Static Stability in Symmetric and Population Games

Igal Milchtaich[*]

November 2016

Static stability in strategic games differs from dynamic stability in only considering the players' incentives to change their strategies. It does not rely on any assumptions about the players' reactions to these incentives and it is thus independent of the law of motion (e.g., whether players move simultaneously or sequentially). This paper presents a general notion of static stability in symmetric N-player games and population games, of which evolutionarily stable strategy (ESS) and continuously stable strategy (CSS) are essentially special cases. Unlike them, the proposed stability concept does not depend on the existence of special structures in the game such as multilinear payoff functions or unidimensional strategy spaces. *JEL Classification:* C72.

Keywords: Static stability, evolutionarily stable strategy, continuously stable strategy, risk dominance, potential games.

1 Introduction

A strategic game is at equilibrium when the players do not have any incentives to act differently than they do. In other words, at an equilibrium point, no player can increase his payoff by unilaterally changing his strategy. Stability differs in referring to the effects – either on the players' incentives or on the actual strategy choices – of starting at *another*, typically close-by, point. Notions of stability that only examine incentives may be broadly classified as static, and those that look at the consequent changes of strategies may be referred to as dynamic. (For a brief review of some additional notions of stability in strategic games, which fit neither of these categories, see Appendix A.) Dynamic stability necessarily concerns a specific law of motion, such as the replicator dynamics. It thus depends both on the game itself, that is, on the payoff functions, and on the choice of dynamics. Static stability, by contrast, depends only on intrinsic properties of the game, and is hence arguably the more basic, fundamental concept. This is not an assessment of the relative importance of the two kinds of stability but of the logical relation between them.

This paper introduces a notion of static stability in symmetric N-player games and population games that is universal in that it does not depend on structures or properties that only certain kinds of games have, such as multilinearity of the payoff function. Stability may be local or global. The former refers to a designated topology on the strategy space, which gives a meaning to a *neighborhood* of a strategy. In this paper, there are no restrictions on the choice of topology, which may in particular be the trivial topology, where the only neighborhood of any strategy is the entire strategy space. The latter choice corresponds to global stability, and it is most natural in the case of finite strategy spaces.

[*] Department of Economics, Bar-Ilan University, Ramat Gan 5290002, Israel
igal.milchtaich@biu.ac.il http://faculty.biu.ac.il/~milchti

The definition of (static) stability in this paper is based on a very simple idea, namely, examination of the incentive to switch to a strategy x from another strategy y in all states in which only these two strategies are used, that is, every player uses either x or y. The effect on a player's payoff of moving from the group of y users to that of x users normally varies with the relative sizes of the two groups. Strategy y is said to be stable if the condition that the change of payoff is *on average* negative holds for all strategies x in some neighborhood of y. This stability condition represents minimal divergence from the (Nash) equilibrium concept. However, the latter does not imply stability and, in general, the reverse implication also does not hold. The paper's focus is on stable equilibrium strategies, which satisfy *both* conditions. For a number of large, important classes of symmetric and population games, it examines the specific meaning of stability (in the above sense) in the class. In some cases, the latter turns out to be equivalent, or essentially so, to an established "native" notion of stability. Evolutionary stability for symmetric $n \times n$ games and continuous stability for games with unidimensional strategy spaces are examples of this. The definition outlined above thus turns static stability from a generic notion to a concrete, well-defined one.

The reliance of static stability solely on incentives makes it particularly suitable for comparative statics analysis, in particular, study of the welfare effects of altruism and spite. Whether people in a group where everyone shares such sentiments are likely to fare better or worse than where people are indifferent to the others' payoffs turns out to depend on the static stability or instability of the corresponding equilibrium strategies (Milchtaich 2012). If the strategies are stable, welfare tends to increase with increasing altruism or decreasing spite, but if they are (definitely) unstable, the effect may go in the opposite direction. Thus, Samuelson's (1983) "correspondence principle", which maintains that conditions for stability coincide with those under which comparative statics analysis leads to what are usually regarded as "normal" conclusions, holds. However, this is so only if 'stability' refers to the notion of static stability presented in this paper. The principle may not hold for other kinds of stability. In particular, dynamic stability under the continuous-time replicator dynamics (Hofbauer and Sigmund 1998) does not guarantee a positive relation between altruism and welfare. Even in a symmetric 3×3 game, a continuously increasing degree of altruism may actually lower the players' (identical material) payoffs if the equilibria involved are dynamically stable, which, as indicated, cannot happen if the equilibrium strategies are statically stable.

2 Symmetric and population games

A *symmetric N-player game* ($N \geq 1$) is a real-valued (payoff[1]) function $g: X^N \to \mathbb{R}$ that is defined on the N-times Cartesian product of a (finite or infinite nonempty) set X, the players' common *strategy space*, and is invariant under permutations of its second through Nth arguments. If one player uses strategy x and the others use $y, z, \dots, w$, in any order, the first player's payoff is $g(x, y, z, \dots, w)$. A strategy y is a (symmetric Nash) *equilibrium strategy* in g, with the *equilibrium payoff* $g(y, y, \dots, y)$, if it is a best response to itself, that is, for every other strategy x,

[1] In this paper, the payoff function and the game itself are identified.

