Journal of Economic Theory **85**, 1–23 (1999)
Article ID jeth.1998.2485, available online at http://www.idealibrary.com on IDEAL®

# Perfect Foresight and Equilibrium Selection in Symmetric Potential Games*

Josef Hofbauer

*Department of Mathematics, University of Vienna, Strudlhofgasse 4, A-1090 Vienna, Austria*
Josef.Hofbauer@univie.ac.at

and

Gerhard Sorger

*Department of Economics, University of Vienna, Brünnerstrasse 72, A-1210 Vienna, Austria*
sorger@econ.bwl.univie.ac.at

Received December 31, 1997; revised October 9, 1998

The equilibrium selection approach of Matsui and Matsuyama (*J. Econ. Theory* **65** (1995), 415–434) which is based on rational players who maximize their discounted future payoff, is analyzed for symmetric two-player games with a potential function. It is shown that the maximizer of the potential function is the unique state that is absorbing and globally accessible for small discount rates. *Journal of Economic Literature* Classification Number: C72. © 1999 Academic Press

## 1. INTRODUCTION

Many game theoretic models of economic and social situations suffer from multiplicity of Nash equilibria. Equilibrium refinements and methods of equilibrium selection are therefore important to obtain less ambiguous predictions from these models. The most prominent and complete theory of equilibrium selection is due to Harsanyi and Selten [6], who define the concept of risk-dominance based on the tracing procedure. Another notable "classical" approach is due to Carlsson and van Damme [4]. More recently, evolutionary/dynamic models have been used for equilibrium selection. The most prominent among those is a stochastic model due to Kandori *et al.* [11] and Young [23], which associates to a large class of

* We thank Carlos Alós-Ferrer and an anonymous referee for useful comments on the paper. Josef Hofbauer thanks the Austrian FWF for sponsoring project P10717-MAT and Gerhard Sorger thanks the Austrian FWF for support under grant SFB F010.



0022-0531/99 $30.00

strategic games (including "weakly acyclic" games as defined in [23]) a unique "long run equilibrium" or "stochastically stable equilibrium." A different approach, based on a deterministic spatio-temporal model, can be found in [8].

In this paper we shall be concerned with two other methods of equilibrium selection. The first one, originally proposed by Matsui and Matsuyama in [10] for $2 \times 2$ matrix games, uses an explicit dynamic context (modeling a population of rational players endowed with perfect foresight) to select those equilibria which are globally accessible and absorbing with respect to the dynamics.[1] The other method is applicable to so-called potential games (see [14]) and selects those equilibria which maximize the potential function. Whereas the maximization of the potential function incorporates some form of collective rationality, the dynamic process described in [10] is based exclusively on individual rationality. Nevertheless, we demonstrate in the present paper that these two selection criteria are equivalent for symmetric two-person potential games. In other words, a (possibly mixed strategy) symmetric Nash equilibrium of a potential game is both globally accessible and absorbing for all small values of the discount rate if and only if the strategy maximizes the potential function.

Admittedly, potential games form a rather narrow class of strategic games. Nevertheless they contain many classes of games which are of particular interest in economics such as pure coordination games, which have been used as simple models for the evolution of conventions (see [23]). Moreover, many games such as symmetric binary choice games or the Cournot oligopoly lead to a potential game after a linear transformation of utilities (see [14]). Potential games are also known as partnership games ([9]), games with identical interests ([15]), team games ([20]), or doubly symmetric games ([22]).

The global maximization of the potential as a criterion for equilibrium selection seems to have been proposed in print only recently by Monderer and Shapley [14]. Interestingly, they point out that this selection is supported by the experimental results in Van Huyck *et al.* [21]. At first glance this way of selecting equilibria looks extremely natural because all players have the same payoff and, hence, there is no antagonism when they individually try to maximize their expected payoff. A maximizer of the potential, provided it is unique, may be considered as the natural focal point of the game. Nevertheless, the method introduced by Matsui and Matsuyama [10] seems to be the only evolutionary dynamic method of equilibrium selection so far which is based on individual rationality and which generally selects the maximizer of the potential in a potential game.

[1] Exact definitions of these terms will be provided in Section 2.

Indeed, the method of selecting the long run equilibria described in [11] and [12] is not consistent with the maximization of the potential, if there are more than two strict equilibria or more than two players (see Example 2 in Section 2.3). Neither is this the case for the risk-dominance concept of Harsanyi and Selten [6], as has been discussed in detail in [20].

The proofs of our main results rest heavily on a close relation between perfect foresight equilibrium paths (as defined in [10]) and optimal paths in an associated optimal control problem. More specifically, global accessibility of a Nash equilibrium which maximizes the potential function is shown by a turnpike theorem for the optimal control problem. The proof that such a Nash equilibrium is absorbing is based on the Hamiltonian structure of the equilibrium conditions. The latter is, of course, a consequence of the close relation between perfect foresight equilibrium paths and optimal solutions of a dynamic optimization problem.

The paper is organized as follows. The following section specifies the class of games under consideration, discusses the dynamic equilibrium selection method from [10], and states our main result. Two examples are given to illustrate this result. Section 3 introduces the associated optimal control problem and shows that every optimal solution of that problem corresponds to a perfect foresight equilibrium path in the context of [10]. Section 4 studies accessibility of Nash equilibria while Section 5 discusses absorbing states. These two sections together contain the proof of our main result. Section 6 presents concluding remarks.

## 2. DEFINITIONS AND MAIN RESULTS

We denote by $\mathbb{R}^n$ the $n$-dimensional real space and by $e_i$, $i \in \{1, 2, ..., n\}$, the $i$th unit vector in $\mathbb{R}^n$. All vectors in this paper are interpreted as column vectors unless they are explicitly written as row vectors. If $x$ is any vector we denote by $x'$ its transpose. By $\Delta^{n-1}$ we denote the $(n-1)$-dimensional simplex in $\mathbb{R}^n$, that is, $\Delta^{n-1} = \{(x_1, x_2, ..., x_n)' \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 1,\ x_i \geqslant 0 \text{ for all } i\}$. For $\bar{x} \in \Delta^{n-1}$ and $\varepsilon > 0$ we denote by $B_\varepsilon(\bar{x})$ the $\varepsilon$-neighborhood of $\bar{x}$ relative to $\Delta^{n-1}$, i.e., $B_\varepsilon(\bar{x}) = \{x \in \Delta^{n-1} \mid \|x - \bar{x}\| < \varepsilon\}$.

### 2.1. *Potential Games*

We consider finite symmetric two-player games in which each player has $n \geqslant 2$ pure strategies. The payoff matrix of such a game will be denoted by $A = (a_{ij}) \in \mathbb{R}^{n \times n}$, where $a_{ij}$ is the payoff received by a player using the pure strategy $i$ against an opponent playing the pure strategy $j$. The pure strategy $i \in \{1, 2, ..., n\}$ is identified with $e_i$, the $i$th vertex of the simplex $\Delta^{n-1}$. An arbitrary element $x = (x_1, x_2, ..., x_n)' \in \Delta^{n-1}$ corresponds to a mixed strategy which assigns the probability $x_i$ to the pure strategy $e_i$.

A pure strategy $e_i$ is a *symmetric* (*pure strategy*) *Nash equilibrium* of the symmetric matrix game determined by $A$ if and only if

$$a_{ii} = e_i' A e_i \geqslant e_j' A e_i = a_{ji} \tag{1}$$

for all $j \in \{1, 2, ..., n\}$. The equilibrium is *strict* if the strict inequality holds in (1) whenever $j \neq i$. In this paper we shall only consider symmetric Nash equilibria. A state $\bar{x} \in \Delta^{n-1}$ is a *symmetric* (*mixed strategy*) *Nash equilibrium* if and only if

$$\bar{x}' A \bar{x} \geqslant x' A \bar{x}$$

for all $x \in \Delta^{n-1}$. The *support* of a mixed strategy $x \in \Delta^{n-1}$ is denoted by $\text{supp}(x) = \{i \mid x_i > 0\}$. If $\bar{x} \in \Delta^{n-1}$ is a (mixed strategy) Nash equilibrium of the game, then each player must be indifferent between the pure strategies in $\text{supp}(\bar{x})$ and she must (weakly) prefer any pure strategy in $\text{supp}(\bar{x})$ to any pure strategy which is not contained in $\text{supp}(\bar{x})$. Formally, these conditions can be written as

$$e_i' A \bar{x} = e_j' A \bar{x} \qquad \text{for all} \quad i, j \in \text{supp}(\bar{x}) \tag{2}$$

and

$$e_i' A \bar{x} \geqslant e_j' A \bar{x} \qquad \text{for all} \quad i \in \text{supp}(\bar{x}) \quad \text{and} \quad j \notin \text{supp}(\bar{x}). \tag{3}$$

It is easy to see that conditions (2) and (3) are not only necessary for $\bar{x}$ to be a symmetric Nash equilibrium, but also sufficient.

