

MATHEMATICS OF OPERATIONS RESEARCH  
Vol. 26, No. 3, August 2001, pp. 611–635  
Printed in U.S.A.

## ON THE RELATION AMONG SOME DEFINITIONS OF STRATEGIC STABILITY

JOHN HILLAS, MATHIJS JANSEN, JOS POTTERS, AND DRIES VERMEULEN

In this paper we examine a number of different definitions of strategic stability and the relations among them. In particular, we show that the stability requirement given by Hillas (1990) is weaker than the requirements involved in the various definitions of stability in Mertens' reformulation of stability (Mertens 1989, 1991). To this end, we introduce a new definition of stability and show that it is equivalent to (a variant of) the definition given by Hillas (1990). We also use the equivalence of our new definition with the definition of Hillas to provide correct proofs of some of the results that were originally claimed (and incorrectly "proved") in Hillas (1990).

**1. Introduction.** The theory of strategic stability is an attempt to answer the question: What are the self-enforcing outcomes of a game? The theory is based on the intuition that the answer to such a question should (1) involve somewhat more than just the conditions of Nash equilibrium, and (2) should not depend on "irrelevant" aspects of the game. While much of the literature on the refinement of equilibrium might be thought of in this way, the term "strategic stability" was introduced by Kohlberg and Mertens (1986) who gave the first analysis systematically based on such an approach.

Kohlberg and Mertens gave a list of requirements that a concept of strategic stability should satisfy. They showed that even quite weak versions of their requirements implied that the solution concept should assign sets of equilibria as solutions to the game. Thus a stability concept is a rule that assigns a collection of subsets of the space of (mixed) strategy profiles to each game in the domain of games under consideration. Since the paper of Kohlberg and Mertens, the list of requirements that a concept of strategic stability should satisfy has been modified and expanded, particularly in the work of Mertens (1987, 1989, 1991, 1992). We shall not be concerned, in this paper, with the justification of these requirements. The interested reader is referred to Kohlberg and Mertens (1986), Mertens (1987, 1989, 1991, 1992), or Hillas and Kohlberg (1994).

Up until now, several attempts have been made to construct a solution concept that satisfies all requirements (see, for example, Kohlberg and Mertens 1986, Mertens 1989, Hillas 1990, McLennan 1989a, and Vermeulen et al. 1997). Of these attempts only Mertens (1989) succeeded completely. All other known versions of strategic stability fail on at least one of the requirements.

**Aim of the paper.** Insight into the relationships between various types of strategic stability is significant from a conceptual (as well as, occasionally, a computational) point of view, since these relations can simplify the derivation of the requirements to a great extent. Some relations between these types are already known. For instance, Govindan (1995) proved that each strategically stable set in the sense of Mertens (1989) contains a fully stable set as they are defined in Hillas (1990). A fairly complete overview of how the most important types hang together is given in Figure 1.

Received September 21, 1999; revised March 8, 2000, and January 4, 2001.

*MSC 2000 subject classification.* Primary: 91A06.

*OR/MS subject classification.* Primary: Games/group decisions.

*Key words.* Game theory, strategic stability.

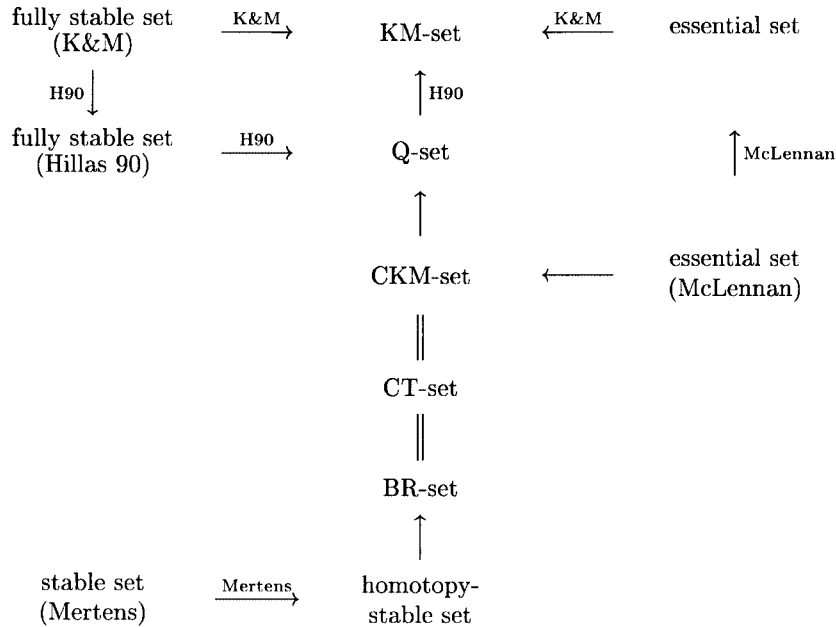


FIGURE 1

A connection between the two most commonly used notions of strategic stability (stable sets in the sense of Mertens 1989 and Hillas 1990, respectively) was not yet established. This brings us to the main goal of this paper. We will show that the reformulation of Mertens in fact yields a stronger form of strategic stability than the one in Hillas (1990). After this fact is established we will elaborate on some results and simplifications of proofs of known results that arise as by-products from our proofs.

**Organization of the paper.** First in §2 we will introduce the notational conventions used in this paper. In §3 we will give the specific versions of the three concepts that are central in this paper. The first one, homotopy stability, is a simplified (weak) version of the reformulation in Mertens (1989). The second one, best reply stability or BR-stability for short, is a slightly simplified version of stability in the sense of Hillas (1990). (We will explain in §6.2 why we use the simplification instead of the original definition.) The third one, CKM-stability, can be seen as a continuous version of the definition of stability in Kohlberg and Mertens (1986). A similar definition can be found in Vermeulen et al. (1997).

In §4 we show that homotopic stability is a stronger requirement than CKM-stability. In §5 we show that CKM-stability and BR-stability are equivalent notions. In the proof we need yet another type of stability, called CT-stability. First we will show that CKM-stability implies CT-stability, and CT-stability implies BR-stability. Since the converse implications are quite trivial, this yields the equivalence of these notions.

Combining the results from §§4 and 5 we see that homotopy stability implies BR-stability. Hence, since homotopy stability is weaker than the definition of stability in Mertens (1989) in terms of homology groups, we get the main result of this paper that stability in the sense of Mertens (1989) implies stability in the sense of Hillas (1990).

Finally, in §6 we will use the equivalence of BR-stability and CKM-stability to give correct proofs of two properties of BR-stability that are already mentioned in Hillas (1990) (in fact the proofs of these properties turn out to be relatively simple for the equivalent notion of CKM-stability) and to prove that BR-stable sets also satisfy abr-invariance.

To give the reader an overview of what is basically accomplished in this paper, most relations between various stability concepts known to us as well as the ones that are proved here are displayed in Figure 1. The relations marked K&M were proved in Kohlberg and Mertens (1986); those marked H90 in Hillas (1990); that marked McLennan in McLennan (1989a); and that marked Mertens in Mertens (1989). The unmarked relations are either obvious or proved in this paper.

A number of stability concepts used in the diagram do not occur in this paper. The definitions of full stability (Kohlberg and Mertens) and essential set can be found in Kohlberg and Mertens (1986). The definitions of fully stable sets and Q-sets (quasi stable sets) can be found in Hillas (1990). In the diagram “A-set→B-set” means that every A-set is also a B-set. If we define an A-stable set (B-stable set) to be a minimal A-set (B-set), then the arrow implies that every A-stable set *contains* a B-stable set.

**2. Preliminaries.** We first specify some notational conventions. Most of our notation is completely conventional. One exception is that in defining open balls in  $\mathbb{R}^n$  and in defining the Hausdorff distance between compact subsets of  $\mathbb{R}^n$  we use the maximum norm rather than the Euclidean norm. For any subset  $A$  of  $\mathbb{R}^n$  we denote the convex hull of  $A$  by  $ch(A)$  and the affine hull of  $A$  by  $ah(A)$ . We are only concerned with the boundaries or interiors of convex sets and we always mean the boundaries or interiors relative to the affine hulls of these sets. That is, if  $A$  is a closed and convex subset of  $\mathbb{R}^n$ , the boundary  $\partial A$  of  $A$  is the collection of those points  $x$  in  $A$  for which every neighborhood  $U$  of  $x$  has a nonempty intersection with  $ah(A) \setminus A$ . Further,  $\overset{\circ}{A} = A \setminus \partial A$  is the relative interior of  $A$ . For a correspondence  $\varphi: X \rightarrow X$  we denote by  $fix(\varphi)$  the set of fixed points of  $\varphi$ .

A *finite n-person game* (in normal form) is a pair  $\Gamma = \langle A, u \rangle$ , where  $A = \times_{i \in N} A_i$  is a product of finite (nonempty) sets and  $u = (u_1, \dots, u_n)$  is an  $n$ -tuple of functions  $u_i: A \rightarrow \mathbb{R}$ . The set  $A_i$  is called the set of *pure strategies* of player  $i$  and  $u_i$  his *payoff function*. We abusively also use  $u_i$  to denote the multilinear extension of  $u_i$  to the space  $\Delta = \times_{i \in N} \Delta(A_i)$  of strategy profiles, i.e., to denote player  $i$ 's *expected payoff function*. (The set  $\Delta(A_i)$  is player  $i$ 's set of mixed strategies, i.e., the set of probability vectors on  $A_i$ .)

In what follows we use the following simplified notation. We omit prefixes and simply call  $\Gamma$  a game. We write  $\Delta_i$  instead of  $\Delta(A_i)$ ,  $\Delta_{-i} = \times_{j \neq i} \Delta_j$  for the set of strategy profiles of the opponents of player  $i$ , and  $(x_{-i} | y_i)$  in  $\Delta$  for the strategy profile in which player  $i$  uses  $y_i \in \Delta_i$  and his opponents use the strategies  $x_{-i}$  in  $\Delta_{-i}$ . The correspondence  $BR_i: \Delta \rightarrow \Delta_i$  associates to a mixed strategy profile  $x$  the set of all player  $i$ 's mixed *best replies* to  $x_{-i}$ . The *best-reply correspondence*  $BR: \Delta \rightarrow \Delta$  is the product of the  $BR_i$ 's, i.e.,  $BR(x) = \times_{i \in N} BR_i(x)$ . Since  $BR_i(x)$  depends only on  $x_{-i}$  the profile of mixed strategies of the other players,  $BR_i$  defines in a natural way a correspondence on  $\Delta_{-i}$  which we also denote by  $BR_i$ . The set of equilibria of  $\Gamma$  is denoted by  $E(\Gamma)$ .

We shall abuse notation by identifying a pure strategy  $a \in A_i$  with the mixed strategy  $e_i^a$  in  $\Delta_i$  that puts all weight on  $a$ . Pure strategy profiles (i.e., elements of  $A$ ) will be denoted by boldface letters to distinguish them from elements of  $A_i$ . Typically we will write  $\mathbf{a} = (a_i)_{i \in N}$  and  $\mathbf{b} = (b_i)_{i \in N}$  for  $\mathbf{a}$  and  $\mathbf{b}$  in  $A$ . The set of *pure best replies* to a strategy profile  $x$  in  $\Delta$  is  $PB_i(x) = \{a \in A_i \mid a \in BR_i(x)\}$ . The *pure best-reply correspondence*  $PB: \Delta \rightarrow A$  is the product of the  $PB_i$ 's, i.e.,  $PB(x) = \times_{i \in N} PB_i(x)$ . Again  $PB_i(x)$  depends only on  $x_{-i}$  the profile of mixed strategies of the other players and defines in a natural way a correspondence on  $\Delta_{-i}$  which we also denote by  $PB_i$ . Notice that  $PB(x) = \{\mathbf{a} \in A \mid \mathbf{a} \in BR(x)\}$ .

Following Kohlberg and Mertens (1986) and Selten (1975), we will say that, for a game  $\Gamma = \langle A, u \rangle$ , an  $n$ -tuple  $\kappa = (\kappa_1, \dots, \kappa_n)$  (where  $\kappa_i = (\kappa_{ia})_{a \in A_i}$  is a vector of nonnegative real numbers) is a *KM-perturbation* of  $\Gamma$  if the set

$$\Delta_i(\kappa_i) = \{x_i \in \Delta_i \mid x_{ia} \geq \kappa_{ia} \text{ for all } a \in A_i\}$$

is nonempty for each player  $i$  in  $N$ .

The KM-perturbation whose coordinates are all equal to zero is denoted by 0 and the space of all KM-perturbations by  $\mathcal{K}$ . For a KM-perturbation  $\kappa \in \mathcal{K}$ , the set  $\Delta(\kappa) = \times_{i \in N} \Delta_i(\kappa_i)$  is called the  $\kappa$ -perturbed strategy space.

A KM-perturbation  $\kappa$  gives rise to the  $\kappa$ -perturbed game  $\Gamma[\kappa]$ , whose space of strategy profiles is the  $\kappa$ -perturbed strategy space  $\Delta(\kappa)$ . The payoff function of player  $i$  is simply the restriction of  $u_i$  to this space of strategy profiles.

For two strategy profiles  $x$  and  $z$  in  $\Delta$  and a KM-perturbation  $\kappa$ , the strategy profile  $z$  is called a  $\kappa$ -perturbed best reply to  $x$  if  $z$  is an element of the  $\kappa$ -perturbed strategy space and  $u_i(x_{-i} | z_i) \geq u_i(x_{-i} | y_i)$  for all players  $i$  in  $N$  and all strategies  $y_i$  in  $\Delta_i(\kappa_i)$ . The set of  $\kappa$ -perturbed best replies to  $x$  is denoted by  $\text{BR}(\kappa, x)$ . A strategy profile  $x$  in  $\Delta$  is called a  $\kappa$ -perturbed equilibrium of  $\Gamma$  if  $x \in \text{BR}(\kappa, x)$ . The set of  $\kappa$ -perturbed equilibria of  $\Gamma$  is denoted by  $E(\Gamma[\kappa])$ .

For a game  $\Gamma = \langle A, u \rangle$ , a strategy profile  $x$  in  $\Delta$  is called a *perfect equilibrium* of  $\Gamma$  if there exists a sequence  $(\kappa^t)_{t \in \mathbb{N}}$  of completely mixed KM-perturbations converging to zero and a sequence  $(x^t)_{t \in \mathbb{N}}$  in  $\Delta$  converging to  $x$  such that  $x^t$  is a  $\kappa^t$ -perturbed equilibrium of  $\Gamma$  for every  $t$  in  $\mathbb{N}$ .

**3. Definitions of the central concepts.** In this section three types of stability are introduced. The first one, called homotopy-stability was introduced by Mertens (1989). The second one, BR-stability, is closely related to the notion of stability introduced in Hillas (1990). The definition of the third type, CKM-stability, can also be found in Vermeulen (1996). We will briefly discuss the definitions and prove some of the more obvious relations between these and other well-known types of stability.

For  $\eta > 0$ ,  $\mathcal{K}^\eta = \{\kappa \in \mathcal{K} \mid \|\kappa\|_\infty \leq \eta\}$ . The real number  $\eta$  is assumed to be small enough to guarantee that  $\mathcal{K}^\eta$  is a hypercube in  $\mathcal{K}$ . To be more precise, let us say that for  $i \in N$  and  $a \in A_i$ ,  $(i, a)$  is called a *pair* and let the set of all such pairs be denoted by  $P$ . The boundary  $\partial\mathcal{K}^\eta$  is now assumed to be equal to the set  $\{\kappa \in \mathcal{K}^\eta \mid \kappa_{ia} \in \{0, \eta\} \text{ for some } (i, a) \in P\}$ . Then the vertices of  $\mathcal{K}^\eta$  are those KM-perturbations whose coordinates are equal to either 0 or  $\eta$ .

