}\}$, $W_2 \cup Z_2 = \mathcal{J}_2$, and, for some j, either:

(a) $W_1 \cap Z_1 = \{j\}$ and $W_2 \cap Z_2 = \emptyset$, or

(b) $W_1 \cap Z_1 = \emptyset$ and $W_2 \cap Z_2 = \{j\}$.

In case (a) 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 otherwise $W_1 \subset \mathcal{J}_1 \setminus \{\tilde{s}\}$ would have $m-1$ elements, and this can only happen if $\sigma = \delta_{\tilde{s}}$. Similarly, if (b), 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 $W_2 \cap Z_2 = \{j\}$. Thus, in both cases (σ, τ) is an endpoint of two $\tilde{s}$-edges.

In general, 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. *Note that in the last of the cases considered above, namely (σ, τ) is not a Nash equilibrium and $\sigma \neq \delta_{\tilde{s}}$, one of the $\tilde{s}$-edges having (σ, τ) as an endpoint is horizontal because it is obtained by dropping an element of W_2 or Z_1, and the other is vertical, resulting from dropping an element of W_1 or Z_2.*

Summarizing, if $(\delta_{\tilde{s}}, \delta_{\tilde{t}})$ is a Nash equilibrium, then it is not an endpoint of any $\tilde{s}$-edge, and otherwise it is an endpoint of precisely one $\tilde{s}$-edge. Any 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, and if it is an endpoint of two $\tilde{s}$-edges, then one of these is horizontal and the other is vertical. Since the maximum degree of any vertex is two, $\tilde{G}_{\tilde{s}}$ consists of isolated points, loops, and line segments. In particular, the (possibly degenerate) 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. Because the set of $\tilde{s}$-vertices is finite, it must eventually arrive at a Nash equilibrium. The *Lemke-Howson algorithm* follows this path.

There is such a path for each $\tilde{s} = 1, \ldots, m$ and, by symmetry, for each $\tilde{s} = m+1, \ldots, m+n$. The path for $\tilde{s}$ coincides with the path for $\tilde{t} \neq \tilde{s}$ if $(\delta_{\tilde{s}}, \delta_{\tilde{t}})$ is a

Nash equilibrium. Otherwise the two paths have distinct starting points, and they cannot share any edges or any vertices that are not Nash equilibria.

We now specialize to the case of an imitation game, so henceforth $n = m$ and $B = I$. For a mixed strategy $\sigma \in \Delta^m$ for the mover, $|\sigma^\circ| + |\overline{\sigma}| \leq m$ because $\overline{\sigma} = \{\, i+m : i \in \operatorname{argmax}_{i'} \sigma_{i'} \,\}$. This implies that (A, I) is in general position if and only if A is in general position in the sense described in Section 3.

Recall that $\mathcal{I} := \{1, \dots, m\}$. Define $d : \mathcal{J}_1 \cup \mathcal{J}_2 \to \mathcal{I}$ by setting

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

We continue to work with a particular $\tilde{s} \in \mathcal{J}_1$. (Symmetry no longer implies that this is without loss of generality, but it will be clear that the analysis for $\tilde{s} \in \mathcal{J}_2$ is similar.) Let $s := d(\tilde{s})$. The Lemke paths algorithm, as it was described in Section 3, has the following graph-theoretic expression. 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.

Our analysis relates the path in $\tilde{G}_{\tilde{s}}$ starting at $(\delta_{\tilde{s}}, \delta_{\tilde{s}+m})$ to the path in G_s starting at δ_s. This is possible because of the simple structure of the imitator's best responses. In particular, if (σ, τ) is a vertex in $\tilde{V}_{\tilde{s}}$ or an element of a horizontal edge in $\tilde{E}_{\tilde{s}}$, then $|\sigma^\circ| + |\overline{\sigma}| = m$, so $\sigma^\circ \cup \{\, j - m : j \in \overline{\sigma} \,\} = \mathcal{J}_1$ and σ is the uniform distribution on $\{\, j - m : j \in \overline{\sigma} \,\}$. 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 $\{\, j - m : j \in Z_2 \,\}$ and the uniform distribution on $\mathcal{J}_1 \setminus W_1$.

The following four results describe in detail how $\tilde{G}_{\tilde{s}}$ and G_s are related.