A symmetric two-player matrix game will be called a *potential game* if the common payoff matrix $A$ for both players is symmetric. In this case we refer to the quadratic form $p_A(x) = (1/2)\, x' A x$ as the *potential function* of the game.[2]

If $A$ is a symmetric matrix, then (2) and (3) are first order necessary conditions for the potential function $p_A(x)$ to attain a local maximum over the set $\Delta^{n-1}$ at $x = \bar{x}$. Since the same conditions are necessary and sufficient for $\bar{x}$ to be a symmetric Nash equilibrium it follows that *every local maximum of the potential function $p_A$ is a symmetric Nash equilibrium of the corresponding potential game*. The converse is not true. For example, a symmetric Nash equilibrium may also be a local minimum of the potential

[2] More generally, a *rescaled potential game* is a symmetric two-player matrix game which is linearly equivalent to a potential game. The $n \times n$ games $A$ and $B$ are linearly equivalent if there exist real numbers $c_1, c_2, ..., c_n$ such that $a_{ij} = b_{ij} + c_j$ holds for all $i, j \in \{1, 2, ..., n\}$. Since the equilibrium selection method of [10], which is the topic of the present paper, is invariant with respect to linear transformations of this form, our results can be applied to all rescaled potential games. Rescaled potential games are characterized by the triangular integrability condition $a_{ij} + a_{jk} + a_{ki} = a_{ik} + a_{kj} + a_{ji}$ for all pairwise different indices $i, j, k \in \{1, 2, ..., n\}$ (see [9, p. 244]). Because this condition is trivially satisfied for all symmetric $2 \times 2$ games, our results are applicable to all symmetric $2 \times 2$ games.

function. We call a vector $\bar{x} \in \Delta^{n-1}$ which satisfies (2) a *critical point* of the potential function. Then it follows trivially that *every symmetric Nash equilibrium of a potential game is a critical point of the potential function* $p_A$. The set of critical points of $p_A$ contains the local maxima, local minima, and saddle points of $p_A$.

If $\bar{x}$ is a critical point of the potential function $p_A$ then we call $p_A(\bar{x})$ a *critical value* of $p_A$. If $x$ and $y$ are two critical points with the same support then it is easy to see that $p_A(x) = (1/2)\, x'Ax = (1/2)\, x'Ay = (1/2)\, y'Ay = p_A(y)$. Critical points with the same support correspond therefore to the same critical value. Since there are only finitely many possible supports, it follows that *the potential function can have at most finitely many different critical values* (although it can have a continuum of critical points).

For every symmetric Nash equilibrium $\bar{x}$ the value of the potential function, $p_A(\bar{x})$, is half of the expected equilibrium payoff for each player. A sensible method of equilibrium selection in potential games is therefore to select those equilibria which maximize the potential function (see [14]).

### 2.2. *Perfect Foresight Equilibrium Paths*

We are now going to discuss the dynamic equilibrium selection method proposed in [10]. This method is not only applicable to potential games but to general (not necessarily symmetric) two-player matrix games. In this subsection it is therefore not necessary to assume that $A$ is a symmetric matrix.

In the approach from [10] one assumes that the game described by the matrix $A$ is played repeatedly in a society consisting of a continuum of identical players. Time $t \in [0, \infty)$ is a continuous variable. At each point in time the players are matched randomly to form pairs, which then play the game anonymously. Players are not able to choose their strategy at every point in time. Instead, it is assumed that each player must make a commitment to a particular pure strategy for an exogenously given (random) time interval. Time instants at which a player can switch between strategies follow a Poisson process with mean arrival rate $p$. These processes are assumed to be independent across players. Without loss of generality we choose the unit of time in such a way that $p = 1$.[3]

Let us denote by $x_i(t)$ the fraction of players who are playing the pure strategy $i$ at time $t$. Of course, we must have

$$x(t) = (x_1(t), x_2(t), ..., x_n(t))' \in \Delta^{n-1} \tag{4}$$

[3] It should be noted that the story by which Matsui and Matsuyama [10] motivate their selection dynamics involves technical problems of two kinds: the random matching process for a continuum of agents and the assumption of a continuum of independent Poisson processes. We refer to [1] for a discussion and partial resolution of these problems.

for all $t \in [0, \infty)$. The vector $x(t)$ describes the strategy distribution in the society at time $t$ and will be called the *state of the society* at time $t$. Since players are matched randomly, $x(t)$ can also be thought of as the mixed strategy against which each player plays at time $t$. It follows that the expected payoff of playing the pure strategy $i$ at time $s$ is given by $e_i' Ax(s)$. It is assumed that all players have perfect foresight so that they correctly anticipate the future evolution of the strategy distribution in the society. Since the time instants at which it is possible to switch between strategies form a Poisson process with mean arrival rate $p = 1$, the period of commitment to a fixed strategy has an exponential distribution with mean 1. Denoting the common discount rate of the players by $\theta > 0$ it follows that the expected discounted payoff of committing to strategy $i$ at time $t$ is given by

$$V_i(t) = \int_0^\infty \int_t^{t+z} e^{-\theta(s-t)} e_i' Ax(s)\, ds\, e^{-z}\, dz,$$

which can be simplified as

$$V_i(t) = \int_t^\infty e^{-(1+\theta)(s-t)} e_i' Ax(s)\, ds. \tag{5}$$

Because of the perfect foresight assumption, a rational player who has the opportunity to switch to a new strategy at time $t$ will switch to a strategy $i \in M(t)$ where

$$M(t) = \arg\max\{V_i(t) \mid i = 1, 2, ..., n\}. \tag{6}$$

Given the assumption that the switching times follow independent Poisson processes with arrival rate 1 it follows that $x_i\colon [0, \infty) \mapsto \mathbb{R}$ is Lipschitz continuous with Lipschitz constant less than or equal to 1. This implies in particular that $x_i(\cdot)$ is differentiable almost everywhere. Because of the way how agents switch between strategies it follows that, for all $t$ where $x_i(\cdot)$ is differentiable, the conditions

$$\begin{aligned} \dot{x}_i(t) &= -x_i(t) && \text{if} \quad i \notin M(t), \\ \dot{x}_i(t) &\in [-x_i(t), 1 - x_i(t)] && \text{if} \quad i \in M(t) \end{aligned} \tag{7}$$

are satisfied.[4] A Lipschitz continuous function $x\colon [0, \infty) \mapsto \mathbb{R}^n$ such that Eqs. (4)–(7) hold is called a *perfect foresight equilibrium path* for the game described by the payoff matrix $A$ and the discount rate $\theta$. The following definitions are generalizations of those presented in [10].

[4] In view of the previous footnote, this should be considered as a motivation rather than a rigorous foundation of (7). The purpose of the present paper is the mathematical analysis of system (4)–(7) and *not* its derivation.

Let $\bar{x} \in \Delta^{n-1}$ be a given state of the society and $x^0 \in \Delta^{n-1}$ a given initial state. The state $\bar{x}$ is *accessible from* $x^0$ if there exists a perfect foresight equilibrium path $x(\cdot)$ satisfying $x(0) = x^0$ and $\lim_{t \to \infty} x(t) = \bar{x}$. The state $\bar{x}$ is *locally accessible* if there exists $\varepsilon > 0$ such that $\bar{x}$ is accessible from every initial state $x^0 \in B_\varepsilon(\bar{x})$. The state $\bar{x}$ is *globally accessible* if it is accessible from every initial state $x^0 \in \Delta^{n-1}$.