Let  $S$  be a closed subset of the graph

$$\mathcal{E} = \{(\kappa, x) \in \mathcal{K} \times \Delta \mid x \text{ is a } \kappa\text{-perturbed equilibrium of } \Gamma\}$$

of the perturbed equilibrium correspondence. For  $\eta > 0$ ,

$$S^\eta = \{(\kappa, x) \in S \mid \kappa \in \mathcal{K}^\eta\}$$

is the part of  $S$  above  $\mathcal{K}^\eta$ , and

$$\partial_v S^\eta = \{(\kappa, x) \in S^\eta \mid \kappa \in \partial\mathcal{K}^\eta\}$$

is the part of  $S$  above  $\partial\mathcal{K}^\eta$ . Usually  $\partial_v S^\eta$  is called the *vertical boundary* of  $S^\eta$ .

**DEFINITION 1.** The restriction  $p^\eta: S^\eta \rightarrow \mathcal{K}^\eta$  to  $S^\eta$  of the canonical projection  $p_{\mathcal{K}}$  from  $\mathcal{K} \times \Delta$  to  $\mathcal{K}$  is called *relatively null-homotopic* if there exist a continuous function  $F^\eta: S^\eta \times [0, 1] \rightarrow \mathcal{K}^\eta$  such that

- (1)  $F^\eta(\kappa, x, 0) = p^\eta(\kappa, x)$  for all  $(\kappa, x) \in S^\eta$ ;
- (2)  $F^\eta(\kappa, x, 1) \in \partial\mathcal{K}^\eta$  for all  $(\kappa, x) \in S^\eta$ ; and
- (3)  $F^\eta(\kappa, x, t) \in \partial\mathcal{K}^\eta$  for all  $t \in [0, 1]$  and all  $(\kappa, x) \in \partial_v S^\eta$ .

The function  $F^\eta$  is called a *homotopy* for  $p^\eta$ .  $F^\eta$  transforms  $p^\eta$  in a continuous way into a function that only takes values in the boundary  $\partial\mathcal{K}^\eta$  of  $\mathcal{K}^\eta$ , while the image of the vertical boundary  $\partial_v S^\eta$  of  $S^\eta$  remains within the boundary  $\partial\mathcal{K}^\eta$  of  $\mathcal{K}^\eta$  during this transformation.

The canonical projection  $p_{\mathcal{K}}$  to  $\mathcal{K}$  is called *locally null-homotopic on  $S$*  if there exists a number  $\eta_0 > 0$  such that  $p^\eta$  is relatively null-homotopic for all  $\eta \leq \eta_0$ .

For a closed set  $S \subset \mathcal{E}$ , let  $\text{vint}S^\eta$  be the set  $S^\eta \setminus \partial_\nu S^\eta$ . This is the set of points  $(\kappa, x)$  in  $S^\eta$  for which  $\kappa$  is completely mixed and  $\kappa_{ia} < \eta$  for all pairs  $(i, a)$ . Let  $\mathcal{S}$  be the collection of all nonempty, closed sets  $S \subset \mathcal{E}$  such that for all  $\eta > 0$ , the set  $\text{vint}S^\eta$  is connected and  $S^\eta = \text{cl}(\text{vint}S^\eta)$ .

DEFINITION 2. A closed set  $T \subset \Delta$  is called *homotopy-stable* if there exists a set  $S \in \mathcal{S}$  such that

- (1)  $T = \{x \in \Delta \mid (0, x) \in S\}$ ; and
- (2)  $p_{\mathcal{K}}$  is not locally null-homotopic on  $S$ .

REMARK 1. It follows immediately from the definition of  $\mathcal{S}$  that homotopy-stable sets are connected and consist only of perfect equilibria.

The definition of stability by Mertens (1989) is in terms of nonvoid maps between homology groups. The sets given by such a definition may depend on the coefficient modules used. In §(E) of that paper, Mertens shows that the union over all possible coefficient modules of such solutions is equivalent to a definition in terms of homotopy. However one needs to require not only that the projection map not be null-homotopic, but also that this remain true for the restriction of the projection map to some subset of the graph of the equilibrium correspondence that is of the same dimension as the space of perturbed games. Thus Definition 2 may be thought of as a weaker approximation of the “right definition.” In particular, any set that satisfies the definition in terms of homology will be homotopy-stable. The interested reader may compare the definition of stability by Mertens with the definition of homotopy-stability and construct a proof of the previous statement using, e.g., Theorem 19.3 of Munkres (1984).

Next, we define the notion of BR-stability. This definition is similar to the definition of stability in Hillas (1990) but differs from the original definition in Hillas’ paper in two respects. First, we omit the part concerning invariance. Second, we do not require minimality (with respect to robustness against perturbations), just connectedness and perfection. Both adjustments are because the requirement of invariance has become better understood during the last few years. Therefore, we prefer to use a more careful selection of the collection of BR-sets in the definition of BR-stability similar to the one employed by Mertens (1989). (Again, a more elaborate discussion of the reasons for these adjustments can be found in §6.2 and Vermeulen and Jansen 1999.)

Let  $\Gamma = \langle A, u \rangle$  be a game. Note that the best-reply correspondence BR of the game  $\Gamma$  is an element of the class  $\mathcal{H}$  of all compact and convex-valued upper-hemicontinuous correspondences  $\varphi : \Delta \rightarrow \Delta$ . For two correspondences  $\varphi, \psi \in \mathcal{H}$  we define

$$d(\varphi, \psi) = \sup\{d_H(\varphi(x), \psi(x)) \mid x \in \Delta\}.$$

DEFINITION 3. A closed set  $S \subset \Delta$  is a *BR-set* if for any neighborhood  $V$  of  $S$  there exists a number  $\eta > 0$  such that  $\text{fix}(\varphi) \cap V$  is nonempty for every  $\varphi \in \mathcal{H}$  with  $d(\text{BR}, \varphi) < \eta$ . A connected BR-set that contains only perfect equilibria of  $\Gamma$  is called *BR-stable*.

McLennan (1989a) introduced a related type of stable sets, which he called essential sets. He used a larger class of correspondences to perturb the best-reply correspondence BR (he used contractible valuedness instead of convex valuedness) but, more importantly, he used a coarser topology on the space of perturbations. (That is, there are more perturbations close to a given game.) He required that all such correspondences whose graph is contained in a sufficiently small neighborhood of the graph of BR should have fixed points close to  $S$ . This notion of closeness of correspondences to BR is even weaker than the requirement that the graph of the correspondence is close to the graph of BR in Hausdorff distance. The condition only requires that the graph of the perturbation be within  $\varepsilon$  of the graph of BR. In effect this yields a notion of stability that is stronger than BR-stability.

Finally we introduce a form of stability, called CKM-stability, that is at first sight weaker than BR-stability. A *CKM-perturbation* of a game  $\Gamma = \langle A, u \rangle$  is a continuous function  $\varepsilon : \Delta \rightarrow \mathcal{K}$  from the strategy space  $\Delta$  to the space  $\mathcal{K}$  of KM-perturbations of  $\Gamma$ .

We also write  $\varepsilon = (\varepsilon_i)_{i \in N}$ . The collection  $\mathcal{C}$  of all CKM-perturbations is endowed with the norm  $\|\varepsilon\| = \max_{x \in \Delta} \|\varepsilon(x)\|_\infty$ . If  $\varepsilon$  takes on only strictly positive values we call it a *completely mixed CKM-perturbation*.

For a CKM-perturbation  $\varepsilon \in \mathcal{C}$ , the  $\varepsilon$ -perturbed best-reply correspondence

$$\text{BR}[\varepsilon]: \Delta \rightarrow \Delta$$

is defined by

$$\text{BR}[\varepsilon](x) = \text{BR}(\varepsilon(x), x).$$

**DEFINITION 4.** A closed set  $S \subset \Delta$  is a *continuous KM-set*—CKM-set for short—if for any neighborhood  $V$  of  $S$  there is a number  $\eta > 0$  such that  $\text{fix}(\text{BR}[\varepsilon]) \cap V$  is nonempty for every  $\varepsilon \in \mathcal{C}$  with  $\|\varepsilon\| < \eta$ . A connected CKM-set that only contains perfect equilibria of  $\Gamma$  is called *CKM-stable*.

**REMARK 2.** The sets defined in Definition 4 would not be changed by requiring only that  $\text{fix}(\text{BR}[\varepsilon]) \cap V$  is nonempty for every completely mixed  $\varepsilon$  with  $\|\varepsilon\| < \eta$ . This follows directly from the upper hemicontinuity of  $\text{fix}: \mathcal{C} \rightarrow \Delta$ , which in turn follows from the upper hemicontinuity of the best-reply correspondence.

It is clear that every BR-set of a given  $\Gamma$  is also a CKM-set of that game. For every CKM-perturbation  $\varepsilon$  of the  $\Gamma$  the  $\varepsilon$ -perturbed best-reply correspondence  $\text{BR}[\varepsilon]$  is an element of  $\mathcal{H}$ . Furthermore, by Lemma 7 in Appendix A we know that there exists a constant  $C > 0$  such that  $d(\text{BR}[\varepsilon], \text{BR}) < C\|\varepsilon\|$  for every CKM-perturbation  $\varepsilon$ . Given these two observations, it is an elementary exercise to complete the proof.

**4. Every homotopy-stable set is CKM-stable.** Our main concern in this section is to prove that every homotopy-stable set is CKM-stable. We first give an intuition as to why this is true and why it is unlikely that the converse is true.

Consider the graphs shown in Figure 2. This diagram is not to be taken too literally. The spaces  $\mathcal{K}^\eta$  and  $\Delta$  are typically of quite high dimension, even for fairly simple games, but are represented here as one-dimensional. (One probably gets a better idea by at least thinking of one additional dimension that is not shown, representing the size of the perturbation. The set  $S^\eta$  and the graph of  $\varepsilon$  are typically of the same dimension as  $\mathcal{K}^\eta$  and  $\Delta$ , respectively. Their boundaries are of one less dimension.)

The graph of  $S^\eta$ , as drawn, does not have any “holes” in it. If there were holes then it would be possible to “stretch out” the set  $S^\eta$ , without moving  $\partial_v S^\eta$  so that it was completely above  $\partial \mathcal{K}^\eta$ . On the other hand, if there are no holes then it is not possible to draw the graph of a continuous function  $\varepsilon$  from  $\Delta$  to  $\mathcal{K}^\eta$  whose graph has an empty intersection with  $S^\eta$ .

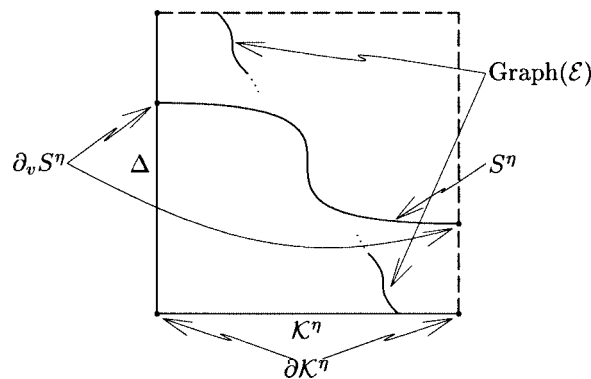


FIGURE 2

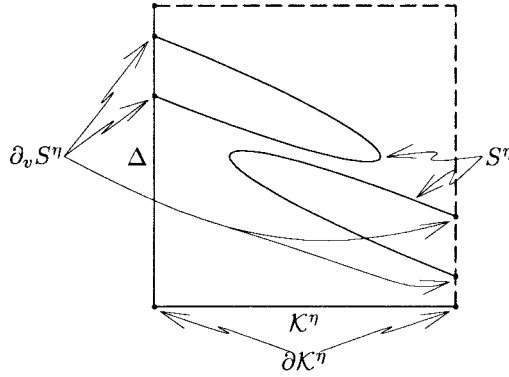


FIGURE 3

And any  $\varepsilon$  whose graph does not have an empty intersection with  $S^\eta$  is such that the  $\varepsilon$ -perturbed best-reply correspondence has fixed points close to the projection of  $S^\eta$  onto  $\Delta$ , and hence—for sufficiently small  $\eta$ —close to the homotopy stable set.

On the other hand, consider the graph in Figure 3. In this case the set  $S^\eta$  does have a “hole” in it and one could clearly stretch it out in a continuous way so that it was completely above  $\partial\mathcal{K}^\eta$ . And yet it should also be clear that, even in this case, it is not possible to find a continuous function  $\varepsilon$  from  $\Delta$  to  $\mathcal{K}^\eta$  whose graph has an empty intersection with  $S^\eta$ . Thus the set  $T = \{x \in \Delta \mid (0, x) \in S\}$  would in this case be a BR-set, and indeed even a BR-stable set, but would not be a homotopy-stable set.

Of course we have not exhibited a game for which the graph of the equilibrium correspondence is like this. However this kind of behavior of the equilibrium correspondence is not ruled out by anything that we currently know. Indeed, it is sufficiently well behaved that there is a strong intuition that there would be games for which the graph of the equilibrium correspondence above the perturbed games would have these general features.

We now turn to proving the theorem.

**THEOREM 1.** *Homotopy-stable sets are CKM-stable.*

**PROOF.** Suppose that  $T \subset \Delta$  is not a CKM-stable set. If  $T$  is not a connected set of perfect equilibria then it is clearly not homotopy-stable by Remark 1. Suppose that  $T$  is not a CKM-set, then there is a neighborhood  $V$  of  $T$  such that for any  $\eta > 0$  there is a completely mixed CKM-perturbation  $\varepsilon$  such that  $\|\varepsilon\| < \eta$  and  $\text{fix}(\text{BR}[\varepsilon]) \cap V$  is empty.

Let  $S \in \mathcal{S}$  be such that  $T = \{x \in \Delta \mid (0, x) \in S\}$ . If there is no such  $S$  then  $T$  is clearly not homotopy-stable and we are done.

Let  $\eta_0$  be small enough that  $S^{\eta_0} \subset \mathcal{K} \times V$ . Now consider an (arbitrary) value  $\eta$  less than  $\eta_0$ . Since  $T$  is not a CKM-set there is a completely mixed CKM-perturbation  $\varepsilon$  with  $\|\varepsilon\| < \eta$  such that  $\text{fix}(\text{BR}[\varepsilon]) \cap V$  is empty.

For any  $(\kappa, x) \in S^\eta$  it must be that  $\varepsilon(x) \neq \kappa$ . For if  $\varepsilon(x) = \kappa$  then  $x$  is in  $E(\Gamma[\kappa]) = E(\Gamma[\varepsilon(x)])$ . Thus  $x$  is in  $\text{BR}(\varepsilon(x), x) = \text{BR}[\varepsilon](x)$  and so  $x$  is a fixed point of  $\text{BR}[\varepsilon]$ . Thus since  $S^\eta \subset S^{\eta_0} \subset \mathcal{K} \times V$ ,  $x$  is in  $\text{fix}(\text{BR}[\varepsilon]) \cap V$ , which was assumed to be empty.

We shall now construct a homotopy with the properties given in Definition 1. Let the function  $\rho^\eta: S^\eta \rightarrow \partial\mathcal{K}^\eta$  be given by defining  $\rho^\eta(\kappa, x)$  to be, for  $(\kappa, x)$  in  $S^\eta$ , the point obtained by extending the line from  $\varepsilon(x)$  through  $\kappa$  until it hits  $\partial\mathcal{K}^\eta$ . If  $(\kappa, x)$  is in  $\partial_v S^\eta$  this is simply the point  $\kappa = \rho^\eta(\kappa, x)$ .

