

92-F-9

**Evolution of Equilibria in the Long Run:  
A General Theory and Applications**

KANDORI, Michihiro  
University of Tokyo  
and  
Rafael ROB  
University of Pennsylvania

September 1992

Evolution of Equilibria in the Long Run:  
A General Theory and Applications\*

KANDORI, Michihiro\*\*  
Faculty of Economics, University of Tokyo

Rafael ROB\*\*\*  
Department of Economics, University of Pennsylvania

July 1991

Working paper version: September 1992

Abstract

We study an adaptive process of learning in the context of societal games played by randomly matched players. Our construction extends the analysis of Kandori, Mailath and Rob (1991) to general  $n \times n$  games. The basic assumptions we employ are: (1) Players adjust their behavior from time to time, picking a best-response against the present strategy configuration played by others. (2) Players mutate or experiment with strategies which need not be best-responses. Experimentation also occurs from time to time, but much less frequently than the best-response adjustment. In this situation, we show that a unique behavior pattern emerges in the long run, even when the underlying game possesses multiple Nash equilibria. In general, the stable pattern, or the long run equilibrium, may or may not correspond to a Nash-equilibrium. We first provide a general algorithm to determine the long run equilibrium, and then we apply the algorithm to pure coordination games and to games with strategic complementarities. For the former, the long run equilibrium turns out to be the Pareto efficient Nash equilibrium. For the latter, the long run equilibrium corresponds to a particular Nash equilibrium, and we present a simple graphical procedure to identify it.

---

\* References to relevant literature as well as useful discussions with Peyton Young are gratefully acknowledged. We would also like to thank the seminar audiences at Harvard, Montreal, North Carolina, Penn, Stony Brook, Tel Aviv, the NBER Decentralization Conference (Arizona) and the TCER Conference (Japan) for their comments.

\*\* Financial support from the NSF (SES-9108351) is gratefully acknowledged.

\*\*\* This project was started while the second author was visiting Princeton University.

## I. Introduction

In this paper we consider an evolutive process of learning in the context of societal games. The purpose of this is to generate a descriptive theory which explains how a Nash equilibrium is reached at in the realistic situation where individuals possess neither full information nor a congruent set of expectations. A second purpose of the investigation is to indicate that the evolutive process will lock into one specific equilibrium even when the game possesses several Nash equilibria.

In doing so we follow our previous work (see Kandori, Mailath and Rob (1991)) which considered the same issues in the context of  $2 \times 2$  games. In the present paper we generalize the analysis to  $n \times n$  games, broadening thereby the applicability of the theory, and exhibiting a richer range of outcomes that occurs in this more general setting. Once the scope of the theory is broadened we show how it applies to specific classes of economic games where the coordination of strategies plays an important role.

A brief review of the formulation might be useful at this point. We consider a population of individuals who are randomly and repeatedly matched to play a two person game. Individuals gradually adapt their behavior over time as they learn that other strategies yield higher payoffs (than the payoff they currently receive). In addition to this individuals experiment with (or "mutate" towards) new strategies independently of their payoff experience. These hypotheses capture the idea that players cannot maximize their payoffs in the traditional sense because (i) They are not aware of their own payoff structure; (ii) They cannot predict the present and future behavior of others; and (iii) It is costly to change actions.

Given these hypotheses, we determine the long-run behavior pattern in the population when the rate of experimentation is much smaller than the rate of the best response adjustment. By making the rate of experimentation arbitrarily small we capture the idea that (i) individuals are more likely to switch towards "proven" strategies than towards strategies whose payoff is uncertain; and (ii) the rate of experimentation is likely to decay over time as individuals approach a stable Nash configuration where their payoff cannot be increased any further. We show how to determine the long-run behavior pattern in the general case (ie, for generally specified games), and show what this behavior pattern looks like for specific classes of games which arise in the Industrial Organization literature.

For generally specified games we provide a complete algorithm for determining the set of long-run states. This algorithm extends the works of Kandori, Mailath and Rob (1991), and Young (1991). The algorithm consists of three subprocedures. First, it identifies a collection of limit sets. Those are the sets towards which the society tends under the best-response adjustment alone (i.e., without mutations). Second, it computes costs of transitions between the various limit sets. Third, it computes a minimum cost spanning-tree among the limit sets. The root(s) of this tree is (are) the long-run state(s). This construction follows essentially the footsteps of Freidlin and Wentzell (1984), except that we develop a discrete-time, discrete state-space formulation where perturbations (or mutations) are specified at the individual-player level (as opposed specifying them at the aggregate or population-level).

We then apply this algorithm to some classes of games possessing multiple equilibria. The first class is games with pure coordination: two players receive positive payoffs if and only if they take the same strategy, and otherwise they get zero. An example is the network problem, where each player is to choose one from several communication network systems. In this situation we show that the Pareto-efficient equilibrium, or the use of the best communication system, is the unique long-run state.

The second class of games we apply our algorithm to are games with strategic complementarities, or supermodular games. Those are games where each player's marginal payoff is increasing with respect to the rival's strategy, and important models in industrial organization and macroeconomics have this feature. An example is the differentiated-product oligopoly model with prices as strategies, which we examine in complete detail. The first result we establish for those games is that the best-response dynamics converges globally. That is, in the absence of mutations, starting from any initial condition, the society will tend towards some pure strategy Nash-equilibrium and will stay there forever. This result is to be contrasted with the well-known Shapley (non-supermodular) example indicating the possibility of limit cycles (Shapley (1964)).

Having established this result we know that the limit sets consist of singletons; namely, the pure-strategy Nash equilibria of the underlying game. Using this fact, we then show how the remaining parts of our algorithm apply to the differentiated-product oligopoly example (a special case of supermodular games). In particular, we present a simple graphical analysis of the best response correspondence to determine the minimum cost

transition between two Nash-equilibria (which is the algorithm's second component). This illustrates how the computation of the minimum cost spanning tree, the algorithm's third component, works in the context of specific games, and provides intuition concerning how the game's payoff structure determines which equilibrium is selected in the long-run.

One essential feature of our construction is that we formulate the model so that the long run behavior of the system does not depend on the details of the adjustment process. For the  $2 \times 2$  case, our previous work (Kandori, Mailath and Rob (1991)) employs the "Darwinian" assumption, which requires that players adjust towards the best response against the current population, and the main results for the  $2 \times 2$  case are derived independently of any further specifications of the adjustment process (for instance, how many players can adjust at each moment of time). The same Darwinian assumption, however, does not exactly pin down the direction of adjustment for general  $n \times n$  games. For example, consider a game with three strategies, and suppose that strategy 1 is the current best response. The Darwinian assumption says that the population with strategy 1 should increase, but it does not specify how the relative populations of strategies 2 and 3 change. This implies that the very basic nature of the adjustment process, and thus the results in the  $n \times n$  case may well depend on such details. To overcome this difficulty, we choose not to commit to any one specific adjustment, and we assume that at each moment any adjustment is possible with some positive probability, as long as it is consistent with the Darwinian assumption. This, we feel, is more natural than assuming that the adjustment always takes a specific form. More interestingly, with this (completely general) specification we can still show that the results are independent of specific details, such as the probability distribution over the possible directions of adjustment. Furthermore, this assumption turns out to simplify our analysis drastically.

The remainder of the paper is organized as follows. The next section introduces the underlying game, the societal game, and the adjustment process. Section III introduces the equilibrium concept and defines long-run states and limit sets. In section IV we analyze the model and show how to go about computing the set of long-run states. Section V applies the general theory to pure coordination games and supermodular games. We first prove global convergence results for them, and then show--in complete detail-- how the results of section IV apply to those games. Section VI concludes the paper.

## II. Formulation

### II.1 The Underlying Static Games.

In this section we describe symmetric, two-person games. We denote each player's strategy set by  $S = \{s_1, s_2, \dots, s_n\}$ , and denote the payoff to a player adopting  $s \in S$  while the opponent is adopting  $s' \in S$  by

$$u(s, s').$$

(In this case the opponent's payoff is  $u(s', s)$ ). We shall sometimes abbreviate  $u(s_i, s_j)$  as  $u_{ij}$ , and refer to "strategy  $i$ " instead of "strategy  $s_i$ ". No confusion will ensue. A mixed strategy is represented by a point in the  $(n-1)$ -dimensional simplex  $\Delta$ , and the payoff of strategy  $i$  against mixed strategy  $\alpha \in \Delta$  is denoted by

$$u(i, \alpha) \equiv \sum_{j=1}^n \alpha_j u_{ij}.$$

The set of pure strategy best-responses against strategy  $\alpha \in \Delta$  is denoted by

$$BR(\alpha).$$

### II.2 The Societal Game and Intertemporal Learning

Next, given a two-person game, the following repeated societal game is considered. Society consists of  $M$  individuals (or "members"). At the beginning of each period each individual chooses a pure strategy  $s \in S$ , and sticks to it for the duration of the period. The configuration of strategy-choices in the society is summarized by the state-vector  $z$ , whose  $i^{\text{th}}$  element,  $z_i$ , represents the number of players with strategy  $i$ . The state space is a finite set

$$Z \equiv \{(z_1, \dots, z_n) \mid z_i \in \{0, 1, \dots, M\}, \sum_{i=1}^n z_i = M\}. \quad (2.1)$$

Let  $|Z|$  denote the number of elements in  $Z$  (actually,  $|Z| = \binom{M+n-1}{M}$ ), and for each  $z \in Z$ , define the set of existing strategies by  $C(z) \equiv \{i \mid z_i > 0\}$ . This is referred to as the carrier of  $z$ .

Once individuals have chosen strategies they are randomly matched and each pair of matched players plays the

above (two-person) game, implementing their preselected strategies. At the end of the period, an individual who chose  $s_i \in S$  collects an average payoff of

$$\begin{aligned}\pi_i(z) &= \frac{1}{M-1} [\sum_{j \neq i} z_j u(s_i, s_j) + (z_i - 1) u(s_i, s_i)] \\ &= \frac{1}{M-1} [\sum_{j=1}^n z_j u(s_i, s_j) - u(s_i, s_i)],\end{aligned}$$

if  $M$  is even. If  $M$  is odd, there is one unmatched player, and we assume that players get zero payoffs if they are not matched. In that case, the number  $M-1$  in the above expression must be replaced with  $M$ . For concreteness, we assume that the population size,  $M$ , is even, but all the results also hold if  $M$  is odd.

Once the period is over and all individuals collect their payoffs, a certain subset of individuals are given the opportunity to change their strategies. Below we specify how this subset is determined. Consider an individual belonging to this subset, and assume he or she chose strategy  $s_i \in S$  in the previous period. In picking a (possibly) new strategy the individual compares the profitability of each available strategy on the assumption that other players will continue to use the strategies they are presently using. In other words, we assume that a player, if she adjusts at all in the next period, takes a myopic best response to the strategy configuration in the current period. We offer the following two different stories that underlie this specification.

First, suppose that there are large adjustment costs of changing actions. Assume further that on infrequent occasions the adjustment cost declines to the point where it pays to switch strategy. For example, if the strategy is to choose a personal computer, a player switches strategies when the machine deteriorates. Thus, only a small number of individuals will adjust their strategies at any given point in time, so that the present strategy profile in the society is likely to persist into the future. In addition to this, if the players' discount factors are not too large (relative to the speed of adjustment), the model is interpreted as one where players are fully rational, i.e., where they maximize the discounted sum of payoffs given a set of rational expectations concerning the future evolution of the game.

As a second possibility consider the case where players are boundedly rational. That is, where they don't even know their own payoff matrix or they just do not have enough computational ability to find the optimal behavior. However, they are informed (each time they are given the opportunity to change strategy)

of the relative success of all existing strategies in the society. In other words, they are informed of the payoff vector  $(\pi_i(z))_{i \in C(z)}$ , and pick a currently highest paying strategy,  $\text{Argmax}_{i \in C(z)} \pi_i(z)$ . This interpretation is more descriptive of the evolution of biological games, where the growth and decline of strategies stem from the mechanics of imitation and reproduction, as opposed to those which stem from intelligent choices (as would be the case under the first interpretation).

The details of the two processes is not the same, however. For instance, one configuration where the dynamics is different is where the entire population is playing the same non-Nash strategy. Then -- under the biological interpretation -- players are stuck with that strategy. But--under the rationality interpretation--this configuration is upset as soon as one player is given the opportunity to change. Despite of these differences the qualitative features of the dynamics are the same once mutations are introduced into the picture (see below). Therefore, our main results apply to both scenarios. For concreteness we will focus our attention only on the first one, which is based on the rationality of the players.