A verbal interpretation of accessibility is as follows. If $\bar{x}$ is accessible from $x^0$ then there exists a belief about the future evolution of the state of the society with the following properties: (i) the belief is feasible and coincides with the true current state $x^0$ at time 0, (ii) the state of the society is believed to approach $\bar{x}$ in the long run, and (iii) if all agents share this common belief and choose their strategies optimally then the believed evolution of the society coincides with the true evolution of the society.

A state $\bar{x} \in \Delta^{n-1}$ is called *absorbing* if there exists $\varepsilon > 0$ such that for all initial states $x^0 \in B_\varepsilon(\bar{x})$ the following is true: if $x(\cdot)$ is a perfect foresight equilibrium path such that $x(0) = x^0$ then it holds that $\lim_{t \to \infty} x(t) = \bar{x}$. The state $\bar{x}$ is called *fragile* if it is not absorbing.

Intuitively, the state $\bar{x}$ is fragile if there exist initial states arbitrarily close to $\bar{x}$ together with a feasible belief about the future evolution of the society such that the following is true: the actual strategy distribution in the society (provided that all agents choose their actions optimally given the belief) coincides with the believed distribution at all dates and does not converge to the state $\bar{x}$. If this condition is not true then $\bar{x}$ is absorbing. In this case every feasible and consistent belief about the future state of the society implies that the state approaches $\bar{x}$ asymptotically whenever the initial state is sufficiently close to $\bar{x}$.

The discount rate $\theta$ can be interpreted as the degree of friction (see [10, p. 421]). The equilibrium selection criterion developed by [10] requires that a Nash equilibrium is globally accessible and the only absorbing state as the friction vanishes, that is, in the limit as $\theta$ approaches zero.

### 2.3. *Main Result and Examples*

In the previous two subsections we have described two selection criteria for symmetric Nash equilibria in two-player matrix games. Whereas the maximization of the potential function incorporates some form of collective rationality, the dynamic process described by Eqs. (4)–(7) is based exclusively on individual rationality. It is therefore not obvious that these two selection criteria are equivalent. The following theorem formalizes this equivalence (its proof is given in Sections 4 and 5 below).

THEOREM 1. *Consider a symmetric two-person potential game, and let A be the associated symmetric payoff matrix. Suppose the potential function*

*$p_A(x) = (1/2)\, x'Ax$ has the unique global maximum $\bar{x}$, i.e., $\{\bar{x}\} = \arg\max\{p_A(x) \mid x \in \Delta^{n-1}\}$. Then $\bar{x}$ is absorbing for all $\theta > 0$ and there exists $\bar{\theta} > 0$ such that $\bar{x}$ is globally accessible for all $\theta < \bar{\theta}$. No point $x \in \Delta^{n-1}$ which is different from $\bar{x}$ satisfies any of these properties.*

Hence we have a twofold characterization of the global maximum $\bar{x}$ of the potential function. Like any other state, the global maximizer $\bar{x}$ may fail to be globally accessible if the discount rate $\theta$ is large. We conjecture that every strict local maximum becomes absorbing for sufficiently large $\theta$. This is intuitively plausible, as for $\theta \to \infty$ the dynamics given by (4)–(7) approaches the best response dynamics (see [7] or [10, Footnote 11]).

We note that Theorem 1 is stated for *symmetric* games where all players are assumed to be of the same type, i.e., there is only one population of players. In contrast, a symmetric game $A$ could also be interpreted as an *asymmetric* game played between two types of agents which form two separate populations. Then the resulting game is still a potential game; the potential function is now the bilinear form $p(x, y) = x'Ay$, where $x$ and $y$ denote the strategy distributions in the two populations. Matsui and Matsuyama [10] consider absorbing and accessible states also in this setting (for $2 \times 2$ games). There is no difficulty to extend Theorem 1 to asymmetric two-person potential games with two populations: again, if $(\bar{x}, \bar{y})$ is a unique global maximizer of the potential function (in which case it must be an equilibrium in pure strategies), then it is globally accessible (for small $\theta$) and absorbing.

In the remainder of this section we discuss two examples which illustrate our results.

EXAMPLE 1. Consider the case of symmetric $2 \times 2$ games as discussed in [10]. Every such game is linearly equivalent (see Footnote 2) to a potential game with the payoff matrix

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

Assume that $a$ and $b$ are strictly positive numbers. In this case both pure strategies are strict Nash equilibria. The potential function is given by $p_A(x_1, x_2) = (ax_1^2 + bx_2^2)/2$. It attains local maxima at the pure strategies $e_1$ and $e_2$ and a global minimum at the mixed strategy $x^* = (x_1^*, x_2^*)' = (b/(a+b), a/(a+b))'$. It is easy to verify that $x(t) = (x_1(t), 1 - x_1(t))'$ with $x_1(t) = x_1(0)\, e^{-t}$ is a perfect foresight equilibrium path converging to $e_2$ whenever $0 \leqslant x_1(0) \leqslant x_1^*(2+\theta)/(1+\theta)$. An analogous argument shows that $x(t) = (1 - x_2(t), x_2(t))'$ with $x_2(t) = x_2(0)\, e^{-t}$ is a perfect foresight equilibrium path converging to $e_1$ whenever $0 \leqslant x_2(0) \leqslant x_2^*(2+\theta)/(1+\theta)$. We show in the Appendix that, together with the constant path $x(t) = x^*$, these are essentially all perfect foresight equilibrium paths for the game when $a > 0$, $b > 0$, and $\theta > 0$.[5] Note that, from all initial values in a certain neighborhood of $x^*$, at least two different perfect foresight equilibrium paths emanate. On one of them there is a discrepancy between maximizing the expected long-term payoff $V_i$ and the expected short-term payoff $e_i' Ax(t)$.

The above arguments show that both $e_1$ and $e_2$ are locally accessible. Assume now that $a > b > 0$ and, hence, $x_1^* < x_2^*$. Then $e_1$ is the unique global maximizer of the potential function. For all $\theta \in (0, (x_2^* - x_1^*)/x_1^*]$ we have $x_2^*(2+\theta)/(1+\theta) \geqslant 1$. As has been shown before, this implies that for all initial states there exists a perfect foresight equilibrium path emanating from this initial state and converging to $e_1$. Hence $e_1$ is globally accessible for small values of $\theta$. Moreover, if $x_1(0) > x_1^*(2+\theta)/(1+\theta)$ the path converging to $e_1$ is the only perfect foresight equilibrium path. This confirms Theorem 1 for this example.

Example 2. Consider the $3 \times 3$ game taken from [20] in which

$$A = \begin{pmatrix} 3 & 0 & 2 \\ 0 & 3-\varepsilon & 2 \\ 2 & 2 & 2+\varepsilon \end{pmatrix}, \tag{8}$$

where $0 < \varepsilon < 1$. This game has three strict equilibria: $e_1$, $e_2$, and $e_3$. Furthermore, it holds that $p_A(e_1) = 3/2$, $p_A(e_2) = (3-\varepsilon)/2$, and $p_A(e_3) = (2+\varepsilon)/2$ so that $p_A(e_1) > \max\{p_A(e_2), p_A(e_3)\}$. Although $e_1$ is the unique maximizer of the potential function, $e_3$ is the risk-dominant equilibrium according to the Harsanyi–Selten theory [6]. The basic reason for this is that $e_3$ is the best reply against the mixed strategy $(e_1 + e_2)/2$ and, hence, $e_1$ and $e_2$ are eliminated even though each of them (payoff- and risk-) dominates $e_3$; see [20] for details. This effect is not unintended in the Harsanyi–Selten theory: one may argue that $e_3$ is the safer option because it guarantees the maximin payoff of 2 and avoids a potential miscoordination between $e_1$ and $e_2$, which may occur for small $\varepsilon$ and which leads to a low payoff of 0. In this sense, the concept of risk-dominance introduced in [6] captures the uncertainty about the other player's rationality and tries to rationally cope with it. In our opinion, there is room for argument whether this is the right way of equilibrium selection in games for which a potential function with a unique maximizer exists. This maximizer presents itself as a natural focal point, like $e_1$ in the present example.