$$g(y, y, \dots, y) \geq g(x, y, \dots, y). \tag{1}$$

A *population game*, as defined in this paper, is formally a symmetric two-player game g such that the strategy space X is a convex set in a (Hausdorff real) linear topological space (for example, the unit simplex in a Euclidean space) and $g(x, y)$ is continuous in y for all $x \in X$. However, the game is interpreted not as representing an interaction between two specific players but as one involving an (effectively) infinite[2] population of individuals who are "playing the field". This means that an individual's payoff $g(x, y)$ depends only on his own strategy x and on the *population strategy* y. The latter may, for example, be the population's *mean* strategy with respect to some nonatomic measure, which attaches zero weight to each individual. In this case, the meaning of the equilibrium condition,

$$g(y, y) = \max_{x \in X} g(x, y), \tag{2}$$

is that, in a *monomorphic* population, where everyone plays y, single individuals cannot increase their payoff by choosing any other strategy. Alternatively, a population game g may describe a dependence of an individual's payoff on the *distribution* of strategies in the population (Bomze and Pötscher 1989), with the latter expressed by the population strategy y. In this case, X consists of *mixed strategies*, that is, probability measures on some underlying space of allowable actions or (pure[3]) strategies, and $g(x, y)$ is linear in x and expresses the expected payoff for an individual whose choice of strategy is random with distribution x. Provided the space X is rich enough, the equilibrium condition (2) now means that the population strategy y is supported in the collection of all best response pure strategies. In other words, the (possibly) *polymorphic* population is in an equilibrium state.

Example 1. *Random matching in a symmetric multilinear game* (Bomze and Weibull 1995; Broom et al. 1997). The N players in a symmetric N-player game g are picked up independently and randomly from an infinite population of potential players. The strategy space X is a convex set in a linear topological space, and g is continuous and is linear in each of its arguments. (This assumption may be relaxed by dropping the linearity requirement for the first argument.) Because of the multilinearity of g, a player's expected payoff only depends on his own strategy x and on the population's mean strategy y. Specifically, the expected payoff is given by

$$\bar{g}(x, y) \stackrel{\text{def}}{=} g(x, y, \dots, y). \tag{3}$$

This defines a population game $\bar{g}: X^2 \to \mathbb{R}$. Clearly, a strategy y is an equilibrium strategy in $\bar{g}$ if and only if it is an equilibrium strategy in the underlying N-player game g.

Another example of a population game is a nonatomic congestion game with a continuum of identical users (Milchtaich 2012, Example 4).

[2] An infinite population may represent the limiting case of an increasingly large population, with the possible effect of each player's actions on each of the other players correspondingly decreasing. Alternatively, it may represent all possible characteristics of players, or *potential* players, when the number of *actual* players is finite.

[3] "Pure" and "mixed" are relative terms. In particular, a pure strategy may itself be a probability vector.

3 Static stability

By far the most well-known kind of static stability in symmetric two-player games and population games is evolutionary stability (Maynard Smith 1982).

Definition 1. A strategy y in a symmetric two-player game or population game g is an *evolutionarily stable strategy* (ESS) or a *neutrally stable strategy* (NSS) if, for every strategy $x \neq y$, for sufficiently small[4] $\epsilon > 0$ the inequality

$$g(y, \epsilon x + (1-\epsilon)y) > g(x, \epsilon x + (1-\epsilon)y) \tag{4}$$

or a similar weak inequality, respectively, holds. An ESS or NSS *with uniform invasion barrier* satisfies the stronger condition obtained by interchanging the two logical quantifiers. That is, for sufficiently small $\epsilon > 0$ (which cannot vary with x), (4) or a similar weak inequality, respectively, holds for all $x \neq y$.

Continuous stability (Eshel and Motro 1981; Eshel 1983) is another kind of static stability, which is applicable to games with a unidimensional strategy space.

Definition 2. In a symmetric two-player game or population game g with a strategy space that is a (finite or infinite) interval in the real line $\mathbb{R}$, an equilibrium strategy y is a *continuously stable strategy* (CSS) if it has a neighborhood where, for every strategy $x \neq y$, for sufficiently small $\epsilon > 0$ the inequality

$$g(x + \epsilon(y - x), x) > g(x, x) \tag{5}$$

holds and a similar inequality with ϵ replaced by $-\epsilon$ does not hold.

In other words, a strategy y that satisfies the "global" condition of being an equilibrium strategy[5] is a CSS if it also satisfies the "local" condition (known as *m*-stability or convergence stability; Taylor 1989; Christiansen 1991) that every nearby strategy x is not a best response to itself, specifically, any small deviation from x towards y, but not in the other direction, increases the payoff.

Yet another static notion of stability in symmetric and (with $N = 2$) population games is local superiority (or strong uninvadability; Bomze 1991).

Definition 3. A strategy y in a symmetric N-player game or population game g is *locally superior* if it has a neighborhood where, for every strategy $x \neq y$,

$$g(y, x, \dots, x) > g(x, x, \dots, x). \tag{6}$$

Local superiority is applicable to any symmetric or population game in which the strategy space is a topological space, so that the notion of neighborhood is well defined.[6] It does not rely on any other properties of the strategy space or of the payoff function – unlike ESS and

---

[4] A condition holds for "sufficiently small" $\epsilon > 0$ if there is some $\delta > 0$ such that the condition holds for all $0 < \epsilon < \delta$.

[5] The original definition of CSS differs slightly from the version given here in requiring a stronger global condition, which is a version of ESS.

[6] A subset of X is said to be a *neighborhood* of a strategy x if its interior includes x (Kelley 1955).

CSS, which would not be meaningful without a linear structure. It is well known (see Section 6) that in the special case of symmetric $n \times n$ games local superiority is in fact equivalent to the ESS condition. However, for games with a unidimensional strategy space (Section 7) local superiority and CSS are not equivalent.

