

# Heterogeneous Interacting Economic Agents and Stochastic Games

P. Jean-Jacques Herings\*      Ronald J.A.P. Peeters†

7th December 2001

## Abstract

Stochastic games offer a rich mathematical structure that makes it possible to analyze situations with heterogeneous and interacting economic agents. Depending on the actions of the economic agents, the economic environment changes from one period to another. We focus on stationary equilibrium, the simplest form of behavior that is consistent with rationality. Since the number of stationary equilibria abound, we present the stochastic tracing procedure, a method to select equilibria. Since stationary equilibria are difficult to characterize analytically, we also present a numerical algorithm by which they can be computed. The algorithm is constructed in such a way that the equilibrium selected by the stochastic tracing procedure is computed. We illustrate the usefulness of this approach by showing how it leads to new insights in the theory of dynamic oligopoly.

*Key words:* Stochastic games, stationary equilibrium, tracing procedure, numerical algorithms.

JEL-code: C62, C63, C72, C73.

## 1 Introduction

Situations with heterogeneous and interacting economic agents typically have the following structure. Initially, the economic environment is in a certain state. Economic agents have to decide what kind of actions to take. The choice of action will influence their own immediate payoff, but also the payoff of others. Moreover, and perhaps more importantly, the actions taken affect the state of the economic environment in the next stage. The resulting economic environment next stage might also be influenced by some random shocks. The action of an economic agent today may affect his future environment in several ways. It might influence his set of options among which to choose next stage, as well as the set of options available to other agents, and it can also change his utility function, or the one of his opponents.

A setting as described above can be analyzed in many ways. It is not uncommon to endow the agents with certain rules of thumb, or to assume that they follow specific forms of learning

---

\*Department of Economics, Universiteit Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands.  
E-mail: P.Herings@algec.unimaas.nl

†Department of Economics, Universiteit Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands.  
E-mail: R.Peeters@algec.unimaas.nl

behavior. Next the development of the economic environment is analyzed, for instance by simulation methods and in a number of special cases by analytical derivations. Common is the assumption of boundedly rational behavior.

Irrespective of the intrinsic interest in these forms of boundedly rational behavior, it is of utmost importance to know how fully rational agents would behave. Apart from the standard argument that deviations from rational behavior can be quite costly, implying that agents have strong incentives to behave as rational as possible, rational behavior also serves as an interesting benchmark. It is of central interest to see to what extent boundedly rational behavior deviates from fully rational behavior, how it deviates from fully rational behavior, and under what circumstances the difference is substantial.

Stochastic games as introduced in Shapley (1953) constitute a rich mathematical structure that exactly incorporates situations as described above. Both boundedly rational and fully rational behavior can be studied in this framework. This paper focuses on the latter. It presents the structure of stochastic games, as well as the most commonly used solution concept, the one of stationary equilibrium. Stationary equilibria describe the simplest form of behavior that is consistent with rationality. A strategy in a stationary equilibrium depends only on the current state of the economic environment, but not on the history leading to it. It thereby captures the notion that bygones are bygones.

Restricting the analysis to stationary equilibria avoids the indeterminateness of equilibria as implied by the folk theorem. It can be shown that the number of stationary equilibria is finite, see Haller and Lagunoff (2000), and even odd, see Herings and Peeters (2000). There is no hope, however, that the number of stationary equilibria, though odd, is small. McLennan (1999) shows that the expected number of Nash equilibria in normal form games explodes when the number of players or the number of actions per player increases. The logic of his arguments carries over to the framework of stochastic games, implying a large number of stationary equilibria on average for stochastic games of modest sizes. To deal with the issue of multiplicity of equilibria, we extend the equilibrium selection method of Harsanyi and Selten (1988) to the framework of stochastic games as in Herings and Peeters (2001). The extension selects a unique stationary equilibrium for almost all stochastic games.

The analysis of stochastic games is hampered by the fact that even simple stochastic games are often difficult to analyze by analytical methods. Closed form solutions for stationary equilibria in a stochastic game are difficult to obtain. Finding the stationary equilibrium selected by the extension of Harsanyi and Selten's equilibrium selection method is even more cumbersome. In these cases, if one wants to go beyond the mere existence of equilibrium, there is no alternative but to resort to numerical methods as has also been argued forcefully in Judd (1997, 1998). We present the numerical algorithm introduced in Herings and Peeters (2000) that computes an approximation of the stationary equilibrium selected by the extension of Harsanyi and Selten's method for almost all stochastic games. We illustrate how the use of numerical algorithms can lead to new economic insights by means of an example in the theory of dynamic oligopoly.

## 2 The stochastic game model

The concept of a finite discounted stochastic game is defined by Shapley (1953).

**Definition 2.1** A *finite discounted stochastic game* is an ordered sextuple

$$\Gamma = \langle N, \Omega, \{S_\omega^i\}_{(i,\omega) \in N \times \Omega}, \{u^i\}_{i \in N}, \pi, \delta \rangle,$$

where  $N$ ,  $\Omega$  and  $S_\omega^i$  are finite non-empty sets,  $u^i$  is a real-valued function on the set  $H = \{(\omega, s_\omega) \mid \omega \in \Omega, s_\omega \in S_\omega\}$  with  $S_\omega = \prod_{i \in N} S_\omega^i$ ,  $\pi$  is a map from  $H$  on  $\Delta(\Omega)$  with  $\Delta(\Omega)$  the family of probability distributions on the space  $\Omega$  and  $\delta$  is an element in the open unit interval.

The game parameters have the following meaning.

- $N = \{1, \dots, n\}$  is the player set.
- $\Omega = \{\omega_1, \dots, \omega_z\}$  is the state space.
- $S_\omega^i = \{s_{\omega_1}^i, \dots, s_{\omega_{m_i}}^i\}$  is the action set of player  $i \in N$  in state  $\omega \in \Omega$ .
- $u^i : H \rightarrow \mathbb{R}$  is the payoff function of player  $i \in N$ ; if in state  $\omega \in \Omega$  the players action choices are  $s_\omega = (s_\omega^1, \dots, s_\omega^n) \in S_\omega$ , then player  $i$  gets an instantaneous payoff of  $u^i(\omega, s_\omega)$ .
- $\pi : H \rightarrow \Delta(\Omega)$  is the transition map: for each  $(\omega, s_\omega) \in H$ , we can identify  $\pi(\omega, s_\omega)$  with the vector  $(\pi(\omega_1 \mid \omega, s_\omega), \dots, \pi(\omega_z \mid \omega, s_\omega))$ , where  $\pi(\bar{\omega} \mid \omega, s_\omega)$  represents the probability that the system jumps to state  $\bar{\omega}$  if in state  $\omega$  the strategy tuple  $s_\omega$  is played.
- $\delta \in (0, 1)$  is the discount factor and is used to discount future payoffs.