[5] More precisely, all perfect foresight equilibrium paths coincide eventually with one of these paths after a possible initial phase of oscillations around $x^*$. For details see the Appendix and Fig. 1.

Using the procedure explained in [23] (and [12]) one can show that $e_1$ is the stochastically stable equilibrium (and the long-run equilibrium in the sense of [11]) for the game defined in (8). However, it is easy to construct other $3 \times 3$ potential games for which the long-run equilibrium does not maximize the potential function. This is for example the case if one replaces $a_{33} = 2 + \varepsilon$ by $a_{33} = 3 - 2\varepsilon$ in (8), whereby $\varepsilon$ is sufficiently small. After this modification, the least efficient equilibrium $e_3$ is the long-run equilibrium. There is nothing wrong with this, as in the model proposed in [11] players are boundedly rational and persistently make mistakes or try suboptimal strategies with a positive probability. Therefore, they may get stuck in a suboptimal equilibrium.[6] In contrast, in the approach introduced by Matsui and Matsuyama [10] a population of rational players endowed with perfect knowledge about the game and perfect foresight about the future evolution of the society—admittedly somewhat heavy assumptions—eventually reaches a coordination upon the strategy $e_1$, because $e_1$ is the unique globally accessible and absorbing state.

## 3. THE ASSOCIATED OPTIMAL CONTROL PROBLEM

In the present section we explore an interesting relation between perfect foresight equilibrium paths in potential games and optimal paths of an infinite horizon optimal control problem. This relation will be used to prove existence of perfect foresight equilibrium paths and, in later sections, to derive results on accessibility or fragility of symmetric Nash equilibria.

The optimal control problem is defined as follows:

$$\text{maximize} \quad \int_0^\infty e^{-\theta t} p_A(x(t))\, dt \tag{9}$$

$$\text{subject to} \quad \dot{x}_i(t) = u_i(t) - x_i(t) \qquad i \in \{1, 2, ..., n\} \text{ for almost all } t \tag{10}$$

$$x_i(0) = x_i^0 \qquad i \in \{1, 2, ..., n\} \tag{11}$$

$$(u_1(t), u_2(t), ..., u_n(t))' \in \Delta^{n-1} \qquad \text{for almost all } t. \tag{12}$$

[6] Robson and Vega-Redondo [16] study a variant of the model from Kandori *et al.* [11] which is based on a true random matching process and which selects the payoff-dominant equilibrium (instead of the risk-dominant one) in certain $n \times n$ common interest games. Thus, in examples like (8) it selects the same equilibrium as the approach from [10]. Nevertheless, this coincidence is superficial. A basic difference is that the approach from [10] (as well as those from [11] and [12]) is invariant under linear equivalence whereas that from [16] is not.

Here $x^0 = (x_1^0, x_2^0, ..., x_n^0)' \in \Delta^{n-1}$ is an arbitrary initial state. The control constraint (12) implies that every solution of (10)–(11) satisfies (4). An absolutely continuous function $x\colon [0, \infty) \mapsto \mathbb{R}^N$ satisfying (10)–(12) is an *optimal solution* if it achieves the maximum in (9) among all feasible solutions to (10)–(12).

In the following theorem we show that optimal solutions to the above problem exist and how they are related to perfect foresight equilibrium paths.

THEOREM 2. *Consider the optimal control problem* (9)–(12) *and assume that $A$ is a symmetric matrix and $\theta$ a positive discount rate.*

1. *There exists an optimal solution to the problem.*
2. *Every optimal solution of the problem is a perfect foresight equilibrium path for the corresponding potential game with payoff matrix $A$.*

*Proof.* Part 1 of the theorem is an immediate consequence of the existence theorem by Baum [3] (see also [19, Theorem 3.15]).

In order to prove part 2 we have to show that every optimal solution of problem (9)–(12) satisfies (4)–(7). We have already mentioned before that (10)–(12) imply (4). It is also easy to see that (10) and (12) ensure that $\dot{x}_i(t) = u_i(t) - x_i(t) \in [-x_i(t), 1 - x_i(t)]$ holds for almost all $t \in [0, \infty)$ and all $i \in \{1, 2, ..., n\}$. To prove the theorem it is therefore sufficient to verify

$$\dot{x}_i(t) = -x_i(t) \qquad \text{for all } i \notin M(t) \text{ and almost all } t \in [0, \infty), \tag{13}$$

where $M(t)$ is defined by (5) and (6).

The current value Hamiltonian function for the optimal control problem (9)–(12) is

$$H(x, u, \lambda_0, \lambda) = (\lambda_0/2)\, x'Ax + \sum_{i=1}^{n} \lambda_i(u_i - x_i),$$

where $\lambda = (\lambda_1, \lambda_2, ..., \lambda_n)' \in \mathbb{R}^n$ is the adjoint variable and $\lambda_0$ is a constant. If $x(\cdot)$ is an optimal solution of (9)–(12) and $u(\cdot)$ the corresponding control path, then there exists a constant $\lambda_0 \in \{0, 1\}$ and an absolutely continuous adjoint function $\lambda\colon [0, \infty) \mapsto \mathbb{R}^n$ such that the following conditions hold.

C1. (10)–(12) holds,

C2. $(\lambda_0, \lambda(t)') \neq 0 \in \mathbb{R}^{n+1}$ for all $t \in [0, \infty)$,

C3. $u(t) \in \arg\max\{H(x(t), u, \lambda_0, \lambda(t)) \mid u \in \Delta^{n-1}\}$ for almost all $t \in [0, \infty)$, C4. $\dot{\lambda}_i(t) = \theta\lambda_i(t) - H_{x_i}(x(t), u(t), \lambda_0, \lambda(t))$ for $i \in \{1, 2, ..., n\}$ and almost all $t \in [0, \infty)$,

C5. $\lim_{t \to \infty} e^{-\theta t}\lambda_i(t) = 0$ for all $i \in \{1, 2, ..., n\}$.

Condition C1 is the feasibility requirement, C2 the normality condition, C3 the maximum principle, C4 the adjoint equation, and C5 is a limiting transversality condition. For a statement and discussion of these conditions we refer to Seierstad and Sydsaeter [19, pp. 244–245].[7]

Condition C3 from above implies that $u(t) = (u_1(t), u_2(t), ..., u_n(t))'$ maximizes the linear function $u \mapsto \lambda(t)' u$ over $\Delta^{n-1}$. It is obvious that the optimal solution of this linear program satisfies $u_i(t) = 0$ for all $i \notin M_\lambda(t)$, where $M_\lambda(t) = \arg\max\{\lambda_i(t) \mid i = 1, 2, ..., n\}$. Because of this property and (10), the proof of (13) is complete if one can show that $M_\lambda(t) = M(t)$ for all $t$, which is trivially the case if

$$\lambda_i(t) = V_i(t) \qquad \text{for all } i \in \{1, 2, ..., n\} \text{ and all } t \in [0, \infty). \tag{14}$$

To see that this condition holds, we first note that the adjoint equation C4 can be written as $\dot{\lambda}_i(t) = (1+\theta)\,\lambda_i(t) - \lambda_0 e_i' A x(t)$. The general solution of this linear differential equation is $\lambda_i(t) = C_i e^{(1+\theta)t} + \lambda_0 V_i(t)$ where $V_i(\cdot)$ is the bounded function defined by (5) and $C_i$ is an integration constant. Because of the boundedness of $V_i(\cdot)$ it is clear that the limiting transversality condition C5 can only hold if $C_i = 0$. Therefore we must have $\lambda_i(t) = \lambda_0 V_i(t)$ for all $i \in \{1, 2, ..., n\}$ and all $t \in [0, \infty)$. If $\lambda_0 = 0$ then it follows from this equation that $\lambda_i(t) = 0$ for all $i \in \{1, 2, ..., n\}$ and all $t \in [0, \infty)$ which contradicts condition C2. Consequently, $\lambda_0 = 1$ and the proof of (14) is complete. ∎

An immediate consequence of Theorem 2 is that there exists a perfect foresight equilibrium path for every initial state $x^0 \in \Delta^{n-1}$.