The next section presents a universal notion of static stability, that is, one which is applicable to all symmetric and population games. It (essentially) gives ESS and CSS as special cases when applied to specific, suitable classes of games.

4 A general framework

Inequality (1) in the equilibrium condition and inequality (6) in the definition of local superiority both concern a player's lack of incentive to use a particular alternative x to his strategy y. In the equilibrium condition, all the other players are using y, and in local superiority, they all use x. Stability, as defined below, differs from both concepts in considering not only the incentives to be first or last to move to x from y but also all the intermediate cases. In the simplest version, described by the followings definition (and extended in Section 4.2), the same weight is attached to all cases. Put differently, stability requires that, when the players move one-by-one to x from y, the corresponding changes of payoff are negative on average.

Definition 4. A strategy y in a symmetric N-player game $g: X^N \to \mathbb{R}$ is *stable*, *weakly stable* or *definitely unstable* if it has a neighborhood where, for every strategy $x \neq y$, the inequality

$$\frac{1}{N}\sum_{j=1}^{N}\left(g(x,\underbrace{x,\dots,x}_{j-1\text{ times}},\underbrace{y,\dots,y}_{N-j\text{ times}}) - g(y,\underbrace{x,\dots,x}_{j-1\text{ times}},\underbrace{y,\dots,y}_{N-j\text{ times}})\right) < 0, \tag{7}$$

a similar weak inequality or the reverse (strict) inequality, respectively, holds.

Stability, as defined here, is a local concept. It refers to neighborhood systems of strategies, or equivalently to a topology on the strategy space X. The topology may be explicitly specified or it may be understood from the context. The latter applies when it is natural to view X as a subspace of a Euclidean space or some other standard topological space, so that its topology is the relative one. For example, if the strategy space is an interval in the real line $\mathbb{R}$, so that strategies are simply (real) numbers, a set of strategies is a neighborhood of a strategy y if and only if, for some $\varepsilon > 0$, every $x \in X$ with $|x - y| < \varepsilon$ is in the set. In a game with a finite number of strategies, it may seem natural to consider the discrete topology, that is, to view strategies as isolated. However, as discussed in Section 5 below, a more useful choice of topology in a finite game is the trivial, or indiscrete, topology. This choice effectively puts topology out of the way, since it means that the only neighborhood of any strategy is the entire strategy space. The trivial topology may be used also with an infinite X. Stability, weak stability or definite instability of a stategy y with respect to the trivial topology automatically implies the same with respect to any other topology. Such a strategy y will be referred to as *globally* stable, weakly stable or definitely unstable, respectively. Note that there can be at most one globally stable strategy.

In some classes of games (see Sections 6, 8 and 9), stability of a strategy automatically implies that it is an equilibrium strategy. In other games, the reverse implication holds. In particular, an equilibrium strategy is automatically *globally* stable in every symmetric game satisfying *symmetric substitutability* (see Milchtaich 2012, Section 6), which is the condition that for all strategies $x, y, z, \dots, w$ with $x \neq y$

$$g(x, x, z, \dots, w) - g(y, x, z, \dots, w) < g(x, y, z, \dots, w) - g(y, y, z, \dots, w).$$

The condition implies that the expression in parentheses in (7) strictly decreases as j increases from 1 to N. If y is an equilibrium strategy, then by (1) the expression is nonpositive for $j = 1$, which means that the whole sum is negative, which proves that y is globally stable.

In general, however, the equilibrium and stability conditions are incomparable: neither of them implies the other. The incomparability is partially due to equilibrium being a global condition: all alternative strategies, not only nearby ones, are considered. However, it holds also with 'equilibrium' replaced by 'local equilibrium' (with the obvious meaning) or if the strategy space has the trivial topology, which obviates the distinction between local and global. A *stable equilibrium strategy* is a strategy that satisfies both conditions. It is not difficult to see that in the special case of symmetric two-player games, where the equilibrium condition can be written as (2) and inequality (7) can be rearranged to read

$$\frac{1}{2}\big(g(x, x) - g(y, x) + g(x, y) - g(y, y)\big) < 0, \tag{8}$$

a strategy y is a stable equilibrium strategy if and only if it has a neighborhood where, for every $x \neq y$, the inequality

$$pg(x, x) + \ (1 - p)g(x, y) < pg(y, x) + (1 - p)g(y, y) \tag{9}$$

holds for all $0 < p \leq 1/2$. This condition means that the alternative strategy x affords a lower expected payoff than y against an uncertain strategy that may be x or y, with the former no more likely than the latter.

Local superiority is similar to stability in being a local condition. Moreover, for equilibrium strategies in symmetric two-player games, locally superiority implies stability, since (with $N = 2$) inequality (8) can be obtained by averaging (1) and (6). The same implication holds also for certain kinds of symmetric games with more than two players (see Section 9). The reverse implication does not hold even for equilibrium strategies in symmetric two-player games (see Section 7).

4.1 Stability in population games

Stability in population games can be defined by a variant of Definition 4 that replaces the number of players using strategy x or y with the size of the subpopulation to which the strategy applies, p or $1 - p$ respectively. Correspondingly, the sum in (7) is replaced with an integral.

Definition 5. A strategy y in a population game $g: X^2 \to \mathbb{R}$ is *stable*, *weakly stable* or *definitely unstable* if it has a neighborhood where, for every strategy $x \neq y$, the inequality

$$\int_0^1 \big(g(x, px + (1-p)y) - g(y, px + (1-p)y)\big)\, dp < 0, \tag{10}$$

a similar weak inequality or the reverse (strict) inequality, respectively, holds.