Consider now an individual who chose strategy  $s_i \in S$  under the population configuration  $z$  ( $i \in C(z)$ ). If this individual is given the opportunity to change his strategy, he first excludes himself from the population and calculates the remaining strategy distribution. Namely, the strategy distribution he expects to face is  $\alpha(z, i) \in \Delta$ , which is defined by

$$\alpha_j(z, i) = z_j / (M-1) \text{ for } j \neq i \text{ and } \alpha_i(z, i) = (z_i - 1) / (M-1). \quad (2.2)$$

The expected payoff of strategy  $k$  against this distribution is given by

$$u_i(k, z) \equiv u(k, \alpha(z, i)).$$

Define the set of pure strategy best responses to strategy distribution  $z$  for the players with strategy  $s_i$  by

$$\beta_i(z) \equiv \text{Argmax}_k u_i(k, z).$$

The mechanics of adjustment in the absence of mutations are stated in the following assumption.



**Assumption A:** In the absence of mutations,

- (i) At the beginning of each period, each player has an opportunity to adjust her action with probability  $\eta \in (0,1)$ , and this is independent across players and over time,
- (ii) If a player is not given the opportunity to adjust, or if the opportunity is given but the player's strategy is already a best-response against the present strategy configuration (ie, if  $i \in \beta_i(z)$ ) then the player sticks to her last period strategy.
- (iii) If adjustment is possible and if  $i \notin \beta_i(z)$ , then the player switches to one of the best responses in  $\beta_i(z)$ , where  $z$  is the population configuration and  $i$  is the player's strategy in the previous period. The choice of strategy in  $\beta_i(z)$  is according to a prespecified distribution which puts a positive weight on each element in  $\beta_i(z)$ . These choices are independent across players and over time.

**Remarks:** 1. The choice of best response in (iii) above when there are multiple best responses is for concreteness' sake only: most of our results hold for any tie breaking rule (allowing for time varying probabilities, dependence, etc). Only the analysis of the differentiated-product Industry (Section V.B.2) uses the full force of assumption A.

2. The speed of adjustment for a player can depend on the current state and on the current strategy for the player. That is, instead of having a single  $\eta$ , we can specify different adjustment rates,  $\eta_i(z) \in (0,1)$  for  $i \in C(z)$ . All of our results remain intact, and exactly the same proofs go through, even under this more general formulation. This is in a sharp contrast to the continuous models (Foster and Young (1990) and Fudenberg and Harris (1991)), where the speed of adjustment plays a crucial role.

Note that condition (ii) says that the players change their strategies only when their current strategies are strictly worse than the best ones. This assures that once a Nash equilibrium is reached, the society stays there until mutations come in, even if the equilibrium is not strict. In particular, when the strategy distribution in the society corresponds to a mixed strategy equilibrium, each player has multiple best responses, but assumption A(ii) implies that the strategy distribution does not drift away unless players mutate. We call the process specified by assumption A the Darwinian dynamics, and denote it by  $P(0)$  (the "0" in the argument

indicates zero mutation rate).

In addition to the Darwinian dynamics,  $P(0)$ , we postulate that new strategies enter into the population with small probabilities. More specifically, we assume that an individual who is expected to play strategy  $k$  "mutates" to strategy  $j$  with probability  $m_j \varepsilon > 0$ , where  $\sum_j m_j = 1$ , and  $m_j, \varepsilon \in (0,1)$ . These mutations are independent across players and over time and occur (if they occur) after the Darwinian adjustment. One way to justify this assumption in economic contexts is to assume the turnover of population. Each player exits from the society with probability  $\varepsilon$  and then her position is replaced with a newcomer (so that the population size stays constant). The newcomer does not know much about the game and the history of play in the society, so she takes each strategy with positive probability, reflecting her prior beliefs. In the case of a rational player who knows the payoff function, it would be better to assume that the newcomer picks only rationalizable strategies with positive probabilities. We will come back to this issue later (see p. 29). One of the striking features of our analysis is that the main result, the choice of long-run equilibria, does not depend on the distribution of mutations,  $m \equiv (m_1, \dots, m_n)$ , as long as that distribution places positive probabilities on all strategies.

Between the Darwinian dynamics,  $P(0)$ , and the mutations we have a composite stochastic adjustment rule:

$$z_i(t+1) = \tilde{z}_i(t+1) + \sum_{j \neq i} (x_{ji}(t) - x_{ij}(t)), \quad i=1, \dots, n, \quad (2.3)$$

where  $\tilde{z}(t+1)$  is the state at  $t+1$  according to  $P(0)$  and  $x_{ij}(t)$  is a binomial random variable denoting the number of mutants departing the  $i^{\text{th}}$  strategy and switching over into the  $j^{\text{th}}$  strategy. The distribution  $\tilde{z}(t+1)$  is called the intended strategy distribution. According to our independence assumption, etc., we have

$$x_{ij}(t) \sim \text{Bin}(\tilde{z}_i(t+1), m_j \varepsilon) \quad (2.4)$$

(i.e.,  $x_{ij}(t)$ 's are binomially distributed random variables).

From (2.3) and (2.4) it follows that the stochastic law governing the adjustment of strategies is given by a  $|Z| \times |Z|$  Markovian transition matrix  $P(\varepsilon) = (p_{zz'}(\varepsilon))$ , where each element  $p_{zz'}(\varepsilon)$  is a polynomial expression in  $\varepsilon$ :

$$p_{zz}(\varepsilon) = \sum_{k=0}^M r_{zz}(k) \varepsilon^k \quad (2.5)$$

Next let us introduce an important concept, the distance between two states.

**Definition 1:** The distance between two states,  $z$  and  $z'$ , is given by

$$d(z, z') = \frac{1}{2} \sum_{i=1}^n |z_i - z'_i|.$$

Note that this represents the minimum number of strategy changes which are necessary to achieve state  $z'$  from state  $z$ . That is, state  $z'$  is achieved from state  $z$  if at least  $d(z, z')$  players switch their strategies in appropriate ways. For example, consider  $z = (3, 2, 2)$  and  $z' = (1, 2, 4)$ , where we have seven players and three strategies. State  $z'$  is achieved in many different ways from state  $z$ : One way is that two players with strategy 1 change their strategies to strategy 3. Another way is that all three players with strategy 1 switch to strategy 3 and at the same time one player with strategy 3 switches to strategy 1. The former involves less changes than the latter, and indeed it involves the minimum number of changes,  $d(z, z') = 2$ , to achieve state  $z'$  from state  $z$ .

Given this definition of distance, we introduce the cost of transition between two states, which plays a crucial role in what follows.

**Definition 2:** The cost of transition between  $z$  and  $z'$  is defined by

$$c(z, z') = \text{Min}_{z'' \in b(z)} d(z', z''), \quad (2.6)$$

where  $b(z)$  is the set of possible intended strategy configurations at  $t+1$  when strategy configuration at time  $t$  is given by  $z$ :  $b(z) \equiv \{ z' \in Z \mid p_{zz'}(0) > 0 \}$ .

The number  $c(z, z')$  defined above measures the minimum number of mutations to achieve state  $z'$  from state  $z$  in one period. To understand this, consider how state  $z'$  is achieved from state  $z$  within one period. Given that yesterday's strategy configuration is  $z$ , today's intended strategy configuration is first chosen from the set  $b(z)$ . Then, some of the players mutate, and this determines today's strategy configuration. So state

$z'$  is achieved with minimum number of mutations (or "least cost"<sup>3</sup>) if (i) the intended strategy configuration that is closest to  $z'$  is chosen first and (ii) then just enough mutations occur to achieve  $z'$ . The expression (2.6) comes from this particular way of transition, and the importance of this definition comes from the fact that this is the most likely way to achieve state  $z'$  from state  $z$  when the mutation rate,  $\epsilon$ , is small.

Given this preparation, we prove the following.

**Proposition 1:** (i) If  $\epsilon > 0$ ,  $p_{zz'}(\epsilon) > 0$  for any pair of states  $z, z' \in Z$ . (ii) The coefficients in expression (2.5) satisfy the following property:  $r_{zz'}(k) = 0$  for  $k < c(z, z')$  and  $r_{zz'}(k) > 0$  for  $k = c(z, z')$ .

**Proof:** Claim (i) is true because a transition from  $z$  to  $z'$  is always feasible by picking a suitable pattern of mutations. Assertion (ii) follows from the fact that  $c(z, z')$  is the minimum number of mutations to switch the society over from  $z$  to  $z'$ , and that mutations are independent across players. ■

As an immediate implication of proposition 1 we have:

$$p_{zz'}(\epsilon) = O(\epsilon^{c(z, z')}). \quad (2.7)$$

That is, the cost of transition  $c(z, z')$  measures how fast the transition probability from  $z$  to  $z'$  tends to zero as the mutation rate ( $\epsilon$ ) goes to zero.

### III. The Equilibrium Concept

Given the above adjustment rule it is clear (from proposition 1 (i)) that the society will perpetually fluctuate among the different states,  $z \in Z$ . Accordingly, we consider a stochastic equilibrium concept, measuring the average proportion of time that the society will spend in each state. Formally, let  $\Delta^{|Z|}$  denote the  $|Z|-1$  dimensional simplex, i.e., let

$$\Delta^{|Z|} = \{\mu \in \mathbb{R}_+^{|Z|} \mid \sum_{z \in Z} \mu_z = 1\}.$$

---

<sup>3</sup>The reason for introducing this terminology is that the problem of determining the set of long-run states is mathematically equivalent to a minimum cost spanning tree problem (see below), and  $c(z, z')$  is part of the data in the latter problem.

Then, we introduce the following concept.

**Definition 3:** Given  $\varepsilon > 0$ ,  $\mu(\varepsilon) \in \Delta^{|Z|}$  is a stationary distribution if

$$\mu(\varepsilon)P(\varepsilon) = \mu(\varepsilon). \quad (3.1)$$

When the mutation rate  $\varepsilon$  is positive, the existence, uniqueness and the global stability of  $\mu(\varepsilon)$  is guaranteed by Proposition 1(i) and by well-known results of probability theory (see Loeve, p. 39). By global stability we mean that the state distribution in the future converges to the stationary distribution starting from any initial state (or any initial distribution over states). Moreover,  $\mu(\varepsilon)$  corresponds to the proportion of time spent at the various states as well (as the time-horizon tends to infinity). Hence, the system is well behaved, and we can make a unique (stochastic) prediction irrespective of the initial condition whenever the mutation rate is strictly positive.

This is in sharp contrast to the case of  $\varepsilon = 0$ , where the linear system of equations (3.1) may very well possess multiple solutions. This will occur whenever the underlying game possesses multiple pure strategy Nash-equilibria, because the distribution which puts probability one to the configuration where all players choose the same pure strategy Nash equilibrium is always a stationary distribution of  $P(0)$ . Moreover, if the Darwinian dynamics,  $P(0)$ , possesses a limit cycle, there is another stationary distribution which puts all probability to the states on the cycle. Unlike the case with positive mutation rate, which distribution is eventually chosen is crucially dependent on the initial condition in this case. The purpose of introducing (small rate of) mutations is to resolve this indeterminacy and select a particular stationary distribution. This idea is reflected in the following definition.

For any  $q \in \Delta^{|Z|}$ , let  $C(q) \equiv \{z \in Z \mid q_z > 0\}$ . Then:

**Definition 4:** (i) The limit distribution  $\mu^* \in \Delta^{|Z|}$  is defined by

$$\mu^* = \lim_{\varepsilon \searrow 0} \mu(\varepsilon).$$

(ii) The set of long-run states is  $L = C(\mu^*)$ .

(The existence of  $\lim_{\varepsilon \rightarrow 0} P(\varepsilon)$  and  $\lim_{\varepsilon \rightarrow 0} \mu(\varepsilon)$  comes from the fact that (i) the elements of  $P(\varepsilon)$  are

polynomials in  $\varepsilon$  (equation 2.5) and (ii) the elements of  $\mu(\varepsilon)$  are the ratios of polynomials in  $\varepsilon$  (see Section IV.1.)

Having provided the basic idea and the definitions, we shall examine our equilibrium selection procedure in more detail. For the dynamics without mutations (i.e.,  $P(0)$ ), we have seen that the equilibrium points and the limit cycles provide different stationary distributions. Those objects are special cases of the limit sets, which are formally defined as follows.