The converse of Theorem 2 does not hold: not every perfect foresight equilibrium path is an optimal solution for problem (9)–(12). To illustrate this point consider Example 1 from the previous section. There are two strict Nash equilibria $e_1$ and $e_2$ and one mixed Nash equilibrium $x^*$. As has already been mentioned, the constant path $x(t) = x^*$ is a perfect foresight equilibrium path. Since $x^*$ is the global minimum of the potential function over $\Delta^1$, the path $x(t) = x^*$ is the global minimum of the optimal control problem (9)–(12) with $x^0 = x^*$. Hence, this perfect foresight equilibrium path cannot be an optimal solution of the corresponding optimal control problem. Alternatively, consider the case $a > b$ in which $e_1$ is the unique

[7] For the problem under consideration one needs the generalization of Theorem 3.16 in [19] which is mentioned in Footnote 27 on p. 244. See also [18, Theorem 7].

global maximizer of the potential function. As we have seen before, there exist perfect foresight equilibrium paths converging to $e_2$. These paths are not optimal solutions of the associated optimal control problem.

## 4. ACCESSIBLE STATES

In this section we derive conditions under which a Nash equilibrium is globally accessible. Note that Theorem 2 allows us to prove global accessibility of a Nash equilibrium by demonstrating that, for every initial state $x^0 \in \Delta^{n-1}$, there exists an optimal solution of problem (9)–(12) which converges to the Nash equilibrium. To accomplish this we first derive a so-called visiting lemma.[8]

LEMMA 1. *Assume* $\{\bar{x}\} = \arg\max\{(1/2)\, x'Ax \mid x \in \Delta^{n-1}\}$, *that is, the function* $p_A(x) = (1/2)\, x'Ax$ *attains its unique maximum over* $\Delta^{n-1}$ *at* $\bar{x}$. *For every* $\varepsilon > 0$ *there exists* $\bar{\theta}(\varepsilon) > 0$ *such that the following is true: if* $x(\cdot)$ *is an optimal solution of problem* (9)–(12) *with* $\theta \leqslant \bar{\theta}(\varepsilon)$ *then* $\liminf_{t\to\infty} \|x(t) - \bar{x}\| \leqslant \varepsilon$.

*Proof.* The proof uses the following two facts which are easily established:

*Fact* 1. For every $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ such that for all $x \in \Delta^{n-1}$ with $\|x - \bar{x}\| > \varepsilon$ it holds that $p_A(x) < p_A(\bar{x}) - \delta(\varepsilon)$.

*Fact* 2. For every initial state $x^0$ there exists a feasible (not necessarily optimal) path $y(\cdot)$ of problem (9)–(12) satisfying $y(0) = x^0$ and $\lim_{t\to\infty} y(t) = \bar{x}$. This path $y(\cdot)$ may be chosen independently of the discount rate $\theta$.

Now assume that the lemma is not correct such that there exists an optimal solution $x(\cdot)$ of problem (9)–(12) with $\liminf_{t\to\infty} \|x(t) - \bar{x}\| > \varepsilon$. This implies that there exists $T_1(\varepsilon) \in [0, \infty)$ such that $\|x(t) - \bar{x}\| > \varepsilon$ for all $t \geqslant T_1(\varepsilon)$. Because of Fact 1 it follows that for all $t \geqslant T_1(\varepsilon)$

$$p_A(x(t)) < p_A(\bar{x}) - \delta(\varepsilon). \tag{15}$$

From Fact 2 and the continuity of the function $p_A(\cdot)$ it follows that there exists $T_2(\varepsilon) \in [0, \infty)$ such that for all $t \geqslant T_2(\varepsilon)$

$$p_A(y(t)) \geqslant p_A(\bar{x}) - \delta(\varepsilon)/2. \tag{16}$$

[8] Similar results have been used to derive turnpike theorems in optimal growth theory (see, e.g., [17]).

Let us define $T(\varepsilon)=\max\{T_1(\varepsilon), T_2(\varepsilon)\}$. Using (15) and (16) we obtain

$$\int_0^\infty e^{-\theta t} p_A(x(t))\,dt-\int_0^\infty e^{-\theta t} p_A(y(t))\,dt$$

$$=\int_0^{T(\varepsilon)} e^{-\theta t}[p_A(x(t))-p_A(y(t))]\,dt$$

$$+\int_{T(\varepsilon)}^\infty e^{-\theta t} p_A(x(t))\,dt-\int_{T(\varepsilon)}^\infty e^{-\theta t} p_A(y(t))\,dt$$

$$<\int_0^{T(\varepsilon)} e^{-\theta t}[p_A(x(t))-p_A(y(t))]\,dt$$

$$+\int_{T(\varepsilon)}^\infty e^{-\theta t}[p_A(\bar{x})-\delta(\varepsilon)]\,dt-\int_{T(\varepsilon)}^\infty e^{-\theta t}[p_A(\bar{x})-\delta(\varepsilon)/2]\,dt$$

$$=\int_0^{T(\varepsilon)} e^{-\theta t}[p_A(x(t))-p_A(y(t))]\,dt-[\delta(\varepsilon)/2]\int_{T(\varepsilon)}^\infty e^{-\theta t}\,dt$$

$$=\int_0^{T(\varepsilon)} e^{-\theta t}[p_A(x(t))-p_A(y(t))]\,dt-[\delta(\varepsilon)/(2\theta)]\,e^{-\theta T(\varepsilon)}.$$

As $\theta$ converges to 0 the first term on the last line remains bounded whereas the second term diverges to $\infty$. Therefore, the last line is negative whenever $\theta$ is sufficiently small, say, $\theta<\bar{\theta}(\varepsilon)$. This, in turn, shows that the feasible path $y(\cdot)$ attains a higher value than the path $x(\cdot)$ so that $x(\cdot)$ cannot be optimal. This contradiction proves the lemma. ∎

We can now state and prove our main result concerning global accessibility of a symmetric Nash equilibrium of a potential game.

Theorem 3. *Assume that the matrix $A$ is symmetric, i.e., that the game is a potential game. Furthermore assume that $\bar{x}$ is the unique maximizer of the potential function $p_A(x)=(1/2)\,x'Ax$ over $\Delta^{n-1}$, i.e., $\{\bar{x}\}=\arg\max\{p_A(x)\mid x\in\Delta^{n-1}\}$. Then there exists $\bar{\theta}>0$ such that $\bar{x}$ is globally accessible whenever $\theta\leqslant\bar{\theta}$.*

*Proof.* It will be shown in Theorem 4 below that $\bar{x}$ is absorbing independently of the discount rate $\theta$. This means that there exists $\varepsilon>0$ such that all perfect foresight equilibrium paths starting from an initial state in $B_\varepsilon(\bar{x})$ converge to $\bar{x}$. From the proof of Theorem 4 one can see that $\varepsilon$ can be chosen independently of $\theta$. Let us fix this value of $\varepsilon$ for the rest of the present proof and assume that $\theta\leqslant\bar{\theta}:=\bar{\theta}(\varepsilon)$ with $\bar{\theta}(\varepsilon)$ as in Lemma 1. Now consider any initial state $x^0\in\Delta^{n-1}$. From Theorem 2(1) we know that there exists an optimal solution $x(\cdot)$ starting in $x^0$, and from Theorem 2(2) it follows that it is a perfect foresight equilibrium path. Lemma 1 shows that there exists $T>0$ such that $x(T)\in B_\varepsilon(\bar{x})$. Since the truncated path $y(t)=x(t+T)$ is obviously also a perfect foresight equilibrium path it follows from our choice of $\varepsilon$ and Theorem 4 that $\lim_{t\to\infty} x(t)=\lim_{t\to\infty} y(t)=\bar{x}$. This shows that $\bar{x}$ is accessible from $x^0$ and, since $x^0$ was chosen arbitrarily, that $\bar{x}$ is globally accessible when $\theta\leqslant\bar{\theta}$. ∎

## 5. ABSORBING STATES

In this section we investigate conditions under which a symmetric Nash equilibrium is absorbing. We first show that only a Nash equilibrium which is a global maximizer of the potential function $p_A(\cdot)$ can remain absorbing as the discount rate approaches 0. This result holds also for states of the society which do not correspond to Nash equilibria.