The difference between stability in the sense of Definition 5 and in the sense of (the two versions of) ESS (Definition 1) boils down to a different meaning of proximity between population strategies. The definition of ESS reflects the view that a population strategy is close to y if the latter applies to a large subpopulation, of size $1-\epsilon$, and another strategy x applies to a small subpopulation, of size ϵ. By contrast, in Definition 5, the subpopulation to which x applies need not be small, but x itself is assumed close to y. The significance of this difference between the definitions is examined in Sections 6 and 9.

If a population game $\bar{g}$ is derived from a symmetric multilinear game g as in Example 1, then, depending on whether y is viewed as a strategy in g or $\bar{g}$, Definition 4 or 5 applies. However, the point of view turns out to be immaterial.

Proposition 1. A strategy y in a symmetric multilinear N-player game g is stable, weakly stable or definitely unstable if and only if it has the same property in the population game $\bar{g}$ defined by (3).

Proof. For $0 \le p \le 1$, and for strategies x, y and

$$x_p = px + (1-p)y,$$

the linearity of g in each of its second through Nth arguments and its invariance to permutations of these arguments give

$$\begin{aligned}\bar{g}(x, x_p) - \bar{g}(y, x_p) &= g(x, x_p, \dots, x_p) - g(y, x_p, \dots, x_p) \\ &= \sum_{j=1}^{N} B_{j-1,N-1}(p)\big(g(x, \underbrace{x, \dots, x}_{j-1 \text{ times}}, \underbrace{y, \dots, y}_{N-j \text{ times}}) - g(y, \underbrace{x, \dots, x}_{j-1 \text{ times}}, \underbrace{y, \dots, y}_{N-j \text{ times}})\big),\end{aligned} \tag{11}$$

where

$$B_{j-1,N-1}(p) = \binom{N-1}{j-1} p^{j-1}(1-p)^{N-j}, \qquad j = 1,2, \dots, N$$

are the Bernstein polynomials. These polynomials satisfy the equalities

$$\int_0^1 B_{j-1,N-1}(p)\, dp = \frac{1}{N}, \qquad j = 1,2, \dots, N. \tag{12}$$

It therefore follows from (11) by integration that the expression obtained by replacing g on the left-hand side of (10) by $\bar{g}$ is equal to the expression on the left-hand side of (7). ∎

In the subsequent sections, Definitions 4 and 5 are applied, or restricted, to a number of specific classes of symmetric and population games and the results are compared with certain “native” notions of stability for these games. The rest of the present section is concerned with an extension of the above framework, which facilitates the capturing of some additional native notions of stability.

4.2 $\bar{p}$-stability

Stability as defined above in a sense occupies the midpoint between equilibrium and local superiority. It takes into consideration a player's incentive to be the first or last to switch to a particular alternative strategy, but attaches to these extreme cases the same weight it attaches to each of the intermediate ones. This uniform distribution of weights may be interpreted as expressing a particular belief of the player about the total number of players who will be using the alternative strategy after he switches to it, with the rest using the original strategy. Namely, the probabilities $p_1, p_2, \dots, p_N$ that the number in question is $1,2, \dots, N$ are all equal,

$$p_j = \frac{1}{N}, \qquad j = 1,2, \dots, N. \tag{13}$$

Thus, unlike local superiority, in which the gain from switching from strategy y to the alternative strategy x is computed under the belief that all the other players are using x, in stability the expected gain is with respect to the probabilities (13), which give the expression on the left-hand side of (7). A straightforward generalization of both concepts is to allow *any* beliefs.

Definition 6. For a probability vector $\bar{p} = (p_1, p_2, \dots, p_N)$, a strategy y in a symmetric N-player game $g: X^N \to \mathbb{R}$ *$\bar{p}$-stable*, *weakly $\bar{p}$-stable* or *definitely $\bar{p}$-unstable* if it has a neighborhood where, for every strategy $x \neq y$, the inequality

$$\sum_{j=1}^{N} p_j(g(x, \underbrace{x, \dots, x}_{j-1 \text{ times}}, \underbrace{y, \dots, y}_{N-j \text{ times}}) - g(y, \underbrace{x, \dots, x}_{j-1 \text{ times}}, \underbrace{y, \dots, y}_{N-j \text{ times}})) < 0, \tag{14}$$

a similar weak inequality or the reverse (strict) inequality, respectively, holds.

If each of the other players switches to x with probability p and stays with y with probability $1 - p$, then, depending on whether the choices are, respectively, perfectly correlated (i.e., identical) or independent,

$$p_j = \begin{cases} 1 - p, & j = 1 \\ 0 & , 1 < j < N \\ p & , j = N \end{cases} \tag{15}$$

or

$$p_j = \binom{N-1}{j-1} p^{j-1}(1-p)^{N-j} = B_{j-1,N-1}(p), \qquad j = 1,2, \dots, N. \tag{16}$$

A strategy y is *dependently-* or *independently-stable* if it is $\bar{p}$-stable with $\bar{p} = (p_1, p_2, \dots, p_N)$ given by (15) or (16), respectively, for *all* $0 < p < 1$.