Such a stochastic game corresponds to a dynamic system which can be in different states and where at certain stages the players can influence the course of the play. We consider the infinite horizon model and the set of stages is assumed to be identical with the set  $\mathbb{N} = \{0, 1, \dots\}$ . Players know the game itself and that this knowledge is common knowledge among all the players. Moreover, the initial state  $\omega^0$  at stage  $k = 0$  is common knowledge to the players.

The game proceeds as follows. At the initial state  $k = 0$  all players  $i$  select at the initial state  $\omega^0$ , simultaneously and independently of each other an action  $s_{\omega^0}^i \in S_{\omega^0}^i$ . Now two things happen, both depending on the current state  $\omega^0$  and the action choices  $s_{\omega^0}^1, \dots, s_{\omega^0}^n$ :

- (a) player  $i$  earns  $u^i(\omega^0, s_{\omega^0})$ ,
- (b) the system jumps to the next state  $\omega^1$  according to the outcome of a chance experiment: the probability that the next state will be  $\bar{\omega}$  equals  $\pi(\bar{\omega} \mid \omega^0, s_{\omega^0})$ .

Subsequently, prior to the next stage  $k = 1$ , all players are informed about the previous actions chosen by the players, and of the new state  $\omega^1$ . At stage  $k = 1$ , the above procedure is repeated, starting from the state  $\omega^1$ .

We assume that the game is of perfect recall, i.e. at each stage each player remembers all past actions chosen by all players and all past states that have occurred. Note that for

a stochastic game, each stage game resembles a normal form game. However, contrary to the situation with normal form games, the game does not consist of a single play, but jumps according to the probability measure  $\pi(\cdot | \omega, s_\omega)$  to the next state and continues dynamically. In choosing an action in a certain state, a player not only takes into account the immediate payoff, but also his opportunities in future states.

Like in normal form games, the players are allowed to randomize their pure actions. A mixed strategy of player  $i$  in state  $\omega$  is a probability distribution on  $S_\omega^i$ . We identify the set of all probability distributions on  $S_\omega^i$  with  $\Sigma_\omega^i$  and a generic element of this set by  $\sigma_\omega^i$ . For  $\sigma_\omega^i \in \Sigma_\omega^i$ , the probability assigned to pure strategy  $s_{\omega j}^i$  is given by  $\sigma_{\omega j}^i$ . The strategy space of the normal form game in state  $\omega$  is therefore equal to  $\Sigma_\omega = \prod_{i \in N} \Sigma_\omega^i$ . Given a mixed strategy combination  $\sigma_\omega \in \Sigma_\omega$  and a strategy  $\bar{\sigma}_\omega^i \in \Sigma_\omega^i$ , we denote by  $(\sigma_\omega^{-i}, \bar{\sigma}_\omega^i)$  the mixed strategy that results from replacing  $\sigma_\omega^i$  by  $\bar{\sigma}_\omega^i$ . If a mixed strategy combination  $\sigma_\omega \in \Sigma_\omega$  is played, then the instantaneous expected payoff of player  $i$  is denoted by  $u^i(\omega, \sigma_\omega)$  and the expected transition to state  $\bar{\omega}$  by  $\pi(\bar{\omega} | \omega, \sigma_\omega)$ .

We end this section with a few examples of economic environments which can be modeled as a stochastic game. The following example is discussed in Filar and Vrieze (1997).

**Example 2.2 (Fishery Game)** The availability of fish is represented by three states: ‘Good’, ‘Medium’, and ‘Bad’. There are two fisheries that have in each state two actions: they can act ‘Greedy’ (G) or ‘Nongreedy’ (NG). In Figure 1, the game data are depicted. In the upper-left

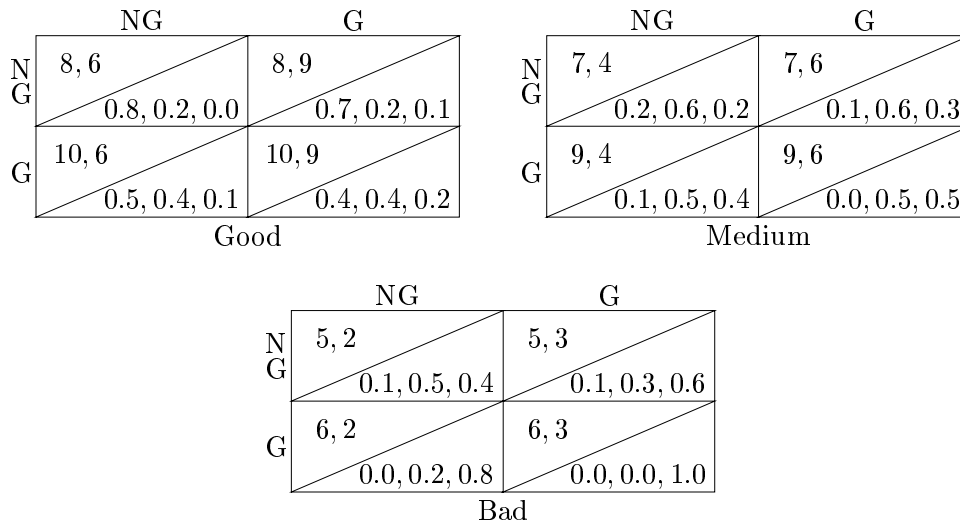


Figure 1: Fishery Game.

corner of each cell, the instantaneous profits to the companies are displayed. The transition probabilities are displayed in the lower-right corner of the cells. For example, when the state is ‘Good’ and Fishery 1 (the row player) acts ‘Greedy’, whereas Fishery 2 (the column player) acts ‘Nongreedy’, Fishery 1’s profit is 10 and Fishery 2’s profit is 6. The probability of staying in the ‘Good’ state equals 0.5, the probability of going to the ‘Medium’ state is 0.4 and the next period the state will be ‘Bad’ with probability 0.1.