LEMMA 2. *Assume that the matrix $A$ is symmetric, i.e., that the game is a potential game. Furthermore, let $\tilde{x}$ be a given state of the society such that $\tilde{x}$ does not maximize the potential function $p_A(\cdot)$, i.e., $\tilde{x}\notin \arg\max\{p_A(x) \mid x\in\Delta^{n-1}\}$. Then there exists $\bar{\theta}>0$ such that for all $\theta\leqslant\bar{\theta}$ the state $\tilde{x}$ is fragile.*

*Proof.* The result follows immediately from Theorem 2(2) and Lemma 1. ∎

Actually, the above proof yields a stronger result than the one that is stated in the lemma because it shows that there exists a perfect foresight equilibrium path starting in $\tilde{x}$ which does not converge to $\tilde{x}$ (the definition of fragile would only require that such a path exists from initial states *arbitrarily close* to $\tilde{x}$).

It remains to prove that a global maximizer of the potential function is absorbing for all possible discount rates. This is the main result of the present section.

THEOREM 4. *Assume that the matrix $A$ is symmetric, i.e., that the game is a potential game. Furthermore assume that $\bar{x}$ is the unique maximizer of the potential function $p_A(x)=(1/2)\,x'Ax$ over $\Delta^{n-1}$, i.e., $\{\bar{x}\}=\arg\max\{p_A(x) \mid x\in\Delta^{n-1}\}$. Then it follows that $\bar{x}$ is absorbing (independently of the discount rate).*

The remainder of this section is devoted to the proof of this result. We need a few preliminary results. Differentiating (5) with respect to $t$ we obtain

$$\dot{V}(t)=(1+\theta)\,V(t)-Ax(t). \tag{17}$$

Because of (4) one can write the state equation (7) also in the form

$$\dot{x}(t) \in X(t), \tag{18}$$

where

$$X(t) = \{u - x(t) \mid u \in \Delta^{n-1},\ u_i = 0 \text{ if } i \notin M(t)\}. \tag{19}$$

Note that $V(\cdot)$ is a continuously differentiable function and $x(\cdot)$ is Lipschitz continuous and therefore differentiable almost everywhere. Now let us define the function $H^*$: $\mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ by

$$H^*(x, V) = p_A(x) + \bar{V} - V'x,$$

where $\bar{V} = \max\{V_i \mid i \in \{1, 2, ..., n\}\}$.[9] The function $H^*$ is Lipschitz continuous and its generalized gradient in the sense of [5] is

$$\partial H^*(x, V) = \{(Ax - V, y - x)' \mid y \in \Delta^{n-1},\ y_i = 0 \text{ if } i \notin M\}, \tag{20}$$

where $M = \arg\max\{V_i \mid i \in \{1, 2, ..., n\}\}$.[10]

LEMMA 3. *Let $A$ be a symmetric matrix and let $(x(\cdot), V(\cdot))'$ be a solution of* (17)–(19) *where $M(\cdot)$ is defined by* (6). *The correspondence $t \mapsto M(t)$ is upper semicontinuous and the function $t \mapsto H^*(x(t), V(t))$ is Lipschitz continuous, non-decreasing, and satisfies*

$$(d/dt)\, H^*(x(t), V(t)) = \theta[\bar{V}(t) - V(t)'\, x(t)] \geqslant 0 \tag{21}$$

*for almost all $t \in [0, \infty)$, where $\bar{V}(t) = \max\{V_i(t) \mid i \in \{1, 2, ..., n\}\}$.*

*Proof.* Upper semicontinuity of $M(\cdot)$ is an immediate consequence of the continuity of $V(\cdot)$. Because $x(\cdot)$, $V(\cdot)$, and $H^*(\cdot, \cdot)$ are Lipschitz continuous it follows that $t \mapsto H^*(x(t), V(t))$ is Lipschitz continuous and therefore differentiable almost everywhere. From the chain rule (see, e.g., [5, Theorem 2.3.9]) and (20) we obtain

$$\begin{aligned}
&(d/dt)\, H^*(x(t), V(t)) \\
&\quad \in \{\alpha' z + \beta' \dot{V}(t) \mid (\alpha, \beta)' \in \partial H^*(x(t), V(t)),\ z \in X(t)\} \\
&\quad = \{[Ax(t) - V(t)]'\, z + [y - x(t)]'\, \dot{V}(t) \mid z \in X(t),\ y \in \Delta^{n-1},\ y_i = 0 \\
&\qquad \text{if } i \notin M(t)\}.
\end{aligned}$$

[9] Using the notation introduced in the proof of Theorem 2 one can show that $H^*(x, V) = \max\{H(x, u, 1, V) \mid u \in \Delta^{n-1}\}$. Thus, the function $H^*$ is the maximized Hamiltonian function of the optimal control problem (9)–(12).

[10] It follows from Eqs. (17)–(20) that the process $(x(\cdot), V(\cdot))'$ satisfies the modified Hamiltonian dynamical system $(\theta V(t) - \dot{V}(t), \dot{x}(t))' \in \partial H^*(x(t), V(t))$.

Substituting for $\dot{V}(t)$ from (17) and for $z \in X(t)$ from (19) we can rewrite this inclusion as

$$\begin{aligned}(d/dt)\, & H^*(x(t), V(t)) \\ & \in \{\theta V(t)' [y - x(t)] - V(t)' (u - y) + x(t)' A(u - y) \mid u, y \in \Delta^{n-1}, \\ & \quad u_i = y_i = 0 \text{ if } i \notin M(t)\}. \qquad (22)\end{aligned}$$

To complete the proof of the lemma it is sufficient to show that for almost all $t \in [0, \infty)$ the set on the right hand side of (22) consists of the single (non-negative) element specified in (21). We establish this by considering the three terms separately.

*Claim* 1. $\theta V(t)' [y - x(t)] = \theta[\bar{V}(t) - V(t)' x(t)] \geqslant 0$ for all $t \in [0, \infty)$ and all $y \in \Delta^{n-1}$ with $y_i = 0$ if $i \notin M(t)$.

Because of $y \in \Delta^{n-1}$ and $y_i = 0$ whenever $i \notin M(t)$ it follows that $V(t)' y = \bar{V}(t)$. Because $x(t) \in \Delta^{n-1}$ we must have $V(t)' x(t) \leqslant \bar{V}(t)$. The claim follows immediately from these two observations.

*Claim* 2. $V(t)' (u - y) = 0$ for all $t \in [0, \infty)$ and all $u, y \in \Delta^{n-1}$ with $u_i = y_i = 0$ if $i \notin M(t)$.

As in the proof of Claim 1 we obtain $V(t)' u = V(t)' y = \bar{V}(t)$ for all $t$. Therefore $V(t)' (u - y) = 0$ and the claim is proved.

*Claim* 3. $x(t)' A(u - y) = 0$ for almost all $t \in [0, \infty)$ and all $u, y \in \Delta^{n-1}$ with $u_i = y_i = 0$ if $i \notin M(t)$.

Since $M(\cdot)$ is upper semicontinuous and takes values in a finite set (the power set of $\{1, 2, ..., n\}$) it follows that for every $t \in [0, \infty)$ there exists $\varepsilon > 0$ such that $M(s) \subseteq M(t)$ for all $s \in (t - \varepsilon, t + \varepsilon)$. Let us call a point $t$ singular if this inclusion is strict, that is, if there exists $\varepsilon > 0$ such that $M(s) \subset M(t)$ for all $s \in (t - \varepsilon, t) \cup (t, t + \varepsilon)$. We first show that the set of singular points is at most countable so that almost all $t \in [0, \infty)$ are non-singular. To this end we define

$$S_k = \{t \in [0, \infty) \mid t \text{ is singular and } |M(t)| = k\}.$$

The set of all singular points is given by $S = \bigcup_{k=1}^{n} S_k$. Consider a sequence $(t_j)_{j=1}^{\infty}$ of singular points such that $t_j \in S_k$ for all $j$ and such that $t = \lim_{j \to \infty} t_j$ exists and satisfies $t \neq t_j$ for all $j$. If $t$ were singular then this would imply that $t \in S_l$ for some $l > k$. It follows that for every $t \in S_k$ there exists $\varepsilon > 0$ such that $(t - \varepsilon, t + \varepsilon) \cap S_k = \{t\}$ which shows that $S_k$ is at most countable. Hence, $S$ is at most countable.