The number of other players using strategy x and the number using y have a symmetric joint distribution if the two numbers are equally distributed (hence, have an equal expectation of $(N-1)/2$), that is,

$$p_j = p_{N-j+1}, \qquad j = 1,2, \dots, N. \tag{17}$$

For $\bar{p} = (p_1, p_2, \dots, p_N)$ satisfying (17), the left-hand side of (14) is equal to the more symmetrically-looking expression

$$G_{\bar{p}}(x,y) \stackrel{\text{def}}{=} \sum_{j=1}^{n} p_j(g(\underbrace{x,\dots,x}_{j\text{ times}},y,\dots,y) - g(\underbrace{y,\dots,y}_{j\text{ times}},x,\dots,x)).$$

Thus, for such $\bar{p}$, a strategy y is $\bar{p}$-stable if and only if it has a neighborhood where it is the unique best response to itself in the symmetric two-player zero-sum game $G_{\bar{p}}: X^2 \to \mathbb{R}$. As a special case, this characterization applies to stability, that is, to $\bar{p}$ given by (13). A strategy y is *symmetrically-stable* if it is $\bar{p}$-stable for all $\bar{p}$ satisfying (17).

For single-player games ($N = 1$), stability and $\bar{p}$-stability of a strategy mean the same thing, namely, strict local optimality: switching to any nearby alternative strategy reduces the payoff. For $N = 2$, stability does not generally imply $\bar{p}$-stability (or vice versa) but the implication does partially hold (specifically, holds whenever $0 < p_2 \leq 1/2$) in the special case of an equilibrium strategy (see (9)). A full appreciation of the differences between stability in the sense of Definition 4 and the varieties based on $\bar{p}$-stability requires looking at multiplayer games. One class of such games is examined in Section 9.

5 Finite games and risk dominance

In every symmetric or population game, every isolated strategy is trivially stable. Therefore, if the strategy space X has the discrete topology, that is, all singletons are open sets, then all strategies are stable. The definition of stability is therefore of interest only for games with non-discrete strategy spaces. This includes games with a finite number of strategies where the topology on X is the trivial one, so that stability and definite instability mean *global* stability and definite instability (see Section 4). The simplest (interesting) such game is a finite symmetric two-player game with only two strategies, strategy 1 and strategy 2, for example, the game with payoff matrix

$$\begin{array}{cc} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} 3 & 4 \\ 1 & 5 \end{pmatrix} \end{array}$$

(where the rows correspond to the player's own strategy and the columns to the opponent's strategy). In this example, both strategies are equilibrium strategies. Strategy 1 is globally stable and strategy 2 is globally definitely unstable, because (using the form (7) of (8))

$$\frac{1}{2}(5-4+1-3) < 0 < \frac{1}{2}(3-1+4-5).$$

The two inequalities, which are clearly equivalent, have an additional meaning. Namely, they express the fact that (1,1) is the *risk dominant* equilibrium (Harsanyi and Selten 1988). It is not difficult to see that this coincidence of global stability and risk dominance holds in general – it is not a special property of the payoffs in this example.

Proposition 2. In a finite symmetric two-player game with two strategies, an equilibrium strategy y is globally stable if and only if the equilibrium (y, y) is risk dominant.

For a pure equilibrium strategy y, risk dominance of (y, y) is equivalent also to global stability of y in the mixed extension g of the finite game, that is, when mixed strategies are allowed. This follows from the fact that global stability of y in the finite game implies that inequality (9) holds for all $0 < p \leq 1/2$, where x is the other pure strategy. Because of the bilinearity of g, the same is then true with x replaced by any convex combination of x and y other than y itself, which proves that y is globally stable in g. However, since in the mixed extension of the finite game, which is a symmetric 2×2 game, the strategy space X is essentially the unit interval, the natural topology on X is not the trivial topology but the usual one. Stability with respect to the latter is a weaker condition than global stability. For example, as shown in the next section, it holds for *both* pure strategies if (as in the above example) the corresponding strategy profiles are strict equilibria.

6 Symmetric $n \times n$ games and evolutionary stability

A symmetric $n \times n$ game is given by a (square) payoff matrix A with these dimensions. The strategy space X, whose elements are referred to as mixed strategies, is the unit simplex in $\mathbb{R}^n$. The interpretation is that there are n possible actions, and a strategy $x = (x_1, x_2, \dots, x_n)$ is a probability vector specifying the probability x_i with which each action i is used ($i = 1,2, \dots, n$). The set of all actions i with $x_i > 0$ is the *support* of x. A strategy is *pure* or *completely mixed* if its support contains only a single action i (in which case the strategy itself may also be denoted by i) or all n actions, respectively. The game (i.e., the payoff function) $g: X^2 \to \mathbb{R}$ is defined by

$$g(x, y) = x^{\mathrm{T}} A y$$

(where superscript T denotes transpose and strategies are viewed as column vectors). Thus, g is bilinear, and $A = \big(g(i, j)\big)_{i,j=1}^{n}$.

A symmetric $n \times n$ game may be viewed either as a symmetric two-player game or as a population game. In the former case, Definition 4 applies, and in the latter, Definition 5 applies. However, by Proposition 1, the two definitions of stability in fact coincide, and the same is true for weak stability and for definite instability. Moreover, as the next two results show, stability is also equivalent to evolutionary stability and to local superiority (see Section 3). It also follows from these results that every (even weakly) stable strategy in a symmetric $n \times n$ game is an equilibrium strategy, and every strict equilibrium strategy is stable.

The following proposition is rather well known (Bomze and Pötscher 1989; van Damme 1991, Theorem 9.2.8; Weibull 1995, Propositions 2.6 and 2.7; Bomze and Weibull 1995).

Proposition 3. For a strategy y in a symmetric $n \times n$ game g, the following conditions are equivalent:[7]

(i) Strategy y is an ESS.
(ii) Strategy y is an ESS with uniform invasion barrier.
(iii) For every strategy $x \neq y$ in some neighborhood of y,

$$g(y, x) > g(x, x). \tag{18}$$

[7] Note that condition (iii) means that y is locally superior and that the first part of (iv) means that it is an equilibrium strategy.