The income of a fishery is determined by the quantity of fish it catches multiplied by the price of the fish, which is in turn determined by the availability of the fish: in state ‘Good’ the price attains the lowest level, whereas in state ‘Bad’ the price is at the highest level. The costs of a fishery represent its investment in fishing-equipment. Furthermore, as the attentive reader might have observed, the parameter values have been chosen such that the instantaneous profits, income minus costs, are independent from the fishing strategy of the other fishery. Thus, in this example, the players are only linked in their joint influence on the dynamics of the system.

In the separate one-shot games, being greedy is always advantageous. However, for an infinite horizon game, as is considered here, this need not to be the case. This shows that dynamic games can usually not be solved by considering the solution to the static stage games.

◁

The following example is studied in more detail in Herings, Peeters and Schinkel (2001).

### **Example 2.3 (Dynamic Competition)**

A finite number of (potential) firms compete dynamically in the market for some homogeneous product. In a given period, each firm is either an incumbent or a potential entrant. In each period, incumbents make a decision about the quantity to produce in that period and about whether to leave the market or to stay. Potential entrants only have to decide whether to enter or not in a given period.

Since all firms are assumed to have identical cost structures, the market structure (the state) in a certain period is determined by the number of firms that are active in production in that period. Tomorrow’s market structure depends on the number of active firms today and the entry and exit decisions made by the firms.

The instantaneous profits to an incumbent firm depend on the current market structure, the firm’s own production quantity and its competitors’ production quantities. For a potential entrant, the instantaneous profits depend on the entry decision. When a potential entrant decides to enter, it has to make some sunk costs, which can be avoided by not entering.

This example of dynamic competition is frequently analyzed from a static perspective. Suppose entry costs are sufficiently low. Without considering the exact specification of the transitional dynamics, it is then argued that firms will enter the industry as long as there are positive economic profits to be made. The firms that are present in the market are engaged in Cournot competition and produce the Cournot quantity corresponding to the number of firms present.

It is straightforward to model this economic environment as a stochastic game. In Section 5 we turn back to this example and analyze it from a full dynamic perspective.

◁

More applications of stochastic games are contained in Filar and Vrieze (1997). For a recent authoritative text the reader is referred to Neyman and Sorin (2001).

### 3 Strategy and equilibrium concepts

A history up to a certain stage  $k$  consists of a specification of all states materialized and all actions taken by all the players in all periods before  $k$ .

**Definition 3.1** A *history* up to a stage  $k$  is a sequence  $h^k = (\omega^0, s_{\omega^0}, \omega^1, s_{\omega^1}, \dots, \omega^{k-1}, s_{\omega^{k-1}})$  that could have occurred up to time  $k$ ,  $k \geq 1$ . Here  $\omega^\kappa$  represents the state and  $s_{\omega^\kappa}$  the actions of the players at stage  $\kappa$ ,  $\kappa = 0, \dots, k-1$ .

Obviously, the set of histories up to time  $k$  equals  $H^k = \prod_{\kappa=0}^{k-1} H$ , i.e. the  $k$ -fold Cartesian product of  $H$ .

A *behavior strategy*  $\sigma^{ik}$  of player  $i$  specifies for each stage  $k$ , each state  $\omega^k$  at time  $k$ , and each history  $h^k$  a probability distribution  $\sigma^{ik}(h^k, \omega^k)$  on the action space  $S_{\omega^k}^i$  of player  $i$  in state  $\omega^k$ . Then  $\sigma^{ik}(s_{\omega^k}^i | h^k, \omega^k)$  is the probability with which player  $i$  chooses action  $s_{\omega^k}^i \in S_{\omega^k}^i$  at time  $k$  if state  $\omega^k$  and history  $h^k$  have occurred. Or formally,

**Definition 3.2** A *behavior strategy*  $\sigma^i$  for player  $i$  is a sequence  $\sigma^{i0}, \sigma^{i1}, \dots$  where  $\sigma^{i0} \in \Sigma^i := \prod_{\omega \in \Omega} S_{\omega}^i$  and  $\sigma^{ik} : H^k \rightarrow \Sigma^i$  for all  $k \geq 1$ .

Since future payoffs are discounted, given initial state  $\omega$  and strategy  $\sigma$ , the stream of expected payoffs to player  $i$  is evaluated by

$$U^i(\omega, \sigma) := \sum_{k=0}^{\infty} \delta^k \cdot U^{ik}(\omega, \sigma),$$

where  $U^{ik}(\omega, \sigma)$  denotes the expected instantaneous utility to player  $i$  at stage  $k$ . Here,  $U^i(\omega, \sigma)$  equals the total discounted expected payoff of player  $i$  when the discount factor equals  $\delta$ , the starting state is  $\omega$  and the strategy tuple  $\sigma$  is played. Since the state and action spaces are assumed to be finite,  $U^i(\omega, \sigma)$  exists.

**Definition 3.3** A strategy tuple  $\sigma$  is an *equilibrium* if and only if  $\sigma^i$  is a best response to  $\sigma^{-i}$  for all  $i \in N$ .

If  $n = 2$  and the payoffs are zero-sum, then the definition characterizes a minimax solution (see Shapley (1953)).

A *stationary strategy* for player  $i$  is a behavior strategy for which  $\sigma^{ik}(h^k, \omega^k)$  is of the form  $\sigma^i(\omega^k)$ , i.e. a stationary strategy is a history and time independent strategy which depends on the state only. In the sequel, a stationary strategy for player  $i$  will be denoted by the symbol  $\rho^i$ . We define  $\rho^i = (\rho_{\omega^1}^i, \dots, \rho_{\omega^z}^i)$ , where  $\rho_{\omega^z}^i$  is a probability measure on the action space  $S_{\omega^z}^i$  for each  $\omega \in \Omega$ , so  $\rho_{\omega^z}^i \in \Sigma_{\omega^z}^i$ . If player  $i$  decides to play the stationary strategy  $\rho^i$ , then every time that the system is in state  $\omega$ , player  $i$  selects his pure action according to  $\rho_{\omega}^i$ . Formally, we have the following definition of a stationary strategy.