To complete the proof of Claim 3 we show that $x(t)' A(u-y)=0$ for all $t \in [0, \infty) \setminus S$. If $t$ is non-singular then there exists a sequence $(t_k)_{k=1}^{\infty}$ with $\lim_{k\to\infty} t_k = t$, $t_k \neq t$ for all $k$, and $M(t_k)=M(t)$ for all $k$. This implies that for all $i, j \in M(t)$ and for all $k$ it must hold that $V_i(t_k)=V_j(t_k)$ and

$$V_i(t)=V_j(t).$$

Therefore we obtain

$$\dot{V}_i(t)=\lim_{k\to\infty}\frac{V_i(t_k)-V_i(t)}{t_k-t}=\lim_{k\to\infty}\frac{V_j(t_k)-V_j(t)}{t_k-t}=\dot{V}_j(t).$$

Using these two equations and (17) it follows that $x(t)' Ae_i = x(t)' Ae_j$ for all $t$ which are non-singular and for all $i, j \in M(t)$. It is easy to see that this implies that $x'(t)\, Au = x'(t)\, Ay$ for all $t$ which are non-singular and for all $u, y \in \Delta^{n-1}$ with $u_i = y_i = 0$ whenever $i \notin M(t)$. This completes the proof of Claim 3.

The monotonicity of $t \mapsto H^*(x(t), V(t))$ follows immediately from (22) and Claims 1–3. This completes the proof of the lemma. ∎

The above lemma allows us to use $H^*(\cdot, \cdot)$ as a Ljapunov function for the differential inclusion (17)–(18), which describes the dynamics of perfect foresight equilibrium paths. We shall also use the fact that

$$H^*(x, V) \geqslant p_A(x) \tag{23}$$

for all $x \in \Delta^{n-1}$ and all $V \in \mathbb{R}^n$. This inequality can be verified by a similar argument as Claim 1 in the proof of Lemma 3.

LEMMA 4. *Let $A$ be a symmetric matrix and let $x(\cdot)$ be a perfect foresight equilibrium path for the potential game defined by $A$. Furthermore, let $x^*$ be an accumulation point of $x(\cdot)$, i.e., $x^* = \lim_{k\to\infty} x(t_k)$ for some sequence of real numbers $(t_k)_{k=1}^{\infty}$ with $\lim_{k\to\infty} t_k = \infty$.*

1. $p_A(x^*) \geqslant p_A(x(0))$.
2. *$x^*$ is a critical point of the potential function $p_A$ on $\Delta^{n-1}$.*

*Proof.* Let $V(\cdot)$ be given by (5). Since $V(\cdot)$ is a bounded function we may assume without loss of generality that $(x^*, V^*)' = \lim_{k\to\infty}(x(t_k), V(t_k))'$ exists, where $(t_k)_{k=1}^{\infty}$ is the sequence mentioned in the lemma. Defining the functions $x^*\colon [0, \infty) \mapsto \mathbb{R}^n$ and $V^*\colon [0, \infty) \mapsto \mathbb{R}^n$ by $(x^*(t), V^*(t))' = \lim_{k\to\infty}(x(t_k+t), V(t_k+t))'$ it follows that $(x^*(\cdot), V^*(\cdot))'$ is a solution of the differential inclusion (17)–(18) through the initial state $(x^*, V^*)'$.[11] We now show that $H(x^*(t), V^*(t))$ is a constant (independent of $t$). If this were not the case then there would exist $t, s \in [0, \infty)$ such that

$$H^*(x^*(t), V^*(t)) < H^*(x^*(s), V^*(s)). \tag{24}$$

Because of Lemma 3 we must have $t < s$. Because $\lim_{k \to \infty} t_k = \infty$ we may assume without loss of generality that $t_{k+1} > t_k + (s - t)$. Using Lemma 3 again we obtain

$$\begin{aligned} H^*(x^*(t), V^*(t)) &= \lim_{k \to \infty} H^*(x(t_k + t), V(t_k + t)) \\ &= \lim_{k \to \infty} H^*(x(t_{k+1} + t), V(t_{k+1} + t)) \\ &\geqslant \lim_{k \to \infty} H^*(x(t_k + s), V(t_k + s)) \\ &= H^*(x^*(s), V^*(s)). \end{aligned}$$

Since this is a contradiction to (24) it follows that $H^*(x^*(t), V^*(t))$ is constant and therefore $(d/dt)\, H^*(x^*(t), V^*(t)) = 0$ for almost all $t \in [0, \infty)$. Because of (21) this implies

$$\bar{V}^*(t) = V^*(t)'\, x^*(t) \tag{25}$$

for almost all $t$ and, consequently, $H^*(x^*, V^*) = H^*(x^*(t), V^*(t)) = p_A(x^*)$. From this equation, (23), and Lemma 3 we get $p_A(x(0)) \leqslant H^*(x(0), V(0)) \leqslant H^*(x^*, V^*) = p_A(x^*)$ and the proof of the first assertion of the lemma is complete.

Now let $I(t) = \operatorname{supp}(x^*(t))$. Because $x^*(\cdot)$ is a solution to (18) we have $\dot{x}_i^*(t) \geqslant -x_i^*(t)$ for almost all $t \in [0, \infty)$ and therefore $x_i^*(s) \geqslant x_i^*(t)\, e^{-(s-t)}$ for all $s \in [t, \infty)$. Obviously this implies that $i \in I(s)$ whenever $i \in I(t)$ and $s > t$. The correspondence $t \mapsto I(t)$ is therefore non-decreasing.

[11] Actually, it is not clear whether the limits in the definitions of the functions $x^*(\cdot)$ and $V^*(\cdot)$ exist. To be more precise, we can define the function $t \mapsto (x^*(t), V^*(t))'$ on an increasing sequence of compact intervals as an accumulation point of the sequence of uniformly Lipschitz continuous functions $t \mapsto (x(t_k + t), V(t_k + t))'$ in the space of continuous functions with the topology induced by uniform convergence on compact intervals. According to the Arcela-Ascoli theorem, such a sequence of functions has at least one accumulation point. Because of convexity, compactness, and upper-semicontinuity of the right hand side of (17)–(18), any such limit functions is a solution of (17)–(18) through the initial state $(x^*, V^*)'$ (see, e.g., [2, Theorem 1.4.1]).

Since it takes values in a finite set (the power set of $\{1, 2, ..., n\}$) it must be piecewise constant. On each of the (finitely many) intervals on which $I(\cdot)$ is constant we have $V_i^*(t) = \bar{V}^*(t)$ for all $i \in I(t)$, since otherwise we would obtain $V^*(t)' x^*(t) < \bar{V}^*(t)$, which is a contradiction to (25). Of course, this implies $V_i^*(t) = V_j^*(t)$ and $\dot{V}_i^*(t) = \dot{V}_j^*(t)$ for all $i, j \in I(t)$. Because of (17) we find $x^*(t)' Ae_i = x^*(t)' Ae_j$ for all $i, j \in I(t)$ and all $t \in [0, \infty)$. This means that $x^*(t)$ is a critical point of $p_A$ for all $t \in [0, \infty)$. In particular, $x^* = x^*(0)$ is a critical point and the proof of the lemma is complete. ∎

*Proof of Theorem* 4. Since $\bar{x}$ is the unique maximizer of the potential function over $\Delta^{n-1}$ it follows that $\bar{x}$ is a critical point of $p_A$. Moreover, because there are only finitely many critical values (see Section 2) there exists $\varepsilon > 0$ such that $p_A(x^c) < p_A(x)$ for all $x \in B_\varepsilon(\bar{x})$ and all critical points $x^c \neq \bar{x}$. Together with Lemma 4 this implies that every perfect foresight equilibrium path $x(\cdot)$ with $x(0) \in B_\varepsilon(\bar{x})$ satisfies $\lim_{t \to \infty} x(t) = \bar{x}$. ∎

## 6. CONCLUDING REMARKS

We have shown that in the class of symmetric two-person potential games the evolutionary dynamic equilibrium selection method proposed by Matsui and Matsuyama in [10] is equivalent to the static equilibrium selection method defined by the maximization of the potential function. This result is intuitively plausible but it is not at all obvious: just recall Example 2 which clearly shows that this appealing property is not shared by the long run equilibria from Kandori *et al.* [11] or the risk-dominance concept from Harsanyi and Selten [6].