Proposition 6.2. *If $\tilde{V}(W, Z) \in \tilde{V}_{\tilde{s}}$, then $V(d(W_2), d(Z_1)) \in V_s$.*

Proof. Let $\tilde{V}(W, Z) = (\sigma, \tau)$. Then $\tau^\circ = W_2$ and $\overline{\tau} = Z_1$, and general position implies that $m = |W_2| + |Z_1| = |d(W_2)| + |d(Z_1)|$. Applying general position again, $S(d(W_2), d(Z_1))$ is zero-dimensional if it is nonempty, and it is nonempty because it contains τ. Thus, $\tau = V(d(W_2), d(Z_1))$ is a vertex.

It remains to show that $\tau \in V_s$, i.e., $\mathcal{I} \setminus \{s\} \subset d(W_2 \cup Z_1)$. We have

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

and $|(W, Z)| = 2m$, so d maps two elements of $W_1 \cup Z_1 \cup W_2 \cup Z_2$ to each element of $\mathcal{I} \setminus \{s\}$. In addition $|W_1| + |Z_2| = m$ and $d(W_1 \cup Z_2) = \mathcal{I}$, so d maps exactly one element of $W_1 \cup Z_2$ to each element of $\mathcal{I} \setminus \{s\}$. Therefore, d maps at least one element of $W_2 \cup Z_1$ to each element of $\mathcal{I} \setminus \{s\}$. ■

The first part of the proof above is applicable to any vertex $(\sigma, \tau) = \tilde{V}(W, Z)$, so such a vertex projects to a vertex $\tau = V(d(W_2), d(Z_1))$. Let π_V be the restriction of this projection to $\tilde{V}_{\tilde{s}}$. Then the last results asserts that $\pi_V(\tilde{V}_{\tilde{s}}) \subset V_s$. The next step is to characterize the preimages of elements of V_s.

Proposition 6.3. *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 all of these will have

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

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

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

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

Then $\sigma^\circ = W_1$ and $\overline{\sigma} = Z_2$, so $\tilde{V}(W, Z) = (\sigma, \tau) \in \tilde{V}_{\tilde{s}}$ as desired.

Let $(\sigma', \tau) = \tilde{V}(W', Z')$ be 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

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

Thus, $\mathcal{J}_1 \setminus \{\tilde{s}\} \subset W_1' \cup Z_1'$ and $\mathcal{J}_2 = W_2' \cup Z_2'$. It cannot be the case that $W_1' \cup Z_1' = \mathcal{J}_1$, because this would imply that $(W', Z') = (W, Z)$. In particular, it cannot be the case that $s \in X$, since then $\tilde{s} \in Z_1$ because $\tilde{s} \in \mathcal{J}_1$. Therefore, $\tilde{s} \in W_1$ and $\tilde{s}' \notin W_1'$. Since $d(\sigma'^{\circ} \cup \overline{\sigma}') = \mathcal{I}$, 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 we have $\tilde{V}(W', Z') = (\sigma', \tau) \in \tilde{V}_{\tilde{s}}$ 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 $\tilde{E}((W_1', W_2), (Z_1, Z_2))$.

Now 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

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

or

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

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. ■

We now turn to the projections of edges in $\tilde{E}_{\tilde{s}}$. Proposition 6.3 describes the situation for vertical edges, and the next two results characterize the projections of horizontal edges.

Proposition 6.4. *If $\tilde{E}(W, Z) \in \tilde{E}_{\tilde{s}}$ is a horizontal edge, then $E(d(W_2), d(Z_1)) \in E_s$.*

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, $|(W, Z)| = 2m - 1$,

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

$|W_1| + |Z_2| = m$, and $d(W_1 \cup Z_2) = \mathcal{I}$. Therefore, $|W_2| + |Z_1| = m - 1$ and $d(W_2 \cup Z_1) = \mathcal{I} \setminus \{s\}$, so $E(d(W_2), d(Z_1)) \in E_s$. ■

Let π_E be the function $\tilde{E}(W, Z) \mapsto E(d(W_2), d(Z_1))$ be the function mapping horizontal edges in $\tilde{E}_{\tilde{s}}$ to E_s.