**Definition 3.4** A *stationary strategy*  $\rho^i$  for player  $i$  is an element of  $\Sigma^i$ .

Suppose that  $\omega$  is the initial state and the players decide to play a stationary strategy tuple  $\rho$ . The total discounted expected payoff of player  $i$  is denoted by  $U^i(\omega, \rho)$ . The instantaneous payoff player  $i$  obtains in stage  $k = 0$  equals  $u^i(\omega, \rho_{\omega})$ . The probability that at the next stage the state will be  $\bar{\omega}$  equals  $\pi(\bar{\omega} | \omega, \rho_{\omega})$ . This gives rise to the following theorem (see e.g. Fink (1964)).

**Theorem 3.5** *When  $\rho$  is a stationary strategy tuple and  $\omega$  is the initial state, the expected payoffs are given by the following recursive formula*

$$U^i(\omega, \rho) = u^i(\omega, \rho) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, \rho_\omega) U^i(\bar{\omega}, \rho).$$

**Definition 3.6** A stationary strategy combination  $\rho \in \Sigma$  is a *stationary equilibrium* if it is a Nash equilibrium in stationary strategies.

In the sequel we restrict ourselves to stationary equilibria. Several motivations for analyzing stationary equilibria can be found in Maskin and Tirole (2001). Stationary strategies prescribe the simplest form of behavior that is consistent with rationality, stationarity captures the notion that ‘bygones are bygones’ more completely than does the concept of subgame-perfect equilibrium, and it embodies the principle that ‘minor causes should have minor effects’, that is, only those aspects of the past that are ‘significant’ should have an appreciable influence on behavior. Some pragmatic motivations are that in applied theory, the focus on stationary strategies allows for clean, unobstructed analysis of the influence of the state variables, that stationary strategies substantially reduce the number of parameters to be estimated in dynamic (econometric) models, and that stationary models can be simulated.

It is a very well-known result that there exists a stationary best response to stationary strategies  $\rho^{-i}$ . Therefore, the following theorem holds.

**Theorem 3.7** *A stationary strategy combination  $\bar{\rho} \in \Sigma$  is a stationary equilibrium if for all  $i$  and all  $\omega$*

$$U^i(\omega, \bar{\rho}) \geq U^i(\omega, \bar{\rho}^{-i}, \rho^i), \quad \rho^i \in \Sigma^i.$$

Independently from one another, Fink (1964), Takahashi (1964), and Sobel (1971) proved the following result in different ways.

**Theorem 3.8** *Every finite discounted stochastic game has a stationary equilibrium.*

Recently, Haller and Lagunoff (2000) showed the following result.

**Theorem 3.9** *For a generic stochastic game, the number of stationary equilibria is finite.*

Alternative proofs for the former two theorems are contained in Herings and Peeters (2000), together with the following additional result.

**Theorem 3.10** *For a generic stochastic game, the number of stationary equilibria is odd.*

Sobel (1971) has the following useful characterization of a stationary equilibrium.

**Theorem 3.11** *A stationary strategy tuple  $\sigma$  constitutes a stationary equilibrium if and only if it is part of a solution to*

$$\begin{aligned} u^i(\omega, \sigma_\omega^{-i}, s_{\omega j}^i) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \pi(\bar{\omega} | \omega, \sigma_\omega^{-i}, s_{\omega j}^i) \mu_{\bar{\omega}}^i + \lambda_{\omega j}^i - \mu_\omega^i &= 0, \\ & (s_{\omega j}^i \in S_\omega^i, \omega \in \Omega, i \in N), \\ \lambda_{\omega j}^i \geq 0, \quad \sigma_{\omega j}^i \geq 0, \quad \lambda_{\omega j}^i \sigma_{\omega j}^i &= 0, \quad (s_{\omega j}^i \in S_\omega^i, \omega \in \Omega, i \in N), \\ \sum_{s_{\omega j}^i \in S_\omega^i} \sigma_{\omega j}^i - 1 &= 0, \quad (\omega \in \Omega, i \in N). \end{aligned}$$

Here,  $\lambda_{\omega j}^i$  is the shadowprice of playing strategy  $s_{\omega j}^i$ , i.e. the disutility from a one-shot deviation at the initial period of an increase in the probability  $\sigma_{\omega j}^i$  by which pure strategy  $s_{\omega j}^i$  is played, and  $\mu_{\omega}^i$  is the expected payoff of player  $i$  when the initial state is  $\omega$ ,  $\sigma^{-i}$  is played by its opponents, and player  $i$  chooses a best response against  $\sigma^{-i}$ .

In contrast to the instantaneous payoffs of a stage game – which is by definition a normal form game – the expected utility property does not hold for the total discounted payoffs in stochastic games.

**Example 3.12** The stochastic game of Figure 2 is a game in which there are two states  $p_h$  and  $p_\ell$  and one player. In each state the player can choose between actions  $q_h$  and  $q_\ell$ . In the upper-left corner of each square, the intermediate payoff of the player is displayed. In the lower-right corner the transition probabilities are given. So, if the player chooses  $q_h$  when the state is  $p_h$ , then he gets an immediate payoff of 1. The next state will be  $p_\ell$  with probability 1. If the player chooses  $q_\ell$  when the state is  $p_\ell$ , then he gets an immediate payoff of 0. The next state will be  $p_h$  with probability 1. Future payoffs are discounted by a factor 0.5. Consider

$q_h$	$\begin{array}{c} 1.0 \\ \diagdown \\ 0.0, 1.0 \end{array}$	$q_h$	$\begin{array}{c} 0.5 \\ \diagdown \\ 0.5, 0.5 \end{array}$
$q_\ell$	$\begin{array}{c} 0.5 \\ \diagdown \\ 0.5, 0.5 \end{array}$	$q_\ell$	$\begin{array}{c} 0.0 \\ \diagdown \\ 1.0, 0.0 \end{array}$
	$p_h$		$p_\ell$

Figure 2: Game in which expected utility property does not hold.