Since  $\mathcal{K}^\eta$  is convex and  $\varepsilon(x)$  is in the interior of  $\mathcal{K}^\eta$  and  $\varepsilon(x) \neq \kappa$ , this point is uniquely defined. Moreover, since  $\varepsilon$  takes values in the interior of  $\mathcal{K}^\eta$  and is continuous it is bounded away from  $\partial\mathcal{K}^\eta$ . Thus  $\rho^\eta$  is continuous.

Then the function  $F^\eta: S^\eta \times [0, 1] \rightarrow \mathcal{K}^\eta$  defined

$$F^\eta(\kappa, x, t) = t\rho^\eta(\kappa, x) + (1-t)p^\eta(\kappa, x)$$

is also continuous. The only part of Definition 1 that might not be immediately obvious is Condition 3. However this too is obvious if one observes that  $\rho^\eta(\kappa, x) = p^\eta(\kappa, x)$  on  $\partial_v S^\eta$ .  $\square$

**5. The equivalence of BR-stability and CKM-stability.** In this section we shall prove the following:

**THEOREM 2.** *Every CKM-set is a BR-set and, conversely, every BR-set is a CKM-set. The same equivalence holds if we replace “set” with “stable set.”*

The part of this theorem stating that every BR-set of a given game  $\Gamma$  is also a CKM-set of that game, was already shown in §3. In this section we will show that the converse statement (every CKM-set is a BR-set) is also true. The second part of the theorem is then a simple consequence of the first part.

One of the tools in the proof of the converse statement is yet another (at first sight somewhat awkward) type of stability, called CT-stability. The perturbations involved are continuous versions of what we shall call T-perturbations. With the help of this notion the proof is split into two parts. First we show that every CKM-set is a CT-set. After that we show that every CT-set is a BR-set. These two facts combine to the desired result.

**5.1. (C)T-perturbations.** The space of CT-perturbations is a subspace of the space of BR-perturbations and it contains the space of CKM-perturbations. Since it is difficult to show directly that a CKM-set is a BR-st, CT-perturbations offer a convenient stepping stone in the proof.

To give the reader an idea of the advantage of the use of CT-perturbations in the proof, consider the following situation. Suppose that we have a CKM-set  $S$  of a game  $\Gamma$ . If we want to show that  $S$  is also a BR-set we have to show that every sufficiently small BR-perturbation  $\varphi$  has a fixed point close to  $S$ . So, if we could construct a (small) CKM-perturbation  $\varepsilon$  such that every fixed point of  $\text{BR}[\varepsilon]$  is also a fixed point of  $\varphi$ , then the proof would be easy. However, such a construction requires a solution for two problems. The first problem is that CKM-perturbations generate perturbed best-reply correspondences that cannot “jump” from one place to another like BR-perturbations can; they can only move continuously. The second problem is that for a given strategy profile  $x$  the set  $\text{BR}[\varepsilon](x)$  is necessarily parallel to  $\text{BR}(x)$  as well as product set with respect to the players, while the set  $\varphi(x)$  does not need to be either of those. So, for our purpose (the approximation of  $\varphi$  by a CKM-perturbation in such a way that the fixed points of the CKM-perturbation are also fixed points of  $\varphi$ ), CKM-perturbations seem to be too rigid in two different ways. CT-perturbations only have the first drawback. The perturbed best-reply correspondence generated by a CT-perturbation still varies in a continuous way, but it is better suited to approximate  $\varphi(x)$  in a given strategy profile  $x$ . Thus, by first proving that every CKM-set is a CT-set and then that every CT-set is a BR-set, we are able to handle these two problems one at a time.

**DEFINITION 5.** A T-perturbation of a game  $\Gamma = \langle A, u \rangle$  is a function  $t: A \rightarrow \Delta$  from the set of pure strategy profiles  $A$  to the space of mixed strategy profiles  $\Delta$ .

For a T-perturbation  $t \in \Delta^A$  and a strategy profile  $x \in \Delta$ ,

$$\text{BR}(t, x) = \text{ch}\{t(\mathbf{a}) \mid \mathbf{a} \in \text{PB}(x)\}$$



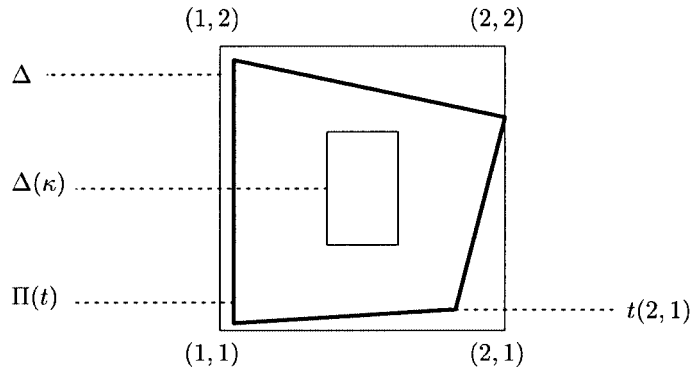


FIGURE 4

is the set of  $t$ -perturbed best replies to  $x$ . Further, the distance between two T-perturbations  $t$  and  $t'$  in  $\Delta^A$  is given by

$$d(t, t') = \max\{\|t(\mathbf{a}) - t'(\mathbf{a})\|_\infty \mid \mathbf{a} \in A\}.$$

Note that for the T-perturbation  $\text{id}$  defined by  $\text{id}(\mathbf{a}) = \mathbf{a}$  for all  $\mathbf{a} \in A$ , we have  $\text{BR}(\text{id}, x) = \text{BR}(x)$  for all  $x \in \Delta$ . So,  $\text{id}$  can be identified with  $\Gamma$ . Therefore, a T-perturbation  $t \in \Delta^A$  is considered to be small if the distance  $d(t, \text{id})$  between  $t$  and  $\text{id}$  is small. Further,

$$\Pi(t) = \text{ch}\{t(\mathbf{a}) \mid \mathbf{a} \in A\}$$

REMARKS. We shall explain some of the intuitions behind the notions defined above. Suppose that we have a game  $\Gamma$  with two players and  $A_1 = A_2 = \{1, 2\}$ . Then the space  $\Delta = \Delta_2 \times \Delta_2$  of strategy pairs is the convex hull of the pure strategy pairs  $\text{id}(1, 1) = (1, 1)$ ,  $\text{id}(1, 2) = (1, 2)$ ,  $\text{id}(2, 1) = (2, 1)$ , and  $\text{id}(2, 2) = (2, 2)$  as indicated in Figure 4. The strategy space of the first player is depicted on the horizontal axis, the vertical axis represents the strategy space of player 2.

For a KM-perturbation  $\kappa = (\kappa_1, \kappa_2)$  of the game  $\Gamma$ , the  $\kappa$ -perturbed strategy space  $\Delta(\kappa)$  is represented by the smaller rectangle. Notice that this space still has a product structure with respect to the players since it is the product of the spaces  $\Delta_1(\kappa_1)$  and  $\Delta_2(\kappa_2)$ .

Now take a T-perturbation  $t \in \mathcal{T}$  of the game  $\Gamma$ . In this two-person example, the T-perturbation  $t$  is a function that assigns a point in  $\Delta$  to each of the four pure strategy pairs  $(1, 1)$ ,  $(2, 1)$ ,  $(2, 2)$ , and  $(1, 2)$ . The point  $t(2, 1)$  is indicated in Figure 1. The polytope  $\Pi(t) \subset \Delta$ , indicated by the fat lines, is the convex hull of the points  $t(1, 1)$ ,  $t(2, 1)$ ,  $t(2, 2)$ , and  $t(1, 2)$ . Notice that the Hausdorff distance between  $\Delta$  and  $\Pi(t)$  is small whenever  $d(t, \text{id}) = \max\{\|t(a, b) - \text{id}(a, b)\|_\infty \mid a, b = 1, 2\}$  is small. Further, notice that  $\Pi(t)$  does not need to have a product structure with respect to the players.

To give an impression of the structure of the  $t$ -perturbed best replies consider Figure 5. Suppose that we have a strategy pair  $x = (x_1, x_2)$  in  $\Delta$  and  $\text{PB}(x) = \{(2, 1), (2, 2)\}$ .  $\text{BR}(x)$  is then the convex hull of the strategy pairs  $\text{id}(2, 1) = (2, 1)$  and  $\text{id}(2, 2) = (2, 2)$  while  $\text{BR}(t, x)$  is the convex hull of  $t(2, 1)$  and  $t(2, 2)$ .

In this example it can clearly be seen that the set of  $t$ -perturbed best replies  $\text{BR}(t, x)$  do not need to be parallel to the set of unperturbed best replies  $\text{BR}(x)$ , in contrast to the situation for the set  $\text{BR}(\kappa, x)$  of  $\kappa$ -perturbed best replies for a KM-perturbation  $\kappa$ . Also observe that  $\text{BR}(t, x)$  is *not* the collection of points in  $\Pi(t)$  in which the payoff vector induced by  $x$  attains its maximum over  $\Pi(t)$  (in the example this would be the one point set  $\{t(2, 2)\}$ ), which is a fundamental difference with the definition of full stability by Kohlberg and Mertens (1986).

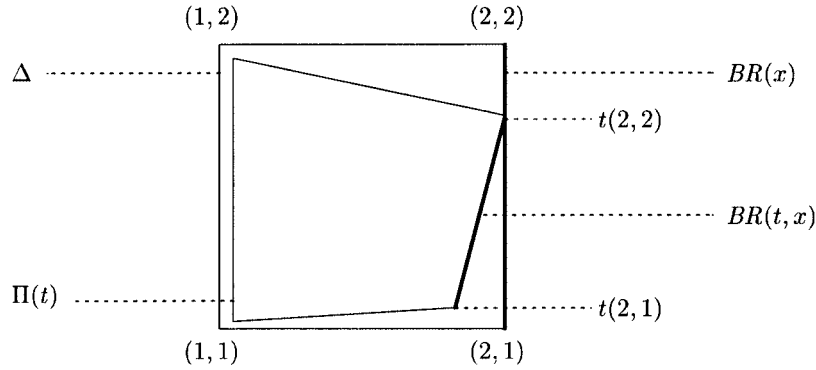


FIGURE 5

Now we turn to the definition of the continuous version of T-perturbations, CT-perturbations. For technical reasons it is convenient to allow CT-perturbations only to take small T-perturbations as values. Formally, we require these values to be in the (compact and convex) neighborhood

$$\mathcal{T} = \{t \in \Delta^A \mid d(t, \text{id}) \leq |A|^{-3}\}$$

of  $\text{id}$  in  $\Delta^A$ . Thus we get

DEFINITION 6. A CT-perturbation is a continuous function  $\tau : \Delta \rightarrow \mathcal{T}$  from the space of strategy profiles  $\Delta$  to the space  $\mathcal{T}$  of well-behaved T-perturbations.

For a CT-perturbation  $\tau$ , the  $\tau$ -perturbed best-reply correspondence  $\text{BR}[\tau] : \Delta \rightarrow \Delta$  is defined by, for all  $x \in \Delta$ ,

$$\text{BR}[\tau](x) = \text{BR}(\tau(x), x).$$

Since the T-perturbation  $\text{id} \in \mathcal{T}$  can be identified with  $\Gamma$ , a CT-perturbation  $\tau$  is considered to be small if  $\|\tau\| = \max_x d(\tau(x), \text{id})$  is small. This naturally leads to

DEFINITION 7. A closed set  $S \subset \Delta$  is a CT-set if for any neighborhood  $V$  of  $S$  there exists a number  $\eta > 0$  such that  $\text{fix}(\text{BR}[\tau]) \cap V$  is not empty for every CT-perturbation  $\tau$  with  $\|\tau\| < \eta$ .

First we turn our attention to the proof that every CKM-set is a CT-set. The proof is designed to solve two technical problems each of which requires some explanation and a fair number of specific lemmas. These two problems are considered in the next two subsections.

The technical problem that is considered in §5.3 concerns the existence of a continuous function  $\kappa$  that assigns to each pair  $(t, x)$  in  $\mathcal{T} \times \Delta$  a KM-perturbation  $\kappa(t, x)$  in such a way that  $x$  is an element of  $\text{BR}(t, x)$  (that is,  $x$  is an equilibrium of the  $t$ -perturbed game) whenever it is an equilibrium of the  $\kappa(t, x)$ -perturbed game. The construction of this function however requires some knowledge of the structure of the graph of the correspondence  $t \mapsto \partial\Pi(t)$ . This structure is studied in the next subsection. The lemmas referring to this graph can be found in Appendix B.

**5.2. The boundary of  $\Pi(t)$ .** Define the graph  $Z$  of the correspondence  $t \mapsto \partial\Pi(t)$  by

$$Z = \{(t, x) \in \mathcal{T} \times \Delta \mid x \in \partial\Pi(t)\}$$

A basic tool in the analysis of the structure of  $Z$  is the supporting hyperplane theorem. First we briefly discuss this theorem. Let  $\Pi$  be a polytope of some  $\mathbb{R}^n$  and let  $x$  be a vector in  $\Pi$ . A vector  $v \in \mathbb{R}^n$  is said to support  $\Pi$  at  $x$  if for all  $y \in \Pi$ ,

$$\langle v, x \rangle \leq \langle v, y \rangle,$$

while the inequality is strict for at least one  $y \in \Pi$ .

**THEOREM 3: SUPPORTING HYPERPLANE THEOREM.** *Let  $\Pi$  be a polytope in  $\mathbb{R}^n$ . A vector  $x \in \Pi$  is an element of  $\partial\Pi$  if and only if there exists a vector that supports  $\Pi$  at  $x$ .*

Obviously a vector  $v$  that supports some polytope  $\Pi$  at a point  $x$  can be chosen to be of unit length. Furthermore, the part of  $v$  perpendicular to  $\text{ah}(\Pi)$  is irrelevant, so we may assume that  $v$  is parallel to  $\text{ah}(\Pi)$ . This standardization has the following special form in case  $\text{ah}(\Pi) = \text{ah}(\Delta)$ .

A vector  $v \in \times_{i \in N} \mathbb{R}^{A_i}$  is called *standard* if

- (1)  $\|v\|_\infty = 1$  and
- (2) for all  $i \in N$ ,  $\sum_{a \in A_i} v_{ia} = 0$ .

If we take  $\Pi = \Pi(t)$  for some T-perturbation  $t \in \mathcal{T}$  we know by Lemma 5 in Appendix A that  $\Pi(t)$  is full-dimensional in  $\Delta$ . Hence,  $\text{ah}(\Pi(t)) = \text{ah}(\Delta)$  and the supporting hyperplane theorem implies

**LEMMA 1.** *Let  $t$  be a T-perturbation in  $\mathcal{T}$  and let  $x$  be a strategy profile in  $\partial\Pi(t)$ . Then there is a standard vector  $v$  that supports  $\Pi(t)$  at  $x$ .*

Using this consequence of the supporting hyperplane theorem we can describe the behavior of those “facets”  $G^{-ia}$  of  $Z$  where player  $i$  is playing his pure strategy  $a \in A_i$  with “sufficiently small” probability. Quotation marks are used here because the subset  $G^{-ia}$  of  $Z$  is actually the union over a (finite) number of facets of  $Z$  and to stress that we work in the product  $\mathcal{T} \times \Delta$  instead of using a “pointwise” approach. Thus, as a formalization of a “facet,” we define—for each pair  $(i, a) \in P$ —the subset  $G^{-ia}$  of  $Z$  by

$$G^{-ia} = \{(t, x) \in Z \mid \text{there is a standard vector } v \\ \text{with } v_{ia} \geq |A|^{-1} \text{ that supports } \Pi(t) \text{ at } x\}.$$

That we need to define the set  $G^{-ia}$  as the union of a number of facets of  $Z$  and not just one facet is because the number of pure strategy profiles  $\mathbf{a}$  with  $\mathbf{a}_i \neq a$  is much higher than the dimension of the facet of  $\Delta$  consisting of those strategy profiles  $x \in \Delta$  that satisfy  $x_{ia} = 0$ . Thus, if we perturb  $\Delta$  by a T-perturbation  $t$ , this facet will generically break into a number of simplices that are all facets of  $\Pi(t)$ .