Proposition 6.5. *π_E is a bijection, and $\pi_V \circ \tilde{e} = e \circ \pi_E$.*

Proof. Suppose $E(X,Y) = \pi_E(\tilde{E}(W,Z))$ for some $\tilde{E}(W,Z) \in \tilde{E}_{\tilde{s}}$. Then it must be the case that $|(W,Z)| = 2m-1$, $W_1 \cup Z_1 \cup W_2 \cup Z_2 = (\mathcal{J}_1 \cup \mathcal{J}_2) \setminus \{\tilde{s}\}$, $|W_1|+|Z_2| = m$, $d(W_1 \cup Z_2) = \mathcal{I}$, $d(W_2) = X$, and $d(Z_1) = Y$. The last two conditions imply that $W_2 = d^{-1}(X) \cap \mathcal{J}_2$ and $Z_1 = d{-}1(Y) \cap \mathcal{J}_1$, after which it is easy to see that the other conditions imply that $W_1 = \mathcal{J}_1 \setminus (Z_1 \cup \{\tilde{s}\})$ and $Z_2 = \mathcal{J}_2 \setminus W_2$. Therefore, π_E is injective. 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 surjective. 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. In particular, when $1 \le \tilde{s} \le m$ and $s = d(\tilde{s})$, then $\delta_{\tilde{s}+m}$ is the unique pure best response to $\delta_{\tilde{s}}$ and the path in $\tilde{G}_{\tilde{s}}$ beginning at $(\delta_{\tilde{s}}, \delta_{\tilde{s}+m})$ projects onto the path in G_s beginning at δ_s. An analogous statement holds when $m+1 \le \tilde{s} \le 2m$.

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 paths of a "symmetric" version of the Lemke-Howson algorithm that only computes symmetric equilibria of the derived symmetric game. The standard Lemke-Howson algorithm always computes, very quickly, an asymmetric pure-strategy equilibrium of the bimatrix game. 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 not negative, 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 7.1. *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 $\mathbf{e} := (1, \ldots, 1)$ 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 $p(y_1), \ldots, p(y_m)$. Because $p(y) - q(y)$ is always a scalar multiple of $\mathbf{e}$, $\varepsilon A'$ and A'' are equivalent. 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.

Proposition 7.2. *If L is a finite dimensional inner product space, then for an open dense subset in L^{2m} of geometric imitation games the lengths of the Lemke paths 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 Proposition 3.7 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. This algorithm and its variants seem interesting in them-

selves and have attractive features that suggest they may be useful in practice. Preliminary tests also show good performance, in terms of both speed and success on problems for which other methods are infeasible. There is certainly a great deal to do in the direction of understanding the behavior of these algorithms.

The study of imitation games has led to other interesting findings. The Lemke paths algorithm has been displayed as the result of projecting the path of the Lemke-Howson algorithm of a suitable imitation game. This allows one to pass from a family of examples in which all Lemke paths are "long" to a family of imitation games whose Lemke-Howson paths are all "long." 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, 2006) 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). In their work Codenotti and Štefanovič (2005); Bonifaci et al. (2005); Codenotti 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.

Bibliography

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

H. F. Bohnenblust and S. Karlin. On a theorem of Ville. In *Contributions to the Theory of Games*, Annals of Mathematics Studies, no. 24, pages 155–160. Princeton University Press, Princeton, N. J., 1950.

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

H. Brézis, L. Nirenberg, and G. Stampacchia. A remark on Ky Fan's minimax principle. *Boll. Un. Mat. Ital. (4)*, 6:293–300, 1972.

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

Xi Chen and Xiaotie Deng. 3-NASH is PPAD-complete. *Electronic Colloquium on Computational Complexity (ECCC)*, (134), 2005a.

Xi Chen and Xiaotie Deng. Settling the complexity of 2-player Nash-equilibrium. *Electronic Colloquium on Computational Complexity (ECCC)*, (140), 2005b.

B. Codenotti, A Saberi, K. Varadarajanz, and Y. Ye. Leontief economies encode nonzero sum two-player games. *Electronic Colloquium on Computational Complexity*, 55, 2005.

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.

K. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou. The complexity of computing a Nash equilibrium. *Electronic Colloquium on Computational Complexity (ECCC)*, (115), 2005.

K. Daskalakis and C. H. Papadimitriou. Three-player games are hard. *Electronic Colloquium on Computational Complexity (ECCC)*, (139), 2005.

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

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

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

P. W. Goldberg and C.H. Papadimitriou. Reducibility among equilibrium problems. *Electronic Colloquium on Computational Complexity (ECCC)*, (090), 2005.

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.

M. Hirsch, C.H. Papadimitriou, and S. Vavasis. Exponential lower bounds for finding Brouwer fixed points. *Journal of Complexity*, 5:379–416, 1989.

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

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, 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. *Computational Complexity*. Adison Wesley Longman, New York, 1994a.

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

C.H. Papadimitriou. On the complexity of the parity argument and other inefficient proofs of existence. *Journal of Computer and System Science*, 48:498–532, 1994b.

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.

R. Savani and B. von Stengel. Hard-to-solve bimatrix games. *Econometrica*, 74: 397–429, 2006.

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

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