the strategies  $\rho^- = (q_\ell, q_h)$  and  $\rho^+ = (q_h, q_\ell)$ , where the first argument is the strategy when the state is  $p_h$  and the second when the state is  $p_\ell$ . Then  $U(p_\ell, \rho^-) = U(p_h, \rho^-) = 1$ ,  $U(p_\ell, \rho^+) = 2/3$  and  $U(p_h, \rho^+) = 4/3$ . However,

$$\begin{aligned}
 U(p_h, \frac{1}{2}\rho^- + \frac{1}{2}\rho^+) &= \frac{6}{5} \neq \frac{7}{6} = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{4}{3} = \frac{1}{2}U(p_h, \rho^-) + \frac{1}{2}U(p_h, \rho^+), \\
 U(p_\ell, \frac{1}{2}\rho^- + \frac{1}{2}\rho^+) &= \frac{4}{5} \neq \frac{5}{6} = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{2}U(p_\ell, \rho^-) + \frac{1}{2}U(p_\ell, \rho^+).
 \end{aligned}$$

The expected utility property does not hold. ◁

## 4 Equilibrium selection

For many normal form games there is a vast multiplicity of equilibria as has been argued convincingly in McLennan (1999). There is no reason to expect that the situation is different for the multiplicity of stationary equilibria in stochastic games. For this reason, it is worthwhile to investigate stationary equilibrium selection theories. In Herings and Peeters (2001), the stochastic tracing procedure, which is based on the linear tracing procedure of Harsanyi and Selten (1988), is proposed to do the selection within the class of stationary equilibria of stochastic games.

The linear tracing procedure as presented in Harsanyi and Selten (1988) models a process of convergent expectations by which rational players will come to adopt, and expect each other to adopt, a particular equilibrium as the solution for a given game. Before applying

the tracing procedure, every player is assumed to have a subjective probability distribution expressing his expectation about the strategic choices of the other players. Each player is assumed to use the same theory to determine his subjective probability distributions, which makes that all players have the same expectations about the other players. This common subjective probability distribution is called the prior.

In the naive Bayesian approach, all players choose best responses to their prior beliefs and would in this way reach a strategy combination that does not constitute an equilibrium in general. In the linear tracing procedure, the information on the best responses is only gradually fed back into the expectations of the players. As the linear tracing procedure proceeds, both the priors and their best responses will gradually change until both converge to some equilibrium of the game.

In Harsanyi and Selten (1988) the linear tracing procedure is defined for normal form games and for extensive form games with a perfect recall information structure. For a normal form game  $\Gamma = \langle N, \{S^i\}_{i \in N}, \{u^i\}_{i \in N} \rangle$  and a prior  $p \in \Sigma$ , the linear tracing procedure is defined by tracing a curve in the set of Nash equilibria of the games  $\Gamma^t = \langle N, \{S^i\}_{i \in N}, \{v^i(t)\}_{i \in N} \rangle$  for  $t \in [0, 1]$ , where  $v^i(t, s) = (1 - t)u^i(p^{-i}, s^i) + tu^i(s)$ . Notice that  $v^i(t, s)$  represents the payoffs to player  $i$  when all his opponents play according to the prior  $p^{-i}$  with probability  $1 - t$ , and according to  $s^{-i}$  with probability  $t$ . Herings and Peeters (2000) show that the curve to be traced is uniquely determined for almost all games, and that the linear tracing procedure selects a unique equilibrium for almost all games.

For extensive form games, Harsanyi and Selten (1988) first transform the game into one in standard form and subsequently define the tracing procedure for that class of games. The tracing procedure is not defined for stochastic games, which are games with instantaneous payoffs and infinite time horizon. Since expected utility in stationary strategies does not hold in stochastic games (see Example 3.12), it is not possible to transform this game into one in standard form.

Herings and Peeters (2001) propose the following extension of the tracing procedure for a stochastic game. Consider some stochastic game  $\Gamma$  and some prior  $p \in \Sigma$ . For every  $t \in [0, 1]$ , the stochastic tracing procedure generates a stationary equilibrium of the stochastic game

$$\Gamma^t = \langle N, \Omega, \{S_\omega^i\}_{\omega \in \Omega, i \in N}, \{\tilde{u}^i(t)\}_{i \in N}, \tilde{\pi}^i(t), \delta \rangle,$$

where the instantaneous payoff function  $\tilde{u}^i(t)$  of player  $i$  is defined by

$$\tilde{u}^i(t; \omega, s_\omega) = (1 - t)u^i(\omega, p_\omega^{-i}, s_\omega^i) + tu^i(\omega, s_\omega)$$

and where the transition mapping  $\tilde{\pi}^i(t)$  is defined by

$$\tilde{\pi}^i(t; \bar{\omega} | \omega, s_\omega) = (1 - t)\pi(\bar{\omega} | \omega, p_\omega^{-i}, s_\omega^i) + t\pi(\bar{\omega} | \omega, s_\omega).$$

The superscript ‘ $i$ ’ in the transition mapping  $\tilde{\pi}^i(t)$  is used to make explicit that for  $t < 1$  players have idiosyncratic beliefs about what the transition probabilities are. This is consistent with the fact that the tracing procedure should be thought of as a reasoning process. In line of Theorem 3.5, the expected payoff of player  $i$  is easily shown to satisfy the recursive relation

$$\tilde{U}^i(t; \omega, \rho) = \tilde{u}^i(t; \omega, \rho_\omega) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \tilde{\pi}^i(t; \bar{\omega} | \omega, \rho_\omega) \tilde{U}^i(t; \bar{\omega}, \rho).$$

The stochastic game  $\Gamma^0$  corresponds to a trivial stochastic game, where all players believe that all their opponents play with probability 1 according to the prior belief. The stochastic game  $\Gamma^1$  coincides with the original stochastic game  $\Gamma$ . A best response against a strategy combination  $\rho^{-i} \in \Sigma^{-i}$  in the stochastic game  $\Gamma^t$  corresponds to a best response against the stationary probability distribution  $(1-t)[p^{-i}] + t[\rho^{-i}]$  on  $S^{-i}$  in the stochastic game  $\Gamma$ .