The next theorem expresses an essential property of the elements of  $G^{-ia}$ . In words, it states that for any element  $(t, x)$  of  $G^{-ia}$ , the points  $t(\mathbf{a})$  with  $\mathbf{a}_i = a$  cannot be used in any convex decomposition of  $x \in \Pi(t)$  into strategy profiles  $t(\mathbf{a})$  with  $\mathbf{a} \in A$ . Note that a completely analogous statement holds for the space  $\Delta$  of strategy profiles and its facets.

From now on we assume that the game  $\Gamma = \langle A, u \rangle$  has at least two players ( $n \geq 2$ ) and that every player has at least two strategies (for all  $i \in N$ ,  $|A_i| \geq 2$ ). Games excluded by these assumptions are either very simple ( $n = 1$ ) or can easily be reduced to games not excluded by the assumptions (by elimination of players with only one pure strategy).

**THEOREM 4.** *Let  $(t, x) \in G^{-ia}$ . Suppose for a subset  $B$  of  $A$  that we can write the strategy profile  $x$  as a positive convex combination of the collection of strategy profiles  $t(\mathbf{b})$  with  $\mathbf{b} \in B$ . Then  $\mathbf{b}_i \neq a$  for all  $\mathbf{b} \in B$ .*

**PROOF.** Take an  $\mathbf{a} \in A$  with  $\mathbf{a}_i = a$ . The theorem is proved if we can show that  $\mathbf{a} \notin B$ .

Since  $(t, x) \in G^{-ia}$ , there is a standard vector  $v$  with  $v_{ia} \geq |A|^{-1}$  that supports  $\Pi(t)$  at  $x$ . In particular, the linear function  $\langle v, \cdot \rangle: y \mapsto \langle v, y \rangle$  attains its minimum over  $\Pi(t)$  at  $x$ .

(a) First we will argue that  $\langle v, \cdot \rangle$  attains its minimum over  $\Pi(t)$  at  $t(\mathbf{b})$  for every  $\mathbf{b} \in B$ . Since  $\langle v, \cdot \rangle$  attains its minimum over  $\Pi(t)$  at  $x$ , it is clear that  $\langle v, x \rangle \leq \langle v, t(\mathbf{b}) \rangle$  for every  $\mathbf{b} \in B$ . However,  $x$  is a positive convex combination of all strategy profiles  $t(\mathbf{b})$  with  $\mathbf{b} \in B$ . So, if (at least) one of the inequalities  $\langle v, x \rangle \leq \langle v, t(\mathbf{b}) \rangle$  would be strict, we could deduce that  $\langle v, x \rangle < \langle v, x \rangle$ , which is clearly impossible. Hence, all these inequalities must in fact be equalities.

(b) On the other hand, we will now show that there is a pure strategy profile  $\mathbf{c} \in A$  with

$$\langle v, t(\mathbf{c}) \rangle < \langle v, t(\mathbf{a}) \rangle.$$

To this end, recall that  $v_{ia} \geq |A|^{-1}$ . Assume that, for all  $b \in A_i$ ,  $v_{ib} > -|A|^{-2}$ . Then we have

$$\sum_{b \in A_i} v_{ib} = v_{ia} + \sum_{b \in A_i \setminus \{a\}} v_{ib} > |A|^{-1} - |A_i| |A|^{-2} \geq 0.$$

This contradicts the assumption that  $v$  is a standard vector. Hence, there is a pure strategy  $c \in A_i$  with  $v_{ic} \leq -|A|^{-2}$ . Define  $\mathbf{c} \in A$  by

$$\mathbf{c}_j = \begin{cases} \mathbf{a}_j & \text{if } j \neq i, \\ c & \text{if } j = i. \end{cases}$$

Then we can write

$$\langle v, t(\mathbf{a}) \rangle - \langle v, t(\mathbf{c}) \rangle = \langle v, t(\mathbf{a}) - t(\mathbf{c}) \rangle = \sum_{(j,b) \in P} v_{jb} (t(\mathbf{a})_{jb} - t(\mathbf{c})_{jb}).$$

First, we will derive a lower bound for each of the terms in this summation. The pairs  $(i, a)$  and  $(i, c)$  will be treated separately for obvious reasons. Observe that these two pairs cannot be identical to each other, since  $v_{ic} \leq -|A|^{-2} < |A|^{-1} \leq v_{ia}$ . So, take a pair  $(j, b)$ .

If  $(j, b) = (i, c)$ . First of all, we know that  $v_{ic} \leq -|A|^{-2}$ . Second, it is easily verified that  $t(\mathbf{c})_{ic} \geq 1 - d(t, \text{id})$  and  $t(\mathbf{a})_{ic} \leq d(t, \text{id})$ . Therefore,

$$(1) \quad v_{ic} (t(\mathbf{a})_{ic} - t(\mathbf{c})_{ic}) \geq |A|^{-2} (1 - 2d(t, \text{id})).$$

If  $(j, b) = (i, a)$ . We know that  $v_{ia} \geq |A|^{-1}$ . Furthermore,  $t(\mathbf{c})_{ia} \leq d(t, \text{id})$  and  $t(\mathbf{a})_{ia} \geq 1 - d(t, \text{id})$ . Therefore,

$$(2) \quad v_{ia} (t(\mathbf{a})_{ia} - t(\mathbf{c})_{ia}) \geq |A|^{-1} (1 - 2d(t, \text{id})).$$

If  $(j, b) \neq (i, a), (i, c)$ . First of all, we know that  $|v_{jb}| \leq 1$  since  $v$  is a standard vector. Furthermore,  $|t(\mathbf{a})_{jb} - t(\mathbf{c})_{jb}| \leq 2d(t, \text{id})$  since  $\mathbf{a}_{jb} = \mathbf{c}_{jb}$ . So, we can calculate that

$$(3) \quad v_{jb} (t(\mathbf{a})_{jb} - t(\mathbf{c})_{jb}) \geq -|v_{jb}| |t(\mathbf{a})_{jb} - t(\mathbf{c})_{jb}| \geq -2d(t, \text{id}).$$

Now, from Inequalities 1, 2, and 3 and because  $\sum_{j \in N} |A_j| \leq |A|$  (all players have at least two pure strategies), we can deduce that the above summation is larger than or equal to

$$|A|^{-2} (1 - 2d(t, \text{id})) + |A|^{-1} (1 - 2d(t, \text{id})) - 2|A|d(t, \text{id}).$$

Using the fact that  $d(t, \text{id}) \leq |A|^{-3}$  (since  $t \in \mathcal{T}$ ) and  $|A| > 2$  (since there are at least two players, each having at least two pure strategies), it is straightforward to calculate that this expression is larger than zero. Hence,  $\langle v, t(\mathbf{c}) \rangle < \langle v, t(\mathbf{a}) \rangle$ .

Since  $t(\mathbf{c})$  is evidently an element of  $\Pi(t)$ , the conclusions of (a) and (b) put next to each other show that  $\mathbf{a}$  cannot be an element of  $B$ .  $\square$

**COROLLARY 1.** *Let  $(t, x) \in G^{-ia}$ . Then  $x_{ia} \leq d(t, \text{id})$ .*

**PROOF.** Since  $x \in \Pi(t)$ , there is a subset  $B$  of  $A$  such that  $x$  can be written as a positive convex combination of all strategy profiles  $t(\mathbf{b})$  with  $\mathbf{b} \in B$ . By Theorem 4 we know that  $\mathbf{b}_i \neq a$  for every  $\mathbf{b} \in B$ . So,  $\text{id}(\mathbf{b})_{ia} = 0$  for every  $\mathbf{b} \in B$ , which implies that  $t(\mathbf{b})_{ia} \leq d(t, \text{id})$  for every  $\mathbf{b} \in B$ . Hence,  $x_{ia} \leq d(t, \text{id})$  since  $x$  is a convex combination of the strategy profiles  $t(\mathbf{b})$  with  $\mathbf{b} \in B$ .  $\square$

**5.3. The construction of  $\kappa$ .** In this subsection we discuss the second technical problem. As said before, we shall construct a continuous function  $\kappa$  that assigns to each pair  $(t, x)$  in  $\mathcal{T} \times \Delta$  a KM-perturbation  $\kappa(t, x)$  in such a way that  $x$  is an element of BR  $(t, x)$  whenever it is an equilibrium of the  $\kappa(t, x)$ -perturbed game. It turns out that  $\kappa$  only needs to satisfy the conditions of the next theorem. (The proofs of the lemmas that are relevant in this subsection can be found in Appendices B and C.)

**THEOREM 5.** *There exists a continuous function  $\kappa$  from  $\mathcal{T} \times \Delta$  to  $\mathcal{K}$  that satisfies the following properties: for all  $(t, x) \in \mathcal{T} \times \Delta$ ,*

- (1)  $\|\kappa(t, x)\|_\infty \leq d(t, \text{id})$ ;
- (2) if  $x \in \Delta(\kappa(t, x))$ , then  $x \in \Pi(t)$ ;
- (3) if  $x \in \Pi(t)$  and  $\kappa(t, x)_{ia} = x_{ia}$ , then  $(t, x) \in G^{-ia}$ .

We briefly discuss the construction of the function  $\kappa$ . First the values of  $\kappa$  are specified on the subspace  $Z$  of  $\mathcal{T} \times \Delta$  as follows. For a given pair  $(i, a)$  Lemma 11 in Appendix B says that  $G^{-ia}$  is a nonempty set. Thus the minimum distance  $d((t, x), G^{-ia})$  from the point  $(t, x)$  to the set  $G^{-ia}$  is a well-defined nonnegative real number for every  $(t, x) \in Z$ . So we can define the function  $\kappa^* : Z \rightarrow \mathcal{K}$  for every  $(t, x) \in Z$  and every pair  $(i, a)$ , by

$$\kappa^*(t, x)_{ia} = \min\{d(t, \text{id}), [x_{ia} - d((t, x), G^{-ia})]_+\}.$$

Note that  $\kappa^*$  is continuous since it is the composition of a number of continuous functions. Next we will extend  $\kappa^*$  to a function  $\kappa$  defined on the whole product space  $\mathcal{T} \times \Delta$ . To get an impression about how this is done, consider, consider the strategy profile  $\hat{x} \in \Delta$  defined by  $\hat{x}_{ia} = |A_i|^{-1}$ . To every  $(t, x) \in \mathcal{T} \times (\Delta \setminus \{\hat{x}\})$  we will assign a point  $z(t, x) \in Z$ . The value of  $\kappa^*$  in  $z(t, x)$  is then used to define the value of  $\kappa$  in  $(t, x)$ . Points  $(t, \hat{x})$  are treated separately. We will first explain the geometrical intuition behind the construction of this function  $z$ .

**REMARK 3.** Consider the perturbed strategy space  $\Delta(\xi)$  of the KM-perturbation  $\xi \in \mathcal{K}$  defined by  $\xi_{ia} = |A_i|^{-1}$ . It is straightforward to check that  $\Delta(\xi)$  is full-dimensional. Furthermore, as indicated in Figure 6, the strategy tuple (pair, in this case)  $\hat{x}$  is an element of the relative interior of  $\Delta(\xi)$  (represented by the smaller square) since  $\hat{x}_{ia} = |A_i|^{-1} > |A_i|^{-1} = \xi_{ia}$ .

Now take a point  $(t, x) \in \mathcal{T}(\Delta \setminus \{\hat{x}\})$ . Since  $x \neq \hat{x}$ , we can consider the half line with origin  $\hat{x}$  through  $x$ , indicated by the line segment from  $\hat{x}$  to  $x$  in Figure 6. Since  $\hat{x} \in \Delta^\circ(\xi) \subset \Pi(t)$  by Lemma 5 in Appendix A, this half line intersects the boundary  $\partial\Pi(t)$  of  $\Pi(t)$  in exactly one point. This point is defined to be  $z(t, x)$ . The unique positive number  $\lambda(t, x)$  for which  $z(t, x) = \hat{x} + \lambda(t, x)(x - \hat{x})$  will also be used in the definition of  $\kappa$ .

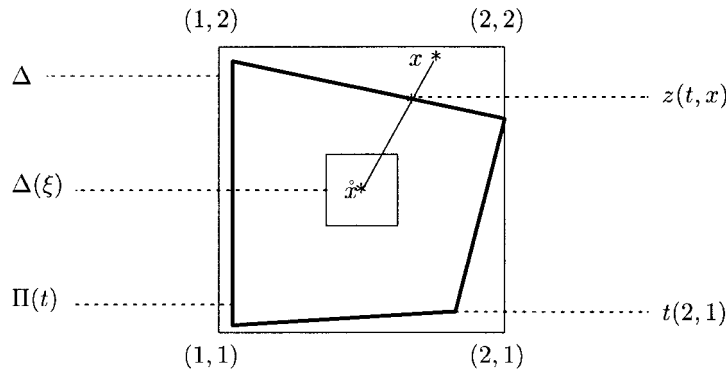


FIGURE 6

Following this idea, we define the function  $\lambda : \mathcal{T} \times (\Delta \setminus \{\hat{x}\}) \rightarrow \mathbb{R}_+$  as follows: For all  $(t, x) \in \mathcal{T} \times (\Delta \setminus \{\hat{x}\})$ ,

$$\lambda(t, x) = \max\{\lambda \geq 0 \mid \hat{x} + \lambda(x - \hat{x}) \in \Pi(t)\}.$$

Obviously,  $\lambda(t, x)$  is a well-defined nonnegative number since  $x \neq \hat{x}$  and  $\Pi(t)$  is a closed and bounded set. The function  $z : \mathcal{T} \times (\Delta \setminus \{\hat{x}\}) \rightarrow \Delta$  can now be given by, for all  $(t, x) \in \mathcal{T} \times \Delta$  with  $x \neq \hat{x}$ ,

$$z(t, x) = \hat{x} + \lambda(t, x)(x - \hat{x}).$$

Note that, since  $\hat{x} + \lambda(x - \hat{x})$  is an element of  $\text{ah}(\Pi(t)) \setminus \Pi(t)$  for all  $\lambda > \lambda(t, x)$ , we have that  $z(t, x)$  is an element of  $\partial\Pi(t)$  and  $(t, z(t, x))$  is an element of  $Z$ . Using these observations it is shown in Lemma 13 of Appendix C that both  $\lambda$  and  $z$  are continuous functions.

Moreover it is shown in this lemma that  $0 < |A|^{-1} - |A|^{-3} \leq \lambda(t, x)\|x - \hat{x}\|_\infty$ . So,  $\lambda(t, x) \geq |A|^{-1} - |A|^{-3}$  since  $\|x - \hat{x}\|_\infty \leq 1$  for all  $x \in \Delta$ . Therefore we can define  $\kappa : \mathcal{T} \times \Delta \rightarrow \mathcal{K}$  by

$$\kappa(t, x) = \begin{cases} \min\{\lambda(t, x)^{-1}, 1\}\kappa^*(t, z(t, x)) & \text{if } x \neq \hat{x}, \\ 0 & \text{if } x = \hat{x}. \end{cases}$$

In Appendix C it is shown that this function  $\kappa$  indeed satisfies the conditions of Theorem 5.

