THE NASH FIELD

Klaus RITZBERGER
Karl VOGELSBERGER

Forschungsbericht/
Research Memorandum No. 263
Februar 1990

Content

Abstract

The present paper starts with a modified definition of regular equilibria. The system of equations used to define regular equilibria induces a globally differentiable structure on the space of mixed strategy combinations. This structure is called the Nash field. The set of zeros of the Nash field which correspond to equilibria is characterized. Then the Nash field is interpreted as an evolutionary process and it is shown that stability properites of such a process correspond to self-enforcement properties of equilibria of the game. Finally, slight perturbations of the Nash field can detect stable components of Nash equilibria.

Zusammenfassung

Die vorliegende Arbeit geht von einer modifizierten Definition regulärer Gleichgewichte aus. Das Gleichungssystem, das zur Definition regulärer Gleichgewichte benutzt wird, induziert eine global differenzierbare Struktur auf den Raum der gemischten Strategiekombinationen. Diese Struktur ist ein Vektorfeld und wird "Nash-Feld" genannt. Die Menge jener Nullstellen des Nash-Feldes, die gleichzeitig Nash-Gleichgewichte bilden, wird vollständig charakterisiert. Weiters läßt sich das Nash-Feld als ein evolutionärer Prozeß verstehen und es wird gezeigt, daß Stabilitätseigenschaften eines solchen evolutionären Prozesses gleichzeitig Robustheitseigenschaften der entsprechenden Gleichgewichte darstellen. Schließlich läßt sich zeigen, daß kleine Störungen des Nash-Feldes dazu benützt werden können, stabile Komponenten von Nash-Gleichgewichten aufzufinden.

The Nash Field

KLAUS RITZBERGER AND KARL VOGELSBERGER

Institute for Advanced Studies, Vienna

December 1989

Abstract. The present paper starts with a modified definition of regular equilibria. The system of equations used to define regular equilibria induces a globally differentiable structure on the space of mixed strategy combinations. This structure is called the Nash field. The set of zeros of the Nash field which correspond to equilibria is characterized. Then the Nash field is interpreted as an evolutionary process and it is shown that stability properties of such a process correspond to self-enforcement properties of equilibria of the game. Finally, slight perturbations of the Nash field can detect stable components of Nash equilibria.

1. INTRODUCTION

A finite normal form game is usually thought of as portraying some kind of social interaction, where strategic behavior is potentially relevant. A solution concept for such an abstract representation of social interaction is a device to predict, from the definitions of the game, the behavior of rational players in such a situation. This is not quite the same thing as "understanding" the structure of the social interaction presumably modelled by the game. "Understanding" the situation may also entail information about the behavior outside the equilibrium. This observation has, especially during the last decade, lead many theorists to study refined notions of equilibria beyond the - still indispensible - solution concept of a Nash equilibrium. The basic property of a Nash equilibrium to be self-enforcing has, thereby, be found to be somewhat vague or not always "strategically stable".

Interestingly, the efforts to refine the Nash equilibrium notion have been pursued quite seperately for the normal form and for extensive form games. As a common denominator, however, both streams of research are strongly inspired by initiating examples in which common sense can identify that something undesirable happens. The mere fact that in such examples common sense can identify what is unsatisfactory about an equilibrium, shows that refinements of Nash equilibrium are drawing upon the structure of interaction within the whole game rather than

This is a first and preliminary version. All comments are welcome. We are indebted to Erwin Amann for helpful comments and discussions on numerous occasions.

exclusively processing equilibrium predictions. It is somewhat puzzling that there seems to have been little effort going into the question of how the structure of the whole game can be represented as a mathematical object. Of course, this is not generally true: Selten's notion of a subgame is an example of how the complete structure of an extensive form game can be characterized.

Also, one may get the feeling that the most widely used - at least within the economics profession - refinement concepts are extensive form concepts (e.g. subgame perfection or sequential equilibrium, [**Selten**, 1965, 1975], [**Kreps and Wilson**, 1982]). This is despite the view expressed by **von Neumann and Morgenstern** [1972, e.g. p.85] that the normal form and the extensive form are essentially equivalent and despite the fact that the normal form is mathematically somewhat more handy. It is even despite the strongly supported view, expressed by **Kohlberg and Mertens** [1986, p.1010], that the set of "strategically stable" equilibria should depend only on the reduced normal form of the game.

One reason for this may be that normal form analysis seems to be unable to capture the essence of backward induction. With respect to this problem it is known, however, that a proper equilibrium [**Myerson**, 1978] of the normal form is sequential in any tree with that normal form [**Kohlberg and Mertens**, 1986, p.1009], [**van Damme**, 1984]. This is one of the reasons, why we feel that it is worthwile to attempt a mathematical description of the structure of interaction within the complete normal form game as additional information to the predictions of the Nash equilibrium. The hope is that, once the structure of the complete game can be represented as a tractable mathematical object, the various refinement concepts will fall into place and will, sometime in the future, boil down to a unified notion of "strategic stability".

The reader may wonder what is meant by the vague phrase "structure of interaction in the whole game" and whether there is not already such a thing, induced by the best-reply correspondences, such that there remains nothing to study. In fact, the correspondence, manufactured from individual best-reply correspondences by taking the product, induces a structure on the space of mixed strategies which may well be viewed as a representation of interactions in the complete game (a programme carried through by **Kalai and Samet** [1984]). The problem with this structure is that it is not a very smooth one. Would it not be nicer to have a structure which contains the same information as the one induced by best-reply correspondences but is, moreover, differentiable? The latter is what the present paper is devoted to.

More specifically, the present paper starts out with a modified definition of a regular equilibrium. The modified definition of a regular equilibrium has the same properties as those already known for the standard definitions. But beyond these, the system of equations used to define a regular equilibrium induces a globally differentiable structure on (a neighbourhood of) the space of mixed strategy combinations. One way to think about this structure is to interpret it as a vector field and call it the "Nash field". Each equilibrium of a given game then corresponds to a zero of this vector field, and the zeros which are equilibria can be quite easily distinguished from those zeros which are not equilibria.

The advantage of the smoothness of the Nash field becomes apparent, when it is given a biological interpretation in terms of an evolutionary process. With this interpretation the rationality requirements on the players of a game can be entirely dispensed with. As a substitute for rational players interacting populations of individuals with sticky behavorial patterns can be introduced. It is shown that this alternative interpretation changes virtually nothing. Full rationality of players is, therefore, by no means an indispensable part of the solution concept of a Nash equilibrium. What is indispensable about the Nash concept are rather only two features: (i) That the choices of rival players or the composition of rival populations is taken as given, and (ii) that payoff functions indeed describe the incentives of players or the reproductive success of behavorial patterns.

An extensive part of the paper is devoted to demonstrating that the smooth Nash field and the best-reply correspondences indeed contain the same information. In particular it is shown that there is really an equivalence between stability and the property of a Nash equilibrium to be self-enforcing also with respect to a neighbourhood of the equilibrium.

The smoothness of the Nash field also allows some new insights into the structure of the graph of the Nash equilibrium correspondence. In particular neither a whole family of strategy perturbations, nor perturbations of the entire payoff vector of the game are necessary to identify stable components of Nash equilibria (by a stable component is meant a closed set of equilibria, such that any close game has an equilibrium close to the set). A simple (one parameter) perturbation of the Nash field will suffice. From a practical perspective this makes it a lot simpler to identify the Stable Set [**Kohlberg and Mertens**, 1986]: Basically every calculation necessary can be run efficiently on a computer.

The paper is organized as follows: Section 2 introduces notation and the modified definition of regular equilibria plus their properties. Section 3 studies the Nash field, reproves the well known result on existence of equilibria and oddness in the case of regularity and shows how to distinguish between equilibria and artificial zeros of the Nash field. Section 4 interprets the Nash field as portraying an evolutionary process and shows that stability of such a process is equivalent to a strong self-enforcement property of Nash equilibria. Moreover, section 4 contains a result on how to identify a stable component of Nash equilibria and relates this to the Stable Set. Section 5 is an illustrative chapter in that it shows how the Nash field can quickly inform the analyst on which equilibrium is risk dominant in the sense of **Harsanyi and Selten** [1988]. Section 6 draws conclusions.

2. Regular Equilibria

2.1 Notation. A *finite n-person normal form game* is a $2n$-tuple $\Gamma = (S_1, \ldots, S_n, u_1, \ldots, u_n)$, where S_i is a finite non-empty set, referred to as the set of *pure strategies* of player $i \in \mathcal{N} = \{1, \ldots, n\}$. Denoting $S = \prod_{i \in \mathcal{N}} S_i$, the set of all pure strategy combinations, each u_i is a mapping $u_i : S \to \Re$, for each $i \in \mathcal{N}$, and is called player i's *payoff function*. A typical element of the space of pure strategies will be written $s = (s_1, \ldots, s_n) \in S$. The cardinality of player i's pure strategy set S_i is denoted by $K_i = |S_i|$ and the cardinality of S is denoted by $K = \prod_{i \in \mathcal{N}} K_i$. For most of what follows it will be convenient to index player i's pure strategies by $k \in \{1, \ldots, K_i\}$, such that $s_i^k \in S_i$ denotes the k'th pure strategy of player $i \in \mathcal{N}$. The set of *mixed strategies* of player $i \in \mathcal{N}$ is the set of probability distributions on S_i. The probability which player i assigns to his k-th pure strategy s_i^k will be denoted by $\sigma_i^k = \sigma(s_i^k)$ and the space of all mixed strategies of player $i \in \mathcal{N}$ will be denoted by Σ_i. Since by deciding upon $(K_i - 1)$ probabilities assigned to his pure strategies, player i already has decided on the probability of the remaining pure strategy (because the probabilities have to add up to unity), the space Σ_i can taken to be $(K_i - 1)$-dimensional, i.e.

$$\Sigma_i = \{\sigma_i : S_i \to \Re_+^{K_i - 1} \mid \sum_{k=1}^{K_i - 1} \sigma_i^k \leq 1\}$$

The set of mixed strategy combinations, Σ, is the product $\Sigma = \prod_{i \in \mathcal{N}} \Sigma_i$. A *completely mixed strategy* for player $i \in \mathcal{N}$ is a probability vector $\sigma_i \in int\, \Sigma_i = \{\sigma_i : S_i \to \Re_{++}^{K_i - 1} | \sum_{k=1}^{K_i - 1} \sigma_i^k < 1\}$ and a completely mixed strategy combination is a mixed strategy combination $\sigma = (\sigma_1, \ldots, \sigma_n) \in int\, \Sigma = \prod_{i \in \mathcal{N}} int\, \Sigma_i$. The strategy combination resulting from $\sigma \in \Sigma$, when σ_i is replaced by $\hat{\sigma}_i$, is denoted by $(\sigma_{-i}, \hat{\sigma}_i) = (\sigma_1, \ldots, \sigma_{i-1}, \hat{\sigma}_i, \sigma_{i+1}, \ldots, \sigma_n) \in \Sigma$. A pure strategy of player $i \in \mathcal{N}$ is identified with the degenerate probability distribution which assigns 1 to the pure strategy selected and zero to all other pure strategies. When in $\sigma \in \Sigma$ player i's strategy $\sigma_i \in \Sigma_i$ is replaced by such a degenerate distribution assigning all the weight to pure strategy $s_i \in S_i$, a shorthand notation frequently used will be $(\sigma_{-i}, s_i) \in \Sigma \backslash int\ \Sigma$. For a given $\sigma_i \in \Sigma_i$ the subset of pure strategies to which σ_i assigns positive probability is called the support of σ_i, denoted

$$supp(\sigma_i) = \{s_i^k \in S_i \mid \sigma_i^k = \sigma_i(s_i^k) > 0\}.$$

Analogously, $supp(\sigma) = \prod_{i \in \mathcal{N}} supp(\sigma_i)$.

The space of all mixed strategy combinations Σ is a compact and convex polyhedron in $\Re^M, M = \sum_{i \in \mathcal{N}} K_i - n$, with dimension M. Since players in a normal form game decide independently on their mixed strategies, the joint probability that the pure strategy combination $s = (s_1^{k_1}, \dots, s_n^{k_n}) \in S,\ k_i \in \{1, \dots, K_i\}, \forall i \in \mathcal{N}$, will be played, given that player $i \in \mathcal{N}$ chooses $\sigma_i \in \Sigma_i$, is given by

$$\sigma(s) = \sigma(s_1^{k_1}, \dots, s_n^{k_n}) = \prod_{i \in \mathcal{N}} \sigma_i^{k_i}.$$

The *expected payoff* to player $i \in \mathcal{N}$, given that $\sigma \in \Sigma$ is played, is

$$U_i(\sigma) = \sum_{s \in S} u_i(s)\sigma(s),$$

i.e. is a multilinear function $U_i : \Sigma \to \Re$.

Since S_i is a finite set for each $i \in \mathcal{N}$, the payoff functions $u_i : S \to \Re$ can only take finitely many values. Collecting these values $u_i(s), s \in S$, in a K-dimensional vector for each $i \in \mathcal{N}$ and collecting these vectors $u_i = (u_i(s))_{s \in S}$ in a nK-dimensional vector $u = (u_i)_{i \in \mathcal{N}}$ makes it possible to identify Γ, for fixed player set and fixed pure strategy sets, with a point $u \in \Re^{nK}$. Writing $G(S_1, \dots, S_n)$ for the set of all normal form games with pure strategy sets $(S_1, \dots, S_n)$, there is, consequently, a one-to-one correspondence between $\Re^{nK}$ and $G(S_1, \dots, S_n)$. The notation for a game $\Gamma \in G(S_1, \dots, S_n)$ with *payoff vector* $u = (u_i)_{i \in \mathcal{N}} = ((u_i(s))_{s \in S})_{i \in \mathcal{N}} \in \Re^{nK}$ will frequently read $\Gamma = \Gamma(u)$. Within $G(S_1, \dots, S_n)$ there is, therefore, a natural way to measure distances between games by measuring the euclidean distance between their payoff vectors in $\Re^{nK}$. Accordingly, the measure of a subset of games in $G(S_1, \dots, S_n)$ is determined by the Lebesgue-measure of the corresponding subset of payoff vectors in $\Re^{nK}$. If it is necessary to stress the dependence of some mapping b on the payoff vector $u \in \Re^{nK}$, subscripts u will be used, i.e. b will be written as b_u.

For a fixed game $\Gamma \in G(S_1, \ldots, S_n)$ define the set of *best replies* of player $i \in \mathcal{N}$ against a stategy combination $\sigma \in \Sigma$ as the correspondence $BR_i : \Sigma \to \Sigma_i$ defined by

$$BR_i(\sigma) = \arg\max_{\hat{\sigma}_i \in \Sigma_i} U_i(\sigma_{-i}, \hat{\sigma}_i).$$

Where this is necessary, the set of pure best replies of player $i \in \mathcal{N}$ against $\sigma \in \Sigma$ will be written as $\overline{BR}_i(\sigma)$. Let the correspondece $BR = \prod_{i \in \mathcal{N}} BR_i$.

A *Nash eqilibrium* (or, an equilibrium) of a game $\Gamma \in G(S_1, \ldots, S_n)$ is a strategy combination $\sigma \in \Sigma$ such that $\sigma \in BR(\sigma)$. Such an equilibrium always exists for any game $\Gamma \in G(S_1, \ldots, S_n)$ [**Nash**, 1951]. The set of equilibria of a game Γ will be denoted by $\mathrm{E}(\Gamma)$. The set $\mathrm{E}(\Gamma)$ can also be viewed as a correspondence mapping $G(S_1, \ldots, S_n)$ into Σ. Using this view, the *graph of the Nash equilibrium correspondence*, denoted $\mathcal{G}(E)$, is defined by

$$\mathcal{G}(E) = \{(\Gamma, \sigma) \in G(S_1, \ldots, S_n) \times \Sigma \,|\, \sigma \in \mathrm{E}(\Gamma)\}.$$

A *strict* equilibrium is a $\sigma \in \mathrm{E}(\Gamma)$ which satisfies $\overline{BR}(\sigma) = \{\sigma\}$ [**Harsanyi**, 1973a]. A *quasi-strict* equilibrium is a $\sigma \in \mathrm{E}(\Gamma)$ which satisfies $\overline{BR}(\sigma) = supp(\sigma)$ ([**Harsanyi**, 1973a]; the terminology is from **van Damme** [1987]).

2.2 Regular Equilibria. In this subsection a modified definition of *regular* equilibrium is introduced and it is shown that a regular equilibrium possesses all robustness properties one can reasonably hope for. In particular, the present definition of regularity has the same implication as the definition introduced in **van Damme** [1987, chp. 2.5]. Moreover, in the example by which van Damme motivates his own deviation from Harsanyi's original definition [**van Damme**, 1987, Fig.2.5.1, p.39], the present definition selects the same equilibrium as van Damme's definition and does not rule out both equilibria, as **Harsanyi's** [1973a] definition would do.