The *stochastic tracing procedure*  $\mathcal{S}(\Gamma, p)$  is defined as the set of pairs  $(t, \rho)$  for which it holds that  $\rho$  is a stationary equilibrium of the stochastic game  $\Gamma^t$ , i.e.

$$\mathcal{S}(\Gamma, p) = \{ (t, \rho) \in [0, 1] \times \Sigma \mid \rho \text{ is a stationary equilibrium of } \Gamma^t \}.$$

The stochastic tracing procedure is said to be *feasible* if there exists a path in  $\mathcal{S}(\Gamma, p)$  connecting a best response against the prior to a stationary equilibrium of the stochastic game  $\Gamma$ , i.e. there exists a continuous function  $\gamma : [0, 1] \rightarrow \mathcal{S}(\Gamma, p)$  such that  $\gamma(0) \in \mathcal{S}(\Gamma, p) \cap (\{0\} \times \Sigma)$  and  $\gamma(1) \in \mathcal{S}(\Gamma, p) \cap (\{1\} \times \Sigma)$ . In general there may be many trajectories  $\gamma([0, 1])$  that link a stationary equilibrium of  $\Gamma^0$  to a stationary equilibrium of  $\Gamma^1$ . If this trajectory is unique, then the stochastic tracing procedure is said to be *well-defined*. If the stochastic tracing procedure is well-defined, then it selects a unique stationary equilibrium of the stochastic game  $\Gamma$ .

In line of Theorem 3.11, Herings and Peeters (2000) prove the following result.

**Theorem 4.1** *An element  $(t, \rho)$  belongs to  $\mathcal{S}(\Gamma, p)$  if and only if it is part of a solution to*

$$\begin{aligned} \tilde{u}^i(t; \omega, \rho_\omega^{-i}, s_{\omega j}^i) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \tilde{\pi}^i(t; \bar{\omega} \mid \omega, \rho_\omega^{-i}, s_{\omega j}^i) \mu_{\bar{\omega}}^i + \lambda_{\omega j}^i - \mu_\omega^i &= 0, \\ & (s_{\omega j}^i \in S_\omega^i, \omega \in \Omega, i \in N), \\ \lambda_{\omega j}^i \geq 0, \quad \rho_{\omega j}^i \geq 0, \quad \lambda_{\omega j}^i \rho_{\omega j}^i = 0, \quad & (s_{\omega j}^i \in S_\omega^i, \omega \in \Omega, i \in N), \\ \sum_{s_{\omega j}^i \in S_\omega^i} \rho_{\omega j}^i - 1 = 0, \quad & (\omega \in \Omega, i \in N). \end{aligned}$$

This characterization is then used to show the following result on the structure of the set  $\mathcal{S}(\Gamma, p)$ .

**Theorem 4.2** *For a generic stochastic game  $\Gamma$  and prior  $p$ , the stochastic tracing procedure  $\mathcal{S}(\Gamma, p)$  consists of a finite number of arcs and loops. Furthermore, all arcs have their boundary points in the subspace with  $t = 0$  or  $t = 1$  and there is a unique boundary point in  $t = 0$ .*

Graphically, the structure of the stochastic tracing procedure can be represented by Figure 3. In this figure, the bullet at  $t = 0$  corresponds to all players playing best responses against the prior. The bullets at  $t = 1$  together represent the set of stationary equilibria of the stochastic game. The unique path going from  $t = 0$  to  $t = 1$  is the curve traced by the stochastic tracing procedure. It selects a unique equilibrium at  $t = 1$ . This observation leads to the following result.

**Theorem 4.3** *For a generic stochastic game  $\Gamma$  and prior  $p$ , the stochastic tracing procedure  $\mathcal{S}(\Gamma, p)$  is well-defined.*

Generic finiteness comes from the fact that all stationary equilibria are geometrically isolated. The generic oddness of the number of stationary equilibria (see Theorem 3.10) comes from the fact that all equilibria not lying on the feasible path are pairwise connected.

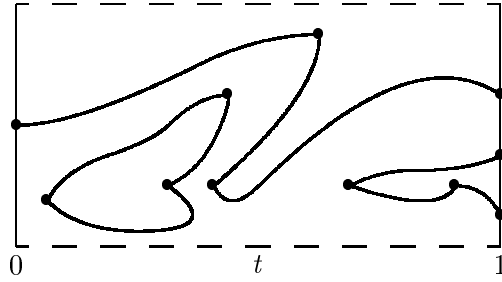


Figure 3: The structure of  $\mathcal{S}(\Gamma, p)$ .

## 5 Equilibrium computation

In the literature on the computation of Nash equilibria in normal form games, a distinction is made between 2-player games on the one hand and 3 or more players on the other hand. For the class of 2-player games exact algorithms are possible. This is caused by the bilinear structure of such games (see, for instance the algorithm of Lemke and Howson (1964)). For stochastic games, this distinction disappears. The system of equations of Theorem 3.11 is not bilinear, even for 2-player stochastic games. The paper of Parthasarathy and Raghavan (1981) presents an example of a 2-player stochastic game with only rational numbers of payoffs and transition probabilities, but with a unique Nash equilibrium involving strategies with only irrational probabilities. This means that the ordered field property does not hold. A straightforward application of a Lemke-Howson type algorithm to 2-player stochastic games is therefore not possible. This shows once more that the class of stochastic games is considerably different from the class of normal form games.

In Breton, Filar, Haurie, and Schultz (1986), Schultz (1986), Filar, Schultz, Thuijsman, and Vrieze (1991), and Breton (1991) the problem of finding discounted stationary equilibria in the general  $n$ -person stochastic game is reduced to that of finding a global minimum in a nonlinear program with linear constraints. Solving this nonlinear program is equivalent to solving a certain nonlinear system for which it is known that the objective value in the global minimum is zero. But, as is noted in Breton (1991), the convergence of an optimization algorithm to the global optimum is not guaranteed.

In Herings and Peeters (2000) an algorithm is developed that is shown to converge to a stationary equilibrium for a generic  $n$ -person stochastic game. This algorithm is the first globally convergent algorithm to solve for an equilibrium in an arbitrary  $n$ -person stochastic game. The algorithm also deals with the equilibrium selection problem in that it computes the stationary equilibrium selected by the stochastic tracing procedure.