**5.4. Every CKM-set is a BR-set.** Finally we come to the proof of Theorem 2. As said before the proof is split into two parts. First we prove Theorem 6:

**THEOREM 6.** *Every CKM-set is a CT-set.*

**PROOF.** Suppose that  $S$  is a CKM-set of  $\Gamma$ . Let  $V$  be a neighborhood of  $S$ . Then there is a number  $\mu > 0$  such that  $\text{fix}(\text{BR}[\varepsilon]) \cap V$  is not empty for every CKM-perturbation  $\varepsilon$  with  $\|\varepsilon\| < \mu$ . Since  $\mathcal{T}$  is a neighborhood of  $\text{id}$ , we can choose  $\mu$  small enough to guarantee that  $\tau(x) \in \mathcal{T}$  for every  $x \in \Delta$  and every CT-perturbation with  $\|\tau\| < \mu$ . Take a CT-perturbation  $\tau$  with  $\|\tau\| < \mu$ . We will show that  $\text{fix}(\text{BR}[\tau]) \cap V$  is not empty.

Consider the function  $\varepsilon : \Delta \rightarrow \mathcal{K}$  defined by, for all  $x \in \Delta$ ,

$$\varepsilon(x) = \kappa(\tau(x), x).$$

Since  $\kappa$  is continuous by Theorem 5 and  $\tau$  is also continuous, we know that  $\varepsilon$  is a CKM-perturbation. Furthermore, by Theorem 5(1) we know that for every  $x \in \Delta$ ,

$$\|\varepsilon(x)\|_\infty = \|\kappa(\tau(x), x)\|_\infty \leq d(\tau(x), \text{id}) \leq \|\tau\| < \mu.$$

Hence,  $\|\varepsilon\| < \mu$  and we can take a strategy profile  $y \in \text{fix}(\text{BR}[\varepsilon]) \cap V$ . Clearly,  $y \in V$ . We will show that  $y \in \text{fix}(\text{BR}[\tau])$ .

Since  $y \in \text{fix}(\text{BR}[\varepsilon])$ , at least we know that  $y \in \Delta(\varepsilon(y)) = \Delta(\kappa(\tau(y), y))$ . So by Theorem 5(2) we know that  $y \in \Pi(\tau(y))$ . So, there is a subset  $B$  of  $A$  such that  $y$  can be written as a positive convex combination of the collection of strategy profile  $\tau(y)$  ( $\mathbf{b}$ ) with  $\mathbf{b} \in B$ .

To show that  $B \subset \text{PB}(y)$ , take an arbitrary  $\mathbf{a} \notin \text{PB}(y)$ . Then there is a player  $i$  for whom  $\mathbf{a}_i$  is not a best reply to  $y_{-i}$ . Furthermore, since  $y \in \text{fix}(\text{BR}[\varepsilon])$ , we know that  $y$  is an  $\varepsilon(y)$ -perturbed best reply to  $y$ . This implies that the strategy profile  $y$  uses the pure strategy  $\mathbf{a}_i \in A_i$  with minimum probability in the  $\varepsilon(y)$ -perturbed strategy space. Hence,  $y_{i\mathbf{a}_i} = \varepsilon_{i\mathbf{a}_i}(y) = \kappa(\tau(y), y)_{i\mathbf{a}_i}$ . So, by Theorem 5(3),  $(\tau(y), y) \in G^{-i\mathbf{a}_i}$ . Then Theorem 4 implies that, for every  $\mathbf{b} \in B$ ,  $\mathbf{b}_i \neq \mathbf{a}_i$  and hence,  $\mathbf{b} \neq \mathbf{a}$ . So, every  $\mathbf{b} \in B$  must be an element of  $\text{PB}(y)$ .

This means that  $y$  can be written as a convex combination of strategy profiles  $\tau(y)$  ( $\mathbf{b}$ ) with  $\mathbf{b} \in B \subset \text{PB}(y)$ . Hence,  $y \in \text{BR}(\tau(y), y) = \text{BR}[\tau](y)$ .  $\square$

Finally we show that every CT-set is a BR-set. The techniques involved in the proof of this statement are similar to the techniques used in the proof of backward induction for CKM-sets and for stable sets in the sense of Mertens. To give this proof we first need to prove three preliminary results. Note that especially Lemma 3 highlights the usefulness of the (odd) choices of  $\mathcal{T}$  and  $\text{BR}(t, x)$ .

LEMMA 2. *Let  $x$  be a strategy profile in  $\Delta$ . Then there exists a number  $\nu(x) > 0$  such that  $\text{BR}(t, y) \subset \text{BR}(t, x)$  for all  $t \in \mathcal{T}$  and  $y \in \Delta$  with  $\|x - y\|_\infty < \nu(x)$ .*

PROOF. Take a strategy profile  $x \in \Delta$ . First of all we can take a real number  $\nu(x) > 0$  such that for all strategy profiles  $y \in B_{\nu(x)}(x)$  we have  $\text{PB}(y) \subset \text{PB}(x)$ . Second, recall that for each  $t \in \mathcal{T}$  and  $z \in \Delta$ ,  $\text{BR}(t, z)$  is the convex hull of the set of points  $t(\mathbf{a})$  with  $\mathbf{a} \in \text{PB}(z)$ . Now using these two facts it is elementary to show that  $\text{BR}(t, y) \subset \text{BR}(t, x)$  for all  $t \in \mathcal{T}$  and  $y \in B_{\nu(x)}(x)$ .  $\square$

LEMMA 3. *Let  $x$  be a strategy profile in  $\Delta$  and let  $C$  be a nonempty compact convex subset of  $\Delta$ . Then there is a T-perturbation  $t$  with  $\text{BR}(t, x) \subset C$  and  $d(t, \text{id}) \leq d_H(C, \text{BR}(x))$ .*

PROOF. Let  $x$  and  $C$  be as described. First note that  $\text{PB}(x)$  is a subset of  $\text{BR}(x)$ . So, by the compactness of  $C$ , we can choose for each pure strategy profile  $\mathbf{a} \in \text{PB}(x)$  a strategy profile  $y(\mathbf{a}) \in C$  with  $\|\mathbf{a} - y(\mathbf{a})\|_\infty \leq d_H(\text{BR}(x), C)$ . Define the T-perturbation  $t$  by

$$t(\mathbf{a}) = \begin{cases} y(\mathbf{a}) & \text{if } \mathbf{a} \in \text{PB}(x), \\ \mathbf{a} & \text{if } \mathbf{a} \notin \text{PB}(x). \end{cases}$$

Then clearly  $d(t, \text{id}) = \max\{\|t(\mathbf{a}) - \mathbf{a}\|_\infty \mid \mathbf{a} \in A\} \leq d_H(\text{BR}(x), C)$ . Finally, using the convexity of  $C$  and the fact that  $y(\mathbf{a}) \in C$  for all  $\mathbf{a} \in \text{PB}(x)$  we get

$$\text{BR}(t, x) = \text{ch}\{t(\mathbf{a}) \mid \mathbf{a} \in \text{PB}(x)\} = \text{ch}\{y(\mathbf{a}) \mid \mathbf{a} \in \text{PB}(x)\} \subset C. \quad \square$$

Further, for two sets  $X, Y \subset \Delta$  and  $\alpha \in [0, 1]$  we write  $\alpha X + (1 - \alpha)Y = \{\alpha x + (1 - \alpha)y \mid x \in X, y \in Y\}$ .

LEMMA 4. *Let  $t^0$  and  $t^1$  be elements of  $\mathcal{T}$  and let  $x$  be a strategy in  $\Delta$ . Then for  $\alpha \in [0, 1]$ ,  $t^\alpha = \alpha t^1 + (1 - \alpha)t^0 \in \mathcal{T}$  and  $\text{BR}(t^\alpha, x) \subset \alpha \text{BR}(t^1, x) + (1 - \alpha)\text{BR}(t^0, x)$ .*

PROOF. Clearly,  $t^\alpha$  is an element of  $\mathcal{T}$ . Take a strategy profile  $y \in \text{BR}(t^\alpha, x)$ . Then  $y$  is a convex combination of the collection of strategy profiles  $t^\alpha(\mathbf{a})$ , where  $\mathbf{a}$  ranges through  $\text{PB}(x)$ , with weights, say,  $(\lambda(\mathbf{a}))_{\mathbf{a} \in \text{PB}(x)}$ . Now  $y^j = \sum_{\mathbf{a} \in \text{PB}(x)} \lambda(\mathbf{a})t^j(\mathbf{a})$  is an element of  $\text{BR}(t^j, x)$  for  $j = 0, 1$  by definition of  $\text{BR}(t^j, x)$ , and  $y = \alpha y^1 + (1 - \alpha)y^0$ .  $\square$

THEOREM 7. *Every CT-set is a BR-set.*

PROOF. Let  $S$  be a CT-set of  $\Gamma$  and let  $V$  be a closed neighborhood of  $S$ . Then there is an  $\eta$  with  $0 < \eta < |A|^{-3}$  such that  $\text{fix}(\text{BR}[\tau]) \cap V$  is nonempty for every CT-perturbation  $\tau$  with  $\|\tau\| < \eta$ . Take a correspondence  $\varphi \in \mathcal{H}$  with  $d(\varphi, \text{BR}) < \eta$ . We will show that  $\text{fix}(\varphi) \cap V$  is nonempty.

(a) Take an arbitrary number  $\gamma > 0$ . Construct the CT-perturbation  $\tau$  as follows. Choose for each  $x \in \Delta$  a number  $\nu(x) < \gamma$  as in Lemma 2. Then  $\{B_{\nu(x)}(x) \mid x \in \Delta\}$  is an open cover of the compact space  $\Delta$ . So we can choose  $x^1, \dots, x^s$ , such that  $B_{\nu(x^1)}(x^1), \dots, B_{\nu(x^s)}(x^s)$  still covers  $\Delta$ . Then by Lemma 9 in Appendix A there is a partition of unity  $\alpha^1, \dots, \alpha^s$  subordinate to this finite cover of  $\Delta$ . Furthermore, note that each  $\varphi(x^k)$  is a nonempty compact and convex set. So, by Lemma 3 there is a T-perturbation  $t^k$  with  $\text{BR}(t^k, x^k) \subset \varphi(x^k)$  and

$$d(t^k, \text{id}) \leq d_H(\varphi(x^k), \text{BR}(x^k)) \leq d(\varphi, \text{BR}) < \eta.$$

Since this last inequality particularly implies that  $d(t^k, \text{id}) < \eta < |A|^{-3}$  we know that  $t^k \in \mathcal{F}$ . So we can define  $\tau : \Delta \rightarrow \mathcal{F}$  by

$$\tau(y) = \sum_{k=1}^s \alpha^k(y) t^k.$$

Since each  $\alpha^k$  is continuous,  $\tau$  is a CT-perturbation. Furthermore, using the triangle inequality and the fact that  $\alpha^1, \dots, \alpha^s$  is a partition of unity we get for each  $y \in \Delta$ ,

$$d(\tau(y), \text{id}) \leq \sum_{k=1}^s \alpha^k(y) d(t^k, \text{id}) \leq \max_k \{d(t^k, \text{id})\} < \eta.$$

Therefore, since  $\|\tau\|$  equals the maximum over the numbers  $d(\tau(y), \text{id})$  where  $y$  ranges through  $\Delta$ , we get that  $\|\tau\| < \eta$ .

(b) Now repeat this procedure for each element  $\gamma^l$  of a sequence  $(\gamma^l)_{l \in \mathbb{N}}$  of positive real numbers converging to zero. This yields a sequence  $(\tau^l)_{l \in \mathbb{N}}$  of CT-perturbations with

$$\tau^l(y) = \sum_{k=1}^{s(l)} \alpha^{lk}(y) t^{lk}$$

and  $\|\tau^l\| < \eta$ . So, for each  $l$  we can take a strategy profile  $y^l \in \text{fix}(\text{BR}[\tau^l]) \cap V$  by the choice of  $\eta$ . We may assume without loss of generality that  $y^l \rightarrow y$  as  $l \rightarrow \infty$  for some  $y \in \Delta$ . Then  $y \in V$ , since  $V$  is closed. We will show that  $y \in \varphi(y)$ .

Take an arbitrary real number  $\mu > 0$ . Since  $\varphi$  is upper hemicontinuous in  $y$ , we can choose a real number  $\rho > 0$ , such that

$$\varphi(z) \subset B_\mu(\varphi(y)) \quad \text{for all } z \in B_\rho(y).$$

Take a natural number  $L$  with  $y^l \in B_{\frac{1}{2}\rho}(y)$  and  $\gamma^l \leq \frac{1}{2}\rho$  whenever  $l \geq L$ . Take a fixed  $l \geq L$  and an index  $k$ ,  $1 \leq k \leq s(l)$  with  $\alpha^{lk}(y^l) > 0$ . Then by the choice of  $\alpha^{lk}$ ,

$$\|x^{lk} - y\|_\infty \leq \|x^{lk} - y^l\|_\infty + \|y^l - y\|_\infty < \nu(x^{lk}) + \frac{1}{2}\rho < \gamma^l + \frac{1}{2}\rho \leq \frac{1}{2}\rho + \frac{1}{2}\rho = \rho.$$

Hence, for every  $l \geq L$  and  $k \in \{1, \dots, s(l)\}$  with  $\alpha^{lk}(y^l) > 0$ ,

$$(4) \quad \text{BR}(t^{lk}, y^l) \subset \text{BR}(t^{lk}, x^{lk}) \subset \varphi(x^{lk}) \subset B_\mu(\varphi(y)).$$

The first inclusion follows from the definition of the function  $\alpha^{lk}$  and the fact that  $\|x^{lk} - y^l\|_\infty < \nu(x^{lk})$ . The second one follows by the construction of  $t^{lk}$ . The third one follows from the inequality  $\|x^{lk} - y\|_\infty < \rho$  and the choice of  $\rho$ . Now, since  $y^l \in \text{BR}[\tau^l](y^l)$ , we can derive for  $l \geq L$  that  $y^l$  is an element of

$$\text{BR}(\tau^l, (y^l), y^l) = \text{BR}\left(\sum_{k=1}^{s(l)} \alpha^{lk}(y^l) t^{lk}, y^l\right) \subset \sum_{k: \alpha^{lk}(y^l) > 0} \alpha^{lk}(y^l) \text{BR}(t^{lk}, y^l) \subset B_\mu(\varphi(y)).$$

The first inclusion follows from Lemma 4. The second one follows from (1) and the fact that  $B_\mu(\varphi(y))$  is a convex set. So,  $y \in \text{cl}(B_\mu(\varphi(y)))$ . Now recall that  $\mu > 0$  was arbitrary. Therefore,  $y \in \text{cl}(\varphi(y))$ . Hence,  $y \in \varphi(y)$ , since  $\varphi(y)$  is closed.  $\square$

This concludes the last step in the proof that every CKM-set is a BR-set. In §4 we have shown that any stable set in the sense of Mertens is a CKM-stable set. In this section we have shown that it is therefore also a BR-set. Hence, since a stable set in the sense of Mertens is certainly a connected subset of the set of perfect equilibria, it is even a BR-stable set.



**6. Properties of BR-stable sets.** In this section we will use the results from the previous sections to derive some properties of BR-stable sets.

As is already stated in the introduction, the original motivation in the search for stability concepts is a list of requirements composed by Kohlberg and Mertens (1986) and Mertens (1989). The list presented here is a somewhat modified and expanded version of the original one.

(1) Existence.

*Every game possesses at least one stable set.*

(2) Connectedness.

*Stable sets are connected.*

(3) Admissibility.

*Every stable set consists of perfect equilibria.*

(4) Backward induction.