At an equilibrium $\sigma \in \mathrm{E}(\Gamma)$ each $\sigma_i \in \Sigma_i$ must maximize the expected payoff $U_i(\sigma)$ for each $i \in \mathcal{N}$ subject to the constraint that $\sigma_i \in \Sigma_i$. Since $U_i(\sigma)$ is linear in $\sigma_i \in \Sigma_i$ this boils down to a problem of constrained, non-negative linear programming:

$$\max_{(\sigma_i^1, \ldots, \sigma_i^{K_i - 1})} \left[\sum_{k=1}^{K_i - 1} \sigma_i^k U_i(\sigma_{-i}, s_i^k) + (1 - \sum_{k=1}^{K_i - 1} \sigma_i^k) U_i(\sigma_{-i}, s_i^{K_i})\right]$$

$$s.t. \sum_{k=1}^{K_i - 1} \sigma_i^k - 1 \leq 0, \quad \sigma_i^k \geq 0, \quad \forall k = 1, \ldots, K_i - 1.$$

By linearity the Kuhn-Tucker complementary-slackness conditions are necessary and suffucient, such that the optimum $\bar{\sigma}_i \in \Sigma_i$ can - after eliminating the Lagrange multiplier - be characterized by

$$(1) \qquad \begin{aligned} U_i(\sigma_{-i}, s_i^k) \leq U_i(\sigma_{-i}, \bar{\sigma}_i), \forall s_i^k \in S_i, \\ \bar{\sigma}_i^k[U_i(\sigma_{-i}, s_i^k) - U_i(\sigma_{-i}, \bar{\sigma}_i] = 0, \forall k = 1, \ldots, K_i - 1. \end{aligned}$$

At an equilibrium, however, $\bar{\sigma}_i = \sigma_i$ must hold, such that the second part of (1) reads

$$(2) \qquad \sigma_i^k[U_i(\sigma_{-i}, s_i^k) - U_i(\sigma)] = 0, \forall k = 1, \ldots, K_i - 1,$$

for all $i \in \mathcal{N}$. This already is the modification to be introduced: While Harsanyi's and van Damme's definitions use pure strategy combinations jointly with the mix of the other players (σ_{-i}, s_i), $s_i \in supp(\sigma_i)$, as the reference point, instead of $U_i(\sigma)$ in (2), plus a condition $\sum_{k=1}^{K_i} \sigma_i^k = 1$, the present definition of regularity uses (a) for each player a mixed strategy space reduced by one dimension and (b) the equilibrium itself as the reference point. This is basically the same definition as it is frequently used in evolutionary game theory.

Formally, let the function $b : \Sigma \to \Re^M$ be defined by

$$(3) \qquad b_i^k(\sigma) = \sigma_i^k[U_i(\sigma_{-i}, s_i^k) - U_i(\sigma)], \ \forall k = 1, \ldots, K_i - 1, \ \forall i \in \mathcal{N}.$$

In contrast to the above mentioned definitions, b is continuously differentiable on a neighbourhood of $\Sigma \subset \Re^M$ (van Damme's definition is in general only continuously differentiable on a neighbourhood of the equilibrium). Let $D_\sigma b(\bar{\sigma})$ denote the Jacobian matrix of b at a point $\bar{\sigma} \in \mathrm{E}(\Gamma)$ and denote by $|D_\sigma b(\bar{\sigma})|$ its determinant.

Definition: An equilibrium $\bar{\sigma} \in \mathrm{E}(\Gamma)$ is said to be *regular*, if and only if $|D_\sigma b(\bar{\sigma})| \neq 0$.

The steps to follow are now basically adjustments of theorems in **van Damme** [1987, chp.2.5] to the present modification. Therefore, most proofs are gathered in the Appendix and the text only contains the statements.

LEMMA 1. *If $\bar{s}_i \notin supp(\sigma_i)$ and $b(\sigma) = 0$, then $[U_i(\sigma_{-i}, \bar{s}_i) - U_i(\sigma)] \in \Re$ is an eigenvalue of the Jacobian matrix $D_\sigma b(\sigma)$.*

(PROOF: see Appendix)

COROLLARY 1. *Every regular equilibrium is quasi-strict.*

PROOF: If $\sigma \in \mathrm{E}(\Gamma)$ is not quasi-strict, then for some $i \in \mathcal{N}$ there exists $s_i^k \notin supp(\sigma_i)$ such that $U_i(\sigma_{-i}, s_i^k) - U_i(\sigma) = 0$ which implies that $|D_\sigma b(\sigma)| = 0$. ∎

COROLLARY 2. *Every strict equilibrium is regular.*

PROOF: Since every strict equilibrium is in pure strategies, for each $i \in \mathcal{N}$ there are $(K_i - 1)$ pure strategies not used at $\sigma \in \mathrm{E}(\Gamma)$ which give the $(K_i - 1)$ corresponding eigenvalues $[U_i(\sigma_{-i}, s_i^k) - U_i(\sigma)] < 0, s_i^k \notin supp(\sigma_i)$. This determines $\sum_{i \in \mathcal{N}} K_i - n = M$ real and negative eigenvalues which are all eigenvalues of $D_\sigma b(\sigma)$, such that $|D_\sigma b(\sigma)| \neq 0$. ∎

The latter result can be sharpened to a rather obvious conclusion:

COROLLARY 3. *A pure strategy equilibrium is regular, if and only if it is strict.*

PROOF: Corollary 2 covers the if part. Since at a pure strategy equilibrium all eigenvalues are known and regularity of the equilibrium implies that there is no zero eigenvalue, one must have $[U_i(\sigma_{-i}, s_i^k) - U_i(\sigma)] < 0, \ \forall s_i^k \notin supp(\sigma_i), \ \forall i \in \mathcal{N}$. ∎

The next step is to show that regular equilibria are *strongly stable* in the sense of **Kojima, Okada and Shindoh** [1985].

THEOREM 1. *Let $\bar{\Gamma} \in G(S_1, \dots, S_n)$ and assume that $\bar{\sigma} \in E(\bar{\Gamma})$ is a regular equilibrium. Let $\bar{u} \in \Re^{nK}$ denote the payoff vector of $\bar{\Gamma}$. Then there exists a neighbourhood $\mathcal{U}$ of $\bar{u}$ in $\Re^{nK}$ and a neigbourhood $\mathcal{V}$ of $\bar{\sigma}$ in $\Re^M$, such that*

- (i) $|E(\Gamma(u)) \cap \mathcal{V}| = 1, \quad \forall u \in \mathcal{U}$, *and*
- (ii) *the mapping $\sigma : \mathcal{U} \to \mathcal{V}$, defined by $\{\sigma(u)\} = E(\Gamma(u)) \cap \mathcal{V}$, is continuous.*

(PROOF: see Appendix)

An equilibrium $\sigma \in \mathrm{E}(\Gamma)$ is said to be *isolated*, if and only if there exists a neigbourhood $\mathcal{V}$ of σ in $\Re^M$, such that $\mathrm{E}(\Gamma) \cap \mathcal{V} = \{\sigma\}$. An equilibrium $\sigma \in \mathrm{E}(\Gamma)$ is said to be *essential*, if and only if every game Γ' in a neighbourhood of Γ (in $\Re^{nK}$) has an equilibrium $\sigma' \in E(\Gamma')$ in a neighbourhood of $\sigma \in \mathrm{E}(\Gamma)$ [**Wu Wen-Tsün and Jiang Jia-He**, 1962].

COROLLARY 4. *Every regular equilibrium is essential and isolated.*

COROLLARY 5. *Every regular equilibrium is strictly perfect [**Okada**, 1981] and proper [**Myerson**, 1978].*

PROOF: The first part follows from Theorem 2.4.3 in **van Damme** [1987, p.34], where it is proved that every essential equilibrium is strictly perfect. Theorems 2.4.7 and 2.3.8 in van Damme [**1987**, p.36 and p.32] ensure that every strongly stable equilibrium is strictly proper and every strictly proper equilibrium is proper. ∎

Except for the rather obvious Corollary 3 all these results are known from **van Damme** [1987, chp.2.5]. The only reason, we list these results here, is to show that the present definition does not change the properties of regular equilibria. The methods of proofs use eigenvalues of the Jacobian matrix, rather than the more straightforward methods of van Damme, for the only reason to facilitate the presentation in the sections to follow. Before turning to these, however, another well known result ([**van Damme**, 1987, chp.2.6]; [**Harsanyi**, 1973a]; [**Wilson**, 1971]) is stated.

THEOREM 2. *Almost all games* $\Gamma \in G(S_1, \dots, S_n)$ *have all equilibria regular.*

(PROOF: see Appendix)

The phrase "almost all" here refers to $G(S_1, \dots, S_n)$, i.e. to $\Re^{nK}$. On the other hand, nearly any non-trivial extensive form will impose certain indifference relations upon the corresponding normal form and will, thereby, give rise to a degenerate normal form game in $G(S_1, \dots, S_n)$. In fact a given extensive form specifies a linear subspace of $G(S_1, \dots, S_n)$ and there is no guarantee whatsoever that in such a subspace games without regular equilibria will not cover a (relatively) open subset. But most of the refinement literature is motivated by extensive form arguments. From this point of view, therefore, Theorem 2 should be interpreted with due care. What, however, may be interesting about Theorem 2 is the proof given in the Appendix: In contrast to the well known method of proof, here the smoothness of b is exploited in order to obtain the Theorem from very powerful insights of Differential Topology.

3. THE NASH FIELD

An advantage of the system of equations defined by (3) is that it induces a globally differentiable structure on a neighbourhood of the polyhedron Σ. This structure contains, in our view, valuable information about the underlying game Γ. One particularily attractive way to view the structure induced by b on Σ is to think of it as a vector field on Σ. If it is appropriate to stress the interpretation of the mapping b defined in (3) as a vector field on Σ, the notation $\vec{b}$ will be used. The vector field $\vec{b}$ (or $\vec{b}_u$, if the dependence on the payoff vector $u \in \Re^{nK}$ is to be stressed) will be called the *Nash field* .

The reason to assign such a prominent name to $\vec{b}$ is that it indeed captures the spirit of the Nash equilibrium as a solution concept. Suppose that for some reason the players consider to play an arbitrary strategy combination $\sigma \in \Sigma$. If now one of the players $i \in \mathcal{N}$ considers to deviate from $\sigma \in \Sigma$ his coordinates of $\vec{b}$, denoted $\vec{b}_i(\sigma) = (b_i^1(\sigma), \ldots, b_i^{K_i-1}(\sigma))$, tell him in which direction he could unilaterally change his mix in order to improve his expected payoff. This can be seen from the following direct calculation:

$$
\begin{aligned}
U_i(\sigma_{-i}, \sigma_i + \vec{b}_i(\sigma)) = & \sum_{k=1}^{K_i-1} (\sigma_i^k + b_i^k(\sigma)) U_i(\sigma_{-i}, s_i^k) + \\
& + (1 - \sum_{k=1}^{K_i-1} (\sigma_i^k + b_i^k(\sigma))) U_i(\sigma_{-i}, s_i^{K_i}) = \\
& = U_i(\sigma) + \sum_{k=1}^{K_i-1} \sigma_i^k [U_i(\sigma_{-i}, s_i^k) - U_i(\sigma)] U_i(\sigma_{-i}, s_i^k) - \\
& - [\sum_{k=1}^{K_i-1} \sigma_i^k (U_i(\sigma_{-i}, s_i^k) - U_i(\sigma))] U_i(\sigma_{-i}, s_i^{K_i}) = \\
& = U_i(\sigma) + \sum_{k=1}^{K_i-1} \sigma_i^k [U_i(\sigma_{-i}, s_i^k)]^2 + \\
& + (1 - \sum_{k=1}^{K_i-1} \sigma_i^k) [U_i(\sigma_{-i}, s_i^{K_i})]^2 - [U_i(\sigma)]^2 = \\
& = U_i(\sigma) + Var(U_i(\sigma_{-i}, s_i) \,|\, \sigma_i) \geq U_i(\sigma), \; \forall i \in \mathcal{N}.
\end{aligned}
$$

If no such direction $\vec{b}_i(\sigma)$ exists, along which the player under consideration can unilaterally improve, for any player, then $\sigma \in \Sigma$ must be a zero of the vector field $\vec{b}$.

Very vaguely speaking, $\vec{b}$ could be viewed as the direction in which players with only bounded rationality, who change their mix only in very small steps (i.e. in a differentiable manner), but still try to improve unilaterally (i.e. assuming that other players stay with their present mix), would take the "state" of the game. In section 4 these claims will be made more precise by showing that $\vec{b}$ not only portrays the behavior of these artificial Nash-players with a certain kind of bounded rationality, but indeed captures the most important aspects of the behavior of fully rational Nash-players (a "Nash-player" should be read as a player, who takes the strategies of all other players as given). In any case the two major ingredients of the Nash equilibrium concept - "improving expected payoff" and "taking the actions of others as given" - are captured by $\vec{b}$.

To be able to interpret b as a vector field $\vec{b}$ on Σ, it is necessary to show that b indeed maps into the tangent space of Σ which, because of the simple structure of Σ, is just $\Re^M$. But beyond this trivial step a much stronger result is available: Let $\sigma(t, \sigma^0)$, $t \in \Re_+$, $\sigma^0 \in \Sigma$, be a solution to the system of differential equations $d\sigma = \vec{b}(\sigma)\,dt$, with $\sigma(0, \sigma^0) = \sigma^0$. In the following Lemma it is shown that no solution $\sigma(t, \sigma^0)$ ever leaves the boundary face of Σ in which it starts. In other words, Σ and each of its boundary faces (the boundary of Σ will be denoted $\partial\Sigma$) are invariant under the operation of $\vec{b}$.

LEMMA 2. *If $\sigma^0 \in \Sigma$, then $\sigma(t, \sigma^0) \in \Sigma$, $\forall t \in \Re_+$ and, moreover, if $\sigma^0 \in \partial\Sigma$, then $\sigma(t, \sigma^0) \in \partial\Sigma$, $\forall t \in \Re_+$.*

PROOF: It suffices to demonstrate the second part of the statement, because, if the latter is true, by continuity no solution path can ever leave Σ, once it starts in Σ. With the understanding that $\sigma(0, \sigma^0) = \bar{\sigma} \in \partial\Sigma$ abbreviate $\sigma(t, \bar{\sigma}) = \sigma(t)$, $\forall t \in \Re_+$. Since $\vec{b}(\sigma)$ is continuously differentiable the differential equation has a unique solution $\sigma(t, \bar{\sigma}) = \sigma(t)$ which satifies the initial condition $\sigma(0) = \bar{\sigma}$. For later reference note that

$$(*)\qquad \begin{aligned} \frac{d}{dt}(1 - \sum_{k=1}^{K_i-1} \sigma_i^k(t)) &= - \sum_{k=1}^{K_i-1} b_i^k(\sigma(t)) = \\ &= (1 - \sum_{k=1}^{K_i-1} \sigma_i^k(t))[U_i(\sigma_{-i}(t), s_i^{K_i}) - U_i(\sigma(t))]. \end{aligned}$$

By uniqueness of the solution satisfying the initial condition it is possible to write $\sigma_i^k(t)$ as

$$\sigma_i^k(t) = \bar{\sigma}_i^k \exp\{\int_0^t [U_i(\sigma_{-i}(v), s_i^k) - U_i(\sigma_{-i}(v), s_i^{K_i})- \\ - \sum_{\substack{h=1 \\ h\neq k}}^{K_i-1} \sigma_i^h(v)(U_i(\sigma_{-i}(v), s_i^h) - U_i(\sigma_{-i}(v), s_i^{K_i}))]dv\,\}\times \\ \times [\,1 + \bar{\sigma}_i^k \int_0^t [U_i(\sigma_{-i}(w), s_i^k) - U_i(\sigma_{-i}(w), s_i^{K_i})]\times \\ \times \exp\{\int_0^w [U_i(\sigma_{-i}(v), s_i^k) - U_i(\sigma_{-i}(v), s_i^{K_i})- \\ - \sum_{\substack{h=1 \\ h\neq k}}^{K_i-1} \sigma_i^h(v)(U_i(\sigma_{-i}(v), s_i^h) - U_i(\sigma_{-i}(v), s_i^{K_i}))]dv\,\}\,dw\,]^{-1}\;.$$

From (*) it follows by integration that

$$(1 - \sum_{k=1}^{K_i-1} \bar{\sigma}_i^k) \exp\{\int_0^t [U_i(\sigma_{-i}(v), s_i^{K_i})- \\ - \sum_{k=1}^{K_i-1} \sigma_i^k(v)\, U_i(\sigma_{-i}(v), s_i^k)]dv\,\} \times [\,1+ \\ + (1 - \sum_{k=1}^{K_i-1} \bar{\sigma}_i^k) \int_0^t U_i(\sigma_{-i}(w), s_i^{K_i}) \exp\{\int_0^w [U_i(\sigma_{-i}(v), s_i^{K_i})- \\ - \sum_{k=1}^{K_i-1} \sigma_i^k(v)\, U_i(\sigma_{-i}(v), s_i^k)]dv\,\}\,dw\,]^{-1} = 1 - \sum_{k=1}^{K_i-1} \sigma_i^k(t)\,.$$

If now $\bar{\sigma} \in \partial\Sigma$, then there exists some $i \in \mathcal{N}$ and $s_i^k \in S_i$, such that either (i) $\bar{\sigma}_i^k = 0$, or (ii) $\bar{\sigma}_i^k = 1$. In case (i) $\bar{\sigma}_i^k = 0$ implies by the above solution $\sigma_i^k(t) = 0,\ \forall t \in \Re_+$, and in case (ii) $\bar{\sigma}_i^k = 1 \implies \bar{\sigma}_i^h = 0,\ \forall h \neq k$, implies $\sigma_i^h(t) = 0,\ \forall t \in \Re_+,\ \forall h \neq k$, such that

$$\sigma_i^k(t) = \exp\{\int_0^t [U_i(\sigma_{-i}(v), s_i^k) - U_i(\sigma_{-i}(v), s_i^{K_i})]dv\,\}\times \\ \times [\,1 + \exp\{\int_0^t (U_i(\sigma_{-i}(v), s_i^k) - U_i(\sigma_{-i}(v), s_i^{K_i}))dv\,\} - 1\,]^{-1} = 1$$

for all $t \in \Re_+$. ▮

REMARK. *The reason why it is possible to solve explicitly for one component $\sigma_i^k(t)$ in the preceeding proof is that in each component one obtains a Riccati-equation for which explicit solutions are known.*

To interpret b as a vector field $\vec{b}$ on Σ provides the opportunity to demonstrate another well known result by a somewhat unusual method. (The method of the following proof is borrowed from **Hofbauer and Sigmund** [1988, pp.166], who apply it to a very similar system of replicator equations.)