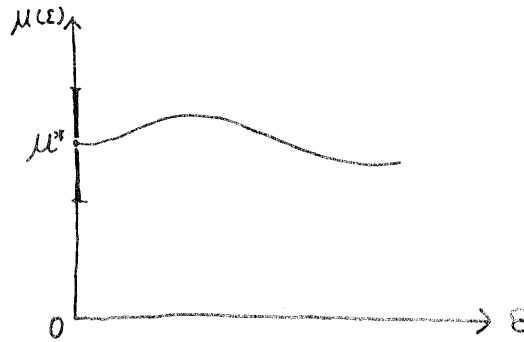
**Definition 5:** A set  $A \subset Z$  is a limit set if, under  $P(0)$ ,

- (i)  $\Pr(z(t+1) \notin A \mid z(t) \in A) = 0$ , and
- (ii) for all  $z, z' \in A$ ,  $\Pr(z(t+k) = z' \mid z(t) = z) > 0$  for some  $k > 0$ .

The collection of all limit sets is denoted  $\Omega$ .

A limit set is "closed" in a probabilistic sense (Definition 5(i)) and its elements are mutually reachable ((ii)). By the definition, it is easy to see that the state under the Darwinian dynamics will eventually converge to one of the limit sets with probability one. It is well known that there is a unique stationary distribution  $\mu_A$  for each limit set  $A$ , and the set of all stationary distributions of  $P(0)$  is the convex hull of those,  $\text{Co}\{\mu_A, A \in \Omega\}$  (see Loeve (1963), p.41). As suggested above, out of those multiple distributions, we select a particular one, by first introducing a positive mutation rate  $\varepsilon$  and then let it tend to zero;  $\mu^* = \lim_{\varepsilon \rightarrow 0} \mu(\varepsilon)$ . The fact that the limit distribution  $\mu^*$  is indeed a stationary distribution of  $P(0)$  is verified by taking the limits of both sides of the equation  $\mu(\varepsilon) = P(\varepsilon)\mu(\varepsilon)$ . (See paranthetical remark after definition 4 above). The nature of our procedure is illustrated in Figure 1. The figure shows that the set of solutions to (3.1) as a correspondence (from the domain of  $\varepsilon$ 's into the simplex  $\Delta^{|z|}$ ) is upper hemicontinuous. Thus, one property of our equilibrium selection is continuity with respect to perturbations. However, unlike the usual equilibrium refinements (Kohlberg and Mertens (1986), for example), the perturbations are not performed in the payoff space, but in the space of explicitly formulated dynamic adjustment processes.

Figure 1



Given the above argument, we have the following proposition which gives the classification of the long run states.

**Proposition 2:** The limit distribution  $\mu^*$  is uniquely decomposed as follows:

$$\mu^* = \sum_{A \in \Omega^*} r_A \mu_A,$$

where  $\Omega^*$  is a subset of  $\Omega$ ,  $\mu_A$  is the unique stationary distribution on the limit set  $A$ , and  $r_A \in (0,1]$  represents the likelihood of the limit set  $A$ . An element of  $\Omega^*$  is called a long run equilibrium.

Note that the uniqueness follows from the fact that limit sets are mutually disjoint by definition.

**Remarks:** 1. This shows that we can further partition the set of long-run states into a collection of components, which are called long run equilibria. The nature of these components is such that once the society has reached a specific component it will fluctuate among the states belonging to it (driven by the forces of Darwinian dynamics), and will depart only upon the occurrence of mutations. The partitioning is trivial, of course, in case of a single long-run state. In case there is more than one long-run state, however, this partitioning gives us an interesting classification of cyclical behavior in the society. First, there are short-run cycles: those occur when the society is at a long-run equilibrium which contains more than one state; in this case the society frequently fluctuates among those states. Second, there are long-run cycles: those occur whenever infrequent mutations drive the society from one component to another. This terminology is justified when  $\epsilon \downarrow 0$ , so that the rate of mutations is indeed much slower than the rate of Darwinian adjustment ( $\eta$ ).

2. It is conceivable that the limit distribution puts positive probabilities to all limit sets, so that the collection of

long run equilibria coincides with the collection of limit sets ( $\Omega^* = \Omega$ ). If this is the case, our procedure does not narrow down the set of equilibria, although the limit distribution still provides the relative likelihood of each limit set. However, in what follows, we will argue that "typically" the limit distribution attaches positive probability to only one limit set; that is, the long run equilibrium is typically unique. This has already been shown by Kandori, Mailath and Rob (1991) for 2x2 symmetric games.

3. Even in cases where a unique long run equilibrium obtains, it may not be a Nash equilibrium. It could be the case that a limit cycle is selected by our procedure. A sufficient condition for this not to happen is that the Darwinian dynamic does not possess any cycles. Below we will show that this is in fact the case for pure coordination games and supermodular games, so that our procedure selects a particular Nash equilibrium in those games (Theorems 2 and 3 below).

#### IV. Analysis

##### IV.1. Generalities

As in our previous paper (Kandori, Mailath and Rob (1991)) we identify long-run states by using a graph-theoretic characterization of the invariant measures  $\mu(\epsilon)$ .

According to this approach,  $\mu$  is a scalar multiple of a vector  $q \in \Delta^{|Z|}$  where

$$q_\zeta = \sum_{h \in H_\zeta} \prod_{(z,z') \in h} p_{zz'} \quad (4.1)$$

and where  $H_\zeta$  is the set of a particular class of directed graphs defined on the state space  $Z$ , called  $\zeta$ -trees. A  $\zeta$ -tree is a set of directed branches (ordered pairs of states), so that every  $z \neq \zeta$  is the origin of exactly one branch, and starting from any such state there is a unique sequence of branches terminating at  $\zeta$ .

Expressions (4.1) and (2.5) make it clear that each  $q_\zeta$  is a polynomial expression in  $\epsilon$  which tends to zero as  $\epsilon \rightarrow 0$  (unless the game possesses a unique globally stable equilibrium). Note that the stationary distribution  $\mu$  is obtained by normalizing  $q$ ,  $\mu = (1/\sum_\zeta q_\zeta)q$ , and in this expression both the denominator and the numerators tend to zero as  $\epsilon \rightarrow 0$ . So the identification of long-run states, which receive positive



probabilities in the limit, will hinge on how fast each  $q_\zeta$  tends to zero. (For instance, if one state,  $\zeta^*$ , can be identified for which  $q_{\zeta^*}$  tends to zero slower than any other  $q_\zeta$ , then  $\zeta^*$  will receive probability one as  $\epsilon \downarrow 0$ .)<sup>4</sup> In other words, in such cases we would have  $\mu_{\zeta^*}^* = 1, \mu_\zeta^* = 0$  for  $\zeta \neq \zeta^*$ )

Notationally, expression (4.1) gives us:

$$q_\zeta = \sum_{v=0}^K a_\zeta(v) \epsilon^v \quad (4.2)$$

where  $K \leq M|Z|$ .

Thus, if  $v_\zeta$  is the speed of convergence of  $q_\zeta$ , i.e. if

$$q_\zeta = O(\epsilon^{v_\zeta}),$$

then expressions (2.5), (4.1) and (4.2) imply

$$v_\zeta = \text{Min}\{v | a_\zeta(v) \neq 0\} = \text{Min}_{h \in H_\zeta} \sum_{(z,z') \in h} c(z,z'). \quad (4.3)$$

We call  $v_\zeta$  the cost of transition to state  $\zeta$ , and it can be thought of as a measure of the difficulty of transiting from every other state to  $\zeta$ . Therefore, the problem of finding the set of long-run states is reduced to the problem of minimizing expression (4.3) over all states  $\zeta \in Z$  and over all trees  $h \in H_\zeta$ :

$$\text{Min}_{\zeta \in Z} \text{Min}_{h \in H_\zeta} \sum_{(z,z') \in h} c(z,z'). \quad (4.4)$$

In other words, the long run states are those with the least cost of transition. (The motivation for calling  $c(z, z')$  "the cost of transition" is now revealed, since (4.4) translates the identification of long run states into an optimum branching program where  $c(z, z')$  is the cost of branch  $(z, z')$ ). Thus we have:

**Theorem 0 (Kandori, Mailath and Rob (1991)):** The set of long run states is the set of solutions to program (4.4).

We also have:

---

<sup>4</sup>This is exactly what occurs in the applications; see Section IV.

**Theorem 1:** The set of long run states does not depend on the speed of adjustment, the direction of adjustment (reflected in  $\eta_i(z)$ ), or the distribution of mutations ( $m$ ).

**Proof:** Expression (2.6) shows that  $c(z, z')$  depends only on  $b(z)$  which is independent of the above parameters. Hence, the solution to (4.4) cannot depend on these parameters either.

We now turn to an algorithm to solve program (4.4).

#### IV.2. An Algorithm for Determining the Set of Long-run States

For simple models such as  $2 \times 2$  games, the problem (4.4) can be solved directly. However, the problem could be highly complicated for more general games. Accordingly, the first step is to show that there is a way to simplify program (4.4). Program (4.4) amounts to finding the minimum cost tree defined on the whole state space,  $Z$ . Instead of considering all points in the state space, however, we can restrict our attention to the collection of limit sets, which were introduced in the last section. This collection is typically much smaller than  $Z$ .

We start with the definition of the cost of transition between two limit sets. For  $z, z' \in Z$ , let  $G(z, z')$  be the set of directed paths from  $z$  to  $z'$ . For two limit sets  $A, A' \in \Omega$ , define the cost of transition from  $A$  to  $A'$  by

$$C(A, A') = \min_{z \in A, z' \in A'} \min_{g \in G(z, z')} \sum_{(\zeta, \zeta') \in g} c(\zeta, \zeta').$$

Now consider the optimum branching problem defined on the collection of limit sets:

$$\min_{A \in \Omega} \min_{h \in H_A} \sum_{(A', A'') \in h} C(A', A''). \quad (4.5)$$

$H_A$  in the above expression refers to the set of  $A$ -trees defined on the collection of limit sets,  $\Omega$ . It is known that the original program (4.4) is equivalent to the much simpler program, (4.5):

**Fact:** Program (4.5) provides the solution to program (4.4). Namely, let  $Z^*$  and  $\Omega^*$  be the sets of solutions to (4.4) and (4.5), respectively. Then,  $Z^* = \cup_{A \in \Omega^*} A$ . Therefore, the long run equilibria are obtained as the

solutions to program (4.5).

**Remark:** This is a discrete version of Freidlin and Wentzell's analysis, which is given in the continuous time and continuous state space framework (1984, Chapter 6). A complete proof for the discrete case is found in Young (1991) (Lemma 2 in his Appendix).

What remains is to solve the reduced program (4.5). This can be broken into two subprograms: (1) the determination of transition costs among limit sets,  $C(A, A')$ ; and (2) an algorithm to solve program (4.5) given the transition costs. Efficient algorithms for handling both problems are available in the combinatorial optimization literature. The first problem is solvable by the Dijkstra algorithm (1959), while the second is solvable by the Chu-Liu/Edmonds algorithm (1965/7). We provide a complete description of the latter in Appendix A, and in the next section we illustrate how it works in particular examples. As to the costs of transition, we shall determine them directly (ie, without relying on the Dijkstra algorithm), using the special structure we have at hand. The next proposition leads towards that goal.

**Definition 6:** For strategy  $i$ , let  $e_i = (0, \dots, M, \dots, 0)$  be the state where all players take strategy  $i$  and let  $BR_i$  be the set of states where  $i$  is a best response for all players:

$$BR_i = \{ z \in Z \mid i \in \beta_j(z) \text{ for all } j \in C(z) \}.$$

$BR_i$  is called the best response region for strategy  $i$ .

**Proposition 3 (The triangular inequality):** Suppose strategy  $i$  constitutes a pure strategy Nash equilibrium.

Then, for all  $x \in Z$  and  $y \in BR_i$ , we have

$$c(e_i, x) \leq c(e_i, y) + c(y, x).$$

**Remark:** The triangular inequality says that an immediate jump from a "Nash state",  $e_i$ , to another state  $x$  is less costly than any gradual transition through the best response region of the starting point,  $(BR_i)$ . Note that in the above inequality, the "destination"  $x$  can be any point but the "intermediate point",  $y$ , must be in  $BR_i$ . (In general, the same is true for any number of intermediate points: repeated application of the triangular inequality trivially shows the corollary below.) The intuition is that if the society goes to an intermediate point  $y$ , the selection pressure pushes the state back to the Nash state  $e_i$ , because strategy  $i$  is the best response at state  $y$ . This is why

going through  $y \in BR_i$  is inefficient. An important assumption in this argument is that mutations are independent, which implies that the cost of transition is just equal to the number of required mutations.