*Every stable set contains a proper equilibrium.*

(5) Independence of inadmissible strategies.

*A stable set  $S$  contains a stable set of the game obtained by deleting a strategy that is not an admissible best reply anywhere in  $S$ .*

(6) Ordinality.

*A stability concept is ordinal.*

(7) The small worlds axiom

*For a game in which there is a set  $I$  (the “insiders”) of players whose payoff does not depend on the strategies of the players outside  $I$  (the “outsiders”) each stable set of the game played by the insiders can be extended to a stable set of the original game.*

We shall briefly comment on some of the requirements. For a more profound argumentation we refer to the above-mentioned papers.

The domain of games we consider here is that of all finite games. Of course, even for single-agent decision problems one would not expect solutions if the strategy sets were not compact or the utility function not continuous. The restriction to finite games gives us a domain in which requirement (1) seems reasonable. See Mertens (1989, p. 582) for a discussion of why one wants a solution defined on all games rather than simply on a generic subset of games.

The restriction to perfect equilibria is a rather strong form of admissibility. The form of the backward induction requirement is justified by the result of Kohlberg and Mertens (1986) and van Damme (1984) that any proper equilibrium “is” a sequential equilibrium. Mailath et al. (1997) and Hillas (1997) prove a partial converse.

The independence of inadmissible strategies is a strengthening of both the requirements of iterated dominance and forward induction occurring in the original list of requirements of Kohlberg and Mertens. This strengthening is used in the “reformulation” papers of Mertens. It is also used in van Damme (1994) under the name of “independence of irrelevant alternatives.”

The ordinality condition is a strengthening of the requirement, known as invariance, that the solution should only depend on the reduced normal form. It was first introduced by Mertens (1987). There Mertens shows that ordinality is implied by invariance and admissible best-reply invariance (abr-invariance). The requirement of abr-invariance means that games that have the same strategy spaces have the same solutions if their admissible best-reply correspondences are the same.

Using the results of Hillas (1990) we can easily prove that BR-stable sets satisfy requirements (1) to (4). It should be noted that although Hillas used an additional condition concerning invariance, his proofs of (1)–(4) are also valid without any alteration for *minimal* BR-sets. Thus, given a game  $\Gamma$ , there exists a minimal BR-set of  $\Gamma$ , and this minimal BR-set is both connected and contained in the set of perfect equilibria of  $\Gamma$ . Hence, this minimal BR-set is a BR-stable set in the sense of this paper and we have verified (1). Concerning (2)

and (3), BR-stable sets of a game  $\Gamma$  are connected and contained in the collection of perfect equilibria of  $\Gamma$  by definition. As for the backward induction requirement, note that every BR-stable set of a game  $\Gamma$  is also a BR-set of this game. Then the proof of Hillas shows that such a set must contain a proper equilibrium of the game  $\Gamma$ .

Concerning the independence of inadmissible strategies, it might be expected that the proof would be similar to the proofs of iterated dominance (Proposition 8) and forward induction (Proposition 9) by Hillas (1990). However, the correspondence  $F$  constructed in the proof of Proposition 8 is not close to BR in Hausdorff distance, which is one of the assumptions in the proof. This effect is caused by the use of the orthogonal projection in the definition of  $F$ . This projection does not respect the values of the best-reply correspondence, and therefore certain strategies may be used in the construction of  $F(\sigma)$  that are not a best reply to  $\sigma$ . It is possible to directly construct extensions of a correspondence  $F'$  that satisfy the three conditions mentioned in the proof of Proposition 8. However, both methods known to us that do this are very intricate, and it is an arduous task to check that they work. The same remarks can be made concerning Proposition 9. The result of the previous section provides a much shorter proof.

**6.1. Independence of inadmissible strategies.** Originally Kohlberg and Mertens required that a stability concept should satisfy two other conditions, namely iterated dominance and forward induction. However, both these conditions are implied by independence of inadmissible strategies and the techniques introduced in this section can also be used to prove the latter requirement. Therefore we will work with the independence of inadmissible strategies in this paper.

A strategy  $y_i$  of player  $i$  is an *admissible* best reply against an element  $x \in \Delta$ —denoted as  $y_i \in \text{BR}_i^a(x)$ —if there is a sequence  $(x^k)_{k \in \mathbb{N}}$  of completely mixed strategy profiles in  $\Delta$  converging to  $x$  such that  $y_i \in \text{BR}_i(x^k_{-i})$  for all  $k$ . The set  $\text{BR}^a(x)$  is defined in the obvious way. For a subset  $S$  of  $\Delta$ ,  $\text{BR}_i^a(S) = \bigcup_{x \in S} \text{BR}_i^a(x)$  is the set of admissible best replies (of Player  $i$ ) against  $S$ . A pure strategy  $b$  of Player  $j$  is called an inadmissible reply against  $S$  if  $b \notin \text{BR}_j^a(S)$ .

Loosely speaking, independence of inadmissible strategies means that a stable set  $S$  of a game  $\Gamma$  remains stable if a pure inadmissible reply against  $S$  is deleted from  $\Gamma$ . To get a formal definition of this property, we need to describe how a pure strategy  $b$  of a player  $j$  can be deleted.

Let  $\Gamma = \langle A, u \rangle$  be a game. Fix a pure strategy  $b \in A_j$  of a player  $j$ , who has at least two pure strategies. The game  $\Gamma'$  induced by the deletion of  $b$  is by definition  $\langle A', u' \rangle$ , wherein

$$A'_i = \begin{cases} A_i & \text{if } i \neq j, \\ A_j \setminus \{b\} & \text{if } i = j, \end{cases}$$

and  $u'_i$  is the restriction of  $u_i$  to  $A'$ . The strategy spaces of  $\Gamma$  and  $\Gamma'$  are denoted by  $\Delta$  and  $\Delta'$ , respectively. From  $y \in \Delta'$  the *lift*  $\tilde{y}$  in  $\Delta$  is obtained by adding zero as the  $b$ th coordinate to  $y_j$ . Further, let  $\pi = (\pi_i)_{i \in N}$  be any continuous function such that (1) for all  $i \neq j$ ,  $\pi_i$  is the identity from  $\Delta_i$  to itself, and (2)  $\pi_j : \Delta(A_j) \rightarrow \Delta(A'_j)$  is such that  $\pi(\tilde{y}) = y$ .

Now let  $S$  be a CKM-stable set of the game  $\Gamma$ . Let  $b$  be an inadmissible reply against  $S$ . Let  $S' \subset \Delta'$  be defined by  $S' = \{y \in \Delta' \mid \tilde{y} \in S\}$ .

**THEOREM 8: INDEPENDENCE OF INADMISSIBLE STRATEGIES.** *The set  $S'$  contains a CKM-stable set of the game  $\Gamma'$  induced by the deletion of  $b$ .*

**PROOF.** It is sufficient to show that  $S'$  is a CKM-set of  $\Gamma'$ . To this end, note that  $S'$  is closed. Moreover, since  $S$  is included in the set of perfect equilibria of  $\Gamma$ , it is easy to see that  $S$  is a subset of  $\{\tilde{y} \mid y \in S'\}$ . Hence,  $S'$  is not empty. Take a neighborhood  $V'$  of  $S'$ . Then  $\pi^{-1}(V')$  is a neighborhood of  $S$ . Since  $b$  is not an element of  $\text{BR}_j^a(S) \supset \text{BR}_j^a(T)$ ,  $\pi^{-1}(V')$

contains a neighborhood  $V$  of  $S$  with  $b \notin \text{BR}_j^a(V)$ . Furthermore, since  $S$  is a CKM-set of the game  $\Gamma$ , there is a number  $\eta > 0$  such that the intersection of  $V$  and  $\text{fix}(\text{BR}[\varepsilon])$  is not empty whenever  $\|\varepsilon\| < \eta$ . Now take a completely mixed perturbation  $\delta'$  of  $\Gamma'$  with  $\|\delta'\| < \eta$ . Define the extension  $\delta$  of  $\delta'$  by, for all  $x \in \Delta$ ,

$$\delta_{ia}(x) = \begin{cases} \delta'_{ia}(\pi(x)) & \text{if } i \neq j \text{ or } a \neq b, \\ 0 & \text{if } i = j \text{ and } a = b. \end{cases}$$

It is evident that  $\delta$  indeed is a CKM-perturbation and that  $\|\delta\| = \|\delta'\|$ . So there is a fixed point  $z$  of  $\text{BR}[\delta]$  contained in  $V$ . Obviously,  $\pi(z) \in V'$ . We will show that  $\pi(z)$  is a fixed point of  $\text{BR}[\delta']$ .

(a) First we will prove that  $z_{jb} = 0$ . Since  $b \notin \text{BR}_j^a(V)$  and  $z \in V$  we know that  $b \notin \text{BR}_j^a(z_{-j})$ . Furthermore,  $z_{-j}$  is completely mixed since  $\delta'$  is completely mixed. So,  $b \notin \text{BR}_j^a(z_{-j}) = \text{BR}_j(z_{-j})$ . Hence, since  $z_j \in \text{BR}_j(z_{-j})$ , we have that  $z_{jb} = 0$ .

(b) Now we will prove that  $\pi(z) \in \text{BR}[\delta'](\pi(z))$ . Because of Lemma 8(1) in Appendix A we know that  $\pi(z) \in \Delta(\delta'(\pi(z)))$ . Take a strategy profile  $x \in \Delta(\delta'(\pi(z)))$ . We will prove that  $\pi_i(z_i)$  is at least as good a reply as  $x_i$  to  $\pi(z)_{-i}$  in  $\Delta_i(\delta'(\pi(z)))$ .

First note that  $z = \pi(z)$ , since  $z_{jb} = 0$ . Furthermore,  $\tilde{x} \in \Delta(\delta(z))$  by Lemma 8(2) in Appendix A. So, since  $z \in \text{BR}[\delta](z)$ , we get

$$\begin{aligned} u'_i(\pi(z)_{-i} \mid \pi_i(z_i)) &= u_i(\widetilde{\pi(z)}_{-i} \mid \widetilde{\pi_i(z_i)}) = u_i(z_{-i} \mid z_i) \\ &\geq u_i(z_{-i} \mid \tilde{x}_i) = u'_i(\pi(z)_{-i} \mid \pi_i(\tilde{x}_i)) = u_i(\pi(z)_{-i} \mid x_i). \end{aligned}$$

Hence,  $\pi(z) \in \text{BR}[\delta'](\pi(z))$  and we have shown that  $S'$  is a CKM-set of  $\Gamma'$  by Remark 2.  $\square$

**6.2. Ordinality.** As said before, in this section we will first give an account of the reasons why we chose to use a slightly simplified version of the original definition of BR-stability in Hillas (1990).

The ordinality requirement is a modern version of the requirement that the solution of a game should be robust against both duplication of strategies and deletion of duplicate strategies. This informal requirement was originally referred to as “invariance,” and by now has several formalizations apart from the strongest one known as ordinality.

The definition of BR-stable sets in Hillas (1990) was specifically designed to satisfy one of the earlier, and weaker, invariance requirements. As it turned out, the robustness requirement by itself against sufficiently small perturbations of the game did not automatically imply this weaker requirement. Therefore, an additional condition was included in the definition stating that the robustness requirement should also be valid for any “equivalent” game. Because of this extra condition the resulting solution concept automatically satisfied the above-mentioned weaker variety of invariance.

However, the example given in Vermeulen and Jansen (1999) shows that the solution concept generated by this definition does not satisfy the stronger invariance requirements. In particular it does not satisfy ordinality. Therefore, since it does not give any additional virtue, we decided to leave out the complicating extra condition referring to equivalent games, and only use the robustness requirement with respect to the game itself.

Secondly, the example in Vermeulen and Jansen (1999) points out that the minimality condition—stable sets were usually defined as sets that were minimal with respect to the robustness requirement—upsets ordinality as well. For this reason we choose to leave out the minimality condition too, and only insist on perfection and connectedness. This way we get a version of BR-stability that satisfies at least some of the invariance requirements, as we will readily show.

As it was already noted in the introduction of this section, Mertens has proven that a solution concept is ordinal if it is both invariant and abr-invariant. Although Vermeulen and Jansen (1999) provides an ordinal selection from the collection of BR-sets, the invariance of BR-stable sets  $\text{sec}$  is still an open problem. (One of the referees conjectures that CKM-stability, and therefore also BR-stability, might fail invariance for basically the same reason—pointed out in Mertens (1991)—why homotopy stability might fail to do so.) Nevertheless, using the result of the previous section we can construct a relatively simple proof of the abr-invariance of BR-stable sets. In order to define abr-invariance, take two games  $\Gamma = \langle A, u \rangle$  and  $\Gamma^* = \langle A, u^* \rangle$  with the same set of strategy profiles  $\Delta = \Delta_A$ . The distinction between notions like (admissible) best replies for the game  $\Gamma$  and the game  $\Gamma^*$  is made by adding a  $*$  as superscript to the notions that refer to  $\Gamma^*$ .

**DEFINITION 8.** The games  $\Gamma$  and  $\Gamma^*$  are said to be *admissible best-reply equivalent* (abr-equivalent) if  $\text{BR}^a(x) = \text{BR}^{*a}(x)$  holds for each strategy profile  $x \in \Delta$

**THEOREM 9: ABR-INVARIANCE.** *Let  $\Gamma$  and  $\Gamma^*$  be two abr-equivalent games. Then a subset of  $\Delta$  is a CKM-stable set of  $\Gamma$  if and only if it is a CKM-stable set of  $\Gamma^*$ .*

**PROOF.** (a) First we will prove that  $\text{BR}(\kappa, x) = \text{BR}^*(\kappa, x)$  for any KM-perturbation  $\kappa$  and completely mixed strategy profile  $x \in \Delta$ . Since  $x$  is completely mixed, we have  $\text{BR}^a(x) = \text{BR}(x)$  and  $\text{BR}^{*a}(x) = \text{BR}^*(x)$ . Then, using the assumption that  $\Gamma$  and  $\Gamma^*$  are abr-equivalent, it follows that

$$\text{BR}(x) = \text{BR}^a(x) = \text{BR}^{*a}(x) = \text{BR}^*(x).$$

From the equality  $\text{BR}(x) = \text{BR}^*(x)$  we can deduce that  $\text{PB}(x) = \text{PB}^*(x)$ . Finally, from this last equality and Lemma 6 from Appendix A it follows that  $\text{BR}(\kappa, x) = \text{BR}^*(\kappa, x)$ .