THEOREM 3. *Every game $\Gamma \in G(S_1, \dots, S_n)$ has at least one Nash equilibrium. If all Nash equilibria of Γ are regular, then their number is finite and odd.*

PROOF: (i) Consider as a perturbation of $\vec{b}(\sigma)$ the vector field $\vec{b}_\varepsilon(\sigma)$, with $\varepsilon > 0$, defined by

$$(4) \quad b_{\varepsilon,i}^k(\sigma) = \sigma_i^k[U_i(\sigma_{-i}, s_i^k) - U_i(\sigma) - K_i\varepsilon] + \varepsilon\,, \ k = 1, \dots, K_i - 1,$$

for all $i \in \mathcal{N}$. For any $\varepsilon > 0$ the vector field $\vec{b}_\varepsilon(\sigma)$ now points inwards at the boundary of Σ, as can be seen from

$$\sigma_i^k = 0 \implies b_{\varepsilon,i}^k(\sigma) = \varepsilon > 0, \quad \text{and}$$
$$\sum_{k=1}^{K_i-1} \sigma_i^k = 1 \implies \sum_{k=1}^{K_i-1} b_{\varepsilon,i}^k(\sigma) = -\varepsilon < 0.$$

Consequently, for given $\varepsilon > 0$ the vector field $\vec{b}_\varepsilon$ must have a zero in *int* Σ. Let such a zero of $\vec{b}_\varepsilon$ be denoted by $\sigma(\varepsilon) \in$ *int* Σ. Any sequence $\{\sigma(\varepsilon)\}_{\varepsilon\downarrow 0}$ must have a cluster point, because Σ is compact. Any element of such a sequence satifies for any $\varepsilon > 0$

$$U_i(\sigma_{-i}(\varepsilon), s_i^k) - U_i(\sigma(\varepsilon)) - \varepsilon K_i = -\frac{\varepsilon}{\sigma_i^k(\varepsilon)} < 0,$$

$\forall k = 1, \dots, K_i - 1,\ i \in \mathcal{N}$, and

$$0 = \sum_{k=1}^{K_i-1} \sigma_i^k(\varepsilon)[U_i(\sigma_{-i}(\varepsilon), s_i^k) - U_i(\sigma(\varepsilon)) - \varepsilon K_i] + \varepsilon(K_i - 1) =$$
$$= -[(1 - \sum_{k=1}^{K_i-1} \sigma_i^k(\varepsilon))(U_i(\sigma_{-i}(\varepsilon), s_i^{K_i}) - U_i(\sigma(\varepsilon)) - \varepsilon K_i) + \varepsilon]$$
$$\implies U_i(\sigma_{-i}(\varepsilon), s_i^{K_i}) - U_i(\sigma(\varepsilon)) - \varepsilon K_i = -\frac{\varepsilon}{1 - \sum_k \sigma_i^k(\varepsilon)} < 0,$$

for all $i \in \mathcal{N}$. Now consider any cluster point σ of $\sigma(\varepsilon)$ as $\varepsilon \to 0$. By continuity the weak inequality

$$U_i(\sigma_{-i}, s_i^k) - U_i(\sigma) \leq 0, \; \forall k = 1, \dots, K_i, \; \forall i \in \mathcal{N}$$

must be satisfied, such that σ is a Nash equilibrium. This completes the first part.

(ii) If all equilibria are regular, then there are only finitely many, because Σ is compact. By the implicit function theorem each equilibrium is, for $\varepsilon > 0$ sufficiently small, continuously approximated by a unique family $\{\sigma(\varepsilon)\}_{\varepsilon>0}$ of zeros of $\vec{b}_\varepsilon$, $\varepsilon > 0$. As the index of an equilibrium does not change as $\varepsilon \to 0$ by regularity, the indices of all zeros would have to sum to the Euler characteristic $+1$ of the polyhedron Σ, if Σ would be a compact manifold (cf. the Poincaré-Hopf Index Theorem [**Guillemin and Pollack**, 1974, p.134]). However, Σ is not a manifold and *int* Σ is not compact. But as a substitute for compactness of the manifold *int* Σ the property that $\vec{b}_\varepsilon$ points inwards at the boundary $\partial\Sigma$ can be used, as demonstrated in **Dierker** [1972, Theorem 1]. Since this still guarantees the invariance of the index sum and the index (defined as the sign of $| - D_\sigma b_\varepsilon(\sigma(\varepsilon))|$) can only take values in $\{-1, +1\}$, the number of equilibria must be odd. ∎

REMARK. *The perturbation of $\vec{b}$ defined by (4) will continue to play an important role in the present paper. It is worth pointing out that zeros of the perturbation are equivalent to the necessary and sufficient conditions for an equilibrium of a game, where payoffs are given by the payoffs in Γ plus ε times the payoffs from the logarithmic game as defined in* **Harsanyi** *[1973a]. This provides the formal link between Harsanyi's analysis and the present one.*

Regularity of an equilibrium has been defined in terms of the non-singularity of the Jacobian matrix $D_\sigma\vec{b}(\bar{\sigma})$, $\bar{\sigma} \in \mathrm{E}(\Gamma)$. If an equilibrium is regular, then one may expect that most of the information about the local behavior of the vector field $\vec{b}$ around the equilibrium can be obtained by studying its linearization $D_\sigma\vec{b}(\bar{\sigma})$. This holds true, provided that all eigenvalues of $D_\sigma\vec{b}(\bar{\sigma})$ have non-zero real parts, by Hartmann's Theorem [**Chillingworth**, 1976, p.208].

At a pure strategy equilibrium the structure of $D_\sigma\vec{b}(\bar{\sigma})$ is particularily simple. In fact it is possible to strenghten Lemma 1 with respect to pure strategy combinations (which are always zeros of the Nash field): At pure strategy combinations not only all eigenvalues of $D_\sigma\vec{b}(\bar{\sigma})$ are known from Lemma 1. Also all eigenvectors correspond to the "edges" of Σ emerging from the "corner" which corresponds to the pure strategy combination.

LEMMA 3. *Let $\bar{\sigma} \in \partial\Sigma$ be the mixed strategy combination corresponding to the pure strategy combination $\bar{s} \in S$. Then all eigenvectors belonging to the eigenvalues $[U_i(\bar{\sigma}_{-i}, s_i^h) - U_i(\bar{\sigma})] \in \Re$, $\forall s_i^h \notin supp(\bar{\sigma}_i)$, for all $i \in \mathcal{N}$, are given by $[(\bar{\sigma}_{-i}, s_i^h) - \bar{\sigma}]$, $s_i^h \notin supp(\bar{\sigma}_i)$, $\forall i \in \mathcal{N}$.*

(PROOF: See Appendix.)

COROLLARY 6. *If $\bar{\sigma} \in E(\Gamma)$ is a strict equilibrium, then it is locally asymptotically stable, i.e. the vector field $\vec{b}$ points locally towards $\bar{\sigma} \in \Sigma$.*

PROOF: Lemma 3 explicitly gives the complete eigensystem of $D_\sigma\vec{b}(\bar{\sigma})$. If $\bar{\sigma} \in \mathrm{E}(\Gamma)$ is strict, then it is regular (Corollary 2) and Hartmann's Theorem is applicable, because all eigenvalues are given by $[U_i(\bar{\sigma}_i, s_i) - U_i(\bar{\sigma})] < 0$, $\forall s_i \notin supp(\bar{\sigma}_i)$, $\forall i \in \mathcal{N}$. ∎

COROLLARY 7. *Every pure strategy equilibrium is near strict in the normal form [***Fudenberg, Kreps and Levine***, 1988, p.357].*

PROOF: At a pure strategy equilibrium $\bar{\sigma} \in \mathrm{E}(\Gamma)$

$$[U_i(\bar{\sigma}_{-i}, s_i) - U_i(\bar{\sigma})] \leq 0, \ \forall s_i \notin supp(\bar{\sigma}_i), \ \forall i \in \mathcal{N}.$$

Hence the required sequence $\{\Gamma^m\}_{m=1}^{\infty}$, $\Gamma^m \to \Gamma$, can be constructed in such a way that all payoffs to a player $i \in \mathcal{N}$ to his pure strategy $s_i \notin supp(\bar{\sigma}_i)$, for which equality holds (instead of a strict inequality), are disturbed downwards. ∎

Corollary 7 is merely a restatement of Proposition 1 in **Fudenberg, Kreps and Levine** [1988]. In the present context, however, this result emerges from the particularily transparent behavior of the Nash field around pure strategy combinations (which are always zeros of $\vec{b}$).

Clearly the Nash field will have more zeros than Γ has Nash equilibria, i.e. not every zero of $\vec{b}$ will correspond to an equilibrium. The next result shows, how easy it is to distinguish between zeros of the Nash field which are equilibria and those which are not.

LEMMA 4. *$\bar{\sigma} \in E(\Gamma) \iff \vec{b}(\bar{\sigma}) = 0$ and $U_i(\bar{\sigma}_{-i}, s_i) \leq U_i(\bar{\sigma})$, $\forall s_i \notin supp(\bar{\sigma}_i)$, $\forall i \in \mathcal{N}$.*

PROOF: If $\bar{\sigma} \in \mathrm{E}(\Gamma)$, then clearly from (1) and (2) one must have $\vec{b}(\bar{\sigma}) = 0$ and $U_i(\bar{\sigma}_{-i}, s_i) \leq U_i(\bar{\sigma})$, $\forall s_i \notin supp(\bar{\sigma}_i)$, $\forall i \in \mathcal{N}$. If the latter is true, then $\vec{b}(\bar{\sigma}) = 0$ implies $U_i(\bar{\sigma}_{-i}, s_i) = U_i(\bar{\sigma})$, $\forall s_i \in supp(\bar{\sigma}_i)$, such that $U_i(\bar{\sigma}_{-i}, \hat{s}_i) \leq U_i(\bar{\sigma})$, $\forall \hat{s}_i \notin supp(\bar{\sigma}_i)$, $\forall i \in \mathcal{N}$, is sufficient for $\bar{\sigma} \in \mathrm{E}(\Gamma)$. ∎

To check whether a zero of $\vec{b}$ is an equilibrium it, therefore, suffices to check the eigenvalues of the Jacobian matrix corresponding to unused pure strategies. That this is true even when for some $i \in \mathcal{N}$

his final strategy $s_i^{K_i} \in supp(\bar{\sigma}_i)$, can be seen from the following argument: Suppose $1 - \sum_{k=1}^{K_i-1} \bar{\sigma}_i^k > 0$, but $U_i(\bar{\sigma}) < U_i(\bar{\sigma}_{-i}, s_i^{K_i})$ in which case all eigenvalues corresponding to unused pure strategies could still be non-positive, but the zero would not be an equilibrium. However, from the definition of expected payoffs, $U_i(\sigma) = \sum_{k=1}^{K_i-1} \sigma_i^k U_i(\sigma_{-i}, s_i^k) + (1 - \sum_{k=1}^{K_i-1} \sigma_i^k) U_i(\sigma_{-i}, s_i^{K_i})$, then $\vec{b}(\bar{\sigma}) = 0$ would imply for any $s_i^k \in supp(\bar{\sigma}_i)$, $k = 1, \ldots, K_i - 1$, that

$$
\begin{aligned}
0 = & \bar{\sigma}_i^k [U_i(\bar{\sigma}_{-i}, s_i^k) - U_i(\bar{\sigma}_{-i}, s_i^{K_i}) - \\
& - \sum_{h=1}^{K_i-1} \bar{\sigma}_i^h (U_i(\bar{\sigma}_{-i}, s_i^h) - U_i(\bar{\sigma}_{-i}, s_i^{K_i}))] = \\
= & \bar{\sigma}_i^k (1 - \sum_{h=1}^{K_i-1} \bar{\sigma}_i^h)[U_i(\bar{\sigma}) - U_i(\bar{\sigma}_{-i}, s_i^{K_i})],
\end{aligned}
$$

i.e. a contradiction. Hence $\vec{b}(\bar{\sigma}) = 0$ and $s_i^{K_i} \in supp(\bar{\sigma}_i)$ together imply $U_i(\bar{\sigma}) = U_i(\bar{\sigma}_{-i}, s_i^{K_i})$. Indeed, therefore, the property of a zero of the Nash field to form an equilibrium boils down to a simple eigenvalue condition on the Jacobian matrix $D_\sigma \vec{b}(\bar{\sigma})$. This condition is, moreover, independent of whether the zero of the Nash field is regular or not. In this sense the problem of determining Nash equilibria reduces to a problem of solving a system of equations and checking the solutions.

As an immediate consequence of Lemma 4, zeros of the Nash field which are not equilibria can only emerge at the boundary of Σ. Now suppose that a zero of $\vec{b}$ lies in the (relative) interior of a boundary face of Σ. Since any boundary face of Σ can be seen as a reduced game, where all unused pure strategies have been deleted, the same logic carries over to the boundary face: In this reduced game the zero of the Nash field must describe an equilibrium. And, by Lemma 2, the Nash field of the reduced game is merely the Nash field of the full game restricted to the boundary face.

Also Lemma 4 implies that at any regular zero of $\vec{b}$ which is not an equilibrium the vector field $\vec{b}$ must locally "point away" from the zero in at least one direction (provided $D_\sigma \vec{b}(\bar{\sigma})$ has no eigenvalue with zero real part). This "divergent" direction is given by the (real) eigenvector belonging to the (real) eigenvalue $[U_i(\bar{\sigma}_{-i}, s_i) - U_i(\bar{\sigma})]$, $s_i \notin supp(\bar{\sigma}_i)$, which must be positive, if the zero $\bar{\sigma}$ of $\vec{b}$ is not an equilibrium. (More precisely: If $\bar{\sigma} \notin \mathrm{E}(\Gamma)$, but $\vec{b}(\bar{\sigma}) = 0$, then $\exists\, i \in \mathcal{N}, s_i \in S_i : U_i(\bar{\sigma}_{-i}, s_i) > U_i(\bar{\sigma})$.)

Finally a direct calculation shows that the trace of the Jacobian matrix

$D_\sigma \vec{b}(\bar{\sigma})$ is given by

$$\text{trace}(D_\sigma \vec{b}(\bar{\sigma})) = \sum_{i\in\mathcal{N}} \sum_{h=1}^{K_i} [U_i(\bar{\sigma}_{-i}, s_i^h) - U_i(\bar{\sigma})],$$

for any $\bar{\sigma} \in \Sigma$. Assuming $\vec{b}(\bar{\sigma}) = 0$ and subtracting all eigenvalues $[U_i(\bar{\sigma}_{-i}, s_i) - U_i(\bar{\sigma})]$, $s_i \notin supp(\bar{\sigma}_i)$, $\forall i \in \mathcal{N}$, corresponding to unused pure strategies from the trace of $D_\sigma \vec{b}(\bar{\sigma})$, then yields the conclusion that all the other eigenvalues have to sum to zero, because $\bar{s}_i \in supp(\bar{\sigma}_i) \implies U_i(\bar{\sigma}_{-i}, \bar{s}_i) - U_i(\bar{\sigma}) = 0$. Because the trace is invariant with respect to a change of coordinates, this implies two conclusions on equilibria in *int* Σ: (i) If there is an eigenvalue with non-zero real part, then there will also be an eigenvalue with positive real part. (ii) If the equilibrium is regular and $M = \dim\Sigma$ is odd, then the equilibrium is unstable in the sense that $D_\sigma \vec{b}(\bar{\sigma})$ has at least one eigenvalue with positive real part.

REMARK. *Since the polyhedron Σ is not a differentiable manifold, one may wonder what is meant by the Jacobian matrix $D_\sigma \vec{b}(\bar{\sigma})$ at the boundary of Σ. But int Σ is a differentiable manifold, such that for $\bar{\sigma} \in$ int Σ the Jacobian matrix $D_\sigma \vec{b}(\bar{\sigma})$ can be defined independently of a special choice of a coordinate system [**Hirsch**, 1976, pp.17]. The Jacobian matrix $D_\sigma \vec{b}(\bar{\sigma})$ on the entire polyhedron Σ is obtained by continuous extension of $D_\sigma \vec{b}(\bar{\sigma})$ to the boundary of Σ. This is always uniquely possible, because $D_\sigma \vec{b}(\bar{\sigma})$ is bounded for any $\bar{\sigma} \in \Sigma$ and, therefore, uniformly continuous [**Dieudonné**, 1969, p.57].*

4. EVOLUTION AND STABILITY

In recent years game theory has strongly influenced certain branches of biology [**Hofbauer and Sigmund**, 1988]. Although in biology there is a tendency to study either symmetric games or two-person games, there seems to be nothing restricting the application of game theory to biology to these special cases. Since we are personally not too familiar with biology, the following interpretation will be cast in more game theoretic terms and missing (or: bounded) rationality, but we feel comfortable with the idea that a biological interpretation should be straightforward.