(iv) For every $x \neq y$, the (weak) inequality $g(y, y) \geq g(x, y)$ holds, and if it holds as equality, then (18) also holds.

An NSS is characterized by similar equivalent conditions, in which the strict inequality (18) is replaced by a weak one.

A completely mixed equilibrium strategy y in a symmetric $n \times n$ game g is said to be *definitely evolutionarily unstable* (Weissing 1991) if the reverse of inequality (18) holds for all $x \neq y$.

Theorem 1. A strategy y in a symmetric $n \times n$ game g is stable or weakly stable if and only if it is an ESS or NSS, respectively. A completely mixed equilibrium strategy is definitely unstable if and only if it is definitely evolutionarily unstable.

Proof. The two inequalities in (iii) and (iv) of Proposition 3 together imply (8), and the same is true with the strict inequalities (8) and (18) both replaced by their weak versions. This proves that a sufficient condition for stability or weak stability of a strategy y is that it is an ESS or NSS, respectively. For a completely mixed equilibrium strategy y, the inequality in (iv) automatically holds as equality for all x, and therefore a similar argument proves that a sufficient condition for definite instability of y is that it is definitely evolutionarily unstable.

In remains to prove necessity. For a stable strategy y, inequality (8) holds for all nearby strategies $x \neq y$. Therefore, y has the property that, for *every* strategy $x \neq y$, for sufficiently small $\varepsilon > 0$

$$\begin{aligned} g(\varepsilon x + (1-\varepsilon)y, \varepsilon x + (1-\varepsilon)y) - g(y, \varepsilon x + (1-\varepsilon)y) \\ + g(\varepsilon x + (1-\varepsilon)y, y) - g(y, y) < 0. \end{aligned} \tag{19}$$

It follows from the bilinearity of g that (19) is equivalent to

$$(2-\varepsilon)\big(g(y,y) - g(x,y)\big) + \varepsilon\big(g(y,x) - g(x,x)\big) > 0. \tag{20}$$

Therefore, the above property of y is equivalent to (iv) in Proposition 3, which proves that y is an ESS. Similar arguments show that a weakly stable strategy is an NSS and that a definitely unstable completely mixed equilibrium strategy is definitely evolutionarily unstable. In the first case, the proof needs to be modified only by replacing the strict inequalities in (18), (19) and (20) by weak inequalities, and in the second case (in which the first term in (20) vanishes for all x), they need to be replaced by the reverse inequalities. ■

7 Games with a unidimensional strategy space and continuous stability

In a symmetric two-player game or population game where the strategy space is an interval in the real line $\mathbb{R}$, stability or instability of an equilibrium strategy, in the sense of either Definition 4 or 5, has a simple, familiar meaning. As shown below, if the payoff function is twice continuously differentiable, and with the possible exception of certain borderline cases, the equilibrium strategy is stable or definitely unstable if, at the (symmetric) equilibrium point, the graph of the best-response function, or reaction curve, intersects the forty-five degree line from above or below, respectively. This geometric characterization of stability and its differential counterpart are also shared by continuous stability (Section 3), which shows that these two notions of static stability are essentially equivalent.

Theorem 2. Let g be a symmetric two-player game or population game with a strategy space X that is a (finite or infinite) interval in the real line $\mathbb{R}$, and y an interior equilibrium strategy (that is, one lying in the interior of X) such that g has continuous second-order partial derivatives[8] in a neighborhood of the equilibrium point (y, y). If

$$g_{11}(y,y) + g_{12}(y,y) < 0, \tag{21}$$

then y is stable and a CSS. If the reverse inequality holds, then y is definitely unstable and not a CSS.

Proof. Using Taylor's theorem, it is not difficult to show that, for x tending to y, the left-hand sides of both (8) and (10) can be expressed as

$$g_1(y,y)(x-y) + \frac{1}{2}\big(g_{11}(y,y) + g_{12}(y,y)\big)(x-y)^2 + o((x-y)^2). \tag{22}$$

Since y is an interior equilibrium strategy, the first term in (22) is zero. Therefore, a sufficient condition for (22) to be negative or positive for all $x \neq y$ in some neighborhood of y (hence, for y to be stable or definitely unstable, respectively) is that $g_{11}(y,y) + g_{12}(y,y)$ has that sign.

Next, consider the CSS condition in Definition 2. It may be possible to determine whether this condition holds by looking at the sign of

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \big(g(x,x) - g(x + \epsilon(y-x), x)\big) = g_1(x,x)(x-y). \tag{23}$$

For x tending to y, (23) is given by an expression that is similar to (22) except that it lacks the factor $1/2$. Therefore, if (21) or the reverse inequality holds, then (23) is negative or positive, respectively, for all $x \neq y$ in some neighborhood of y. In the first or second case, (5) holds or does not hold, respectively, for $\epsilon > 0$ sufficiently close to 0 and the converse is true for $\epsilon < 0$. Therefore, in the first case, y is a CSS, and in the second case, it is not a CSS. ■

The connection between inequality (21) and the slope of the reaction curve can be established as follows (Eshel 1983). If y is an interior equilibrium strategy, then it follows from the equilibrium condition (2) that $g_1(y,y) = 0$ and $g_{11}(y,y) \leq 0$. If the last inequality is in fact strict, then by the implicit function theorem there is a continuously differentiable function f from some neighborhood of y to the strategy space, with $f(y) = y$, such that $g_1(f(x), x) = 0$ and $g_{11}(f(x), x) < 0$ for all strategies x in the neighborhood. Thus, strategy $f(x)$ is a local best response to x. By the chain rule, at the point y

$$f'(y) = -\frac{g_{12}(y,y)}{g_{11}(y,y)}.$$

[8] Partial derivatives are denoted by subscripts. For example, g_{12} is the mixed second-order partial derivative of g.