**Corollary:** Suppose strategy  $i$  constitutes a pure strategy Nash equilibrium. Then, for all  $x \in Z$  and  $y^0, \dots, y^n \in BR_i$ , we have

$$c(e_i, x) \leq c(e_i, y^0) + \sum_{k=1}^n c(y^{k-1}, y^k) + c(y^n, x).$$

**Proof of Proposition 3:** Note that  $c(e_i, x) = M - x_i$  and  $c(e_i, y) = M - y_i$ , where  $M$  is the number of players. This is true because  $c(a, b)$  is the minimum number of mutations to achieve state  $b$  from state  $a$  in one period. Since  $c(x, y)$  is nonnegative, if  $y_i \leq x_i$ , we are done. So suppose  $y_i > x_i$ . Then, we have

$$c(y, x) = \text{Mind}_{z \in b(y)}(x, z) \geq \text{Min}_{z \in b(y)} |x_i - z_i| \geq y_i - x_i,$$

where the last inequality follows from the fact that  $z_i \geq y_i > x_i$  for all  $z \in b(y)$ . (Recall that  $b(y)$  is the support of tomorrow's state distribution in the absence of mutations, given that today's state is  $y$ . Since  $y$  lies in the best response region of strategy  $i$ , the number of players with strategy  $i$  should be weakly increasing. This is why we have  $z_i \geq y_i$  for all  $z \in b(y)$ .) Hence we have  $c(e_i, x) = M - x_i = (M - y_i) + (y_i - x_i) \leq c(e_i, y) + c(y, x)$ . ■

## V. Applications: Supermodular and Pure Coordination Games

In this section we show how evolution selects a unique long-run equilibrium for two classes of games with multiple (static) equilibria. For these games the "relevant" (see below) equilibria are in pure strategies, and the Darwinian dynamics always converge to one of them; hence, limit cycles are ruled out. We start by defining the games, and listing a few of their properties.

### V.A.1. Supermodular Games

The set of strategies is assumed to be strictly and completely ordered, and the indexing of strategies is such that

$$s_{i+1} > s_i, \quad i \leq i < n.$$

Supermodularity is defined by the requirement that for any pair  $1 \leq i < k \leq n$  the differences  $u_{kj} - u_{ij}$

are strictly increasing in  $j$ .

Three leading examples of supermodular games are: (i) An oligopoly game (suitably discretized) where the strategic variable is price. In particular, the Hotelling differentiated product model (once firms' locations are fixed at the extremities of the line). Our results will be illustrated for this case. (ii) The classical two-firm Cournot model with linear demand and linear costs. Here the strategic variable is quantity, but the natural ordering over strategies has to be reversed as to satisfy supermodularity (this trick will not work for the  $N$  firm case with  $N > 2$ ). (iii) A macro-economic coordination game where the strategic variable is search effort. In particular, the Diamond Mortensen (1982) matching model. A variety of other examples is given in Milgrom and Roberts (1990).

The fact that a player's payoff increases in the opponent's strategy often leads to multiple equilibria in supermodular games (acquiring special significance in the macroeconomic literature). In our framework a necessary and sufficient condition for this to occur is that the game is not dominant solvable. In other words, after iteratively eliminating all strictly dominated strategies, the game still possesses more than one strategy. This fact is stated in the following proposition, and is proven in Appendix B.

**Proposition 4:** (i) Suppose all strictly dominated strategies have been iteratively removed from the game. Then, both  $(s_1, s_1)$  and  $(s_n, s_n)$  are Nash Equilibria (NE, for short). (ii) No asymmetric NE in pure strategies exists:  $(s_i, s_j)$  is a NE only if  $s_i = s_j$ . (iii) If  $s_t$  is a Best Response (BR, for short) to  $s_i$  and if  $s_m$  is a BR to  $s_j$ , where  $j > i$ , then  $s_m \geq s_t$ .

(With the exception of 4(ii) these properties are well-known; the proof in the appendix is short, and provided for the reader's ease of reference).

According to Proposition 4(ii), the set of pure strategy NE is a subset of the "main diagonal"  $\{(s_i, s_i)\}_{i=1}^n$ . Let us denote this set by  $N$ , i.e., let

$$N \equiv \{1 \leq i \leq n \mid (s_i, s_i) \text{ is a pure-strategy NE}\}.$$

The next proposition extends part (iii) of proposition 4 to the case where we compare across mixed strategies (as opposed to comparing across pure strategies). In this case a similar monotonicity property obtains

when mixed strategies are (partially) ordered according to first-order stochastic-dominance<sup>5</sup>. This simple property will play an important role in the sequel<sup>6</sup>.

**Proposition 5:** (Monotonicity of Best-Responses over  $\Delta$ ):

Assume  $\alpha' \succ \alpha$ , where  $\succ$  refers to first-order stochastic-dominance. Then we have

$$\text{Min BR}(\alpha') \geq \text{Max BR}(\alpha).$$

Proof: Let  $i = \text{Max BR}(\alpha)$ . For any  $j < i$ , we have

$$\begin{aligned} u(i, \alpha') - u(j, \alpha') &= \sum_{k=1}^n \alpha'_k [u_{ik} - u_{jk}] \\ &> \sum_{k=1}^n \alpha_k [u_{ik} - u_{jk}] \\ &= u(i, \alpha) - u(j, \alpha) \geq 0, \end{aligned}$$

where the first inequality follows from the presumed stochastic dominance of  $\alpha'$  over  $\alpha$  (see footnote 5) and from the strict supermodularity assumption; the second inequality follows from the optimality of  $i$  relative to  $\alpha$ . Hence, Any element in  $\text{BR}(\alpha')$  must be no smaller than  $i$ , establishing the claim. ■

### V.A.2. Pure Coordination Games

Pure coordination games have positive payoffs along the main diagonal and zero payoffs elsewhere,  $u_{ij} = 0$  for  $i \neq j$ . For convenience, let us order the strategies in such a way that  $u_{jj} > u_{ii} > 0$  for  $j > i$ . Clearly,

---

<sup>5</sup>The definition is as follows. Let  $\alpha = (\alpha_1, \dots, \alpha_n)$  and  $\alpha' = (\alpha'_1, \dots, \alpha'_n)$  be two probability distributions over  $S$ . We say that  $\alpha'$  stochastically dominates  $\alpha$  if the cumulative distributions of  $\alpha$  and  $\alpha'$ , denoted  $F(i) = \sum_{k \leq i} \alpha_k$  and  $G(i) = \sum_{k \leq i} \alpha'_k$ , are such that  $F(i) \geq G(i)$  for all  $i$  with strict inequality for at least one  $i$ . In this case, the expected value of any increasing function under  $\alpha'$  is larger than the expected value of the same function under  $\alpha$ .

<sup>6</sup>Independent work by Krishna (1991) contains a similar result.

each strategy constitutes a pure strategy Nash equilibrium, and the strategies are Pareto-ranked, where strategy 1 generates the least efficient equilibrium and strategy  $n$  the most efficient one. It is easy to verify that this game is not supermodular unless  $n=2$ . For example, the marginal payoff  $u_{1j} - u_{2j}$  increases when  $j$  changes from 1 to 2, but it decreases when  $j$  is further increased from 2 to 3.

Examples of pure coordination games arise in the network externality/product compatibility literature. Consider, for instance, a situation where each player is to choose one from  $n$  different communication network systems (or computer, or home videogame systems). Two players can communicate if and only if they are in the same network, and the quality of the network systems are different, providing different payoff to the players. In such a game, there are (at least) as many equilibria as the number of strategies, where all players choose the same strategy (in addition, there are mixed strategy equilibria).

### V.B. Global Convergence

We now provide global convergence results for the best-response dynamics,  $P(0)$ ; ie, we show that starting from an arbitrary initial condition, the system will converge to one of the game's pure-strategy equilibria and stay there thereafter (which equilibrium ends up being the terminal-point depends, of course, on where the dynamic process gets started). These results (especially the one for supermodular games) are of interest by their own right, since they counter recent literatures on economic dynamics which emphasize the ease with which one can generate complicated behavior. For instance, Shapley's example (1964), or the literature on chaotic behavior in models with overlapping generations exhibit such behavior. This situation does not occur in the games we consider here.

#### V.B.1. Supermodular Games

First we show that in symmetric strict supermodular games, the states which mimic mixed strategy equilibria are unstable in a strong sense: they are not even stationary points of the Darwinian dynamics.

**Proposition 6** : Suppose all players are taking best responses under  $z$ , that is,  $i \in \beta_i(z)$  for all  $i \in C(z)$ . In symmetric strict supermodular games this implies that only one strategy is played under  $z$ .

**Proof** : Suppose to the contrary that we have  $i < j$ ,  $i, j \in C(z)$  and  $i \in \beta_i(z)$ ,  $j \in \beta_j(z)$ . Since  $i < j$ , the strategy

distribution facing a player with strategy  $i$  stochastically dominates the one for a player with strategy  $j$ :  $\alpha(z_i) \prec \alpha(z_j)$ . Then, Proposition 5 implies that  $\text{Min } \beta_i(z) \geq \text{Max } \beta_j(z)$ , a contradiction. ■

Now we are ready to show the global convergence result.

**Theorem 2:** For symmetric strict supermodular games, the set of limit sets is in one-to-one correspondence with the collection of all pure strategy Nash equilibria. Therefore, for any initial state  $z^0 \in Z$ , the system under the Darwinian dynamics converges to one of the pure strategy Nash states with probability one.

**Remarks:** Similar global convergence results for supermodular games are also found in Milgrom and Roberts (1991) and Krishna (1991). The case of  $2 \times 2$  games was treated earlier by Miyasawa (1961).

Milgrom and Roberts consider supermodular games with a unique Nash equilibrium, while we do not place such a restriction. When a supermodular game has a unique Nash equilibrium, the game is dominant-solvable (see Proposition 4), so a wide class of adjustment processes naturally converge to the unique Nash equilibrium. Krishna, on the other hand, obtains the global convergence result for supermodular games with multiple equilibria, but considers somewhat different adjustment processes, namely, the fictitious play dynamics. The basic assumption is that each player takes the best response against the empirical distribution of strategies in the entire past history. Since the model is more complicated, Krishna employs an additional assumption that a player's payoff function is concave with respect to her own strategy. We do not need such an assumption in Theorem 2.

In similar vein, Monderer and Shapley (1991) analyze fictitious play dynamics, and prove a global convergence result for the class of potential games. Their results are not applicable here because there are examples showing that supermodular games are not a special case of potential games, or vice versa (and also because their results pertain to fictitious play dynamics, not to best-response dynamics as we analyze here)<sup>7</sup>. See Sella (1992).

---

<sup>7</sup>Even for games that belong to the join one would still have to establish (i) that a pure strategy supermodular games possess an ordinal potential, and (ii) that the linear extension of such potential is still a potential for the mixed extension of the game. To the best of our knowledge this is still an open problem at this time (what has to be shown is convergence for all better response paths, not just the best-response one, as shown in theorem 2).



Proof of Theorem 2: By Assumption A (ii) on the nature of the Darwinian dynamics, a singleton set is a limit set if and only if all players are taking best responses. Proposition 6 shows that those states correspond to pure strategy Nash equilibria. Next we will show that starting from any other state, the process converges to a pure strategy Nash equilibrium with some positive probability. By the definition of limit sets, this is sufficient to show that there are no other limit sets.

Take a non-Nash state  $z^0$  and consider a particular trajectory which emanates from it. To define the trajectory, let us first introduce some notation. For state  $z$ , define the set of sub-optimal strategies by  $S(z) = \{i \in C(z) \mid i \notin \beta_i(z)\}$ . Now consider a trajectory  $\{z^k, k=0,1,\dots\}$  such that as long as  $S(z^k)$  is non-empty, only one player adjusts at a time. Furthermore, we require the same adjustment to occur whenever the same state is achieved. Note that this realizes with positive probability because of the independence of stochastic adjustments across players (Assumption A (i)).

Since the state space is finite, this trajectory forms a cycle (which might be a degenerated singleton set). Let  $X$  be the set of strategies whose populations change on this cycle, and let  $x = \text{Max } X$ . Suppose a player with strategy  $i$  is adjusting to strategy  $x$  at point  $z^T$  on this cycle. Then we have

$$\alpha(z^T, i) = \alpha(z^{T+1}, x) \prec \alpha(z^{T+1}, j) \text{ for all } j < x.$$

Recall that under the population configuration  $z$ , a player with strategy  $k \in C(z)$  faces strategy distribution  $\alpha(z, k)$  (see (2.2)). The above equality follows from the fact that only one player is adjusting from  $i$  to  $x$ , and the stochastic dominance follows from  $j < x$ . Given this stochastic dominance and Proposition 5, at  $z^{T+1}$ , any player with strategy  $j < x$  has the best responses which are all no smaller than  $x$ . On the other hand, a player with strategy  $x$  is already taking a best response, because a best response to  $\alpha(z^T, i) = \alpha(z^{T+1}, x)$  is  $x$ . By the definition of  $x$ , it must be that the player who moves at time  $T+1$  adjusts to  $x$ . Proceeding inductively, we conclude that the cycle is actually a singleton  $\{z^*\}$  with  $z^*_x = M$  and  $S(z^*) = \emptyset$ , which is a pure strategy Nash equilibrium with strategy  $x$ . ■

As an immediate corollary, we have the following.