Think of a player now as a large population of individuals, who do not possess consciousness or at least are not rational in the traditional sense. There are n, $n \geq 1$, such populations, indexed by $i \in \mathcal{N} = \{1, \ldots, n\}$, and each population is characterized by having a finite number $K_i > 1$ of possible options of behavior. Hence pure strategies of a population $i \in \mathcal{N}$ are identified with a specific type of behavior according to which individuals in a given population can act. All of the populations interact and their interaction determines the reproductive success of the types of behavior within each population. More precisely: Let σ_i^k, $i \in \mathcal{N}$, $k = 1, \ldots, K_i$, denote the ratio of the number of individuals in population $i \in \mathcal{N}$ which act according to $s_i^k \in S_i$ to the total number of individuals in population $i \in \mathcal{N}$, $\sum_{k=1}^{K_i} \sigma_i^k = 1$, $\sigma_i^k \geq 0$, $k = 1, \ldots, K_i$. Let Σ_i denote the set of all vectors $\sigma_i = (\sigma_i^1, \ldots, \sigma_i^{K_i - 1})$, $\sigma_i^{K_i} = 1 - \sum_{k=1}^{K_i - 1} \sigma_i^k$, $i \in \mathcal{N}$, describing the composition of a given population $i \in \mathcal{N}$, and let $\Sigma = \prod_{i \in \mathcal{N}} \Sigma_i$. For each $i \in \mathcal{N}$ there is a multilinear function $U_i : \Sigma \to \Re$, such that for some given $\sigma_{-i} \in \Sigma_{-i}$ the value $U_i(\sigma_{-i}, s_i^k)$ describes the reproductive success of the k-th type of behavior, $s_i^k \in S_i$, within population $i \in \mathcal{N}$: Suppose at some given point in time the composition of each of the n populations is described by the $M = \sum_{i \in \mathcal{N}} K_i - n$ dimensional vector $\sigma = (\sigma_1, \ldots, \sigma_n) \in \Sigma$. Populations interact and, as an outcome of this interaction, the *rate of change* of σ_i^k, $k = 1, \ldots, K_i - 1$, $i \in \mathcal{N}$, given $\sigma \in \Sigma$, is proportional to the relative success of type k behavior in population $i \in \mathcal{N}$ as compared to the average behavior of population $i \in \mathcal{N}$, i.e.

$$
\begin{aligned}
\frac{\dot{\sigma}_i^k}{\sigma_i^k} &= U_i(\sigma_{-i}, s_i^k) - \sum_{k=1}^{K_i - 1} \sigma_i^k U_i(\sigma_{-i}, s_i^k) - \\
&\quad - (1 - \sum_{k=1}^{K_i - 1} \sigma_i^k) U_i(\sigma_{-i}, s_i^{K_i}) = \\
&= U_i(\sigma_{-i}, s_i^k) - U_i(\sigma), \quad k = 1, \ldots, K_i - 1, \ \forall i \in \mathcal{N}
\end{aligned}
\tag{3'}
$$

where $\dot{\sigma}_i^k = \frac{d}{dt} \sigma_i^k$, $t \in \Re_+$. Individuals within a population are not assumed rational here. Rather the expected payoffs to a particular type of behavior within a population are compared to the expected payoff which the population will obtain on average (given its current composition). If some particular type of behavior yields a higher payoff than the average, then this will lead to imitating behavior on the part of individuals, who are currently using worse types of behavior, such that the percentage of the successful type of behavior within the population expands. On the other hand, if a particular type of behavior within a population is less successful than the average, then individuals sticking to this worse strategy will start to die out, such that the percentage of individuals behaving in this unsuccessful way will shrink. Although no rationality in the traditional sense is assumed here, this may be viewed as an instance of bounded rationality, where some smooth revision of sticky behavioral patterns within a population is portrayed.

In section 3 it has been shown that, if a path determined by (3') is followed unilaterally by only one particular population, then this population eventually increases its average payoff. This feature captures the spirit of the Nash concept that the composition of rival populations is taken as given. Consequently, there is not only no strong notion of rationality required here, but there is also no assumtion of "optimal fitness" in a Darwinian sense: "Fitness" only holds relative to $\sigma_{-i} \in \Sigma_{-i}$, i.e. relative to the current compositions of the rival populations.

Viewed in this way each game $\Gamma \in \mathrm{G}(S_1, \ldots, S_n)$ can be associated with an evolutionary process determined by (3'). This process does not assume sophisticated optimization capabilities on the part of the players. Rather "players" are viewed as populations, the compositions of which change over time according to the criteria specified via the payoff functions. From (3') it is obvious that the evolution of such a system is determined by the Nash field. And this section is devoted to showing that there is an intimate relationship between the evolutionary process specified in (3') and the properties of the traditional game Γ with fully rational players. In particular, it will be shown that stability of the evolutionary process is indeed equivalent to robustness of the corresponding Nash equilibrium against slight deviations of some players from the equilibrium.

From the preceeding section it is known that a solution to (3') will, in general, have more stationary points than the game Γ has Nash equilibria: Lemma 4 implies that stationary points of (3') (rest points) and Nash equilibria agree on *int* Σ. Only at the boundary of Γ stationary points of (3') have to satisfy the extra conditions in Lemma 4 to qualify as Nash equilibria. Also Lemmas 3 and 4 strongly support the conjecture that rest points of (3') which are not Nash equilibria cannot be stable. These observations make it worthwhile to carry through the programme of associating the evolutionary process in (3') with properties of the traditional game Γ by studying stability properties of (3') around an equilibrium.

Denote by $\sigma(t, \sigma^0)$ a solution path of the system of differential equations (3'), with $\sigma(0, \sigma^0) = \sigma^0$, $t \in \Re_+$. Then define a stationary point $\bar{\sigma} \in \Sigma$ of (3') as a point satisfying $\sigma(t, \bar{\sigma}) = \bar{\sigma}$, $\forall t \in \Re_+$. Denote by $\mathrm{F}(\Gamma)$ the set of stationary points of (3') associated with a game $\Gamma \in \mathrm{G}(S_1, \ldots, S_n)$. A stationary point $\bar{\sigma} \in F(\Gamma)$ is said to be *locally asymptotically stable*, if and only if

$$\exists \mathcal{U},\ \mathcal{U} \text{ open},\ \mathcal{U} \cap \mathrm{F}(\Gamma) = \{\sigma\} : \sigma(t, \sigma^0) \xrightarrow{t \to \infty} \bar{\sigma},\ \forall \sigma^0 \in \mathcal{U} \cap \Sigma.$$

The first step in the analysis is to study the region, where $\mathrm{F}(\Gamma)$ and $\mathrm{E}(\Gamma)$ agree, i.e. *int* Σ.

LEMMA 5. *Suppose $\bar{\sigma} \in F(\Gamma) \cap int\ \Sigma$. Then $\bar{\sigma}$ cannot be locally asymptotically stable.*

PROOF: (The idea of the following proof is borrowed from **Amann and Hofbauer** [1985], who apply the following idea to a somewhat simpler replicator equation. We are grateful to Erwin Amann for making us aware of this idea.)

Let $\bar{\sigma} \in \mathrm{F}(\Gamma) \cap int\ \Sigma$ and consider, instead of $\vec{b}$, the modified vector field $\vec{\beta}$ on $int\ \Sigma$ defined by

$$\vec{\beta}(\sigma) = \frac{1}{P(\sigma)}\vec{b}(\sigma), \quad P(\sigma) = \prod_{i \in \mathcal{N}} (1 - \sum_{k=1}^{K_i - 1} \sigma_i^k) \prod_{k=1}^{K_i - 1} \sigma_i^k.$$

On the interior of Σ one must have $P(\sigma) > 0$, such that multiplying $\vec{b}$ by $P(\sigma)^{-1}$ does not alter the orbits of (3'): By reparametrization with $t = f(s)$, where f is chosen such that $f'(s) = P(\sigma(f(s)))^{-1}$, for a solution $\sigma(t)$ of (3') one obtains $d\sigma(f(s)) = \vec{b}(\sigma)f'(s)ds$; from $P(\sigma)^{-1} > 0$, $\forall \sigma \in int\ \Sigma$, the function f is strictly monotone, such that the velocity of movements along orbits is changed, but not orbits themselves. In particular, the differential equation $\dot{\sigma} = \vec{\beta}(\sigma)$ has the same rest points and the same qualitative behavoir around rest points in $int\ \Sigma$ as (3').

Now calculate the divergence of the vector field $\vec{\beta}$ on $int\ \Sigma$. By the chain rule,

$$\frac{\partial \beta_i^k(\sigma)}{\partial \sigma_i^k} = P(\sigma)^{-1}\frac{\partial b_i^k(\sigma)}{\partial \sigma_i^k} - P(\sigma)^{-2}\frac{\partial P(\sigma)}{\partial \sigma_i^k} b_i^k(\sigma) \quad \text{and}$$

$$\frac{\partial P(\sigma)}{\partial \sigma_i^k} = \frac{1 - \sum_{l=1}^{K_i - 1} \sigma_i^l - \sigma_i^k}{\sigma_i^k(1 - \sum_{l=1}^{K_i - 1} \sigma_i^l)} P(\sigma).$$

Summing over $k = 1, \ldots, K_i - 1$ and $i \in \mathcal{N}$ yields

$$\begin{aligned}
\operatorname{div}\vec{\beta}(\sigma) &= \sum_{i\in\mathcal{N}} \sum_{k=1}^{K_i-1} P(\sigma)^{-1}\Big[\frac{\partial b_i^k(\sigma)}{\partial \sigma_i^k} - \frac{1-\sum_{l=1}^{K_i-1}\sigma_i^l - \sigma_i^k}{\sigma_i^k(1-\sum_{l=1}^{K_i-1}\sigma_i^l)} b_i^k(\sigma)\Big] = \\
&= P(\sigma)^{-1} \sum_{i\in\mathcal{N}} \sum_{k=1}^{K_i-1} [U_i(\sigma_{-i}, s_i^k) - U_i(\sigma) - \\
&\quad - \sigma_i^k(U_i(\sigma_{-i}, s_i^k) - U_i(\sigma_{-i}, s_i^{K_i})) - \\
&\quad - \frac{1-\sum_{l=1}^{K_i-1}\sigma_i^l - \sigma_i^k}{1-\sum_{l=1}^{K_i-1}\sigma_i^l}(U_i(\sigma_{-i}, s_i^k) - U_i(\sigma))] = \\
&= P(\sigma)^{-1}[\operatorname{div}\vec{b}(\sigma) - \sum_{i\in\mathcal{N}} \sum_{k=1}^{K_i} (U_i(\sigma_{-i}, s_i^k) - U_i(\sigma))] = \\
&= 0
\end{aligned}$$

Now assume that $\bar{\sigma} \in \mathrm{F}(\Gamma) \cap \mathit{int}\, \Sigma$ is a locally asymptotically stable rest point of (3'); then it must be a locally asymptotically stable rest point of $\dot{\sigma} = \vec{\beta}(\sigma)$. Denote by $\hat{\sigma}(t, \sigma^0)$ a solution to $\dot{\sigma} = \vec{\beta}(\sigma)$, with $\hat{\sigma}(0, \sigma^0) = \sigma^0 \in \mathit{int}\, \Sigma$. Let $\mathcal{U} \subset \mathit{int}\, \Sigma$ be a neighbourhood of $\bar{\sigma}$ such that $\sigma^0 \in \mathcal{U} \Rightarrow \hat{\sigma}(t, \sigma^0) \xrightarrow{t\to\infty} \bar{\sigma} \in \mathcal{U}$. Assign to $\mathcal{U}$ a volume $V = \int_{\mathcal{U}} d\sigma \neq 0$ and set $V(0) = V$. Define

$$\mathcal{U}(t) = \{\sigma \in \Sigma \,|\, \sigma = \hat{\sigma}(t, \sigma^0),\, \sigma^0 \in \mathcal{U}\}.$$

Then by Liouville's theorem the volume $V(t)$ of $\mathcal{U}(t)$ is given by

$$\dot{V}(t) = \int_{\mathcal{U}(t)} \operatorname{div}\vec{\beta}(\sigma) d\sigma = 0.$$

But then $V(t) = \text{constant} = V(0) \neq 0$, $\forall t \in \Re_+$. On the other hand, $\lim_{t\to\infty} \mathcal{U}(t) = \bar{\sigma}$ by the assumption of local asymptotic stability and the construction of $\mathcal{U}$. The latter, however, implies $\lim_{t\to\infty} V(t) = 0$, i.e. a contradiction to continuity of $V(t)$. ∎

Each boundary face of Σ can be thought of as a smaller game of its own, derived from Γ by deleting all unused pure strategies. Moreover, by Lemma 2, the Nash field of these reduced games is just the Nash field of Γ restricted to the corresponding boundary faces of Σ, and each boundary face is invariant with respect to the operation of the Nash field. As a consequence of this, the (relative) interior of each boundary face satisfies the requirements of Lemma 5. We have shown:

COROLLARY 8. *Any locally asymptotically stable stationary point of (3') is a pure strategy combination.*

Clearly, by Lemma 3, a locally asymptotically stable stationary point of (3') must not only be pure, but must also be a Nash equilibrium. From Corollary 6, it must even be a strict Nash equilibrium. (A formal proof of these claims can be extracted from the proof of Proposition 1 below.)

Intuitively it is tempting to conjecture that these findings have to do with the well known instability of mixed equilibria ([**van Damme**, 1987, p.19]; [**Harsanyi**, 1973b]): At a mixed equilibrium some player can choose another mix, than the one prescribed by the equilibrium, without risking losses of expected payoff, given that the other players stay with their equilibrium mix; if, however, the player under consideration indeed deviates from his equilibrium mix, then he disturbs the equilibrium property. That is to say that mixed equilibria are self-enforcing only with respect to themselves, but not necessarily even with respect to a neighbourhood. Intuitively speaking, an equilibrium is self-enforcing with respect to a neighbourhood $\mathcal{O}$, if and only if

$$\exists \mathcal{O},\ \mathcal{O}\,\text{open},\ \mathcal{O} \cap \mathrm{E}(\Gamma) = \{\bar{\sigma}\} :\ \bar{\sigma} \in \mathrm{BR}(\sigma^0),\ \forall \sigma^0 \in \mathcal{O} \cap \Sigma.$$

The next proposition says that this self-enforcing property with respect to a neighbourhood is in fact equivalent to local asymptotic stability of the evolutionary process in (3').

PROPOSITION 1. *An equilibrium $\bar{\sigma} \in E(\Gamma)$ is a locally asymtotically stable stationary point of (3'), if and only if*

$$\exists \mathcal{O},\ \mathcal{O}\,\mathit{open},\ \mathcal{O} \cap E(\Gamma) = \{\bar{\sigma}\} :\ \bar{\sigma} \in BR(\sigma^0),\ \forall \sigma^0 \in \mathcal{O} \cap \Sigma.$$

PROOF: (i) First assume that $\bar{\sigma} \in \mathrm{E}(\Gamma)$ is locally asymptotically stable. By Corollary 8 it must then be pure. Since $\bar{\sigma} \in \mathrm{E}(\Gamma)$, one has $U_i(\bar{\sigma}_{-i}, s_i) \leq U_i(\bar{\sigma})$, $\forall s_i \in S_i$, $\forall i \in \mathcal{N}$. Suppose for some $i \in \mathcal{N}$ there exists $s_i \notin supp(\bar{\sigma}_i)$, such that $U_i(\bar{\sigma}_{-i}, s_i) - U_i(\bar{\sigma}) = 0$. Then consider, for $\alpha \in [0, 1]$, the family of strategy combinations $[\alpha\bar{\sigma} + (1-\alpha)(\bar{\sigma}_{-i}, s_i)]$. Since both $\bar{\sigma}$ and $(\bar{\sigma}_{-i}, s_i)$ are pure strategy combinations

$$\begin{aligned}
&\bar{\sigma}_j^l[\alpha U_j(\bar{\sigma}_{-j}, s_j^l) + (1-\alpha)U_j(\bar{\sigma}_{-ij}, s_i, s_j^l) - \\
&\quad - \alpha U_j(\bar{\sigma}) - (1-\alpha)U_j(\bar{\sigma}_{-i}, s_i)] = \\
&= (1-\alpha)\bar{\sigma}_j^l[U_j(\bar{\sigma}_{-ij}, s_i, s_j^l) - U_j(\bar{\sigma}_{-i}, s_i)] = \\
&= 0 \quad \forall l = 1, \ldots, K_j - 1, \quad \forall j \in \mathcal{N} \setminus \{i\}.
\end{aligned}$$

Also

$$\begin{aligned}(\alpha\bar{\sigma}_i^k + (1-\alpha)\delta_i^k(s_i))[U_i(\bar{\sigma}_{-i}, s_i^k) - \alpha U_i(\bar{\sigma}) - (1-\alpha)U_i(\bar{\sigma}_{-i}, s_i)] = \\ = (\alpha\bar{\sigma}_i^k + (1-\alpha)\delta_i^k(s_i))[U_i(\bar{\sigma}_{-i}, s_i^k) - U_i(\bar{\sigma})] = \\ = 0 \quad \forall k = 1, \ldots, K_i - 1\end{aligned}$$

where $\delta_i^k(s_i) = 1$, if $s_i = s_i^k$, and $\delta_i^k(s_i) = 0$, if $s_i \neq s_i^k$, from the assumption $U_i(\bar{\sigma}_{-i}, s_i) = U_i(\bar{\sigma})$. But then, for some α sufficiently close to 1, one obtains $\vec{b}(\alpha\bar{\sigma} + (1-\alpha)(\bar{\sigma}_{-i}, s_i)) = 0$, i.e. another zero of $\vec{b}$ arbitrarily close to $\bar{\sigma}$. But this contradicts the hypothesis of local asymptotic stability, because a zero which is not isolated cannot be locally asymptotically stable (more precisely: for $\alpha \in [0,1]$ sufficiently close to 1 an orbit starting from $\alpha\bar{\sigma} + (1-\alpha)(\bar{\sigma}_{-i}, s_i)$ does not converge to $\bar{\sigma}$, but remains where it starts). Hence $U_i(\bar{\sigma}_{-i}, s_i) < U_i(\bar{\sigma})$, $\forall s_i \notin supp(\bar{\sigma}_i)$, $\forall i \in \mathcal{N}$ ($\bar{\sigma} \in \mathrm{E}(\Gamma)$ is strict).