(b) Now let  $S$  be a CKM-stable set of  $\Gamma$ . By symmetry we know it is sufficient to prove that  $S$  is a CKM-stable set of  $\Gamma^*$ . First of all, note that  $S$  is connected. Furthermore, since  $S$  is contained in the set of perfect equilibria  $\Gamma$  and perfect equilibria are abr-invariant, we know that  $S$  is contained in the set of perfect equilibria of  $\Gamma^*$ . So we only need to prove that  $S$  is a CKM-set of  $\Gamma^*$ . Take a neighborhood  $V$  of  $S$ . Since  $S$  is a CKM-set of  $\Gamma$  there exists a number  $\eta > 0$  such that  $\text{fix}(\text{BR}[\varepsilon]) \cap V$  is not empty for every completely mixed CKM-perturbation  $\varepsilon$  with  $\|\varepsilon\| < \eta$ . Take such a completely mixed CKM-perturbation  $\varepsilon$  with  $\|\varepsilon\| < \eta$ . We will show that  $\text{fix}(\text{BR}^*[\varepsilon]) \cap V$  is not empty. By the choice of  $\eta$  we can choose a strategy profile  $z$  in  $\text{fix}(\text{BR}[\varepsilon]) \cap V$ . Obviously,  $z \in V$ . We will show that  $z \in \text{fix}(\text{BR}^*[\varepsilon])$ . To this end, observe that

$$z \in \text{BR}[\varepsilon](z) = \text{BR}(\varepsilon(z), z).$$

Furthermore, the KM-perturbation  $\varepsilon(z)$  is completely mixed, since  $\varepsilon$  is completely mixed. Therefore, since clearly  $z \in \Delta(\varepsilon(z))$ ,  $z$  is also completely mixed. So by (a),  $z \in \text{BR}(\varepsilon(z), z) = \text{BR}^*(\varepsilon(z), z)$ . So,  $z \in \text{BR}^*[\varepsilon](z)$  and  $z$  is a fixed point of  $\text{BR}^*[\varepsilon]$ . Hence,  $S$  is a CKM-set of  $\Gamma^*$  by Remark 2.  $\square$

**Appendix A.** Define the fixed KM-perturbation  $\xi \in \mathcal{H}$  by  $\xi_{ia} = |A|^{-1}$ . It is straightforward to check that  $\Delta(\xi)$  is full-dimensional. Furthermore,

**LEMMA 5.** *For any  $t$  in  $\mathcal{T}$  the set  $\Delta(\xi)$  is a subset of  $\Pi(t)$ .*

**PROOF.** Suppose that  $\Delta(\xi) \not\subset \Pi(t)$ . Then we can take a strategy profile  $y \in \Delta(\xi) \setminus \Pi(t)$ . Since  $\Pi(t)$  is compact and not empty we can take a strategy profile  $x$  in  $\Pi(t)$  whose distance in maximum norm to  $y$  is minimal over  $\Pi(t)$ . Since  $x \in \Pi(t)$  and  $y \notin \Pi(t)$ ,  $x \neq y$ . So we can define  $z \in \Delta$  by

$$z = y + \lambda_0(y - x) \quad \text{with} \quad \lambda_0 = \max\{\lambda \geq 0 \mid y + \lambda(y - x) \in \Delta\}.$$

Then  $\|z - y\|_\infty \geq |A|^{-1}$  since at least one of the coordinates of  $z$  must be equal to zero, while the fact that  $y \in \Delta(\xi)$  implies that  $y_{ia} \geq \xi_{ia} = |A|^{-1}$  for all  $i \in N$  and  $a \in A_i$ . Now it is also clear that  $\lambda_0 > 0$  and we can deduce that  $\|z - x\|_\infty = (1 + \lambda_0^{-1})\|z - y\|_\infty > |A|^{-1}$ . Furthermore, by the construction of  $z$  and the convexity of  $\Pi(t)$ ,  $x$  is a strategy profile in  $\Pi(t)$  whose distance in maximum norm to  $z$  over  $\Pi(t)$  is also minimal. So, for all  $x' \in \Pi(t)$ ,  $\|z - x'\|_\infty \geq \|z - x\|_\infty > |A|^{-1}$ . This implies that  $d_H(\Pi(t), \Delta) > |A|^{-1}$  since  $z \in \Delta$ . On the other hand,  $d_H(\Pi(t), \Delta) = d_H(\Pi(t), \Pi(\text{id})) \leq d(t, \text{id}) \leq |A|^{-3}$  since  $t \in \mathcal{T}$ . Contradiction.  $\square$

For a KM-perturbation  $\kappa \in \mathcal{K}$  and a pure strategy profile  $\mathbf{a} \in A$ , define the strategy profile  $d^{\mathbf{a}}(\kappa) \in \Delta$  by

$$d^{\mathbf{a}}(\kappa) = \begin{cases} \kappa_{ia} & \text{if } a \neq \mathbf{a}_i, \\ 1 - \sum_{a \neq \mathbf{a}_i} \kappa_{ia} & \text{if } a = \mathbf{a}_i. \end{cases}$$

Using this notation we have the following.

LEMMA 6. For every  $\kappa \in \mathcal{K}$  and  $x \in \Delta$ ,  $BR(\kappa, x) = \text{ch}\{d^{\mathbf{a}}(\kappa) \mid \mathbf{a} \in PB(x)\}$ .

It is straightforward to show this once we note that  $\Delta(\kappa)$  equals the convex hull of the collection of points  $d^{\mathbf{a}}$  with  $\mathbf{a} \in A$ .

LEMMA 7. There exists a constant  $C > 0$  such that for all  $\kappa, \mu \in \mathcal{K}$  and  $x \in \Delta$ ,

$$d_H(BR(\kappa, x), BR(\mu, x)) \leq C\|\kappa - \mu\|_\infty.$$

PROOF. Take two KM-perturbations  $\kappa, \mu \in \mathcal{K}$  and a strategy profile  $x \in \Delta$ . Then using the previous lemma we can calculate that

$$\begin{aligned} d_H(BR(\kappa, x), BR(\mu, x)) &= d_H(\text{ch}\{d^{\mathbf{a}}(\kappa) \mid \mathbf{a} \in PB(x)\}, \text{ch}\{d^{\mathbf{a}}(\mu) \mid \mathbf{a} \in PB(x)\}) \\ &\leq d_H(\{d^{\mathbf{a}}(\kappa) \mid \mathbf{a} \in PB(x)\}, \{d^{\mathbf{a}}(\mu) \mid \mathbf{a} \in PB(x)\}) \\ &\leq |A| \cdot \|\kappa - \mu\|_\infty. \end{aligned}$$

So we can take  $C = |A| > 0$ .  $\square$

LEMMA 8. Let  $\Gamma$  be a game and let  $z \in \Delta$ . Then for the extension  $\varepsilon$  of a CKM-perturbation  $\varepsilon'$  of the game  $\Gamma'$  the following holds:

- (1) if  $y \in \Delta(\varepsilon(z))$  and  $y_{jb} = 0$ , then  $\pi(y) \in \Delta(\varepsilon'(\pi(z)))$ ;
- (2) if  $y \in \Delta(\varepsilon'(\pi(z)))$ , then  $\tilde{y}$  is an element of  $\Delta(\varepsilon(z))$ .

PROOF. (1) Let  $y \in \Delta(\varepsilon(z))$  and  $y_{jb} = 0$ . Then for  $a \in A'_i$ ,  $\pi(y)_{ia} = y_{ia} \geq \varepsilon_{ia}(z) = \varepsilon'_{ia}(\pi(z))$ .

(2) Let  $y \in \Delta(\varepsilon'(\pi(z)))$ . For  $(i, a) \in P$  with  $i \neq j$ , or  $i = j$  and  $a \neq b$ ,  $\tilde{y}_{ia} = y_{ia} \geq \varepsilon'_{ia}(\pi(z)) = \varepsilon_{ia}(z)$ . Otherwise,  $\tilde{y}_{jb} = 0 = \varepsilon_{jb}(z)$ .  $\square$

LEMMA 9: PARTITION OF UNITY. Let  $x^1, \dots, x^s \in \Delta$  and  $\nu_1 > 0, \dots, \nu_s > 0$  be such that  $B_{\nu_1}(x^1), \dots, B_{\nu_s}(x^s)$  covers  $\Delta$ . Then there are continuous functions  $\alpha^1, \dots, \alpha^s$  from  $\Delta$  to  $[0, 1]$  such that  $\sum_{k=1}^s \alpha^k(y) = 1$  for all  $y \in \Delta$  and for each  $k : \alpha^k(y) > 0 \Leftrightarrow y \in B_{\nu_k}(x^k)$ .

PROOF. Define for  $k \in \{1, \dots, s\}$  the function  $\beta^k : \Delta \rightarrow \mathbb{R}$  by  $\beta^k(y) = (1 - \|x^k - y\|_\infty \times (\nu_k)^{-1})_+$ . Define for  $k \in \{1, \dots, s\}$  the function  $\alpha^k : \Delta \rightarrow [0, 1]$  by  $\alpha^k(y) = \beta^k(y) / (\sum_{k=1}^s \beta^k(y))$ . It is straightforward to show that  $\alpha^1, \dots, \alpha^s$  have the properties mentioned.  $\square$

**Appendix B.** This section of the Appendix contains the (proofs of the) lemmas needed in §5.3 and Appendix C.

LEMMA 10.  $Z$  is a closed subset of  $\mathcal{T} \times \Delta$ .

PROOF. Take a sequence  $(t^k, x^k)_{k \in \mathbb{N}}$  in  $Z$  that converges to  $(t, x) \in \Delta^A \times \Delta$ . Observe that  $(t, x) \in \mathcal{T} \times \Delta$ , since  $\mathcal{T}$  is a closed subset of  $\Delta^A$  and  $t^k \in \mathcal{T}$  for all  $k$ .

So we only have to prove that  $x \in \partial\Pi(t)$ . First observe that  $x \in \Pi(t)$  since  $x^k \in \Pi(t^k)$  for all  $k$  and  $t^k \rightarrow t$ . Now take an arbitrary number  $\eta > 0$ . The proof is complete if we can show that there is a vector  $y \in \text{ah}(\Pi(t)) = \text{ah}(\Delta)$  with  $\|x - y\|_\infty \leq \eta$  that is not contained in  $\Pi(t)$ .

For each  $k \in \mathbb{N}$  there is a standard vector  $v^k$  that supports  $\Pi(t^k)$  at  $x^k$  by Lemma 1. Since  $v^k$  is standard, it is parallel to  $\text{ah}(\Delta)$ , so we can define the vector  $y^k \in \text{ah}(\Delta)$  by  $y^k = x^k - \frac{1}{2}\eta v^k$ . We will show that (a)  $\|x - y^k\|_\infty \leq \eta$  and (b)  $y^k \notin \Pi(t)$  for large  $k$ .

(a) Since  $v^k$  is standard we know that  $\|x^k - y^k\|_\infty = \frac{1}{2}\eta\|v^k\|_\infty = \frac{1}{2}\eta$ . Furthermore,  $\|x - x^k\|_\infty \leq \frac{1}{2}\eta$  for large  $k$  because  $x^k \rightarrow x$ . Hence,  $\|x - y^k\|_\infty \leq \eta$  for large  $k$  by the triangle inequality.

(b) Take a fixed  $k \in \mathbb{N}$  and an arbitrary strategy profile  $z \in \Pi(t^k)$ . Then  $\langle v^k, z - x^k \rangle \geq 0$  since  $v^k$  supports  $\Pi(t^k)$  at  $x^k$ . Furthermore,  $\|v^k\|_2 \geq 1$  since  $v^k$  is a standard vector. Therefore,

$$\begin{aligned} \|z - y^k\|_2^2 &= \|z - x^k + \frac{1}{2}\eta v^k\|_2^2 = \|z - x^k\|_2^2 + \eta \langle v^k, z - x^k \rangle + \frac{1}{4}\eta^2 \|v^k\|_2^2 \\ &\geq \frac{1}{4}\eta^2 \|v^k\|_2^2 \geq \frac{1}{4}\eta^2. \end{aligned}$$

Thus  $\|z - y^k\|_2 \geq \frac{1}{2}\eta$ . Let  $D = \sum_{i \in N} |A_i|$ . The last inequality then implies that

$$\|z - y^k\|_\infty \geq D^{-1/2} \|z - y^k\|_2 \geq \frac{1}{2} D^{-1/2} \eta.$$

Thus  $d_H(y^k, \Pi(t^k)) \geq \frac{1}{2} D^{-1/2} \eta$  since  $z$  was an arbitrary strategy profile in  $\Pi(t^k)$ . Now the triangle inequality for the Hausdorff distance yields

$$d_H(y^k, \Pi(t)) \geq d_H(y^k, \Pi(t^k)) - d_H(\Pi(t), \Pi(t^k)) \geq \frac{1}{2} D^{-1/2} \eta - d_H(\Pi(t), \Pi(t^k)).$$

And this is true for any  $k$  in  $\mathbb{N}$ . Thus, since  $\Pi(t^k) \xrightarrow{d_H} \Pi(t)$ , for any sufficiently large  $k$  we have  $d_H(y^k, \Pi(t)) > 0$ .  $\square$

LEMMA 11. For every pair  $(i, a)$ ,  $G^{-ia}$  is a closed set. Further,  $(t, x) \in G^{-ia}$  for every strategy profile  $x \in \Delta$  with  $x_{ia} = 0$  and  $T$ -perturbation  $t \in \mathcal{T}$  with  $x \in \Pi(t)$ .

PROOF. To prove that  $G^{-ia}$  is closed, take a sequence  $(t^k, x^k)_{k \in \mathbb{N}}$  in  $G^{-ia}$  with  $(t^k, x^k) \rightarrow (t, x)$ . We will prove that  $(t, x) \in G^{-ia}$ .

Clearly,  $(t, x) \in Z$  since  $(t^k, x^k) \in G^{-ia} \subset Z$  for all  $k \in \mathbb{N}$ , and  $Z$  is closed by Lemma 10. To show that there is a standard vector  $v$  with  $v_{ia} \geq |A|^{-1}$  that supports  $\Pi(t)$  at  $x$ , take standard vectors  $(v^k)_{k \in \mathbb{N}}$  such that  $v^k$  supports  $\Pi(t^k)$  at  $x^k$  and  $v_{ia}^k \geq |A|^{-1}$ . Assume, without loss of generality, that  $v^k \rightarrow v$ . Obviously  $v$  is a standard vector with  $v_{ia} \geq |A|^{-1}$ . To prove that  $v$  supports  $\Pi(t)$  at  $x$ , take a  $y \in \Pi(t)$ . Since  $\Pi(t^k) \xrightarrow{d_H} \Pi(t)$ , there is a sequence  $(y^k)_{k \in \mathbb{N}}$  that converges to  $y$  with  $y^k \in \Pi(t^k)$  for all  $k$ . Since  $v^k$  supports  $\Pi(t^k)$  at  $x^k$ , we know that  $\langle v^k, y^k \rangle \geq \langle v^k, x^k \rangle$ . Then the continuity of the inner product yields  $\langle v, y \rangle \geq \langle v, x \rangle$ . Finally, observe that the linear function  $\langle v, \cdot \rangle$  is not constant on  $\Delta$  since  $v$  is standard. Then the fact that  $\Pi(t)$  is full-dimensional in  $\Delta$  by Lemma 5 implies that there is a strategy profile in  $\Pi(t)$  for which the inequality is strict.

To prove that  $G^{-ia}$ , take a strategy profile  $x \in \Delta$  with  $x_{ia} = 0$  and a  $T$ -perturbation  $t \in \mathcal{T}$  with  $x \in \Pi(t)$ . We will show that  $(t, x) \in G^{-ia}$ . To this end, define the standard vector  $v$  by

$$v_{jb} = \begin{cases} -(|A_j| - 1)^{-1} & \text{if } j = i \text{ and } b \neq a, \\ 1 & \text{if } j = i \text{ and } b = a, \\ 0 & \text{if } j \neq i. \end{cases}$$

Then  $x - \eta v \in \text{ah}(\Delta) \setminus \Delta$  for  $\eta > 0$ , since  $v$  is standard,  $x_{ia} = 0$  and  $v_{ia} = 1$ . However,  $\text{ah}(\Delta) \setminus \Delta$  is a subset of  $\text{ah}(\Pi(t)) \setminus \Pi(t)$  by Lemma 5. So, since  $x \in \Pi(t)$ , we get that  $(t, x) \in Z$ . Further it is easily checked that  $v$  supports  $\Delta \supset \Pi(t)$  at  $x$ , and  $v_{ia} = 1 \geq |A|^{-1}$ . Hence,  $(t, x) \in G^{-ia}$ .  $\square$

In analogy to the idea that the sets  $G^{-ia}$  are the “facets” of the space of points  $(t, x)$  with  $x \in \Pi(t)$  we will show that the boundary  $Z$  of this space is the union of these “facets.”