Using the following transformation of variables as proposed in Garcia and Zangwill (1981)

$$\rho_{\omega_j}^i(\alpha) = [\max\{0, \alpha_{\omega_j}^i\}]^2, \quad \text{and} \quad \lambda_{\omega_j}^i(\alpha) = [\max\{0, -\alpha_{\omega_j}^i\}]^2,$$

for all  $s_{\omega_j}^i \in S_{\omega_j}^i$ , the system from Theorem 4.1 is transformed into the system of equations

$$\tilde{u}^i(t; \omega, \rho_{\omega}^{-i}(\alpha), s_{\omega_j}^i) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \tilde{\pi}^i(t; \bar{\omega} | \omega, \rho_{\bar{\omega}}^{-i}(\alpha), s_{\omega_j}^i) \mu_{\bar{\omega}}^i + \lambda_{\omega_j}^i(\alpha) - \mu_{\omega}^i = 0,$$

$$(s_{\omega_j}^i \in S_{\omega}^i, \omega \in \Omega, i \in N),$$

$$\sum_{s_{\omega_j}^i \in S_{\omega}^i} \rho_{\omega_j}^i(\alpha) - 1 = 0, \quad (\omega \in \Omega, i \in N).$$

The transformation of variables enables us to omit all inequality constraints. Next, it is possible to define a homotopy function which maps each element  $(t, \alpha, \mu)$  into the right-hand side of the system above, i.e. to define the homotopy function  $\mathcal{H}$  as

$$\mathcal{H}(t, \alpha, \mu) = \begin{pmatrix} \tilde{u}^i(t; \omega, \rho_{\omega}^{-i}(\alpha), s_{\omega_j}^i) + \delta \cdot \sum_{\bar{\omega} \in \Omega} \tilde{\pi}^i(t; \bar{\omega} | \omega, \rho_{\bar{\omega}}^{-i}(\alpha), s_{\omega_j}^i) \mu_{\bar{\omega}}^i + \lambda_{\omega_j}^i(\alpha) - \mu_{\omega}^i, \\ (s_{\omega_j}^i \in S_{\omega}^i, \omega \in \Omega, i \in N), \\ \sum_{s_{\omega_j}^i \in S_{\omega}^i} \rho_{\omega_j}^i(\alpha) - 1, \quad (\omega \in \Omega, i \in N) \end{pmatrix}.$$

The homotopy function  $\mathcal{H}$  is continuously differentiable and the set of zeros of the homotopy-function,  $\mathcal{H}^{-1}(\{0\})$ , coincides with the set  $\mathcal{S}(\Gamma, p)$ .

**Theorem 5.1** *It holds that  $(t, \alpha, \mu) \in \mathcal{H}^{-1}(\{0\})$  if and only if  $(t, \rho(\alpha)) \in \mathcal{S}(\Gamma, p)$ .*

The following result presents the structure of  $\mathcal{H}^{-1}(\{0\})$ .

**Theorem 5.2** *The set  $\mathcal{H}^{-1}(\{0\})$  consists of finitely many differentiable arcs and loops. All arcs start and end in the subspaces where  $t = 0$  and  $t = 1$ . Loops have no points in common with those subspaces. There is exactly one arc that starts in  $t = 0$  and that ends in  $t = 1$ . All other arcs start and end in  $t = 1$  and connect two points inducing stationary equilibria of the stochastic game  $\Gamma$ .*

In Figure 4 the structure of the set of zeros of the homotopy-function is graphically represented. The structure of  $\mathcal{H}^{-1}(\{0\})$  is even simpler than the one of  $\mathcal{S}(\Gamma, p)$ . Not only are

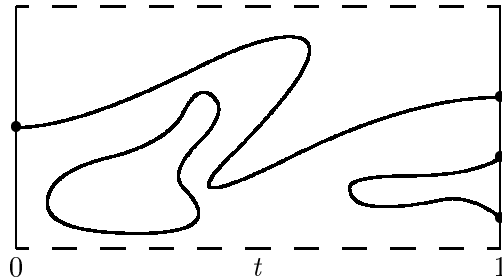


Figure 4: The structure of  $\mathcal{H}^{-1}(\{0\})$ .

complications like bifurcations, spirals, higher dimensional solutions sets, diverging behavior, etc., excluded. The arcs and loops in  $\mathcal{H}^{-1}(\{0\})$  are differentiable everywhere. It is the transformation of variables that smoothed out the kinks.

The stochastic game  $\Gamma^0$  naturally decomposes into  $n$  mutually independent and separate Markov decision problems, one for each player. Generically, Markov decision problems yield a unique optimal pure stationary strategy. The combination of all optimal strategies (for each player one) induces the starting point of our algorithm. This point can be determined

analytically since there are finitely many pure stationary strategies in each decision problem. We only have to compute the total discounted payoffs for each pure stationary strategy of a player and observe which one generates the highest payoff.

Starting at the unique point  $(0, \alpha^0, \mu^0) \in \mathcal{H}^{-1}(\{0\})$  at  $t = 0$  and following the path described by the zeros of  $\mathcal{H}$ , we end up in a point  $(1, \tilde{\alpha}, \tilde{\mu}) \in \mathcal{H}^{-1}(\{0\})$ . This point generates the stationary equilibrium  $\rho(\tilde{\alpha})$  of  $\Gamma$  selected by the stochastic tracing procedure. The differentiability of the homotopy-path makes it possible to apply standard path-following techniques.

There are several methods available to numerically trace homotopy-paths: (1) predictor-corrector methods, and (2) simplicial methods. Predictor-corrector methods approximately follow ‘exact’ solution curves, whereas simplicial methods follow approximate solution curves exactly. For more theory on path-tracking methods the reader is referred to Allgower and Georg (1990, 1993, 1997) and Garcia and Zangwill (1981).

In Herings and Peeters (2000), the software-package HOMPACT is used, a FORTRAN77 program (see Watson, Billups and Morgan (1987)).<sup>1</sup> Once the starting point is determined, the numerical process starts by following the homotopy-path from that point on.

For Example 2.3, Herings, Peeters, and Schinkel (2001) use the algorithm to study the case where there are two firms, no fixed costs and no entrance costs. Demand functions are assumed to be linear. The market is sufficiently big to admit two firms. Two firms involved in Cournot competition are both able to make positive profits. One would expect that in such an economic environment the market will be served by both firms in all periods. This repeated Cournot equilibrium is a stationary equilibrium indeed. However, when the discount factor is large enough ( $\delta > 0.8$ ), there is also an equilibrium, which induces an alternating monopoly. At even periods one firm serves the market and at odd periods the other.

The alternating monopoly equilibrium allows firms to take bi-periodically the full monopoly profit. It can be shown that equilibrium profits are higher in the alternating monopoly equilibrium than in the repeated Cournot equilibrium, implying that firms have strong incentives to coordinate on the alternating monopoly equilibrium.

In a situation where both firms have already entered the market, coordination on the alternating monopoly equilibrium is still possible. In the state where both firms are present, each one will leave the market next period with a small probability. As soon as one firm has left, the alternating monopoly gets into action.