**Theorem 2':** Each long run equilibrium of a supermodular game corresponds to a pure strategy Nash

equilibrium.

### V.B.2. Pure Coordination Games

Consider now the pure coordination game where  $u_{jj} > u_{ii} > 0$  for  $j > i$  and  $u_{ij} = 0$  for  $i \neq j$ . Then we have:

**Proposition 7:** (1) Any mixed strategy configuration is unstable, i.e., if  $|C(z)| > 1$  then at least for one  $i \in C(z)$ ,  $i \notin \beta_i(z)$ . (2) The collection of limit sets is  $\{\{e_1\}, \dots, \{e_n\}\}$ . Therefore, the Darwinian dynamic for a pure coordination game converges to a pure strategy Nash equilibrium with probability one.

**Proof:** (1) Assume  $i, j \in C(z)$  with  $i < j$ , and assume  $i \in \beta_i(z)$ ,  $j \in \beta_j(z)$ . Then we have:

$$\alpha_j(z, j)u_{jj} \geq \alpha_i(z, j)u_{ii} > \alpha_i(z, i)u_{ii} \geq \alpha_j(z, i)u_{jj} > \alpha_j(z, j)u_{jj}$$

(where the strict inequalities follow from the definition of  $\alpha(z, k)$ , (2.2); and the weak inequalities follow from the presumed optimality of  $s_i$  and  $s_j$  relative to  $z$ ). So this contradiction establishes that for at least one strategy,  $i \in C(z)$ ,  $i \notin \beta_i(z)$ .

(2) Note that each pure strategy configuration,  $e_i$ , satisfies (as a singleton) the definition of a limit set (definition 5). Therefore it remains to show that starting from any other state there is a positive probability path leading into one of these singleton sets. Let  $z \in Z$ . By (1) there exists an  $i \in C(z)$  so that  $i \notin \beta_i(z)$ . Let  $k \in \beta_j(z)$ , and let a player taking strategy  $i$  switch over by herself to strategy  $s_k$ , and let that happen before anybody else adjusts their behavior (this occurs with positive probability). Once that happens,  $s_k$  becomes a BR for each and every individual in the society. This is shown as follows. First, given that one player switched optimally from strategy  $i$  to strategy  $k$ , the players with strategy  $k$  must be taking best responses. Next consider a player who is not taking strategy  $k$ . Compared to the players with strategy  $k$ , this player faces a strategy distribution which puts more probability to  $k$  and no more probability to all other strategies. Given the structure of pure coordination games, this implies that strategy  $k$  is even more desirable for this player than for those who are already taking strategy  $k$ . Therefore, we conclude that  $k$  should be the best response for all players. Given this, we let the continuation path be one where everyone switches to  $s_k$ . That, too, occurs with positive

probability, so the proof is complete. ■

### V.C. Costs of Transition between Limit Sets; Optimum Branching Algorithm

We now show how to compute long-run equilibria for the above two classes of games. We start with pure coordination games since the computations are illustrated more straightforwardly for them.

#### V.C.1. Pure Coordination Games

The first thing is to compute costs of transition  $C_{ij}$  between the limit sets,  $e_i$  and  $e_j$ . Assume the society is initially clustered at  $e_i$ . Then the minimum number of mutations ( $m$ ) needed to switch it over into the basin of attraction of  $e_j$  is determined by:

$$\frac{mu_{jj}}{M-1} \geq \frac{(M-m-1)u_{ii}}{M-1}.$$

This represents an immediate jump to escape the best response region of  $i$ , and the triangular inequality (Proposition 3) guarantees that no gradual escape is less costly than this jump. Thus,

$$C_{ij} = m = \left\lceil \frac{(M-1)u_{ii}}{u_{ii} + u_{jj}} \right\rceil \quad (5.1)$$

(where  $\lceil x \rceil$  stands for the smallest integer weakly exceeding  $x$ ).

Note that we mutate individuals taking  $s_i$  into  $s_j$ . Any other mutation is useless since it cannot possibly (directly or indirectly) raise the attractiveness of  $s_j$ .

Next observe from (5.1) that, for large  $M$ ,

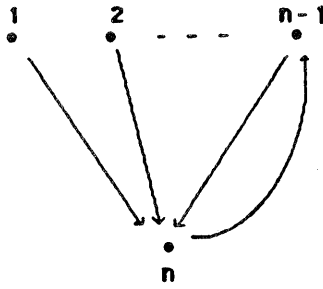
$$C_{in} < C_{ik}, \text{ for } i, k \neq n; \text{ and } C_{n, n-1} < C_{n, j}, j < n-1. \quad (5.2)$$

We need "large  $M$ " because of an integer problem. More precisely, it is easy to see that (5.2) holds if

$$\frac{u_{nn}}{(u_{nn} + u_{n-1, n-1})} - \frac{u_{n-1, n-1}}{(u_{n-1, n-1} + u_{nn})} \geq 1/(M-1).$$

Therefore, the first step of the optimum branching algorithm which establishes the most likely transition from each state, results in the following system of branches (Figure 2).

Figure 2



According to (5.2) the longest branch among those is of length  $C_{n,n-1}$ . Therefore we drop it, and are left with an  $n$ -tree. This terminates the algorithm. The conclusion in this case is that the Pareto-superior equilibrium prevails as a unique long-run state. Our results are summarized as follows:

**Theorem 3:** In a pure coordination game, the unique long run equilibrium is the Pareto efficient Nash equilibrium, if  $u_{nn}/(u_{nn} + u_{n-1,n-1}) - u_{n-1,n-1}/(u_{n-1,n-1} + u_{nn}) \geq 1/(M-1)$ .

### V.C.2. Differentiated-product Industry

We now turn to a special case of supermodular games with separable payoff functions:

$$u(s,s') = f(s) + g(s)s'.$$

The basic assumption is that the payoff is linear in the opponent's strategy. For this class of games, the best-response of a player depends only on the average strategy in the society, and this drastically simplifies the analysis (by reducing the minimum cost problem to a one-dimensional case). An example of such a game is differentiated-product oligopoly with a separable demand function:

$$Q^i = X(p^i) + Y(p^j), \quad (5.3)$$

where  $Q^i$  is the quantity sold by firm  $i$  when it charges price  $p^i$ , and its rival charges  $p^j$ .  $X(p^i)$  is a monotonically decreasing function, representing consumers' response to variations in own price.  $Y(p^j)$  is a monotonically increasing function, representing consumers' response to variations in cross price.

The strategic variable here is price, and we consider a discrete strategy set,

$$S \equiv \{p_1, \dots, p_n\},$$

Note that superscripts denote individual firms whereas subscripts denote individual strategies. Assume linear costs, and let  $c$  denote the per-unit cost of production. Then, under the random matching scenario this leads to a payoff of

$$u_{ij} = (p_i - c)[X(p_i) + Y(p_j)], \quad (5.4)$$

and to an expected payoff of

$$\pi_i(z) = (p_i - c)[X(p_i) + \bar{y}], \quad (5.5)$$

where

$$\bar{y} \equiv \frac{1}{M-1} \sum_k Y(p^k),$$

with  $k$  ranging over all remaining  $M-1$  firms in the industry. It is straightforward to verify from (5.4) that the underlying game is supermodular. Also, equation (5.5) makes it clear that the expected payoff to a firm posting price  $p_i$  depends only on the average,  $\bar{y}$ , defined above.

This embeds the differentiated-product oligopoly model into the framework of randomly matched firms. Alternatively, we can consider a more conventional formulation where all firms interact in one global market. In this case we can just write

$$Q^i(p^1, \dots, p^M) = X(p^i) + \frac{1}{M-1} \sum_{j \neq i} Y(p^j), \quad (5.6)$$

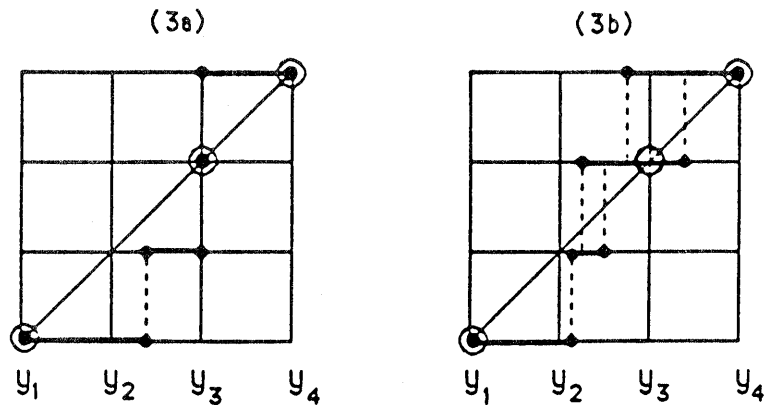
expressing the idea that the quantity sold by firm  $i$  depends on the price that each and every competitor chooses. Clearly, the expected payoff generated by (5.6) is as in (5.5); so our analysis applies to either formulation.

Given that the function  $Y(\cdot)$  is one-to-one, we can think of  $Y$  as the strategy for each player: the new strategy set is  $S' = \{y_1, \dots, y_n\}$ , where  $y_i = Y(p_i)$  and each firm chooses a  $y_i$  (instead of a  $p_i$ ).

As indicated above, an important property of this model is that a firm's best response only depends on the average strategy of the opponent. This is a consequence of the separable demand function. So let  $BR(y) \subset S'$  be the set of best responses when  $y$  is the average of the opponents' strategies. The supermodularity and Proposition 5 imply that  $BR$  is increasing in a particular sense. Figure 3 depicts two "increasing" best response

correspondences, (3a) and (3b). But Proposition 5 excludes the latter case, where the steps have some overlaps. Given that case (3b) is excluded, it follows that generically all pure strategy Nash equilibria are strict: the non-strict Nash equilibrium  $y_3$  disappears by slight perturbation of figure (3a). Note that if (3b) were the case, it would have not been possible to rule out  $y_3$  by (slightly) perturbing the best response correspondence. In what follows, we assume that all pure strategy Nash equilibria are strict.

Figure 3



Let  $N \subset S'$  be the set of Nash equilibria. We now show how to compute the cost of transition between two equilibria  $y'$  and  $y''$ , where  $y' < y''$ . To do this consider a parametrized family of correspondences,

$$B(y,a,i) = BR(ay_i + (1-a)y), \quad a \in [0,1],$$

where  $i$  is either 1 or  $n$ . The motivation of this formula is given in Remark 1 following the next Theorem.

**Theorem 4:** Let  $y'$  and  $y''$  be two Nash strategies such that  $y' < y''$ . If the population size,  $M$ , is sufficiently large, the cost of transition between  $y'$  and  $y''$  is determined as follows:

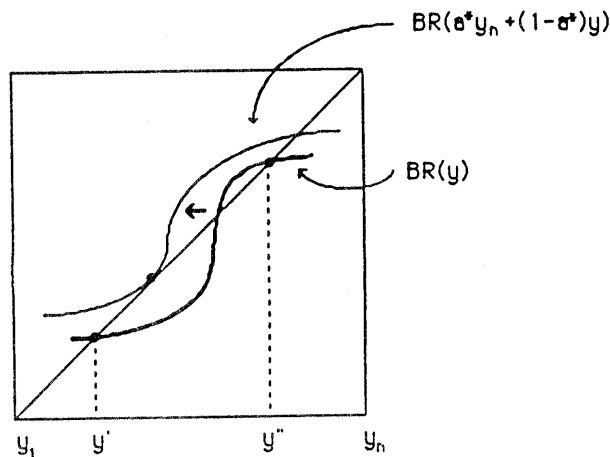
- (i)  $C(y',y'')$  is either  $k$  or  $k+1$ , where  $k$  is the minimum integer such that all elements in  $B(y,k/(M-1),n)$  are strictly greater than  $y$  for all  $y \in [y',y'']$ .
- (ii)  $C(y'',y')$  is either  $k$  or  $k+1$ , where  $k$  is the minimum integer such that all elements in  $B(y,k/(M-1),1)$  are strictly smaller than  $y$  for all  $y \in [y',y'']$ .