LEMMA 12.  $\{G^{-ia} \mid i \in N, a \in A_i\}$  covers  $Z$ .

PROOF. Take  $(t, x) \in Z$ . Then  $x \in \partial\Pi(t)$  by the definition of  $Z$ . So, by Lemma 1, there is a standard vector  $v$  that supports  $\Pi(t)$  at  $x$ .

We only have to prove that there is a pair  $(i, a)$  with  $v_{ia} \geq |A|^{-1}$ . Since  $v$  is a standard vector, we know that  $\|v\|_\infty = 1$ . So, there is a pair  $(j, b)$  with  $|v_{jb}| = 1$ . If  $v_{jb} = 1$ , we choose  $(i, a) = (j, b)$ . If  $v_{jb} = -1$ , the assumption that  $v_{jb} < |A|^{-1}$  for all  $a \in A_j$  leads to

$$\sum_{a \in A_j} v_{ja} = v_{jb} + \sum_{a \in A_j \setminus \{b\}} v_{ja} < -1 + |A||A|^{-1} = 0,$$

contradicting the fact that  $v$  is a standard vector. Hence, also in this case, there is a pair  $(j, a)$  with  $V_{ja} \geq |A|^{-1}$ .  $\square$

**Appendix C.** In this section of the Appendix it is shown that the functions  $\lambda$  and  $z$  are continuous and that the function  $k$  defined §5.3 indeed satisfies the condition in Theorem 5.

LEMMA 13. Both  $\lambda$  and  $z$  are continuous. Furthermore, for any  $(t, x) \in \mathcal{T} \times \Delta$  with  $x \neq \hat{x}$ ,  $|A|^{-1} - |A|^{-3} \leq \lambda(t, x) \|x - \hat{x}\|_\infty$ .

PROOF. Since  $z(t, x) = \hat{x} + \lambda(t, x)(x - \hat{x})$ , the continuity of  $z$  is a direct consequence of the continuity of  $\lambda$ . To prove the continuity of  $\lambda$ , take a sequence  $(t^k, x^k)_{k \in \mathbb{N}}$  in  $\mathcal{T} \times (\Delta \setminus \{\hat{x}\})$  converging to a point  $(t, x)$  with  $x \neq \hat{x}$ . Then  $\|x^k - \hat{x}\|_\infty \geq \frac{1}{2} \|x - \hat{x}\|_\infty$  for large  $k$ . So,

$$1 \geq \|z(t^k, x^k) - \hat{x}\|_\infty = \lambda(t^k, x^k) \|x^k - \hat{x}\|_\infty \geq \frac{1}{2} \lambda(t^k, x^k) \|x - \hat{x}\|_\infty$$

for large  $k$ . Hence the sequence  $\lambda(t^k, x^k)_{k \in \mathbb{N}}$  is bounded since  $\|x - \hat{x}\|_\infty$  is larger than zero. This implies that  $\lambda(t^k, x^k)_{k \in \mathbb{N}}$  converges to  $\lambda(t, x)$  if and only if every convergent subsequence of  $\lambda(t^k, x^k)_{k \in \mathbb{N}}$  converges to  $\lambda(t, x)$ .

To prove the latter statement, take an arbitrary subsequence  $\lambda(t^l, x^l)_{l \in \mathbb{N}}$  that converges to some number  $\mu$ . We will show that  $\mu = \lambda(t, x)$ . To this end, note that  $\hat{x} \in \Delta(\xi) \subset \Pi(t)$  since  $\hat{x}_{ia} = |A_i|^{-1} > |A|^{-1} = \xi_{ia}$  for all pairs  $(i, a)$ . This implies that the halfline emanating from  $\hat{x}$  through  $x$  intersects  $\partial\Pi(t)$  only once. Now, on one hand we know that  $z(t, x)$  is an element of this intersection. On the other hand, recall that  $(t^l, z(t^l, x^l)) \in Z$  for all  $l$ . Then also  $(t, \hat{x} + \mu(x - \hat{x})) \in Z$  since  $(t^l, z(t^l, x^l))_{l \in \mathbb{N}}$  converges to  $(t, \hat{x} + \mu(x - \hat{x}))$  and  $Z$  is closed by Lemma 10. Therefore,  $\hat{x} + \mu(x - \hat{x})$  is also an element of this intersection. Hence,  $z(t, x) = \hat{x} + \mu(x - \hat{x})$  and then  $\lambda(t, x) = \mu$  since  $x \neq \hat{x}$ .

Take a point  $(t, x) \in \mathcal{T} \times \Delta$  with  $x = \hat{x}$ . Since  $(t, z(t, x)) \in Z$  we know by Lemma 12 that there is a pair  $(i, a)$  such that  $(t, z(t, x)) \in G^{-ia}$ . Then, using Corollary 1 and the fact that  $t \in \mathcal{T}$ , we get  $z(t, x)_{ia} \leq d(t, \text{id}) \leq |A|^{-3}$ . Then we can deduce that  $|A|^{-1} - |A|^{-3} \leq \hat{x}_{ia} - z(t, x)_{ia} \leq \|z(t, x) - \hat{x}\|_\infty = \lambda(t, x) \|x - \hat{x}\|_\infty$ .  $\square$

Finally in this Appendix we show that the function  $\kappa$  defined in §5.3 satisfies the conditions of Theorem 5.

PROOF. Of Theorem 5(0) Continuity. Take a sequence  $(t^k, x^k)_{k \in \mathbb{N}}$  in  $\mathcal{T} \times \Delta$  converging to a point  $(t, x)$ . Assume that  $x \neq \hat{x}$ . Then  $x^k \neq \hat{x}$  for large  $k$ . So, in this case, the continuity of  $\kappa$  in  $(t, x)$  follows from the continuity of  $\lambda$ ,  $\kappa^*$  and  $z$  in the point  $(t, x)$ . If on the other

hand  $x = \hat{x}$ , we may assume that  $x^k \neq \hat{x}$  for all  $k$ . So,  $\lambda(t^k, x^k) \rightarrow \infty$  by Lemma 13. Now the fact that  $\kappa^*$  is bounded implies that

$$\kappa(t^k, x^k) = \min\{\lambda(t^k, x^k)^{-1}, 1\} \kappa^*(t^k, z(t^k, x^k)) \rightarrow 0 = \kappa(t, \hat{x}) = \kappa(t, x).$$

To prove the Properties (1) to (3), take a fixed point  $(t, x) \in \mathcal{T} \times \Delta$ .

(1)  $\|\kappa(t, x)\|_\infty \leq d(t, \text{id})$ .

If  $x = \hat{x}$  then  $\kappa(t, x) = 0$ , so clearly  $\|\kappa(t, x)\|_\infty \leq d(t, \text{id})$ . If  $x \neq \hat{x}$  then the definitions of  $\kappa$  and  $\kappa^*$  show that  $\|\kappa(t, x)\|_\infty \leq \|\kappa^*(t, z(t, x))\|_\infty \leq d(t, \text{id})$ .

(2) If  $x \in \Delta(\kappa(t, x))$ , then  $x \in \Pi(t)$ .

Assume that  $x \notin \Pi(t)$ . We will show that  $x \notin \Delta(\kappa(t, x))$ .

(a) Since  $x \notin \Pi(t)$ , we know that  $x \neq \hat{x}$  by Lemma 5. So,  $z(t, x)$  is defined and  $(t, z(t, x)) \in Z$ . Then by Lemma 12 we can take a pair  $(i, a)$  such that  $(t, z(t, x)) \in G^{-ia}$ . We will first show that  $\kappa(t, x)_{ia} = z(t, x)_{ia}$ . To this end, note that  $\lambda(t, x) < 1$  since, by assumption,  $x \notin \Pi(t)$ . So,

$$\kappa(t, x)_{ia} = \min\{\lambda(t, x)^{-1}, 1\} \kappa^*(t, z(t, x))_{ia} = \kappa^*(t, z(t, x))_{ia}.$$

Next, since  $(t, z(t, x)) \in G^{-ia}$ , we know that  $d((t, z(t, x)), G^{-ia}) = 0$ . So,

$$\begin{aligned} \kappa^*(t, z(t, x))_{ia} &= \min\{d(t, \text{id}), [z, (t, x)_{ia} - d((t, z(t, x)), G^{-ia})]_+\} \\ &= \min\{d(t, \text{id}), z(t, x)_{ia}\}. \end{aligned}$$

Now by Corollary 1 we know that  $z(t, x)_{ia} \leq d(t, \text{id})$ . Hence, together with the displayed equalities this yields  $\kappa(t, x)_{ia} = z(t, x)_{ia}$ .

(b) To complete the proof, note that  $z(t, x)_{ia} \leq d(t, \text{id}) \leq |A|^{-3} < |A|^{-1} \leq \hat{x}_{ia}$ . So,  $\hat{x}_{ia} - z(t, x)_{ia} > 0$ . Therefore, since  $\hat{x} - x = \lambda(t, x)^{-1}(\hat{x}_{ia} - z(t, x))$ , we get that  $\hat{x}_{ia} - x_{ia} > 0$ . Secondly, since  $x \notin \Pi(t)$ , we know that  $1 - \lambda(t, x) > 0$ . Hence, using these last two inequalities and (a), we get that

$$\kappa(t, x)_{ia} = z(t, x)_{ia} = x_{ia} + (1 - \lambda(t, x))(\hat{x}_{ia} - x_{ia}) > x_{ia}.$$

Hence,  $x \notin \Delta(\kappa(t, x))$ .

(3) If  $x \in \Pi(t)$  and  $\kappa(t, x)_{ia} = x_{ia}$ , then  $(t, x) \in G^{-ia}$ .

Suppose that  $x \in \Pi(t)$  and  $\kappa(t, x)_{ia} = x_{ia}$ , for some pair  $(i, a)$ . We will prove that  $(t, x) \in G^{-ia}$ .

(a) We will first show that  $\lambda(t, x) = 1$  and  $z(t, x) = x$ . Since  $\kappa(t, x)_{ia} = x_{ia}$ , it can easily be seen that  $x_{ia} < \hat{x}_{ia}$ . This implies that  $x \neq \hat{x}$ . So,  $\lambda(t, x)$  is defined and  $\lambda(t, x) \geq 1$  since  $x \in \Pi(t)$ . Assume that  $\lambda(t, x) > 1$ . Then on one hand, since  $z(t, x) = \hat{x} + \lambda(t, x)(x - \hat{x})$  and  $x_{ia} < \hat{x}_{ia}$  it implies that  $z(t, x)_{ia} < x_{ia}$ . On the other hand it implies that

$$x_{ia} = \kappa(t, x)_{ia} = \lambda(t, x)^{-1} \kappa^*(t, z(t, x))_{ia} \leq \kappa^*(t, z(t, x))_{ia} \leq z(t, x)_{ia}$$

where the last inequality easily follows from the definition of  $\kappa^*$ . Contradiction. Hence,  $\lambda(t, x) = 1$ , and  $z(t, x) = \hat{x} + \lambda(t, x)(x - \hat{x}) = \hat{x} + (x - \hat{x}) = x$ .

(b) If  $x_{ia} = 0$ , then we know that  $(t, x) \in G^{-ia}$  by Lemma 11. So we may assume that  $x_{ia} > 0$ . Since  $x \neq \hat{x}$ ,  $\lambda(t, x) = 1$  and  $x = z(t, x)$  by (a), our assumption that  $\kappa(t, x)_{ia} = x_{ia}$  implies that

$$\begin{aligned} x_{ia} &= \kappa(t, x)_{ia} = \min\{\lambda(t, x)^{-1}, 1\} \kappa^*(t, z(t, x))_{ia} \\ &= \kappa^*(t, x)_{ia} = \min\{d(t, \text{id}), [x_{ia} - d((t, x), G^{-ia})]_+\}. \end{aligned}$$

So,  $x_{ia} \leq [x_{ia} - d_H((t, x), G^{-ia})]_+$ . Since  $x_{ia} > 0$ , this implies that  $d_H((t, x), G^{-ia})$  equals zero. Hence,  $(t, x) \in G^{-ia}$ , since  $G^{-ia}$  is closed by Lemma 11.  $\square$



## References

- van Damme, Eric. 1984. A relation between perfect equilibria in extensive form games and proper equilibria in normal form games. *Internat. J. Game Theory* **13** 1–13.
- van Damme, Eric. 1994. Strategic equilibrium. Robert J. Aumann and Sergiu Hart, eds., *Handbook of Game Theory with Economic Applications*, Vol. 3. Elsevier.
- Govindan, Srihari. 1995. Every stable set contains a fully stable set. *Econometrica* **63** 191–193.
- Hillas, John. 1990. On the definition of the strategic stability of equilibria. *Econometrica* **58** 1365–1390.
- . 1997. On the relation between perfect equilibria in extensive form games and proper equilibria in normal form games. *Res. Memorandum*.
- , Elon Kohlberg. 1994. Foundations of strategic equilibrium. Robert J. Aumann and Sergiu Hart, eds., *Handbook of Game Theory with Economic Applications*, Vol. 3. Elsevier.
- Kakutani, Shizuo. 1941. A generalization of Brouwer's fixed-point theorem. *Duke Math J.* **8** 457–459.
- Kohlberg, Elon, Jean-François Mertens. 1986. On the strategic stability of equilibria, *Econometrica* **54** 1003–1038.
- Mailath, George, Larry Samuelson, Jeroen Swinkels. 1997. How proper is sequential equilibrium? *Games and Econom. Behaviour* **18** 193–218.
- McLennan, Andrew. 1989a. Consistent conditional beliefs in noncooperative game theory. *Internat. J. Game Theory* **18** 141–174.
- . 1989b. Fixed points of contractible valued correspondences. *Internat. J. Game Theory* **18** 175–184.
- Mertens, Jean-François. 1987. Ordinality in non-cooperative games. CORE discussion paper 8728, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.
- . 1989. Stable equilibria—a reformulation, part I: Definition and basic properties. *Math. Oper. Res.* **14** 575–624.
- . 1991. Stable equilibria—a reformulation, part II: Discussion of the definition and further results. *Math. Oper. Res.* **16** 694–753.
- . 1992. The small worlds axiom for stable equilibria. *Games Econom. Behaviour* **4** 553–564.
- Munkres, James R. 1984. *Elements of Algebraic Topology*. Addison-Wesley Publishing Company, CA.
- Nash, John. 1950. Equilibrium points in  $N$ -person games. *Proc. Nat. Acad. Sci.* **36**, Washington, D.C., 48–49.
- Selten, Reinhard. 1975. Reexamination of the perfectness concept for equilibrium points in extensive games. *Internat. J. Game Theory* **4** 25–55.
- Vermeulen, A. J., M. J. M. Jansen. 1999. An ordinal selection of stable sets in the sense of Hillas. *Res. Memorandum* RM 99020, Maastricht University, The Netherlands.
- , J. A. M. Potters, M. J. M. Jansen. 1997. On stable sets of equilibria. A. O. Parthasarathy, ed., *Game Theory and Applications to Economics and Operations Research*. Kluwer Academic Publishers, The Netherlands, 133–148.
- Vermeulen, Dries. 1996. Stability in noncooperative game theory. Ph.D thesis, University of Nijmegen, The Netherlands.
- J. Hillas. Department of Economics, University of Auckland, Private Bag 92019, Auckland, New Zealand; e-mail: j.hillas@auckland.ac.nz
- M. Jansen. Department of Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands; e-mail: m.jansen@ke.unimaas.nl
- J. Potters. Department of Mathematics, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands; e-mail: potters@sci.kun.nl
- D. Vermeulen. Department of Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands; e-mail: d.vermeulen@ke.unimaas.nl