Since all U_i's are continuous on Σ, there $\exists \mathcal{O}$, $\mathcal{O}$ open, $\mathcal{O} \cap \mathrm{E}(\Gamma) = \{\bar{\sigma}\}$ with $\sigma^0 \in \mathcal{O} \cap \Sigma \Rightarrow U_i(\sigma_{-i}^0, s_i) \leq U_i(\sigma_{-i}^0, \bar{\sigma}_i)$; $\forall s_i \notin supp(\bar{\sigma}_i)$, $\forall i \in \mathcal{N}$. Since by Corollary 8 $\bar{\sigma} \in \mathrm{E}(\Gamma)$ is pure, this implies that $\bar{\sigma} \in BR(\sigma^0)$, $\forall \sigma^0 \in \mathcal{O} \cap \Sigma$, as required.

(ii) Now assume that there exists a neigbourhood $\mathcal{O}$ of $\bar{\sigma}$, with $\mathcal{O} \cap \mathrm{E}(\Gamma) = \{\bar{\sigma}\}$, such that $\bar{\sigma} \in BR(\sigma^0)$, $\forall \sigma^0 \in \mathcal{O} \cap \Sigma$. Define the function $V_{\bar{\sigma}} : \mathcal{O} \rightarrow \Re_+$ by

$$V_{\bar{\sigma}}(\sigma) = -\sum_{i \in \mathcal{N}} [(1 - \sum_{k=1}^{K_i - 1} \bar{\sigma}_i^k) \ln(1 - \sum_{k=1}^{K_i - 1} \sigma_i^k) + \sum_{k=1}^{K_i - 1} \bar{\sigma}_i^k \ln \sigma_i^k] \geq 0,$$

which is continuously differentiable on the interior of the boundary face in which $\bar{\sigma}$ lies. Moreover,

$$\frac{d}{dt} V_{\bar{\sigma}}(\sigma(t)) = \sum_{i \in \mathcal{N}} [U_i(\sigma(t)) - U_i(\sigma_{-i}(t), \bar{\sigma}_i)]$$

such that

$$\bar{\sigma} \in BR(\sigma),\ \sigma \in \mathcal{O} \cap \Sigma \Rightarrow \frac{d}{dt} V_{\bar{\sigma}}(\sigma) = \dot{V}_{\bar{\sigma}}(\sigma) \leq 0.$$

Now suppose that for some $\sigma \in \mathcal{O} \cap \Sigma$, $\sigma \neq \bar{\sigma}$, the equality $\dot{V}_{\bar{\sigma}}(\sigma) = 0$ would hold. Then $U_i(\sigma_{-i}, \bar{\sigma}_i) = U_i(\sigma)$, $\forall i \in \mathcal{N}$, such that $\bar{\sigma}_i \in BR_i(\sigma) \Rightarrow \sigma_i \in BR_i(\sigma)$, $\forall i \in \mathcal{N}$, implying that $\sigma \in BR(\sigma)$ or $\sigma \in \mathrm{E}(\Gamma)$. The latter, however, contradicts $\mathcal{O} \cap \mathrm{E}(\Gamma) = \{\bar{\sigma}\}$. Hence $\dot{V}_{\bar{\sigma}}(\sigma) <$

0, $\forall\sigma \in \mathcal{O} \cap \Sigma\backslash\{\bar{\sigma}\}$. This, now, combines to a demonstration, that $V_{\bar{\sigma}}$ is a Lyapunov function on $\mathcal{O}$, implying that $\bar{\sigma} \in \mathrm{E}(\Gamma)$ is a locally asymptotic stable stationary point of (3'). ∎

Proposition 1 says that wether an equilibrium is self-enforcing only as an singleton point or also with respect to a neighbourhood can be read off directly from the Nash field. That this is indeed a reflection of the classical instability problem of mixed equilibria is shown in the next two Lemmas:

LEMMA 6. *If no player $i \in \mathcal{N}$ has (a pair of) equivalent strategies, then for any $\bar{\sigma} \in \Sigma$ there exists a sequence $\{\sigma^k\}_{k=1}^{\infty}$, $\sigma^k \in \Sigma$, $\sigma^k \to \bar{\sigma}$, such that $BR_i(\sigma^k)$ contains only one pure strategy $\forall k$, $\forall i \in \mathcal{N}$.*

(PROOF: See Appendix)

LEMMA 7. *If no player $i \in \mathcal{N}$ has (a pair of) equivalent strategies, then for any $\bar{\sigma} \in E(\Gamma)$,*

$$\exists \mathcal{O},\ \mathcal{O}\ \text{open},\ \bar{\sigma} \in \mathcal{O} \cap E(\Gamma) : \bar{\sigma} \in BR(\sigma^0),\ \forall \sigma^0 \in \mathcal{O} \cap \Sigma$$

implies that $\bar{\sigma}$ is a pure strategy combination.

(PROOF: See Appendix)

In fact Lemma 7 could be stated without proof, because it is an immediate consequence of Proposition 1, when combined with Lemma 5 and Corrollary 8. The Appendix, however, contains an independent proof of Lemma 7 in order to show that the instability of mixed equilibria can be arrived at without reference to the Nash field. But, still, the Nash field provides a compact way to summarize all the information on the stability/self-enforcement properties of Nash equilibria.

The instability problem of mixed equilibria has led **Harsanyi** [1973b] to introduce games with randomly disturbed payoffs. In this context Harsanyi shows that, when an ordinary game is viewed as the limit of a family of games with randomly disturbed payoffs (when the disturbances approach zero), players may be forced to use a mixed strategy combination, because this may be their unique best reply to the uncertainty about the deviations of opponents induced by small payoff-disturbances. It also turns out that the - possibly mixed - equilibria which survive this test for small payoff disturbances are the regular equilibria [**Harsanyi**, 1973b, Lemma 9].

(Insert Figure 1 about here.)

But is a mixed regular equilibrium also sensible, if there is an alternative pure regular equilibrium? Consider the game schematically depicted in Figure 1. This is a 2×2 game of the "Battle-of-the-sexes" variety. In Figure 1 the unit square depicts Σ, the bold lines graph the best-reply correspondences of the two players and the arrows portray the Nash field. Now suppose there are small, but non-zero, payoff disturbances in Harsanyi's sense and what Figure 1 depicts, is player 1's (who controls, say, the horizontal axis) perception of the game, given the observation of his payoff vector. Suppose player 1 considers playing the mixed equilibrium C. Because there are payoff disturbances, he must take into account that his opponent may have a slightly different perception of the game. In fact, for an atomless distribution of payoff disturbances, the probability, given player 1's observation, that player 2 has exactly the same perception of the game is zero. But if player 2 is lead, by his slightly deviant perception of the game, to choose something which does not exactly match point C, then player 1 better responds with a pure strategy. If player 2 realizes this, he may not even consider to play a mix, but rather confine himself to figuring out which pure strategy he should choose. For the pure strategy equilibria (A and B) no analogous argument goes through, because they are strict and Propsition 1 applies. Hence in the presence of strict equilibria, the instability problem, instead of being resolved by payoff disturbances, provides strong support for ruling out the mixed equilibrium, even if it is regular. The unstable behavior of the Nash field around the mixed equilibrium reflects this information.

The question arises, of what will happen, if the unique Nash equilibrium of some game is mixed. An example of this is illustrated in Figure 2. Again in Figure 2 the unit square is Σ and the bold lines graph the best-reply correspondences. Point C is the unique completely mixed equilibrium and the arrows portray the Nash field. This time the orbits of (3') would display persistent oscillation[1].

(Insert Figure 2 about here.)

What Figure 2 portrays is a 2×2 game of the "Matching-pennies" variety. Indeed many people, who are not prepared to accept mixed equilibria without a second thought, would respond, when confronted with such a game, with being puzzled by the circularity induced among pure strategy choices. In our view this is a natural reaction to a "Matching-pennies" game, when the whole game, instead of just the equilibrium, is taken into consideration. The orbits of (3') in a "Matching-pennies"

[1]A warning is added here: In three dimensions it is easy to construct analogous "Matching-pennies" games, where the orbits are not closed and the unique equilibrium is unstable.

game are described by Lotka-Volterra equations. Hence the evolutionary process associated with a "Matching-pennies" game could easily be given an interpretation of the prey-predator type. From a biological point of view a prediction of a stationary mix C would seem a little surprising, given a prey-predator type situation.

This is to say that, in our view, the Nash field, even if there is no strict equilibrium, contains information about the structure of the interaction which is meant to be modelled by the game. Why should all social interactions which can be modelled by a normal form game be stable situations? In our view, it is a virtue of game theory that it can model situations which are inherently unstable. All we are suggesting is that one takes account of the type of situation which is modelled. And the tool to take account of the type of situation (which the game reflects) is the Nash field.

Although regular equilibria may be viewed as a refinement concept for Nash equilibria, the Nash field is not a refinement. It simply adds extra information to the (possibly) sharp, but less informative, predictions of the Nash equilibrium. However, a smart use of the Nash field can at least help with equilibrium selection. One instance of this occurs, when a game has a regular equilibrium. For predictive purposes a regular equilibrium is certainly preferable to an irregular one. But not all games have regular equilibria.

Recall from **Kohlberg and Mertens** [1986, Proposition 1] that the set of Nash equilibria of any game consists of finitely many connected components. Moreover, each game has at least one component which is such that any other game in a neighbourhood has an equilibrium in the neighbourhood of the component. The next and, possibly, major result of the present paper shows that this component can be identified via slightly perturbing the Nash field.

For a given game $\Gamma \in \mathrm{G}(S_1, \dots, S_n)$ and a given $\varepsilon > 0$ let (as in Theorem 3) the ε-perturbation of the Nash field, denoted $\vec{b}_\varepsilon$, be defined by

$$b^k_{\varepsilon,i}(\sigma) = b^k_i(\sigma) + \varepsilon(1 - \sigma^k_i K_i), \quad k = 1, \dots, K_i - 1, \ \forall i \in \mathcal{N}.$$

Denote by $\vec{b}_\varepsilon^{-1}(0) \subset int\, \Sigma$ the preimage of zeros ($0 \in \Re^M$) of the ε-pertubation.

THEOREM 4. *Let $C \subset E(\Gamma)$ be a (connected) component of Nash equilibria. If $C \subset E(\Gamma)$ is such that for all neighbourhoods $\mathcal{O}$ of C there exists an $\varepsilon > 0 : \vec{b}_\varepsilon^{-1}(0) \cap \mathcal{O} \neq \emptyset$, then there exists a neighbourhood $\mathcal{U}$ of Γ in $G(S_1, \dots, S_n)$ such that $E(\Gamma') \cap closure(\mathcal{O}) \neq \emptyset$ for all $\Gamma' \in \mathcal{U}$.*

PROOF: Identify a game $\Gamma \in G(S_1, \ldots, S_n)$ with its payoff vector $u \in \Re^{nK}$ and denote by $b(u, \sigma) = \vec{b}_u(\sigma)$ the mapping which assigns to each game $\Gamma = \Gamma(u)$, $u \in \Re^{nK}$, its Nash field. Then define the mapping

$$g : \Re^{nK} \times int\ \Sigma \times \Re_{++} \to \Re^M$$

which assigns to each game all of the ε-perturbation of its Nash field, by

$$g_i^k(u, \sigma, \varepsilon) = b_i^k(u, \sigma) + \varepsilon(1 - \sigma_i^k K_i), \quad k = 1, \ldots, K_i - 1, \ \forall i \in \mathcal{N}.$$

Since $\Re^{nK}$, $int\ \Sigma$ and $\Re_{++}$ are differentiable manifolds, so is their product, such that g is a continuously differentiable mapping of smooth manifolds.

By definition of g, no zero of g can emerge at the boundary of Σ. Then Lemma A.2 (in the Appendix) implies that the Jacobian matrix $D_{(u,\sigma,\varepsilon)}g(\bar{u}, \bar{\sigma}, \bar{\varepsilon})$ is surjective, i.e. has maximal rank (=M), at any point, where $g(\bar{u}, \bar{\sigma}, \bar{\varepsilon}) = 0$. In particular, the quoted Lemma shows that $D_u g(\bar{u}, \bar{\sigma}, \bar{\varepsilon}) = D_u b(\bar{u}, \bar{\sigma})$ has full rank.

As a consequence, $0 \in \Re^M$ is a regular value of g. The Preimage Theorem [**Guillemin and Pollack**, 1974, p.21] then implies that $g^{-1}(0)$ is a differentiable manifold of dimension $\dim g^{-1}(0) = nK + 1$. The very same argument shows that each of the preimages for fixed $\varepsilon > 0$, i.e. $g^{-1}(0) \cap (\Re^{nK} \times \Sigma \times \{\varepsilon\})$, is a differentiable manifold of dimension nK.

Since for any game $\Gamma \in G(S_1, \ldots, S_n)$ and any $\varepsilon > 0$ there exists a zero of the ε-perturbation in $int\Sigma$ (cf. Theorem 3), the intersection of the manifold $\{\Gamma\} \times int\Sigma \times R_{++}$ with $g^{-1}(0)$ is non-empty for any game Γ. We will show now that the intersection $(\{\Gamma\} \times int\Sigma \times \Re_{++}) \cap g^{-1}(0)$ is in fact transversal: The tangent space to $g^{-1}(0)$ is just the kernel of the linear map $D_{(u,\sigma,\varepsilon)}g(\bar{u}, \bar{\sigma}, \bar{\varepsilon})$ [**Guillemin and Pollack**,1974, p.24] at a point $(\bar{u}, \bar{\sigma}, \bar{\varepsilon}) \in (\{\Gamma\} \times int\Sigma \times \Re_{++}) \cap g^{-1}(0)$, while $\{\Gamma\} \times int\Sigma \times \Re_{++}$ coincides with its tangent space, because it simply forms a set parallel to $\Sigma \times \Re_{++}$ through $\Gamma = \Gamma(\bar{u})$ with dimension $M + 1$. Since $D_u g(\bar{u}, \bar{\sigma}, \bar{\varepsilon}) = D_u b(\bar{u}, \bar{\sigma})$ has full rank (=M) (and is a linear map) by Lemma A.2 (in the Appendix), the kernel of $D_{(u,\sigma,\varepsilon)}g(\bar{u}, \bar{\sigma}, \bar{\varepsilon})$ spans $\Re^{nK} \times \Re$. Consequently, the intersection $(\{\Gamma\} \times int\Sigma \times \Re_{++}) \cap g^{-1}(0)$ is transversal. Again, the very same argument shows that this also holds for fixed $\varepsilon > 0$: The intersection $(\{\Gamma\} \times int\Sigma \times \{\varepsilon\}) \cap g^{-1}(0)$ is transversal.

It follows that $dim((\{\Gamma\} \times int\Sigma \times \Re_{++}) \cap g^{-1}(0)) = 1$ and for fixed $\varepsilon > 0$ analogously $dim((\{\Gamma\} \times int\Sigma \times \{\varepsilon\} \cap g^{-1}(0)) = 0$ [**Guillemin and Pollack**, 1974, p.30]. Hence for each $\Gamma = \Gamma(u)$, $u \in \Re^{nK}$, there exists a sequence $\{(u, \sigma(\varepsilon), \varepsilon)\}_{\varepsilon \downarrow 0}$ with $(u, \sigma(\varepsilon), \varepsilon) \in closure((\{\Gamma\} \times int\Sigma \times \Re_{++}) \cap g^{-1}(0))$ and with isolated elements for any $\varepsilon > 0$. Since Σ is

compact, the sequence must have a cluster point. By the argument in Theorem 3 such a cluster point forms a Nash equilibrium of Γ.
This combines to a demonstration that the differentiable and nK-dimensional manifolds $(\Re^{nK} \times int\Sigma \times \{\varepsilon\}) \cap g^{-1}(0)$, $\varepsilon > 0$, converge pointwise to at least some part of the graph $\mathcal{G}(\mathrm{E})$. But by Theorem 1 of Kohlberg and Mertens [**1986**] the graph $\mathcal{G}(\mathrm{E})$ is itself homeomorphic to $\Re^{nK}$. Hence $(\Re^{nK} \times int\Sigma \times \{\varepsilon\}) \cap g^{-1}(0)$, $\varepsilon > 0$, approximates all of $\mathcal{G}(\mathrm{E})$.

It follows that for all games $\Gamma \in G(S_1, \ldots, S_n)$ there exists a neighbourhood $\mathcal{U}$ of Γ in $G(S_1, \ldots, S_n)$, such that the closure of the projection of $(\mathcal{U} \times \Sigma \times \{\varepsilon\}) \cap g^{-1}(0)$ into Σ converges to $\mathrm{E}(\Gamma)$ as $\varepsilon \to 0$, i.e.

$$closure(proj_\Sigma\,(\mathcal{U} \times \Sigma \times \{\varepsilon\}) \cap g^{-1}(0)) \xrightarrow{\varepsilon \to 0} \mathrm{E}(\Gamma)$$

(where convergence is meant in the sense of the Hausdorff distance for compact sets), because $(\mathcal{U} \times \Sigma \times \{\varepsilon\}) \cap g^{-1}(0)$ converges pointwise to $\mathcal{G}(\mathrm{E}) \cap (\mathcal{U} \times \Sigma)$.