**Remarks:** 1. For the sake of intuition, suppose that the strategy space is a continuum  $[y_1, y_n]$  and that the best response correspondence is a function  $BR(y)$ . Consider a continuous version of the Darwinian dynamics  $dy/dt = BR(y) - y$ , and consider a transformation of this equation

$$dy/dt = BR(ay_n + (1-a)y) - y. \quad (5.7)$$

Consider Figure 4, where  $y'$  and  $y''$  are locally stable under the original dynamics. As we perturb the dynamics by increasing the parameter  $a$ , the function  $BR(ay_n + (1-a)y)$  shifts to the left in the figure. So there is a value of  $a$ ,  $a^*$ , at which  $BR(ay_n + (1-a)y)$  is just tangent to the  $45^\circ$  line between  $y'$  and  $y''$ . Once  $a$  exceeds  $a^*$ , the dynamics given by the system (5.7) drastically changes -- a path connecting  $y'$  and  $y''$  emerges. In other words, the system experiences a bifurcation at  $a = a^*$ . Aside from an integer-number problem, the cost of transition from  $y'$  to  $y''$  is proportional to  $a^*$ :  $C(y', y'') \approx a^* M$ . This is basically because the following path, which can arise with a positive probability under assumption A, turns out to realize the minimum cost transition from  $y'$  to  $y''$ . First, a fraction  $a^*$  of the total population mutates into the largest strategy  $y_n$ . Then, those "mutants" stick to  $y_n$ , while others follow their best responses, until  $y''$  becomes the best response for all individuals. Once this happens, all players (including original mutants) adjust to  $y''$ . The proof of theorem 4 formalizes this intuition.

Figure 4



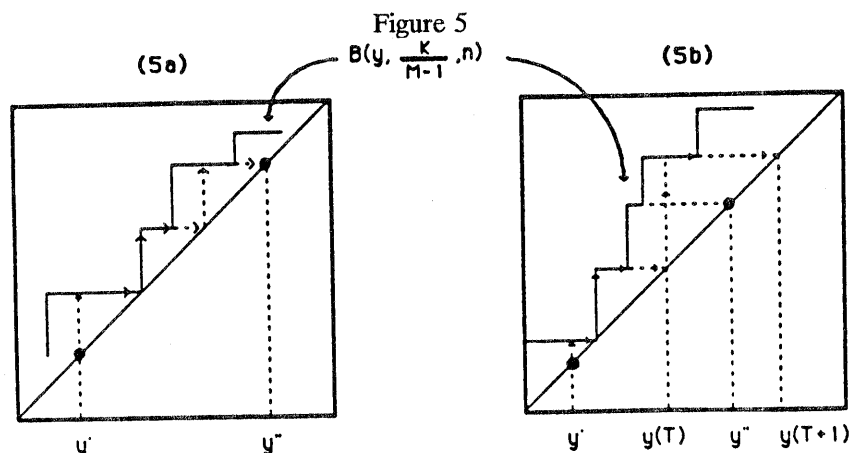
2. Theorem 4 shows that the cost of transition crucially depends on the smallest and largest available strategies. In some games, those may be naturally defined, but in other cases, there is no "obvious way" of specifying them.

For example, in the differentiated-product oligopoly case, it is not obvious how to choose a minimal and maximal price. One way to overcome this ambiguity is to assume that mutants' behavior satisfies some form of rationality. Suppose mutations are caused by the actions of uninformed newcomers, and assume further that newcomers are rational. That is, newcomers know the payoff structure of the game, but not the strategy configuration facing them. Then, they should never play any non-rationalizable strategy. If we take this point of view, we should first eliminate all non-rationalizable strategies from the model before applying our algorithm. Once we do that, Proposition 4 shows that the maximal and minimal strategies in supermodular games are pure strategy Nash equilibria. Therefore (under the assumption of rationalizable mutations),  $y_1$  and  $y_n$  in Theorem 4 are pegged as the smallest and largest pure strategy Nash equilibria.

Proof of Theorem 4: We will show (i); the proof for (ii) is identical. The proof is in two steps.

Step 1: We first show that  $y''$  is achieved from  $y'$  with  $k$  or  $k+1$  mutations. Consider the state where all players adopt strategy  $y'$ . Suppose  $k$  is the integer defined in Theorem 4(i), and consider the situation where  $k$  players simultaneously mutate to the largest strategy,  $y_n$ . Call those "mutants" and other players "nonmutants". Starting from this mixed population configuration,  $y''$  is achieved -- without any further mutations -- by the following adjustment path, which materializes with positive probability under the Darwinian adjustment. The reader should consult Figure 5 to follow the argument. On this adjustment path, except for the last step of adjustment, the mutants stick to  $y_n$ , and only the nonmutants adjust to their best responses. In particular, all nonmutants simultaneously adjust to the largest best response. This means that the strategy for the nonmutants at stage  $t$ , denoted  $y(t)$ , follows  $y(t+1) = \text{Max } B(y(t), k/(M-1), n)$ . By the definition of  $k$ , the right hand side is strictly greater than  $y(t)$  when  $y(t) \in [y', y'']$ . Hence we have two cases: In Case (i),  $y(T) = y''$  for some  $T$  (Figure 5a). In Case (ii), the process  $\{y(t)\}$  "overshoots"  $y''$ ; that is, there is  $T$  such that  $y(T+1) > y'' > y(T)$  (Figure 5b). We examine each of these cases in turn.





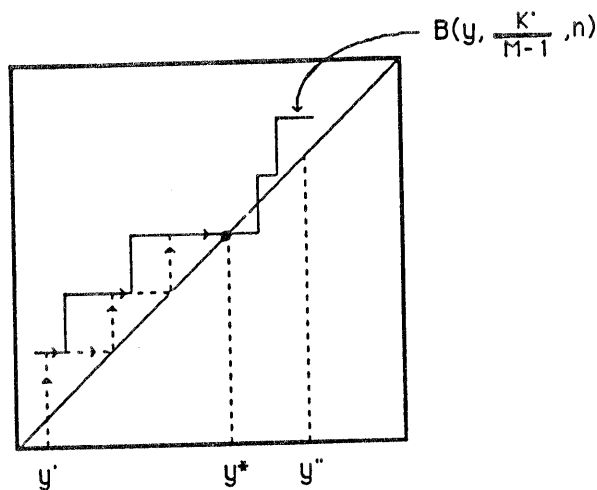
Case (i): consider the stage where nonmutants are playing  $y(T-1)$ . By construction the nonmutants' best response is  $y(T)=y''$ . If it is also the best response for the mutants, then all players switch to  $y''$ , and the adjustment is completed. Otherwise, let one nonmutant mutate into strategy  $y_n$ . Then,  $y''$  becomes a best response for the mutants, because they face distribution  $[(M-k-1)/(M-1)]y(T+1) + (k/(M-1))y_n$ , for which  $y(T)=y''$  is a best response by assumption. The nonmutants face a slightly larger average strategy after the additional mutation than before. Given that  $y''$  is a strict Nash equilibrium and  $y(T)=y''$  is the largest best response for the nonmutants, the original distribution for the nonmutants  $[(M-k-1)/(M-1)]y(T+1) + (k/(M-1))y_n$  must lie in  $[a,b]$ , where  $[a,b]$  is the non-degenerated interval of average strategy for which  $y''$  is a best response. If  $M$  is sufficiently large, the new distribution for the nonmutants lies in  $(a,b)$ , so that  $y''$  remains the best response for them. Therefore, all players can simultaneously adjust to  $y''$ , and the adjustment is completed with  $k+1$  mutations.

Case (ii): Recall that  $[a,b]$  is the range of average strategy for which  $y''$  is a best response. At stage  $T-1$ , let only a particular fraction of the nonmutants simultaneously adjust to the best response  $y(T)$ , so that the average strategy lies in  $(a,b)$  after the adjustment. This is possible when  $M$  is large, and furthermore, each player also faces an average strategy in  $(a,b)$ . So  $y''$  becomes a best response for all players, and they can simultaneously adjust to  $y''$ .

Step 2: On the other hand, if the total number of mutations are less than  $k$ , then we will prove that  $y''$  cannot be achieved from  $y'$  by the Darwinian adjustments. This is shown by contradiction. Suppose in the transition

from  $y'$  to  $y''$ , we have  $k' < k$  mutations. Since  $k' < M$ , at some stage in the transition, at least one player should be adjusting to  $y''$ . Consider the following modification of the transition path. Instead of having  $k'$  mutations over time, let  $k'$  players simultaneously mutate into the largest strategy  $y_n$  at the beginning. As before, those players are called mutants, and the other set of players are nonmutants. In the modified adjustment, the mutants do not adjust and stick to  $y_n$ . In contrast, the nonmutants simultaneously adjust to their largest best responses. This generates a process  $y(t+1) = \text{Max } B(y(t), k'/(M-1), n)$ , and  $k' < k$  implies that the nonmutants are eventually stuck with some strategy  $y^* < y''$  (Figure 6).

Figure 6



However, by Proposition 5, the strategy distribution in each step  $t$  on the new path stochastically dominates the one on the original path. This implies that the best response of any nonmutant at any stage on the original adjustment path is less than or equal to  $y^*$ . This in turn implies that the nonmutants can never adjust to  $y'' > y^*$  on the original path, which is a contradiction. ■

**Theorem 5:** The cost of transition between two Nash equilibria can be calculated by the costs of transitions between all adjacent Nash equilibria. In other words, consider  $y', y'' \in N$ , where  $y' < y''$ , and let  $y_0 = y' < y_1 < \dots < y_m = y''$  be the Nash equilibria in between  $y'$  and  $y''$ . Then, we have

$$C(y', y'') = \text{Max}_{1 \leq i \leq m} C(y_{i-1}, y_i).$$

Proof: Directly follows from Theorem 4. ■

Therefore, Theorem 5 simplifies the problem of computing costs of transition between the various limit sets: we need only compute costs between adjacent Nash equilibria.

The cost of transition  $C(y',y'')$  can be readily determined by the diagram of best response correspondence. Consider the two adjacent equilibria  $y'$  and  $y''$  in Figure 7. By definition,  $C(y',y'')$  is an increasing function of the slope  $\theta$ . More precisely, for large  $M$ ,  $C(y',y'')$  is approximated by  $aM$ , where  $\theta = 1/(1-a)$ . The reader can see this relationship in Figure 8. Similarly,  $C(y'',y')$  is an increasing function of  $\theta'$ . So those slopes represent the costs of transition, and  $C(y',y'') > C(y'',y')$  if and only if  $\theta > \theta'$ .

Figure 7

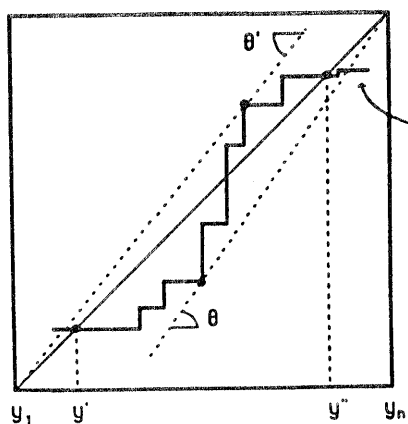
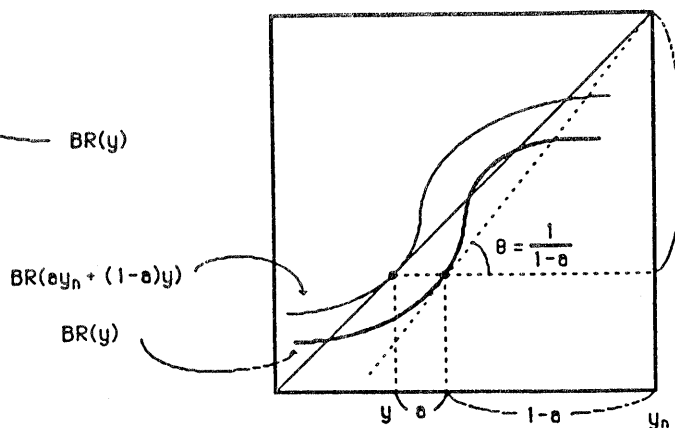
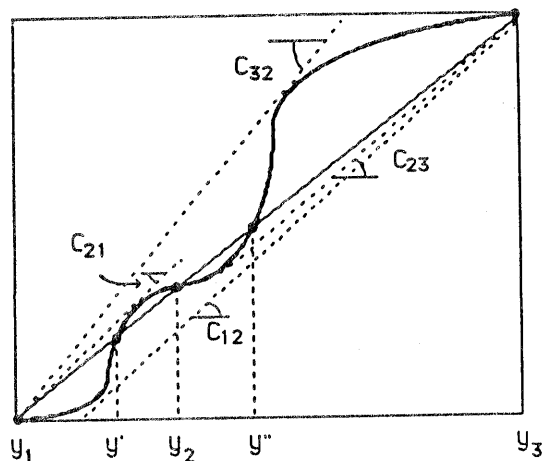


Figure 8



Now we complete the procedure to determine the long run equilibrium in an example. Consider the game in figure 9 with (a discretized version of) the following best response mapping.