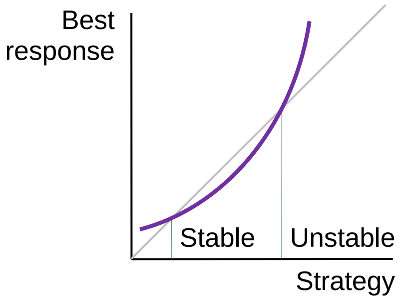

Figure 1. An equilibrium strategy is stable (and a CSS) or definitely unstable (and not a CSS) if, at the equilibrium point, the reaction curve (thick line) intersects the forty-five degree line (thin) from above or below, respectively.

Therefore, (21) holds (so that y is stable) or the reverse inequality holds (y definitely unstable) if and only if the slope of the function f at y is less or greater than 1, respectively.[9] In the first case, the reaction curve (see Figure 1), which is the graph of f, intersects the forty-five degree line from above (which means that the (local) fixed point index is $+1$; see Dold 1980). In the second case, the intersection is from below (and the fixed point index is -1).

In a population or symmetric two-player game with a unidimensional strategy space, an equilibrium strategy y that is locally superior is said to be a *neighborhood invader strategy* (Apaloo 1997). As shown (see Section 4), such a strategy is necessarily stable. However, unlike for symmetric $n \times n$ games (Section 6), the converse is false. This can be seen most easily by considering the differential condition for local superiority of an equilibrium strategy y, which differs from (21) in that the second term $g_{12}(y, y)$ is multiplied by 2 (Oechssler and Riedel 2002). Since, by the equilibrium condition, the first term $g_{11}(y, y)$ is necessarily nonpositive, this makes the condition more demanding than (21).

8 Potential games

A symmetric N-player game $g: X^N \to \mathbb{R}$ is called an (exact) potential game if it has an (exact) *potential*, which is a symmetric function (that is, a function that is invariant under permutations of its N arguments) $F: X^N \to \mathbb{R}$ such that, whenever a single player changes his strategy, the change in the player's payoff is equal to the change in F. Thus, for any $N + 1$ (not necessarily distinct) strategies $x, x', y, z, \dots, w$,

$$F(x, y, z, \dots, w) - F(x', y, z, \dots, w) = g(x, y, z, \dots, w) - g(x', y, z, \dots, w). \qquad (24)$$

It follows immediately from the definition that the potential is unique up to an additive constant. It also follows that a necessary condition for the existence of a potential is that the total change of payoff of any two players who change their strategies one after the other does not depend on the order of their moves: for any $N + 2$ strategies $x, x', y, y', z, \dots, w$,

$$\begin{aligned} &g(x, y, z, \dots, w) - g(x', y, z, \dots, w) + g(y, x', z, \dots, w) - g(y', x', z, \dots, w) \\ &\qquad = g(y, x, z, \dots, w) - g(y', x, z, \dots, w) + g(x, y', z, \dots, w) - g(x', y', z, \dots, w). \end{aligned}$$

[9] This geometric condition for static stability is weaker than the corresponding one for dynamic stability, which requires the *absolute value* of slope to be less than 1 (Fudenberg and Tirole 1995).

It is not difficult to show that this condition is also sufficient (see Monderer and Shapley 1996, Theorem 2.8, which however refers to general, not necessarily symmetric, games, for which the potential function is also not symmetric). Moreover, if g is the mixed extension of a finite game (which means that it is a symmetric $n \times n$ game; see Section 6 for definition and notation), then g is a potential game if and only if the above condition holds for (any $N + 2$) *pure* strategies (Monderer and Shapley 1996, Lemma 2.10). In this case, the potential, like the game g itself, is multilinear.

Example 2. *Symmetric* 2×2 *games.* Every symmetric 2×2 game g, with pure strategies 1 and 2, is a potential game, since it is easy to see that it satisfies the above condition for pure strategies. It is moreover not difficult to check that the following bilinear function is a potential for g:

$$F(x, y) = (g(1,1) - g(2,1))x_1 y_1 + (g(2,2) - g(1,2))x_2 y_2. \tag{25}$$

The potential F of a symmetric potential game g may itself be viewed as a symmetric N-player game, indeed, a doubly symmetric one.[10] It follows immediately from (24) that F and g have exactly the same equilibrium strategies, stable and weakly stable strategies, and definitely unstable strategies. Stability and instability in this case have a strikingly simple characterization, which follows immediately from the observation that the sum in (7) is equal the difference $F(x, x, \dots, x) - F(y, y, \dots, y)$ divided by N.

Theorem 3. In a symmetric N-player game with a potential F, a strategy y is stable, weakly stable or definitely unstable if and only if it is a strict local maximum point, a local maximum point or a strict local minimum point, respectively, of the function $x \mapsto F(x, x, \dots x)$.[11]

The following simple result illustrates the theorem. It also makes use of Theorem 1 and Example 2.