Suppose now that for some fixed game $\Gamma \in G(S_1, \ldots, S_n)$ the component $\mathrm{C} \subset \mathrm{E}(\Gamma)$ is such that $\forall \mathcal{O} \subset \Re^M$, $\mathcal{O}$ open, $\mathcal{O} \cap \mathrm{E}(\Gamma) \supset \mathrm{C} : \exists \varepsilon > 0 : \vec{b}_\varepsilon^{-1}(0) \cap \mathcal{O} \neq \emptyset$. Then for this $\varepsilon > 0$: $(\{\Gamma\} \times \mathcal{O} \times \{\varepsilon\}) \cap g^{-1}(0) \neq \emptyset$ and this implies $(\{\Gamma\} \times \mathcal{O} \times \Re_{++}) \cap g^{-1}(0) \neq \emptyset$. But because $dim((\{\Gamma\} \times \mathcal{O} \times \{\varepsilon\}) \cap g^{-1}(0)) = 1$ by the transversality argument, it follows that $(\{\Gamma\} \times \mathcal{O} \times \{\varepsilon'\}) \cap g^{-1}(0) \neq \emptyset$, for all $\varepsilon' > 0$. Hence one may consider the intersection $(\{\Gamma\} \times \mathcal{O} \times [0, \bar{\varepsilon}]) \cap g^{-1}(0)$ for some $\bar{\varepsilon} > 0$ sufficiently large. Then transversality implies that there exists a neighbourhood $\mathcal{V} \subset \mathcal{U}$ of Γ in $\mathrm{G}(S_1, \ldots, S_n)$, such that $(\{\Gamma'\} \times \mathcal{O} \times [0, \bar{\varepsilon}]) \cap g^{-1}(0) \neq \emptyset$, for all $\Gamma' \in \mathcal{V}$ [**Guillemin and Pollack**, 1974,p.35]. Also by transversality for fixed $\varepsilon \in (0, \varepsilon]$ each element of $(\{\Gamma'\} \times \Sigma \times \{\varepsilon\}) \cap g^{-1}(0)$ is an isolated point $(u', \sigma'(\varepsilon), \varepsilon)$, $\Gamma' = \Gamma(u')$, $\forall \Gamma' \in \mathcal{V}$, and one may consider sequences $\{(u', \sigma'(\varepsilon), \varepsilon)\}_{\varepsilon \downarrow 0}$ with $(u', \sigma'(\varepsilon), \varepsilon) \in (\{\Gamma'\} \times \Sigma \times \{\varepsilon\}) \cap g^{-1}(0)$ along constant $\Gamma' = \Gamma(u') \in \mathcal{V}$. By choosing $\Gamma' \in \mathcal{V}$ it is ensured that the elements $(u', \sigma'(\varepsilon), \varepsilon)$ stay in $\mathcal{V} \times \mathcal{O} \times (0, \bar{\varepsilon})$ for $\varepsilon \in (0, \bar{\varepsilon})$ sufficiently large. As ε shrinks, the coordinate $\sigma'(\varepsilon)$ of such a sequence could leave $\mathcal{O}$ for some or all $\Gamma' \in \mathcal{V} \setminus \{\Gamma\}$. But if this is the case, then, for some $\hat{\varepsilon} \in (0, \bar{\varepsilon}]$ sufficiently small the projection $proj_\Sigma(\mathcal{V} \times \Sigma \times \{\varepsilon\}) \cap g^{-1}(0)$ into Σ will strictly contain $\mathcal{O}$ and, therefore, C for all $\varepsilon \in (0, \hat{\varepsilon})$. But, since $\mathcal{V} \subset \mathcal{U}$ and $(\mathcal{V} \times \Sigma \times \{\varepsilon\}) \cap g^{-1}(0)$ converges to $\mathcal{G}(\mathrm{E}) \cap (\mathcal{V} \times \Sigma)$, this implies that, as $\varepsilon \to 0$, the set C will be strictly contained in a connected component of Nash equilibria of Γ, in contradiction to the hypothesis on C.

Therefore, since C has been chosen a full connected component, each sequence $\{(u', \sigma'(\varepsilon), \varepsilon)\}_{\varepsilon \downarrow 0}$ must satisfy for all $\varepsilon > 0: \; (u', \sigma'(\varepsilon), \varepsilon) \in$

$(\{\Gamma'\} \times \mathcal{O} \times \{\varepsilon\}) \cap g^{-1}(0)$. Because each cluster point of such a sequence is in $\mathrm{E}(\Gamma')$, it follows that $\mathrm{E}(\Gamma') \cap closure(\mathcal{O}) \neq \emptyset$, $\forall \Gamma' \in \mathcal{V}$. ∎

In the course of demonstrating Theorem 4, it has been shown that:

COROLLARY 9. *The graph of the Nash equilibrium correspondence, $\mathcal{G}(E)$, can be arbitrarily closely approximated by a differentiable manifold of dimension nK.*

The approximation in Corollary 9 is a one-parameter approximation. Also note that the reason, why Theorem 4 refers to a component, rather than to a single point, is that as $\varepsilon \to 0$ a piece of $g^{-1}(0) \cap (\Re^{nK} \times \Sigma \times \{\varepsilon\})$ may become "vertical", but still form a component of Nash equilibria. This nicely illustrates the type of bifurcation phenomena one may find, when walking along $\mathcal{G}(\mathrm{E})$: Before some point ceases to be an equilibrium, he becomes part of a component of equilibria which connects him with a point which is going to take over the role of an equilibrium (which is just the UHC property of $\mathcal{G}(\mathrm{E})$).

It may be of interest to observe that the ε-perturbation of the Nash field represents biologically an immigration. In this view it is not too surprising that stability against immigration implies stability against payoff disturbances. Observing that $C \subset \mathrm{E}(\Gamma)$ in Theorem 4 is hyperstable, it is easy to see that the assumption on $\mathrm{C} \subset \mathrm{E}(\Gamma)$ in Theorem 4 are sufficient for property (S) in the definition of the Stable Set [**Kohlberg and Mertens**, 1986, p.1027] to be satisfied:

PROPERTY (S). *S is a closed set of Nash equilibria of Γ satisfying: For any $\varepsilon > 0$ there exists some $\delta_0 > 0$ such that for any completely mixed strategy combination $(\sigma_1, \ldots, \sigma_n)$ and for any $\delta_1, \ldots, \delta_n$, with $0 < \delta_i < \delta_0$, $\forall i \in \mathcal{N}$, the perturbed game, where every strategy $s_i \in S_i$ of player $i \in \mathcal{N}$ is replaced by $(1-\delta_i)s_i + \delta_i\sigma_i$, has an equilibrium ε-close to S.*

PROPOSITION 2. *If $C \subset E(\Gamma)$ satisfies the assumption of Theorem 4, then it satisfies Property (S).*

PROOF: The idea of the proof is to use the technique of the proof of Theorem 2.4.3. in **van Damme**, [1987, p.34], i.e. to associate close games with any perturbed game.

Let $\varepsilon > 0$ be given and assume that the assumtions on $\mathrm{C} \subset \mathrm{E}(\Gamma)$ of Theorm 4 are satisfied. For any $\delta = (\delta_1, \ldots, \delta_n)$ and any completely mixed strategy combination $\sigma \in int\ \Sigma$ denote by $(\Gamma, (\delta, \sigma))$ the corresponding perturbed game. Also for any $\delta = (\delta_1, \ldots, \delta_n)$ and $\sigma \in int\ \Sigma$ define the normal form game $\Gamma^{(\delta,\sigma)} \in \mathrm{G}(S_1, \ldots, S_n)$ by

$$u_i^{(\delta,\sigma)}(s) = (1 - \delta_i)u_i(s_{-i}, s_i) + \delta_i U_i(s_{-i}, \sigma_i),$$

for all $s = (s_{-i}, s_i) \in S$ and all $i \in \mathcal{N}$. From definition, for any $\hat{\sigma}_{-i} \in \Sigma_{-i}$, then

$$U_i^{(\delta, \sigma)}(\hat{\sigma}_{-i}, s_i) = (1 - \delta_i)U_i(\hat{\sigma}_{-i}, s_i) + \delta_i U_i(\hat{\sigma}_{-i}, \sigma_i),$$

for all $s_i \in S_i$ and all $i \in \mathcal{N}$, and for any $\hat{\sigma} \in \Sigma$, then

$$U_i^{(\delta, \sigma)}(\hat{\sigma}_{-i}, \hat{\sigma}_i) = (1 - \delta_i)U_i(\hat{\sigma}_{-i}, \hat{\sigma}_i) + \delta_i U_i(\hat{\sigma}_{-i}, \sigma_i),$$

for all $i \in \mathcal{N}$. Consequently,

$$\begin{aligned} U_i(\hat{\sigma}_{-i}, s_i^h) < U_i(\hat{\sigma}_{-i}, s_i^k) \quad &\Longleftrightarrow \\ \Longleftrightarrow \quad U_i^{(\delta,\sigma)}(\hat{\sigma}_{-i}, s_i^h) &< U_i^{(\delta,\sigma)}(\hat{\sigma}_{-i}, s_i^k), \end{aligned}$$

for all $\hat{\sigma}_{-i} \in \Sigma_{-i}$ and all s_i^h and s_i^k in S_i, $\forall i \in \mathcal{N}$. From the definition of $\Gamma^{(\delta,\sigma)}$ it follows that, as $\delta \to 0$, $\Gamma^{(\delta,\sigma)} \to \Gamma$. Therefore, if δ is sufficient small, each $\Gamma^{(\delta,\sigma)}$ will have an equilibrium $\bar{\sigma}(\delta, \sigma) \in \mathrm{E}(\Gamma^{(\delta,\sigma)})$ close to $\mathrm{C} \subset \mathrm{E}(\Gamma)$, by Theorem 4.

It follows that, if $\exists\, s_i^k \in S_i$:

$$\begin{aligned} &U_i^{(\delta,\sigma)}(\bar{\sigma}_{-i}(\delta, \sigma), s_i^h) < U_i^{(\delta,\sigma)}(\bar{\sigma}_{-i}(\delta, \sigma), s_i^k), \quad s_i^k \neq s_i^h \quad \Longleftrightarrow \\ &\Longleftrightarrow \quad (1 - \delta_i)U_i(\bar{\sigma}_{-i}(\delta, \sigma), s_i^h) + \delta_i U_i(\bar{\sigma}_{-i}(\delta, \sigma), \sigma_i) < \\ &< (1 - \delta_i)U_i(\bar{\sigma}_{-i}(\delta, \sigma), s_i^k) + \delta_i U_i(\bar{\sigma}_{-i}(\delta, \sigma), \sigma_i), \quad s_i^h \neq s_i^k, \end{aligned}$$

then $\bar{\sigma}_i^h(\delta, \sigma) = 0$ by (2.1.9) in **van Damme**, [1987, p.23]. Hence $(1 - \delta_i)\bar{\sigma}_i^h(\delta, \sigma) + \delta_i\sigma_i^h = \delta_i\sigma_i^h$, for any $s_i^h \in S_i$ which satisfies the above condition. But by Lemma 2.2.2. in **van Damme** , [1987, p.26], this is sufficient to guarantee that $[(1-\delta_1)\bar{\sigma}_1(\delta, \sigma)+\delta_1\sigma_1, \ldots, (1-\delta_n)\bar{\sigma}_n(\delta, \sigma)+\delta_n\sigma_n] \in \mathrm{E}((\Gamma, (\delta, \sigma)))$. Since $[(1-\delta_1)\bar{\sigma}_1(\delta, \sigma)+\delta_1\sigma_1, \ldots, (1-\delta_n)\bar{\sigma}_n(\delta, \sigma)+\delta_n\sigma_n]$ converges by construction to some point in $\mathrm{C} \subset \mathrm{E}(\Gamma)$, as $\delta \to 0$, it follows that any perturbation $(\Gamma, (\delta, \sigma))$ has an equilibrium close to $\mathrm{C} \subset \mathrm{E}(\Gamma)$. ∎

To define the Stable Set **Kohlberg and Mertens**, [1986, p.1027] require on top of Property (S) that the set S be minimal with respect to Property (S), i.e. there is no other closed set satisfying Property (S) properly contained in S. Proposition 2 only says that $C \subset \mathrm{E}(\Gamma)$ contains a Stable Set. In **van Damme**, [1987, pp.267] it is shown that a Stable Set may not satisfy connectedness. Corollary 9 supports van Damme's view that this is undesirable. The minimality condition may require to select a point, where the differentiable approximation of the graph gets a "kink" in the limit.

To summarize on how the Nash field can help with equilibrium selection: If a game has strict equilibria, these seem preferable to any non-strict ones (the Nash field may even help to compare strict equilibria: see section 5). For games without strict, but with regular equilibria, the latter seem an arguable choice. For games without regular equilibria a slight (one-parameter) perturbation of the Nash field can help identifying a component which will be robust in the sense that close games will have equilibria close to the component. The latter seems to be the least one could ask from a selection outcome.

5. Risk Dominance in 2×2 Games

In developing a complete theory of equilibrium selection **Harsanyi and Selten** [1988] introduce the concept of Risk Dominance. The intuitive argument for this criterion works as follows: Suppose all players in a given game are certain that one of two possible equilibria will be played, but they are uncertain as to which of the two. In this state of confusion the players enter a process of expectation formation. Starting from a prior distribution over the actions of other players each player tries to improve his forecast of the behavior of his rivals by taking into account what a given vector of prior distributions over the actions will lead his opponents to do. Once a player has figured out what the responses of the other players to the priors will be, he adjusts his estimate and again calculates the consequences of this new distribution over the other players' actions. Where this process ends, the risk dominant equilibrium is located.

For the class of 2×2 games with two strict equilibrium points **Harsanyi and Selten** [1988, chp. 3.9] have formalized the notion of Risk Dominance in three axioms and they have shown that the risk dominant equilibrium is fully characterized by possessing the larger Nash-product. (For other games Harsanyi and Selten formalize Risk Dominance by the tracing-procedure.)

Translating the Nash-product property into terms of the Nash field gives a nice illustration of the information contained in the Nash field. First, it is quite commonplace that all 2×2 games with two strict equilibria are more or less of the type illustrated in Figure 1 in section 4. In Figure 1 the unit square is Σ, the bold lines are the graphs of the best reply correspondences of the two players and the points A, B and C are the three equilibria, two of which (A and B) are strict. The arrows portray the behavior of the Nash field: A and B are, as strict equilibria, locally asymptotically stable, while C is a saddle point. This already is sufficient to restrict attention to A and B. It is now tempting to argue that in Figure 1 knowledge of the Nash field already is sufficient to select

A as the "better" equilibrium, because "A absorbs a larger part of Σ than B does". But can this be made more precise? The answer is in the affirmative, if one takes into account the Nash-products of the two equilibria A and B.

Letting A being associated with the pure strategy combination (s_1^1, s_2^1) and B being associated with the pure strategy combination (s_1^2, s_2^2), the Nash-products, $NP(.)$, are given by

$$NP(A) = [u_1(s_1^1, s_2^1) - u_1(s_1^2, s_2^1)][u_2(s_1^1, s_2^1) - u_2(s_1^1, s_2^2)],$$
$$NP(B) = [u_1(s_1^2, s_2^2) - u_1(s_1^1, s_2^2)][u_2(s_1^2, s_2^2) - u_2(s_1^2, s_2^1)].$$

This shows that the Nash-products equal the determinants of the Jacobian matrix of the Nash field at the corresponding equilibria. But the determinant is just the volume of the image of the unit cube under the linear mapping $D_\sigma b(\bar{\sigma})$. In other words: The determinant of $D_\sigma b(\bar{\sigma})$ is the coefficient of contraction of (oriented) volume, in the sense that the volume of any figure is contracted by a factor of $-|D_\sigma b(\bar{\sigma})|$. This makes it precise, what was meant by the somewhat vague phrase "A absorbs more of Σ than B does". For the class of 2×2 games with two strict Nash equilibria it thus turns out that Risk Dominance is equivalent to the condition that the preferred equilibrium has the larger absolute value of the determinant of $D_\sigma b(\bar{\sigma})$ at the equilibrium.

It is worth mentioning that the above logic holds true for all games with a strict Nash equilibrium: The determinant of $D_\sigma b(\bar{\sigma})$ at the strict equilibrium $\sigma \in \mathrm{E}(\Gamma)$ always equals a somewhat generalized Nash-product,

$$|D_\sigma b(\bar{\sigma})| = \prod_{i \in \mathcal{N}} \prod_{s_i \notin supp(\bar{\sigma}_i)} [U_i(\bar{\sigma}_{-i}, s_i) - U_i(\bar{\sigma})] =$$
$$= (-1)^M \prod_{i \in \mathcal{N}} \prod_{s_i \notin supp(\bar{\sigma}_i)} [U_i(\bar{\sigma}) - U_i(\bar{\sigma}_{-i}, s_i)]$$

and always mesures the coefficient by which the strict equilibrium locally "absorbs its neighbourhood" (in a very similar sense as in **Kalai and Samet** [1984]).

6. Conclusions

The present paper has demonstrated that the structure of interaction of players in a given game at any point in the space of mixed strategies can be represented by a smooth vector field, called the Nash field. This finding, in our view, resolves a number of problems with the Nash equilibrium concept for finite normal form games. First, the rationality assumptions imposed on traditional players turn out to be a simple matter of conveniance rather than substance: Even if players are not rational, but smoothly revise their behavorial patterns in an evolutionary way, the Nash equilibrium as a solution concept remains valid (because a zero of the Nash field which is not an equilibrium can never be stable).

Beyond this observation the Nash field provides a convenient way of representing the type of strategic situation, presumably modelled by the game, as a tractable mathematical object. If the conflict modelled by the game is endangered by inherent instabilities, this may be reflected by the presence of only mixed equilibria and the unstable behavior of the Nash field around these equilibria (which may be more or less dramatically unstable). On the other hand, for example, an equilibrium resolution of an unanimity game is quite robust against disturbances and this is reflected in the presence of strict equilibria which form sinks of the Nash field.

Third, the Nash field reduces the problem of determing Nash equilibria to solving a system of equations and inspecting the solutions. Section 5 illustrates that even some equilibrium selection procedures are reduced to inspecting solutions of a system of equations.