Figure 9



In this game, there are three Nash equilibria,  $y_1$ ,  $y_2$ , and  $y_3$ . Assume that the "unstable" equilibria  $y'$  and  $y''$  are not pure strategy Nash equilibria in the discretized version of this model.

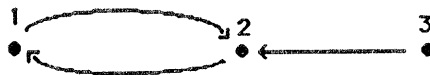
The first step in determining the long run equilibrium is to calculate the costs of transition between adjacent Nash states,  $C_{ij} \equiv C(y_i, y_j)$ , for  $j=i\pm 1$ . As noted above these costs are determined from the slopes in figure 9. Second, we compute the costs of transition between non- adjacent Nash states, using theorem 5 (ie, taking the maximum over all inbetween costs). Third, we apply the optimum branching algorithm to the resulting cost data. For the particular example we consider here (see figure 9), the cost data satisfies:

$$C_{31} = C_{32} > C_{13} = C_{12} > C_{23} > C_{21}. \quad (5.8)$$

(Note that for the three equilibria case only the ordering of the four- tuple  $C_{ij}$  for  $j=i\pm 1$  matters, not the numerical values that  $C_{ij}$ 's assume).

**Step 1:** Establish a shortest branch leaving each node. State 1 has two outgoing branches with the same cost, and we can choose either one. To provide a better illustration of the optimum branching algorithm, let us choose (1,2). This results in Figure 10.

Figure 10



**Step 2:** Drop the longest branch from the set of branches established in step 1:  $\text{Max}(C_{12}, C_{21}, C_{32}) = C_{32}$

**Step 3:** This generates a cycle  $((1,2), (2,1))$ . Determine the longest branch along the cycle:  $\max(C_{12}, C_{21}) = C_{12}$ .

Adjust the cost of outgoing branches from this cycle whose origin is different from the origin of the longest branch. This results in:

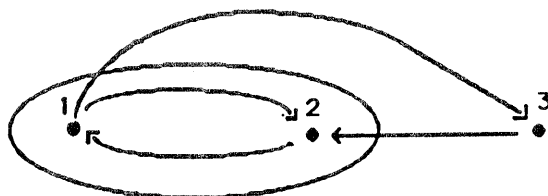
$$\tilde{C}_{23} = C_{23} + C_{12} - C_{21}$$

(The cost of branches starting at 3 remains as is.)

The states  $\{1,2\}$  are collapsed into a single state. This gives us a two-state set of nodes. We now return to step 1.

**Step 1a:**

Figure 11



(note: the branch  $\tilde{C}_{13}$  is picked over  $\tilde{C}_{23} = C_{23} + C_{12} - C_{21}$  since  $C_{13} = C_{12}$  and  $C_{23} > C_{21}$ ).

**Step 2a:** Drop  $\max(C_{32}, C_{13}) = C_{32}$

This terminates the algorithm's first phase. Now we have to open loops. The loop  $\{1,2\}$  is not a root component so it is opened according to "feasibility". This means that we eliminate a branch so that the resulting graph forms a tree. That is, we drop  $(1,2)$  and are left with Figure 12.

Figure 12



Net result: The unique long-run state is  $e_3$  (i.e., the entire society plays strategy 3). A rough intuition for this result can be gleaned from figure 9 which shows that strategy  $y_3$  is contained in a much larger basin of attraction than  $y_1$  or  $y_2$  (in the geometrical sense defined by theorem 4). Therefore, there is a much higher probability of returning to  $y_3$  following a perturbation than to any other state.

As stated above, this illustration is for the case where the 4-tuple,  $(C_{12}, C_{21}, C_{23}, C_{32})$ , satisfies (5.8). We now give a sufficient condition to determine the long-run equilibrium for separable supermodular games. This condition generalizes the idea that the equilibrium surrounded by the "highest bumps" is the long-run equilibrium. When the game is such that this condition is not satisfied, the long-run equilibrium can be computationally determined as in the illustration above.

**Theorem 6:** Consider a supermodular game with separable payoffs. If there is an  $i$  so that

$$C_{i,i-1}, C_{i,i+1} > C_{j,j-1}, C_{j,j+1} \quad (5.9)$$

for all  $j \neq i$ , then  $i$  is the unique long run equilibrium<sup>8</sup>.

**Proof:** Suppose (5.9) holds but  $k \neq i$  is a long run equilibrium. Let  $h^*$  be the optimal  $k$ - tree. Delete the outgoing branch at  $i$  from this tree, and add the branch from  $k$  to  $i$ . The resulting graph is an  $i$ - tree, denoted  $h^{**}$ . By theorem 5 and inequality (5.9), we have  $C(h^{**}) < C(h^*)$ , a contradiction. ■

---

<sup>8</sup>For  $i=1$  delete  $C_{i,i-1}$  from the left- hand side of 5.8. For  $i=n$  delete  $C_{i,i+1}$ .

## VI. Concluding Remarks

The basic motivation underlying this paper is that if one is to use the Nash equilibrium concept, one has to explain why and which Nash equilibrium is going to emerge in realistic situations. The line of attack we have taken here is learning-theoretic: each player adapts (ie, switches to a higher paying strategy) to an evolving environment, and the evolution itself is driven by similar adaptations by this player's rivals. We justify such gradual adjustments either by adjustment costs or by bounded rationality of players. We then investigate whether a Nash equilibrium eventually emerges in such a system, and provide a positive answer for important classes of games. From the perspective of full rationality, the adjustment process we specify is "ad-hoc", of-course. But the very nature of the problem we attack, that is, to explain how a Nash equilibrium emerges (without assuming that players outguess each other's actions in the first place), inevitably calls for some non-equilibrium or "ad-hoc" assumptions.

A major difference between the adaptive system we examine and the traditional ones (the Cournot-tatonnement literature: Cournot (1838), Seade (1980) and Moulin (1984); or the fictitious play literature: Miyasawa (1961), Shapley (1964), Krishna (1991), and Monderer and Shapley (1991)) is that we allow for mutations or experimentations to occur, interpreting them as imperfect information concerning which actions constitute a best response or simply as human errors. This additional element makes a big difference in that it allows us to select a unique stable behavior pattern in games with multiple equilibria. An interesting feature of this selection is that it might, for certain games, pick a cyclical behavior pattern which corresponds to none of the underlying game's Nash equilibria. So while our approach is capable of substantiating a Nash equilibrium in some games, it also points towards totally different behavior patterns in others. We hope to investigate this feature more systematically in the future.

## Appendix A

Optimum Branching Algorithm

We present an algorithm to find the long run equilibria. Let us first describe the problem again. We have a finite state space,  $Z$ , which is interpreted as the set of nodes of directed graphs. The set of (directed) branches  $\{ (i,j) \mid i,j \in Z, i \neq j \}$  is denoted  $B(Z)$ . For branch  $(i,j)$ ,  $i$  and  $j$  are called the origin and the end of the branch. Associated to each branch  $(i,j)$  is the cost (of transition)  $c(i,j) \geq 0$ . A loop is a collection of branches  $\{(i_0,i_1),(i_1,i_2),\dots,(i_{T-1},i_T),(i_T,i_0)\}$ , where  $i_t, t=0,\dots,T$  are all distinct. For  $z \in Z$ , a z-tree  $h$  is a collection of branches such that: (i)  $z$  is not the origin of any branch in  $h$ , (ii) any other node  $z' \in Z$  has exactly one branch in  $h$  whose origin is  $z'$ , and (iii)  $h$  does not form any loop. In other words,  $h$  is a tree directed into root  $z$ . A  $z$ -tree is sometimes called spanning directed tree or (spanning) arborescence. We denote the set of all  $z$ -trees by  $H_z$ . The cost of a  $z$ -tree,  $h$ , is defined as  $c(h) = \sum_{(i,j) \in h} c(i,j)$ . The problem we have to solve is to find the  $z$ -tree with the minimum cost:

$$\mathbf{Min}_{z \in Z} \mathbf{Min}_{h \in H_z} c(h). \quad (\text{A.1})$$

The analysis in the main body of this paper shows that the roots of the optimum  $z$ -trees are the long run states.

Since the number of  $z$ -trees are finite, program (A.1) is a combinatorial optimization problem. This problem is well known in Operations Research and is called Optimum Branching. We present an algorithm provided by Chu and Liu (1965) to solve this problem. Before going into the formal description, we first present the basic idea of the algorithm. First, for each node, we choose a minimum cost outgoing branch. This identifies the easiest way to escape from each state. Thereby we generate a collection of branches. Then we delete the most costly branch among them. If this forms a  $z$ -tree, clearly it is the solution to the program (A.1). Otherwise, it has some loop(s). The next thing is to "retract" each loop into a single node, and to redefine the costs of transition from/to all retracted nodes. This process is repeated until we get a tree. Then, we "open"



each retracted node to recover a z-tree defined on the original state space,  $Z$ .

The algorithm successively defines  $(Z_k, c_k, b_k)$ ,  $k=1,2,\dots$ , where  $Z_k$  is a set of nodes,  $c_k$  defines the costs of transition in  $Z_k$ , and  $b_k: B(Z_k) \rightarrow B(Z_{k-1})$  is a function which keeps track of the identities of branches. The algorithm starts with  $k = 1$ , setting  $Z_1 = Z$ ,  $c_1 = c$ , and defining  $b_1$  as the identity map.

Step 1: For each  $i \in Z_k$ , choose a minimum cost outgoing branch  $(i,j)$ :  $c_k(i,j) = \text{Min}_{z \neq i} c_k(i,z)$ . This gives a collection of  $|Z_k|$  branches. Delete one branch whose cost is the largest among them. The resulting collection of  $|Z_k|-1$  branches defined on  $Z_k$  is denoted  $f_k$ . If  $f_k$  forms a tree, then go to Step 3. Otherwise, go to Step 2.

Step 2: Directed graph  $f_k$  must have some loop(s). So let  $L_k$  be the set of loops in  $f_k$ . We create a new set of nodes by  $Z_{k+1} = X_k \cup L_k$ , where  $X_k$  is the set of nodes in  $Z_k$  which are not on any loop in  $f_k$ . Note that a loop in  $Z_k$  is treated as a single node in  $Z_{k+1}$ . We then define the costs of transition in the new space  $Z_{k+1}$  as follows. Take a branch in the new space. If both the origin and the end of this branch are single nodes in the original space, the cost remains the same:  $c_{k+1} = c_k$ . If the origin is a single node  $i \in X_k$  but the end is a loop  $\ell \in L_k$ , the cost is simply defined by the closest point on  $\ell$ :  $c_{k+1}(i,\ell) = \text{Min}_j c_k(i,j)$ , where  $j$  varies over the points on loop  $\ell$ . When the origin is a loop  $\ell$ , we need to modify the cost somewhat judiciously. Take a branch  $(x,y) \in B(Z_k)$  which is emanating from a point  $x$  on the loop  $\ell$ , and let  $z$  be the immediate successor of  $x$  on the loop (i.e.  $(x,z) \in \ell$ ) and let  $m(\ell)$  be the maximum cost of branches on the loop (i.e.  $m(\ell) = \text{Max}_{(a,b) \in \ell} c_k(a,b)$ ). Then modify the cost of transition as

$$d_k(x,y) = c_k(x,y) - c_k(x,z) + m(\ell). \quad (\text{A.2})$$

The motivation for this formula comes from the operations in Step 3 and will be explained later. The cost of transition from loop  $\ell$  to some other state is determined by taking the closest points with respect to the modified costs (A.2): If the end point is a single node  $i$ , then  $c_{k+1}(\ell,i) = \text{Min}_x d_k(x,i)$ , where  $x$  runs over all states on loop

$\ell$ . If the end point is another loop  $\ell'$ , then we have  $c_{k+1}(\ell, \ell') = \text{Min}_{x,y} d_k(x,y)$  where  $x$  and  $y$  vary over the states on  $\ell$  and  $\ell'$  respectively. Finally, given a branch  $(i,j)$  in the new space  $Z_{k+1}$ , let  $b_{k+1}(i,j) = (i,j)$  if both  $i$  and  $j$  are single nodes in the original space  $Z_k$ . Otherwise, let  $b_{k+1}(i,j)$  be the branch in the original space which attains  $c_{k+1}(i,j)$ . Thus we have defined  $(Z_{k+1}, c_{k+1}, b_{k+1})$ , and we go back to Step 1 with  $k$  increased by one (and with the cardinality of the set of nodes strictly decreased  $|Z_{k+1}| < |Z_k|$ ).