The alternating monopoly equilibrium has been found by applying the numerical methods as outlined before. These methods can also be used to verify the robustness of the alternating monopoly equilibrium to variations in fixed costs, entry costs, elasticities of demand, multiple potential entrants, firms with asymmetric cost structures, and so on.

## 6 Summary

Stochastic games provide a rich framework in which to study dynamic interaction between heterogeneous economic agents. This paper focuses on the analysis of fully rational behavior

---

<sup>1</sup><http://www.netlib.org/hompack/index.html>

in this setting. We incorporate the paradigm that bygones are bygones, i.e. we restrict ourselves to the study of stationary equilibria.

We argue that the restriction to stationary equilibria, even though reducing the infinite number of equilibria predicted by the folk theorem to a finite, odd number, will in general not lead to uniqueness of equilibrium. Moreover, equilibria in stochastic games are notoriously hard to obtain in analytical form.

This paper shows how the problems of equilibrium selection and equilibrium determination can be handled. It extends the equilibrium selection theory of Harsanyi and Selten (1988) to stochastic games. The resulting equilibrium selection method is called the stochastic tracing procedure. The paper also presents a globally convergent numerical algorithm to compute the equilibrium selected by the stochastic tracing procedure.

The paper makes a strong case for the use of numerical methods in general. Not only can numerical methods be used to verify the robustness of equilibria that have been characterized analytically to deviations from the simplifying assumptions that permitted such characterizations. We illustrate by means of an example in the framework of dynamic oligopoly, how numerical methods can be used to challenge conventional economic wisdom.

## References

1. Allgower EL and K Georg (1990), *Numerical Continuation Methods: An Introduction*, Springer-Verlag, Berlin.
2. Allgower EL and K Georg (1993), "Continuation and Path Following," *Acta Numerica*, 1-64.
3. Allgower EL and K Georg (1997), "Numerical Path Following," *Handbook of Numerical Analysis*, 5, pp 3-208.
4. Breton M (1991), "Algorithms for Stochastic Games," in TES Raghavan, TS Ferguson, T Parthasarathy, and OJ Vrieze (eds.) *Stochastic Games and Related Topics*, Series C: Game Theory, Mathematical Programming and Operations Research, pp 45-57.
5. Breton M, JA Filar, A Haurie, and TA Schultz (1986), "On the computation of Equilibria in Discounted Games," in T Basar (ed.) *Dynamic Games and Applications in Economics*, Lecture Notes on Economics and Mathematical Systems 205, Springer-Verlag, Berlin.
6. Damme E van (1999), "Game Theory: The Next Stage," in AP Kirman and LA Gérard-Varet (eds.) *Economics Beyond the Millenium*, Oxford University Press, pp 184-214.
7. Filar JA, TA Schultz, F Thuijsman, and OJ Vrieze (1991), "Nonlinear Programming and Stationary Equilibria in Stochastic Games," *Mathematical Programming*, 50, 227-237.
8. Filar JA and OJ Vrieze (1997), *Competitive Markov Decision Processes*, Springer-Verlag, New York.

9. Fink AM (1964), "Equilibrium in a Stochastic  $n$ -Person Game," *Journal of Science of Hiroshima University Series A-I*, 28, 89-93.
10. Garcia CB and WI Zangwill (1981), *Pathways to Solutions, Fixed Points, and Equilibria*, Prentice Hall, Englewood Cliffs.
11. Haller H and R Lagunoff (2000), "Genericity and Markovian Behavior in Stochastic Games," *Econometrica*, 68, 1231-1248.
12. Harsanyi JC and R Selten (1988), *A General Theory of Equilibrium Selection in Games*, MIT Press, Cambridge, Massachusetts.
13. Herings PJJ and RJAP Peeters (2000), "Stationary Equilibria in Stochastic Games: Structure, Selection, and Computation," *METEOR Research Memorandum 00/031*, Universiteit Maastricht, Maastricht.
14. Herings PJJ and RJAP Peeters (2001), "Equilibrium Selection in Stochastic Games," *METEOR Research Memorandum 01/019*, Universiteit Maastricht, Maastricht.
15. Herings PJJ, RJAP Peeters, and MP Schinkel (2001), "Intertemporal Market Division: A Case of Alternating Monopoly," *METEOR Research Memorandum 01/021*, Universiteit Maastricht, Maastricht.
16. Judd KL (1997), "Computational Economics and Economic Theory: Substitutes or Complements?," *Journal of Economic Dynamics and Control*, 21, 907-942.
17. Judd KL (1998), *Numerical Methods in Economics*, MIT Press, Cambridge MA.
18. Lemke CE and JT Howson Jr. (1964), "Equilibrium Points of Bimatrix Games," *SIAM Journal on Applied Mathematics*, 12, 413-423.
19. Maskin E and J Tirole (2001), "Markov Perfect Equilibrium, I: Observable Actions," *Journal of Economic Theory*, 100, 191-219.
20. McLennan A (1999), "The Expected Number of Nash Equilibria of a Normal Form Game," mimeo.
21. Neyman A and S Sorin (2001), *Stochastic Games, Proceedings of the 1999 NATO Summer Institute on Stochastic Games held at Stony Brook*, forthcoming.
22. Parthasarathy T and TES Raghavan (1981), "An Orderfield Property for Stochastic Games when One Player Controls Transition Probabilities," *Journal of Optimization Theory and Applications*, 33, 375-392.
23. Schultz TA (1986), *Mathematical Programming and Stochastic Games*, Ph.D. Thesis, The John Hopkins University, Baltimore, Maryland.
24. Shapley LS (1953), "Stochastic Games," *Proceedings of the National Academy of Sciences*, 39, 1095-1100.

25. Sobel MJ (1971), "Non-cooperative Stochastic Games," *The Annals of Mathematical Statistics*, 42, 1930-1935.
26. Takahashi M (1964), "Equilibrium Points of Stochastic, Noncooperative  $n$ -Person Games," *Journal of Science of Hiroshima University Series A-I*, 28, 95-99.
27. Watson LT, SC Billups, and AP Morgan (1987), "HOMPACK: A Suite for Codes of Globally Convergent Homotopy Algorithms," *ACM Transactions on Mathematical Software*, 13, 281-310.