Fourth, one may hope at least that various refinement concepts have counterparts in the behavior of the Nash field. In the present paper this was shown for the defining property of the Stable Set and for Risk Dominance in 2×2 games. Recalling that sequential equilibria are induced by proper equilibria of the corresponding normal form, this kind of analysis could potentially be extended to getting a geometric understanding of backward induction in the normal form.

Finally, an inquiry into slightly perturbed Nash fields has shown that the graph of the Nash equilibrium correspondence is, in a specific sense, "smoother" than one may expect at first glance. It can at least be smoothly approximated. Hopefully, this provides also a "smoothing" of the path towards a unified notion of strategic stability.

APPENDIX

LEMMA 1. *If $\bar{s}_i \notin supp(\sigma_i)$ and $b(\sigma) = 0$, then $[U_i(\sigma_{-i}, \bar{s}_i) - U_i(\sigma)] \in \Re$ is an eigenvalue of the Jacobian matrix $D_\sigma b(\sigma)$.*

PROOF: First suppose $s_i^k \notin supp(\sigma_i)$ is such that $k < K_i$. Since $s_i^k \notin$

$supp(\sigma_i) \iff \sigma_i^k = 0$, all off-diagonal elements in the row of $D_\sigma b(\sigma)$ corresponding to $s_i^k \in S_i$ are zero and the diagonal element is given by

$$\frac{\partial}{\partial \sigma_i^k} b_i^k(\sigma) = U_i(\sigma_{-i}, s_i^k) - U_i(\sigma)$$

and, therefore, is an eigenvalue of $D_\sigma b(\sigma)$. Next suppose $s_i^k \notin supp(\sigma_i)$ is such that $k = K_i$. Then subtract $[U_i(\sigma_{-i}, s_i^{K_i}) - U_i(\sigma)]$ from all diagonal elements of $D_\sigma b(\sigma)$ and sum the rows corresponding to s_i^h, $h = 1, \dots, K_i - 1$. This yields for the columns corresponding to s_i^l, $l = 1, \dots, K_i - 1$,

$$\begin{aligned} &\sum_{h=1}^{K_i-1} \sigma_i^h [U_i(\sigma_{-i}, s_i^{K_i}) - U_i(\sigma_{-i}, s_i^l)] + \\ &+ U_i(\sigma_{-i}, s_i^l) - U_i(\sigma) - [U_i(\sigma_{-i}, s_i^{K_i}) - U_i(\sigma)] = \\ &= (1 - \sum_{h=1}^{K_i-1} \sigma_i^h)[U_i(\sigma_{-i}, s_i^l) - U_i(\sigma_{-i}, s_i^{K_i})] = 0, \end{aligned}$$

because $s_i^{K_i} \notin supp(\sigma_i) \iff (1 - \sum_{h=1}^{K_i-1} \sigma_i^h) = 0$. For the columns corresponding to s_j^l, $j \in \mathcal{N} \backslash \{i\}$, $l = 1, \dots, K_j - 1$, this operation yields

$$\begin{aligned} &\sum_{h=1}^{K_i-1} \sigma_i^h [U_i(\sigma_{-ij}, s_i^h, s_j^l) - U_i(\sigma_{-ij}, s_i^h, s_j^{K_j}) - U_i(\sigma_{-j}, s_j^l) + \\ &+ U_i(\sigma_{-j}, s_j^{K_j})] = (1 - \sum_{h=1}^{K_i-1} \sigma_i^h)[U_i(\sigma_{-j}, s_j^l) - U_i(\sigma_{-j}, s_j^{K_j}) - \\ &- U_i(\sigma_{-ij}, s_i^{K_i}, s_j^l) + U_i(\sigma_{-ij}, s_i^{K_i}, s_j^{K_j})] = 0, \end{aligned}$$

$(\sigma_{-ij}, \hat{\sigma}_i, \hat{\sigma}_j) = (\sigma_1, \dots, \sigma_{i-1}, \hat{\sigma}_i, \sigma_{i+1}, \dots, \sigma_{j-1}, \hat{\sigma}_j, \sigma_{j+1}, \dots, \sigma_n)$. It follows that these $K_i - 1$ rows are linear dependent and, therefore, $[U_i(\sigma_{-i}, s_i^{K_i}) - U_i(\sigma)] \in \Re$ is an eigenvalue. ∎

THEOREM 1. *Let $\bar{\Gamma} \in G(S_1, \dots, S_n)$ and assume that $\bar{\sigma} \in E(\bar{\Gamma})$ is a regular equilibrium. Let $\bar{u} \in \Re^{nK}$ denote the payoff vector of $\bar{\Gamma}$. Then there exists a neighbourhood $\mathcal{U}$ of $\bar{u}$ in $\Re^{nK}$ and a neighbourhood $\mathcal{V}$ of $\bar{\sigma}$ in $\Re^M$, such that*

(i) $|E(\Gamma(u)) \cap \mathcal{V}| = 1$, $\forall u \in \mathcal{U}$, and

(ii) *the mapping $\sigma: \mathcal{U} \to \mathcal{V}$, $u \mapsto \sigma(u)$, where $\sigma(u)$ is the unique equilibrium of $\Gamma = \Gamma(u)$ in $\mathcal{V}$, is continuous.*

PROOF: Define the mapping $\bar{b}: \Re^{nk} \times \Re^M \to \Re^M$ by

$$\bar{b}_i^k(u, \sigma) = \sigma_i^k[U_i(\sigma_{-i}, s_i^k) - U_i(\sigma)] = b_{u,i}^k(\sigma),$$

$\forall k = 1, \ldots, K_i - 1$, $\forall i \in \mathcal{N}$. Since $\bar{\sigma} \in E(\bar{\Gamma}) = E(\Gamma(\bar{u}))$, one has $\bar{b}(\bar{u}, \bar{\sigma}) = 0$, and since $\bar{\sigma}$ is regular, $|D_\sigma \bar{b}(\bar{u}, \bar{\sigma})| \neq 0$. Then by the implicit function theorem there exists a neighbourhood $\mathcal{U}_0$ of $\bar{u}$ in $\Re^{nK}$ and a unique function $\sigma: \mathcal{U}_0 \to \Re^M$ such that σ is differentiable on $\mathcal{U}_0$, $\sigma(\bar{u}) = \bar{\sigma}$, and $\bar{b}(u, \sigma(u)) = 0$, $\forall u \in \mathcal{U}_0$.

Choose an open neighbourhood $\mathcal{V}$ of $\bar{\sigma}$ such that

$$\begin{aligned} &\sigma^{-1}(\mathcal{V}) \subseteq \mathcal{U}_0 \\ &\bar{\sigma}_i^k > 0 \implies \sigma_i^k > 0, \ \forall \sigma \in \mathcal{V}, \ \forall k = 1, \ldots, K_i - 1, \\ &\sum_{h=1}^{K_i-1} \bar{\sigma}_i^h < 1 \implies \sum_{h=1}^{K_i-1} \sigma_i^h < 1, \ \forall \sigma \in \mathcal{V}, \end{aligned}$$

for all $i \in \mathcal{N}$. By continuity of σ, $\sigma^{-1}(\mathcal{V})$ is an open set in $\mathcal{U}_0$, with $\bar{u} \in \sigma^{-1}(\mathcal{V})$. Define the continuous mapping

$$\lambda_{i,k}: \sigma^{-1}(\mathcal{V}) \to \Re, \ u \mapsto (u, \sigma(u)) \mapsto U_i(\sigma_{-i}(u), s_i^k) - U_i(\sigma(u)).$$

It follows from continuity of $\lambda_{i,k}$ that the set

$$W_{i,k} = \lambda_{i,k}^{-1}(-\infty, 0) \cap \sigma^{-1}(\mathcal{V})$$

is open. Consider now the following intersection

$$\mathcal{U} = \bigcap_{i \in \mathcal{N}} \ \bigcap_{s_i^k \notin supp(\bar{\sigma}_i)} W_{i,k}.$$

As a finite intersection of open sets, $\mathcal{U}$ is open and, because $\bar{u} \in W_{i,k}$ for all (i, k) such that $\bar{\sigma}_i^k = 0$ ($\bar{\sigma}$ is a regular and, therefore, quasi-strict equilibrium of $\bar{\Gamma} = \Gamma(\bar{u})$), it follows that $\mathcal{U}$ is an open neighbourhood of $\bar{u}$.

Next we show that for all $u \in \mathcal{U}$, $\sigma(u)$ is an equilibrium of $\Gamma(u)$:

$$\forall u \in \mathcal{U}: \ \sigma_i^k(u) = 0 \Rightarrow \bar{\sigma}_i^k = 0 \Rightarrow U_i(\sigma_{-i}(u), s_i^k) - U_i(\sigma(u)) < 0.$$

Moreover, $\bar{b}(u, \sigma(u)) = 0$ implies that $\sigma_i^k(u) > 0 \Rightarrow U_i(\sigma_{-i}(u), s_i^k) - U_i(\sigma(u)) = 0$, such that $\sigma(u)$ is an equilibrium of $\Gamma(u)$. ∎

THEOREM 2. *Almost all games $\Gamma \in G(S_1, \dots, S_n)$ have all equilibria regular.*

PROOF: First "slice" the polyhedron Σ in the following way: Set $\Sigma_M = int\,\Sigma$. Then for each $0 < m < M$ let Σ_m be the set of all interiors of all boundary faces of Σ with dimension m and denote by Σ^m a typical element of Σ_m, $\Sigma^m \in \Sigma_m$. Finally, let Σ_0 be the set of all "corners" of Σ (pure strategy combinations) and again denote by $\Sigma^0 \in \Sigma_0$ a typical point representing a pure strategy combination. For each $0 \leq m \leq M$ every $\Sigma^m \in \Sigma_m$ is a differentiable manifold without boundary of dimension m. Each of the sets Σ_m is finite.

Next identify $G(S_1, \dots, S_n)$ with the space of payoff vectors $u \in \Re^{nK}$ and let the dependence of the mapping b, defined in (3), on the payoff vector be expressed by writing b_u for it. Let $\mathcal{C}^\infty(\Sigma, \Re^M)$ be the set of all mappings taking Σ to $\Re^M$ which are infinitely often continuously differentiable. Define the mapping

$$b: \Re^{nK} \to \mathcal{C}^\infty(\Sigma, \Re^M) \quad \text{by} \quad b(u) = b_u .$$

(The same symbol b is used here as in (3) to avoid extra notation, because no confusion can arise.) Analogously denote by $b_u|\Sigma^m: \Sigma^m \to \Re^m$ the mapping b_u restricted to the boundary face Σ^m and define $b|\Sigma^m: \Re^{nK} \to \mathcal{C}^\infty(\Sigma^m, \Re^m)$ by setting $b|\Sigma^m(u) = b_u|\Sigma^m$. Also let the evaluation map $b^{ev}|\Sigma^m: \Re^{nK} \times \Sigma^m \to \Re^m$ be defined by $(u, \sigma) \mapsto b_u|\Sigma^m(\sigma)$. To ensure that these definitions yield something well defined, the following two Lemmas are needed:

LEMMA A.1. *For each $0 \leq m \leq M$ and any $u \in \Re^{nK}$*

$$Image(b^{ev}|\Sigma^m) \subseteq \Re^m , \; \forall \Sigma^m \in \Sigma_m .$$

PROOF: Lemma A.1 is a weak version of Lemma 2 and follows from the implications $\sigma_i^k = 0 \implies b_i^k(\sigma) = 0$ and $\sigma_i^k = 1 \implies b_i^k(\sigma) = 0$. ∎

LEMMA A.2. *For any $0 \leq m \leq M$ and each $\Sigma^m \in \Sigma_m$ the derivative $D_{(u,\sigma)} b^{ev}|\Sigma^m(\bar{u}, \bar{\sigma})$ is surjective, i.e. has rank m, at each point $(\bar{u}, \bar{\sigma}) \in \Re^{nK} \times \Sigma^m$ such that $b^{ev}|\Sigma^m(\bar{u}, \bar{\sigma}) = 0$.*

PROOF: Since $D_{(u,\sigma)} b^{ev}|\Sigma^m = [D_u b^{ev}|\Sigma^m, D_\sigma b^{ev}|\Sigma^m]$ it suffices now to show that $D_u b^{ev}|\Sigma^m(\bar{u}, \bar{\sigma})$ has rank m. Calculating partial derivatives yields

$$\frac{\partial b_i^k|\Sigma^m}{\partial u_i(s_{-i}, s_i^l)}(\bar{u}, \bar{\sigma}) = (\delta_{kl} - \bar{\sigma}_i^l)\bar{\sigma}_i^k \prod_{j \in \mathcal{N}\setminus\{i\}} \bar{\sigma}_j(s_{-i}) ,$$

where $\delta_{kl} = 0$, if $k \neq l$, and $\delta_{kl} = 1$, if $k = l$. The partial derivatives of $b_i^k|\Sigma^m$ with respect to the payoffs of any other player $j \neq i$ are zero. Therefore, it suffices to consider $D_{u_i} b_i^{ev}|\Sigma^m(\bar{u}, \bar{\sigma})$, where $u_i = [(u_i(s_{-i}, s_i))_{s_{-i} \in S_{-i}}]_{s_i \in S_i}$ is player i's payoff vector. Consider linear combinations of rows of $D_{u_i} b_i^{ev}|\Sigma^m(\bar{u}, \bar{\sigma})$: Along any column corresponding to $(s_{-i}, s_i^k) \in S$ the linear combination with weights α_l, $l = 1, \dots, K_i - 1$, over an arbitrary subset of rows yields

$$\bar{\sigma}_i^k \prod_{j \in \mathcal{N} \setminus \{i\}} \bar{\sigma}_j(s_{-i})[\alpha_k(1 - \bar{\sigma}_i^k) - \sum_{l \neq k} \alpha_l \bar{\sigma}_i^l] =$$
$$= \bar{\sigma}_i^k \prod_{j \in \mathcal{N} \setminus \{i\}} \bar{\sigma}_j(s_{-i})[\alpha_k - \sum_l \alpha_l \bar{\sigma}_i^l] .$$

If $s_i^k \notin supp(\bar{\sigma}_i)$, then this trivially equals zero. If $(s_i, s_i^k) \in supp(\bar{\sigma})$, an assumption that the rows are linear dependent would imply $\alpha_k = \alpha = \sum_l \alpha_l \bar{\sigma}_i^l$ with $\alpha \neq 0$, such that

$$\alpha_k - \sum_l \alpha_l \bar{\sigma}_i^l = \alpha(1 - \sum_l \bar{\sigma}_i^l) .$$

But the RHS of this equation can only equal zero, if the summation is over the entire support of $\bar{\sigma}_i$. This implies that

$$\text{rank}(D_{u_i} b_i^{ev}|\Sigma^m(\bar{u}, \bar{\sigma})) = |supp(\bar{\sigma}_i)| - 1 .$$

But by construction $\sum_{i \in \mathcal{N}} |supp(\bar{\sigma}_i)| - n = \dim\Sigma^m = m$, such that the Lemma follows. ∎

(Proof of Theorem 2 continued): Pick a $\Sigma^m \in \Sigma_m$, $0 \leq m \leq M$. The map $b^{ev}|\Sigma^m$ is infinitely often differentiable and by Lemma A.2 the $0 \in \Re^m$ is a regular value of $b^{ev}|\Sigma^m$. Then the parametric transversality theorem [**Hirsch**, 1976,p.79] states that the set

$$V_{\Sigma^m} = \{u \in \Re^{nK} \mid 0 \in \Re^m \text{ is a regular value of } b|\Sigma^m\}$$

is dense in $\Re^{nK}$.

Next let W be defined as the set of all $u \in \Re^{nK}$, such that the corresponding game $\Gamma = \Gamma(u)$ has only quasi- strict equilibria. From Theorem 2 in **Harsanyi** [1973a] it follows that the complement of W in $\Re^{nK}$ is a closed set with Lebesgue measure zero. Hence W is dense in $\Re^{nK}$ (Suppose not: Then there exists an open set $\mathcal{O}$ contained in the complement of the closure of W and a compact set $\mathcal{Q} \subset \mathcal{O}$. But then, since the measure of $\mathcal{Q}$ is non-zero, this must also be true for the measure of $\mathcal{O}$.

Because $\mathcal{O}$ is a subset of the complement of the closure of W it must be contained in the complement of W - a contradiction to Harsanyi's Theorem 2.). Now consider the set of $u \in \Re^{nK}$ which have only quasi-strict equilibria and all regular equilibria in every Σ^m, $\forall 0 \leq m \leq M$. Each u in this set satifies

$$u \in \bigcap_{0 \leq m \leq M} \bigcap_{\Sigma^m \in \Sigma_m} V_{\Sigma^m} \cap W = \bigcap V_{\Sigma^m} \cap W\,.$$

The Baire-Theorem [**Hirsch**, 1976, p.213] implies that $\bigcap V_{\Sigma^m} \cap W$ is dense in $\Re^{nK}$.