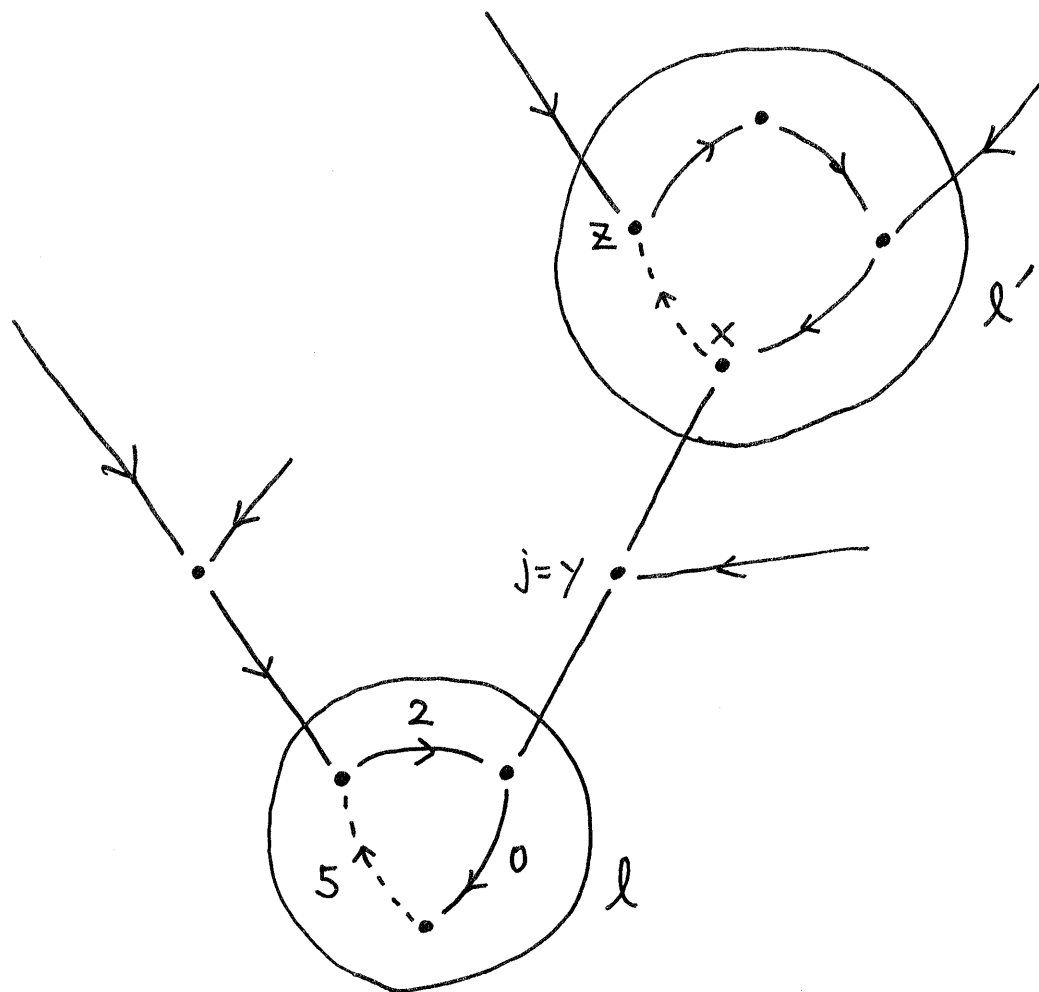
Step 3: Now we have a tree  $f_{k^*}$  defined on  $Z_{k^*}$  after  $k^*$  iteration of Step 1 and Step 2. If  $k^* = 1$ ,  $f_{k^*}$  is the optimum tree. Otherwise, we will recover a tree in the original state space  $Z$  by "opening" the retracted nodes. Let  $g_{k^*} = f_{k^*}$ . We will now define  $g_{k^*-1}, \dots, g_2, g_1$  successively. For each  $k$ ,  $g_k$  is going to be a tree on  $Z_k$ , and  $g_1$  is an optimum tree.

Given  $g_k$ , let  $g_{k-1}$  be defined in the following way. First, all branches in  $g_k$ , properly identified as branches on  $Z_{k-1}$ , belong to  $g_{k-1}$ :  $\{b_k(i,j) \mid (i,j) \in g_k\} \subset g_{k-1}$ . Next, we open the retracted nodes in  $Z_k$ . The reader is advised to consult Figure A.1 below. If a retracted node  $\ell \in L_{k-1}$  (a loop in  $Z_{k-1}$ ) is the root of  $g_k$ , then we delete a branch with the maximum cost from loop  $\ell$  and put all other branches into  $g_{k-1}$ . If loop  $\ell' \in L_{k-1}$  is not the root of  $g_k$ , we eliminate a particular branch from the loop so that the resulting graph forms a tree. In particular, in this case there must be a unique outgoing branch  $(\ell', j)$  in graph  $g_k$ . Let  $(x,y) = b_k(\ell', j)$ , and let  $z$  be the immediate successor of  $x$  in loop  $\ell'$ . Delete  $(x,z)$  from the loop and put all other branches on  $\ell'$  into  $g_{k-1}$ . After opening all the loops in this manner, we get a tree  $g_{k-1}$  defined on  $Z_{k-1}$ . Repeat this process until we get  $g_1$ , which is an optimum tree.

Figure A.1 explains how to open loops, and the meaning of the modification of costs (A.2) becomes clear now. If we choose to add an outgoing branch from loop  $\ell'$ , emanating from point  $x$  on the loop and leading to  $y$ , we incur the cost of transition from  $x$  to  $y$ ,  $c_k(x,y)$ . This is the first term in formula (A.2). Now we have two outgoing branches from  $x$ , one is the new one  $(x,y)$  and the other is the original one  $(x,z)$  on loop  $\ell'$ . Since each node (except the root) in a tree should have exactly one outgoing branch, we eliminate the original branch  $(x,z)$  on the loop in Step 3 so that the resulting graph forms a tree. Therefore, we save the cost

of the original outgoing branch on the loop  $c_k(x,z)$ , which is the second term in formula (A.2). The third term  $m(\ell)$  is the opportunity cost of excluding the possibility that loop  $\ell$  is the root. Note that if we add outgoing branch  $(x,y)$  to loop  $\ell$ ,  $\ell$  cannot be the root. On the other hand, if we added no outgoing branch to loop  $\ell$ , it would become the root component, and according to Step 3, a branch with the maximum cost  $m(\ell)$  would be eliminated from  $\ell$ . Hence we could save  $m(\ell)$  by not adding the outgoing branch  $(x,y)$ , and this explains the presence of the third term. For an elegant proof of the optimality of the algorithm, see Chu and Liu (1965).

Figure A.1



The dotted arrows are to be eliminated. Retracted node  $l$  is the root component and the numbers besides arrows indicate the costs of branches.

## Appendix B

Proof of proposition 4

(i) Since no strategy is strictly dominated, for any pair,  $1 \leq i \neq k \leq n$ , the vector

$$(u_{kj} - u_{ij})_{j=1}^n \quad (1.4)$$

contains components with different signs. Given supermodularity, this is equivalent to: For any pair  $1 \leq i < k \leq n$ , the vector (1.4) changes sign exactly once, starting out nonpositive at  $j = 1$ , and ending up nonnegative at  $j = n$ . Letting  $i=1$  and letting  $k$  vary over the range  $2 \leq k \leq n$ , we get from (1.4) above:

$$u_{k1} - u_{11} \leq 0, \quad 2 \leq k \leq n.$$

Thus,  $(s_1, s_1)$  is a NE. Similarly, letting  $k=n$  and letting  $i$  vary over  $1 \leq i \leq n-1$  establishes that  $(s_n, s_n)$  is a strict NE as well.

(ii) Assume that  $(s_i, s_j)$  is a NE for  $i \neq j$ .

$$\text{Then} \quad u_{ij} \geq u_{jj}, \quad (N.1)$$

$$\text{and} \quad u_{ji} \geq u_{ii}. \quad (N.2)$$

Assume, without loss of generality, that  $j > i$ . Then  $u_{jk} - u_{ik}$  increases in  $k$  by the supermodularity.

In particular,

$$u_{ji} - u_{ii} < u_{jj} - u_{ij}$$

But this contradicts (N.1) and (N.2) above.

(iii) If not, then  $s_\ell > s_m$ . So by the supermodularity

$$u_{\ell i} - u_{mi} < u_{\ell j} - u_{mj} \quad (S.1)$$

On the other hand,  $s_\ell$  is a BR to  $s_i$ . So

$$u_{\ell i} - u_{mi} \geq 0,$$

whereas  $s_m$  is a BR to  $s_j$ . So

$$u_{\ell j} - u_{mj} \leq 0$$

But the last two inequalities contradict (S.1). So the proof is complete. ■

## REFERENCES

- Canning, D., (1991) "Average Behavior in Learning Models," Mimeo. Pembroke College, Cambridge.
- Chu, Y.J. and T.H. Liu, (1965) "On the Shortest Arborescence of a Directed Graph", *Scientia Sinica*, 14, pp.1396-1400.
- Cournot, A., (1838) Recherches sur les Principes Mathématiques de la Théorie des Richesses, translated by N.T. Bacon as *Researches into the Mathematical Principles of the Theory of Wealth*. London: Hafner, 1960.
- Diamond, P., (1982) "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy*, 90,881-894.
- Dijkstra, E.W. (1959) "A Note on Two Problems in Connexion with Graphs," *Numerische Mathematik* 1, 269-71.
- Edmonds, J., (1967) "Optimum Branchings", *Journal of Research of The National Bureau of Standards*, 71B, 233-240.
- Ellison, G., (1991) "Learning, Local Interaction, and Coordination," Mimeo. MIT.
- Foster, D. and Young, P. (1990) "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology*, 38, 219-232.
- Freidlin, M.I. and Wentzell A.D. (1984) Random Perturbations of Dynamical Systems, New-York: Springer Verlag.
- Fudenberg, D. and Harris C. (1991) "Evolutionary Dynamics with Aggregate Shocks," Mimeo. MIT.
- Gilboa, Y., and Matsui, A. (1991) "Social Stability and Equilibrium," *Econometrica* 59, 859-867.
- Kandori, M. (1991) "Independent Mutations and Long Run Equilibria in Games: A Continuous Time Formulation," Mimeo. Princeton University.
- Kandori, M., Mailath, G. and Rob, R. (1991) "Learning, Mutation and Long Equilibria in Games", forthcoming in *Econometrica*.
- Kohlberg E. and J.F. Mertens (1986) "On the Strategic Stability of Equilibria", *Econometrica*, 54, 1003-1038.
- Krishna, V.J. (1991) "Learning in Games with Strategic Complementarities," Mimeo. Harvard Business School.
- Loeve, M. (1963) Probability Theory, New-York: Springer-Verlag.
- Matsui, A., and K. Matsuyama (1991) "An Approach to Equilibrium Selection," Mimeo. Hoover Institution, Stanford.
- Matsui, A., and Rob, R. "Evolution, Rationality and Equilibrium Selection in Societal Games", Mimeo, University of Pennsylvania.
- Maynard Smith, J., Price, G.R. (1973) "The Logic of Animal Conflict." *Nature*, London 246, 15-18.

Milgrom, P. and Roberts J. (1990) "Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities" *Econometrica* 58, 1255-1278.

---

(1991) "Adaptive and Sophisticated Learning in Repeated Normal Form Games," *Games and Economic Behavior*, 3, 82-100.

Miyasawa, K. (1961) "On the Convergence of the Learning Process in a 2x2 Non-zero-sum game," *Economic Research Program, Research Memorandum No. 33, Princeton University.*

Monderer, D., and L. Shapley (1991) "Potential Games", Mimeo UCLA and Technion, Haifa.

Mortensen, D.T. (1982) "Property Rights and Efficiency in Mating, Racing and Related games," *American Economic Review* 72, 968-979.

Moulin, H. (1984) "Dominance Solvability and Cournot Stability," *Mathematical Social Sciences*, 83-102.

Seade, J. (1980) "The Stability of Cournot Revisited", *Journal of Economic Theory*, 23, 15-27.

Sella, A. (1992), "Learning Processes in Games", M.Sc. thesis, The Technion, Haifa, Israel.

Shapley, L (1964) "Some Topics in Two-Person Games," in *Advances in Game Theory, Annals of Mathematical Studies*, 5, 1-28.

Topkis, D. M. (1979) "Equilibrium Points in Nonzero-Sum n-Person Submodular Games," *SIAM Journal of Control and Optimization*, 17, 773-787.

Vives, X.(1990) "Nash Equilibrium with Strategic Complementarities" *Journal of Mathematical Economics*, 19, 305-321.

Young, P. (1991) "The Evolution of Conventions", Mimeo. University of Maryland.