We believe that our result (i.e., that the global maximizer of the potential function is globally accessible and absorbing) holds for general $N$-person potential games. This would include the results in [10] for asymmetric $2 \times 2$ coordination games. An open question concerns the local maxima of the potential function, or Maynard Smith's [13] evolutionarily stable strategies (ESS) for more general games (without potential function).[12] We conjecture that every ESS is locally accessible (independently of the discount rate) and that every ESS is absorbing if the discount rate is sufficiently high. Indeed, it is not very difficult to verify these two properties when the ESS is a strict Nash equilibrium in pure strategies. However, for the general case of mixed strategy Nash equilibria we have not been able to prove (or disprove) these conjectures.

[12] In a symmetric potential game every strict local maximum of the potential function is an ESS and conversely (see Hofbauer and Sigmund [9, p. 228] or Weibull [22, Proposition 2.14]).

# APPENDIX

For the $2 \times 2$ game from Example 1, we can rewrite Eqs. (7) and (17) as

$$\begin{aligned} \dot{p}(t) &= I_+(v(t)) - p(t), \\ \dot{v}(t) &= (1+\theta)\, v(t) + \hat{p} - p(t), \end{aligned} \tag{26}$$

where $p(t) = x_2(t) = 1 - x_1(t)$, $v(t) = [V_2(t) - V_1(t)]/(a+b)$, $\hat{p} = x_2^* = a/(a+b)$, and $I_+(\cdot)$ is the indicator function of the set $(0, \infty)$. Figure 1 shows the phase portrait of the piecewise linear differential equation (26) for the values $a = 0.6$, $b = 0.4$, and $\theta = 0.2$ which is obtained by glueing together the two families of solution curves

$$v = \frac{C}{(1-p)^{1+\theta}} + \frac{p}{2+\theta} - \frac{\hat{p}}{1+\theta} + \frac{1}{(1+\theta)(2+\theta)} > 0$$

and

$$v = \frac{D}{p^{1+\theta}} + \frac{p}{2+\theta} - \frac{\hat{p}}{1+\theta} < 0.$$

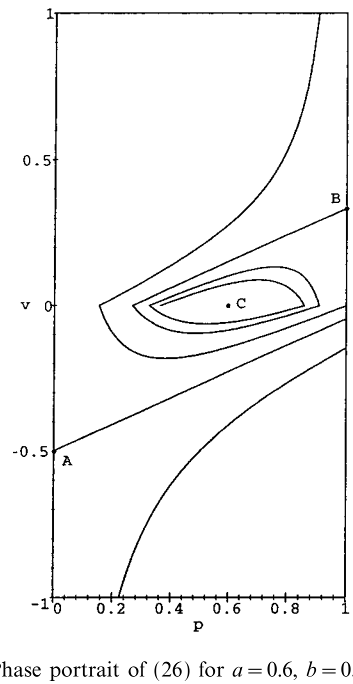


**FIG. 1.** Phase portrait of (26) for $a = 0.6$, $b = 0.4$, and $\theta = 0.2$.

Here $C$ and $D$ are integration constants. Note, that for most initial conditions, the solution is unbounded, as $v(t) \to \pm\infty$. There are only five bounded solutions which are the perfect foresight equilibrium paths. These are the three stationary solutions $(p, v) = (0, -\hat{p}/(1+\theta))$, $(p, v) = (1, (1-\hat{p})/(1+\theta))$, and $(p, v) = (\hat{p}, 0)$, as well as their stable manifolds. In Fig. 1 the stationary solutions are labeled $A$, $B$, and $C$, respectively. The stable manifold of $(0, -\hat{p}/(1+\theta))$ (point $A$) is a line segment extending across the whole $p$-axis (indicating the global accessibility of $e_1$), while the stable manifold of $(1, (1-\hat{p})/(1+\theta))$ (point $B$) winds infinitely often around the unstable steady state $(\hat{p}, 0)$ (point $C$). Hence, starting at $p = \hat{p}$, there are infinitely many perfect foresight equilibrium paths, each of which leads to $e_2$ in the long run, and there is a unique path (the optimal one constructed in Section 3), which leads to the maximizer of the potential function $e_1$.

## REFERENCES


1. C. Alós-Ferrer, "Dynamical Systems with a Continuum of Randomly Matched Agents," Working Paper 9801, Department of Economics, University of Vienna, 1998.
2. J. P. Aubin and A. Cellina, "Differential Inclusions," Springer-Verlag, Berlin, 1984.
3. D. F. Baum, Existence theorems for Lagrange control problems with unbounded time domain, *J. Opt. Theory Appl.* **19** (1976), 89–116.
4. H. Carlsson and E. van Damme, Global games and equilibrium selection, *Econometrica* **61** (1993), 989–1018.
5. F. H. Clarke, "Optimization and Nonsmooth Analysis," Wiley, New York, 1983.
6. J. C. Harsanyi and R. Selten, "A General Theory of Equilibrium Selection in Games," MIT Press, Cambridge, MA, 1988.
7. J. Hofbauer, Stability for the best response dynamics, Preprint, Department of Mathematics, University of Vienna, 1995.
8. J. Hofbauer, Equilibrium selection via travelling waves, *in* "Vienna Circle Institute Yearbook 5/97: Game Theory, Experience, Rationality. Foundations of Social Sciences, Economics, and Ethics. In Honor of John C. Harsanyi" (W. Leinfellner and E. Koehler, Eds.), pp. 245–259, Kluwer, Dordrecht–Boston–London, 1998.
9. J. Hofbauer and K. Sigmund, "The Theory of Evolution and Dynamical Systems," Cambridge Univ. Press, Cambridge, UK, 1988.
10. A. Matsui and K. Matsuyama, An approach to equilibrium selection, *J. Econ. Theory* **65** (1995), 415–434.
11. M. Kandori, G. J. Mailath, and R. Rob, Learning, mutation, and long run equilibria in games, *Econometrica* **61** (1993), 29–56.
12. M. Kandori and R. Rob, Evolution of equilibria in the long run: a general theory and applications, *J. Econ. Theory* **65** (1995), 383–414.
13. J. Maynard Smith, "Evolution and the Theory of Games," Cambridge Univ. Press, Cambridge, UK, 1982.
14. D. Monderer and L. Shapley, Potential games, *Games Econ. Behav.* **14** (1996), 124–143.
15. D. Monderer and L. Shapley, Fictitious play property for games with identical interests, *J. Econ. Theory* **68** (1996), 258–265.
16. A. J. Robson and F. Vega-Redondo, Efficient equilibrium selection in evolutionary games with random matching, *J. Econ. Theory* **70** (1996), 65–92.

17. J. A. Scheinkman, On optimal steady states of $n$-sector growth models when utility is discountered, *J. Econ. Theory* **12** (1976), 11–30.
18. A. Seierstad, "Dynamic Programming Approach to Infinite Horizon Deterministic Continuous Time Control Problems," Memorandum 15, Department of Economics, University of Oslo, 1997.
19. A. Seierstad and K. Sydsaeter, "Optimal Control Theory with Economic Applications," North-Holland, Amsterdam, 1987.
20. E. van Damme, Equilibrium selection in team games, *in* "Understanding Strategic Behavior: Essays in Honor of Reinhard Selten" (W. Güth, Ed.), Springer-Verlag, Berlin, 1996.
21. J. Van Huyck, R. Battalio, and R. Beil, Tactic coordination games, strategic uncertainty, and coordination failure, *Amer. Econ. Rev.* **80** (1990), 234–248.
22. J. W. Weibull, "Evolutionary Game Theory," MIT Press, Cambridge, MA, 1995.
23. H. P. Young, The evolution of conventions, *Econometrica* **61** (1993), 57–84.