Finally, let $u \in \bigcap V_{\Sigma^m} \cap W$ and $\bar{\sigma}$ a zero of $b_u|\Sigma^m$. By elementary operations on determinants the following decomposition is obtained

$$|D_\sigma b_u(\bar{\sigma})| = \prod_{i \in \mathcal{N}} \prod_{s_i^k \notin supp(\bar{\sigma}_i)} (U_i(\bar{\sigma}_{-i}, s_i^k) - U_i(\bar{\sigma}))\, |D_\sigma b_u|\Sigma^m(\bar{\sigma})|\,,$$

if $\bar{\sigma} \in \Sigma^m$. Since all equilibria of $\Gamma(u)$ are quasi-strict and the determinant in the above decomposition is non-zero, all equilibria are regular. This holds for all u in the dense set $\bigcap V_{\Sigma^m} \cap W$. ∎

LEMMA 3. *Let $\bar{\sigma} \in \Sigma$ be the mixed strategy combination corresponding to the pure strategy combination $\bar{s} \in S$. Then all eigenvectors of $D_\sigma \vec{b}(\bar{\sigma})$ belonging to the eigenvalues $[U_i(\bar{\sigma}_{-i}, s_i^h) - U_i(\bar{\sigma})] \in \Re$, $\forall s_i^h \notin supp(\bar{\sigma}_i)$, $\forall i \in \mathcal{N}$, are given by $[(\bar{\sigma}_{-i}, s_i^h) - \bar{\sigma}] \in \Re^M$, $\forall s_i^h \notin supp(\bar{\sigma}_i)$, for all $i \in \mathcal{N}$.*

PROOF: Since $\bar{\sigma} \in \Sigma$ is pure (a "corner" of Σ) the vectors in the statement are $\sum_{i \in \mathcal{N}} K_i - n = M$ in number and all are linear independent, because they form the "edges" of the polyhedron Σ. It remains to verify that these vectors are eigenvectors.

First observe that for any $\hat{\sigma} \in \Sigma$, $\hat{\sigma} \neq \bar{\sigma}$, the product $D_\sigma \vec{b}(\bar{\sigma})(\hat{\sigma} - \bar{\sigma})$

is given by, for any $\bar{\sigma} \in \Sigma$ with $\vec{b}(\bar{\sigma}) = 0$,

$$[D_\sigma \vec{b}(\bar{\sigma})(\hat{\sigma} - \bar{\sigma})]_i^k = \sum_{h=1}^{K_i-1} (\frac{\partial}{\partial \sigma_i^h} b_i^k(\bar{\sigma}))(\hat{\sigma}_i^h - \bar{\sigma}_i^h) +$$
$$+ \sum_{j \in \mathcal{N} \setminus \{i\}} \sum_{l=1}^{K_j-1} (\frac{\partial}{\partial \sigma_j^l} b_i^k(\bar{\sigma}))(\hat{\sigma}_j^l - \bar{\sigma}_j^l) =$$
$$= \sum_{h=1}^{K_i-1} \bar{\sigma}_i^k [U_i(\bar{\sigma}_{-i}, s_i^{K_i}) - U_i(\bar{\sigma}_{-i}, s_i^h)](\hat{\sigma}_i^h - \bar{\sigma}_i^h) +$$
$$+ \sum_{j \in \mathcal{N} \setminus \{i\}} \sum_{l=1}^{K_j-1} \bar{\sigma}_i^k [U_i(\bar{\sigma}_{-ij}, s_i^k, s_j^l) - U_i(\bar{\sigma}_{-ij}, s_i^k, s_j^{K_j}) -$$
$$- U_i(\bar{\sigma}_{-j}, s_j^l) + U_i(\bar{\sigma}_{-j}, s_j^{K_j})](\hat{\sigma}_j^l - \bar{\sigma}_j^l) =$$
$$= \bar{\sigma}_i^k [U_i(\bar{\sigma}) - U_i(\bar{\sigma}_{-i}, \hat{\sigma}_i)] + \sum_{j \in \mathcal{N} \setminus \{i\}} \bar{\sigma}_i^k [U_i(\bar{\sigma}_{-ij}, s_i^k, \hat{\sigma}_j) - U_i(\bar{\sigma}_{-j}, \hat{\sigma}_j)]$$

for all $s_i^k \in supp(\bar{\sigma}_i)$, and by

$$[D_\sigma \vec{b}(\bar{\sigma})(\hat{\sigma} - \bar{\sigma})]_i^h = [U_i(\bar{\sigma}_{-i}, s_i^h) - U_i(\bar{\sigma})]\hat{\sigma}_i^h$$

for any $s_i^h \notin supp(\bar{\sigma}_i)$, $\forall i \in \mathcal{N}$.

Now let $\hat{\sigma} = (\bar{\sigma}_{-i}, s_i^h)$ for some $s_i^h \notin supp(\bar{\sigma}_i)$. Then for $s_i^k \in supp(\bar{\sigma}_i)$ one obtains ($h \neq k$)

$$[D_\sigma \vec{b}(\bar{\sigma})((\bar{\sigma}_{-i}, s_i^h) - \bar{\sigma})]_i^k = -[U_i(\bar{\sigma}_{-i}, s_i^h) - U_i(\bar{\sigma})]\bar{\sigma}_i^k\,,$$

and for $s_i^l \notin supp(\bar{\sigma}_i)$, $l \neq h$, one obtains

$$[D_\sigma \vec{b}(\bar{\sigma})((\bar{\sigma}_{-i}, s_i^h) - \bar{\sigma})]_i^l = 0.$$

For $j \in \mathcal{N} \setminus \{i\}$ and $s_j^l \in supp(\bar{\sigma}_j)$ one obtains

$$[D_\sigma \vec{b}(\bar{\sigma})((\bar{\sigma}_{-i}, s_i^h) - \bar{\sigma})]_j^l = \bar{\sigma}_j^l [U_j(\bar{\sigma}) - U_j(\bar{\sigma})] +$$
$$+ \sum_{\mathcal{N} \setminus \{i,j\}} \bar{\sigma}_j^l [U_j(\bar{\sigma}_{-j}, s_j^l) - U_j(\bar{\sigma})] +$$
$$+ \bar{\sigma}_j^l [U_j(\bar{\sigma}_{-ij}, s_i^h, s_j^l) - U_j(\bar{\sigma}_{-i}, s_i^h)] = 0,$$

because $\bar{\sigma} \in \Sigma$ has been assumed pure, such that for $s_j^l \in supp(\bar{\sigma}_j)$ one has $(\bar{\sigma}_{-ij}, s_i^h, s_j^l) = (\bar{\sigma}_{-i}, s_i^h)$. Finally, for $s_j^l \notin supp(\bar{\sigma}_j)$ the same operation yields again zero, because this row only contains one non-zero element (at the diagonal) which gets multiplied by zero. Calculating now

$$[D_\sigma \vec{b}(\bar{\sigma})((\bar{\sigma}_{-i}, s_i^h) - \bar{\sigma})]_i^h = U_i(\bar{\sigma}_{-i}, s_i^h) - U_i(\bar{\sigma})$$

completes the verification that $((\bar{\sigma}_{-i}, s_i^h) - \bar{\sigma})$, $s_i^h \notin supp(\bar{\sigma}_i)$, is indeed the desired eigenvector. ∎

LEMMA 6. *If no player $i \in \mathcal{N}$ has (a pair of) equivalent strategies, then for any $\bar{\sigma} \in \Sigma$ there exists a sequence $\{\sigma^k\}_{k=1}^{\infty}$, $\sigma^k \in \Sigma$, $\sigma^k \to \bar{\sigma}$, such that $BR_i(\sigma^k)$ contain only one pure strategy $\forall k$, $\forall i \in \mathcal{N}$.*

PROOF: Let $\sigma \in \Sigma$ and observe that $s_i \notin BR_i(\sigma) \iff$

$$\iff U_i(\sigma_{-i}, s_i) < \max_{\bar{s}_i \in S_i} U_i(\sigma_{-i}, \bar{s}_i).$$

Since the LHS of the inequality is continuous in $\sigma_{-i} \in \Sigma_{-i}$ by definition and the RHS is continuous by the maximum theorem, there exists a neighbourhood $\mathcal{O}_\sigma$ of σ such that

$$\begin{aligned} &\sigma^0 \in \mathcal{O}_\sigma \cap \Sigma \quad \text{and} \quad s_i \notin BR_i(\sigma) \implies \\ &\implies U_i(\sigma^0_{-i}, s_i) < \max_{\hat{s}_i \in S_i} U_i(\sigma^0_{-i}, \hat{s}_i) \implies \\ &\implies s_i \notin BR_i(\sigma^0). \end{aligned}$$

Hence $BR_i(\sigma) \supset BR_i(\sigma^0)$, $\forall \sigma^0 \in \mathcal{O}_\sigma \cap \Sigma$, $\forall i \in \mathcal{N}$. If now $BR_i(\bar{\sigma})$, for some $\bar{\sigma} \in \Sigma$, contains only one pure strategy, $\forall i \in \mathcal{N}$, the Lemma is established. Therefore, assume that, for some $i \in \mathcal{N}$, $BR_i(\bar{\sigma})$ contains more than one pure strategy. Let $\{\sigma^k\}_{k=1}^{\infty}$, $\sigma^k \in \Sigma$, $\sigma^k \to \bar{\sigma}$, be a sequence converging to $\bar{\sigma} \in \Sigma$. By the above argument $\{\sigma^k\}_{k=1}^{\infty}$ can be chosen such that $BR_i(\sigma^k) \supset BR_i(\sigma^{k+1})$, $\forall k$, which implies that the number of pure strategies in $BR_i(\sigma^k)$ is decreasing in k. Also, once $BR_i(\sigma^{k'})$, for some $k' \geq 1$, contain only one pure strategy, the $BR_i(\sigma^k)$, $\forall k \geq k'$, contain only one pure strategy, and the desired sequence has been found.

By Lemma 5 of **Kalai and Samet** [1984] for every neighbourhood $\mathcal{O}_{\bar{\sigma}}$ of $\bar{\sigma}$ there exists some $\hat{\sigma} \in \mathcal{O}_{\bar{\sigma}}$, such that, if $s_i \in BR_i(\hat{\sigma}) \cap BR_i(\bar{\sigma})$ and $s_i' \in BR_i(\hat{\sigma}) \cap BR_i(\bar{\sigma})$, $s_i \neq s_i'$, then s_i and s_i' are equivalent. Consequently, if no player has equivalent strategies, then there must always be some $\sigma^1 \in \mathcal{O}_{\bar{\sigma}}$, such that $BR_i(\sigma^1) = \{s_i\}$, for some $s_i \in BR_i(\bar{\sigma})$. Taking this $\sigma^1 \in \mathcal{O}_{\bar{\sigma}}$ as the first element of $\{\sigma^k\}_{k=1}^{\infty}$ and choosing the sequence such that $BR_i(\sigma^k) \supset BR_i(\sigma^{k+1})$ establishes the desired conclusion. ∎

LEMMA 7. *If no player $i \in \mathcal{N}$ has (a pair of) equivalent strategies, then for any $\bar{\sigma} \in E(\Gamma)$,*

$$\exists \mathcal{O},\ \mathcal{O} \text{ open},\ \bar{\sigma} \in \mathcal{O} \cap E(\Gamma) : \bar{\sigma} \in BR(\sigma^0),\ \forall \sigma^0 \in \mathcal{O} \cap \Sigma,$$

implies that $\bar{\sigma}$ is a pure strategy combination.

PROOF: Suppose $\bar{\sigma} \in \mathrm{E}(\Gamma)$ is mixed, i.e. $\exists i \in \mathcal{N} : \{s_i, s_i'\} \subset supp(\bar{\sigma}_i)$, with $s_i \neq s_i'$, and $\exists \mathcal{O}$, $\mathcal{O}$ open, $\bar{\sigma} \in \mathcal{O} \cap \mathrm{E}(\Gamma) : \bar{\sigma} \in BR(\sigma^0)$, $\forall \sigma^0 \in \mathcal{O} \cap \Sigma$. From the definition of an equilibrium $supp(\bar{\sigma}_i) \subset BR_i(\bar{\sigma})$. By Lemma 6 there exists a sequence $\{\sigma^k\}_{k=1}^{\infty}$, $\sigma^k \to \bar{\sigma}$, such that $BR_i(\sigma^k)$ contains only one pure strategy, say s_i^k, for all k. For sufficiently large k one must have $\sigma^k \in \mathcal{O} \cap \Sigma$, but $\bar{\sigma}_i \notin BR_i(\sigma^k) = \{s_i^k\}$, a contradiction. Hence $\bar{\sigma} \in \mathrm{E}(\Gamma)$ cannot be mixed. ∎

REFERENCES

Amann E. and J. Hofbauer, *Permanence in Lotka-Volterra and replicator equations*, in "Lotka-Volterra Approach to Cooperation and Competition in Dynamic Systems," W. Ebeling and M. Peschel (eds.), Akademie-Verlag, Berlin, 1985.

Chillingworth D.R.J., "Differential topology with a view to applications," Pitman Publishing, 1976.

Dierker E., *Two Remarks on the Number of Equilibria of an Economy*, Econometrica **40** (1972), 951-953.

Dieudonné J., "Foundations of Modern Analysis," Academic Press, 1969.

Fudenberg D., D.M. Kreps and D.K. Levine, *On the Robustness of Equilibrium Refinements*, Journal of Economic Theory **44** (1988), 354-380.

Guillemin V. and A. Pollack, "Differential Topology," Prentice-Hall, 1974.

Harsanyi J.C., *Oddness of the Number of Equilibrium Points: A New Proof*, International Journal of Game Theory **2** (1973a), 235-250.

————, *Games with Randomly Disturbed Payoffs: A New Rational for Mixed-Strategy Equilibrium Points*, International Journal of Game Theory **2** (1973b), 1-23.

Harsanyi J.C. and R. Selten, "A General Theory of Equilibrium Selection in Games," MIT-press, 1988.

Hirsch M.W., "Differential Topology," Springer-Verlag, 1976.

Hofbauer J. and K. Sigmund, "The Theory of Evolution and Dynamical Systems," Cambridge University Press, 1988.

Kalai E. and D. Samet, *Persistent Equilibria in Strategic Games*, International Journal of Game Theory **13** (1984), 129-144.

Kohlberg E. and J.-F. Mertens, *On the Strategic Stability of Equilibria*, Econometrica **54** (1986), 1003-1037.

Kojima M., A. Okada and S. Shindoh, *Strongly Stable Equilibrium Points of N-Person Noncooperative Games*, Math. of Operations Research **10** (1985), 650-663.

Kreps D.M. and R. Wilson, *Sequential Equilibrium*, Econometrica **50** (1982), 863-894.

Myerson R.B., *Refinement of the Nash Equilibrium Concept*, International Journal of Game Theory **7** (1978), 73-80.

Nash J., *Non-Cooperative Games*, Annals of Mathematics **54** (1951), 286-295.

Okada A., *On Stability of Perfect Equilibrium Points*, International Journal of Game Theory **10** (1981), 67-73.

Selten R., *Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit*, Zeitschrift für die gesamte Staatswissenschaft **12** (1965), 301-324 and 667-689.

________, *Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games*, International Journal of Game Theory **4** (1975), 25-55.

van Damme E., "Stability and Perfection of Nash Equilibria," Springer-Verlag, 1987.

____________, *A Relation between Perfect Equilibria in Extensive Form Games and Proper Equilibria in Normal Form Games*, International Journal of Game Theory **13** (1984), 1-13.

von Neumann J. and O. Morgenstern, "Theory of Games and Economic Behavior," Princeton University Press, (1944) 1972.

Wilson R., *Computing Equilibria of n-Person Games*, Siam Journal of Applied Mathematics **21** (1971), 80-87.

Wu Wen-Tsün and Jiang Jia-He, *Essential Equilibrium Points of n-Person Non-Cooperative Games*, Scientia Sinica **11** (1962), 1307-1322.

Keywords. normal form games, evolution, stability

Klaus Ritzberger, Institute for Advanced Studies, Department of Economics, Stumpergasse 56, A-1060 Vienna, Austria
Karl Vogelsberger, Institute for Advanced Studies, Department of Mathematics, Stumpergasse 56, A-1060 Vienna, Austria

FIGURE 1

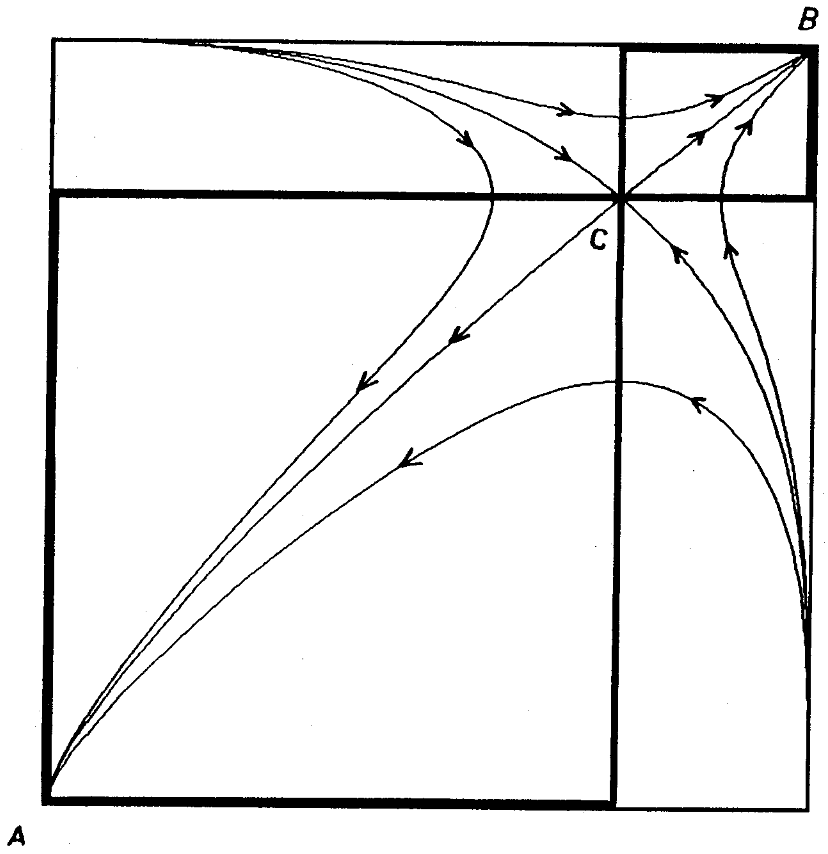

FIGURE 2