Corollary 1. In a symmetric 2×2 game g with pure strategies 1 and 2, a (mixed) strategy is an ESS or an NSS if and only if it is a strict local maximum point or a local maximum point, respectively, of the quadratic function $\Phi: X \to \mathbb{R}$ defined by

$$\Phi(x) = \frac{1}{2}(g(1,1) - g(2,1))x_1^2 + \frac{1}{2}(g(2,2) - g(1,2))x_2^2. \tag{26}$$

8.1 Potential in population games

For population games, which represent interactions involving many players whose individual actions have negligible effects on the other players, the definition of potential may be naturally adapted by replacing the difference on the left-hand side of (24) with a derivative.

Definition 7. For a population game $g: X^2 \to \mathbb{R}$, a continuous function $\Phi: X \to \mathbb{R}$ is a *potential* if for all $x, y \in X$ and $0 < p < 1$ the derivative on the left-hand side of the following equality exists and the equality holds:

---

[10] A symmetric game is *doubly symmetric* if it has a symmetric payoff function, in other words, if the players' payoffs are always equal.

[11] Of course, if $(y, y, \dots, y)$ is a global maximum point of F itself, then y is in addition an equilibrium strategy.

$$\frac{d}{dp}\Phi(px+(1-p)y) = g(x, px+(1-p)y) - g(y, px+(1-p)y). \tag{27}$$

Example 3. *Symmetric* 2×2 *games, viewed as population games.* It is easy to check that, for every such game g, with pure strategies 1 and 2, the function Φ defined by (26) is a potential. Note that, unlike the function F defined in (25), Φ is a function of one variable only.

Example 3 and Corollary 1 hint at the following general result.[12] As for symmetric games, stability and instability (here, in the sense of Definition 5) of a strategy y in a population game with a potential Φ is related to y being a local extremum point of the potential.

Theorem 4. In a population game g with a potential Φ, a strategy y is stable, weakly stable or definitely unstable if and only if it is a strict local maximum point, local maximum point or strict local minimum point of Φ, respectively. In the first two cases, y is in addition an equilibrium strategy. If the potential Φ is strictly concave, an equilibrium strategy is also a strict *global* maximum point of Φ, and necessarily the game's unique stable strategy.

Proof. By (27), the left-hand side of (10) can be written as

$$\int_0^1 \frac{d}{dp}\Phi(px+(1-p)y)\,dp.$$

This integral equals $\Phi(x)-\Phi(y)$, which proves the first part of the theorem. It also follows from (27), in the limit $p\to 0$, that for all x and y

$$\left.\frac{d}{dp}\right|_{p=0^+}\Phi(px+(1-p)y) = g(x,y) - g(y,y). \tag{28}$$

If y is a local maximum point of Φ, then the left-hand side of (28) is nonpositive, which proves that y is an equilibrium strategy.

To prove the last part of the theorem, consider an equilibrium strategy y. For any strategy $x\neq y$, the right-, and therefore also the left-, hand side of (28) is nonpositive. If Φ is strictly concave, this implies that the left-hand side of (27) is negative for all $0<p<1$, which proves that y is a strict global maximum point of Φ. ∎

Since by definition a potential is a continuous function, an immediate corollary of Theorem 4 is the following result, which concerns the existence of strategies that are (at least) weakly stable. The result sheds light on the difference in this respect between symmetric 2×2 games and, for example, 3×3 games. The former, which as indicated are potential games, always have at least one NSS, whereas for the latter, it is well known that this is not so (a counterexample is a variant of the rock–scissors–paper game where a draw yields a small positive payoff for both players; see Maynard Smith 1982, p. 20).

---

[12] Conversely, Example 3 and Theorem 4 below together provide an alternative proof for Corollary 1.

Corollary 2. If a population game g with a potential Φ has a compact strategy space, then it has at least one weakly stable strategy. If in addition the number of such strategies is finite, they are all stable.

The term potential is borrowed from physics, where it refers to a scalar field whose gradient gives the force field. Force is analogous to incentive here. The analogy can be taken one step further by presenting the payoff function g as the differential of the potential Φ. This requires Φ to be defined not only on the strategy space X (which by definition is a convex set in a linear topological space) but on its *cone* $\hat{X}$, which is the collection of all space elements that can be written as a strategy x multiplied by a positive number t. For example, if strategies are probability measures, Φ needs to be defined for all positive, non-zero finite measures. The *differential* of the potential can then be defined as its directional derivative, that is, as the function $d\Phi: \hat{X}^2 \to \mathbb{R}$ given by[13]

$$d\Phi(\hat{x}, \hat{y}) = \left.\frac{d}{dt}\right|_{t=0^+} \Phi(t\hat{x} + \hat{y}). \tag{29}$$

The differential exists if the (right) derivative in (29) exists for all $\hat{x}, \hat{y} \in \hat{X}$.

Proposition 4. Let $g: X^2 \to \mathbb{R}$ be a population game and $\Phi: \hat{X} \to \mathbb{R}$ a continuous function (on the cone of the strategy space). A sufficient condition for the restriction of Φ to X to be a potential for g is that the differential $d\Phi: \hat{X}^2 \to \mathbb{R}$ exists, is continuous in the second argument and satisfies

$$d\Phi(x, y) = g(x, y), \qquad x, y \in X. \tag{30}$$

If this condition holds, then g is necessarily linear in the first argument x.

Proof (an outline). Using elementary arguments, the following can be established.

Fact. A continuous real-valued function defined on an open real interval is continuously differentiable if and only if it has a continuous right derivative.

Suppose that $d\Phi$ satisfies the specified condition. Replacing $\hat{y}$ in (29) with $p\hat{x} + \hat{y}$ gives

$$d\Phi(\hat{x}, p\hat{x} + \hat{y}) = \left.\frac{d}{dt}\right|_{t=p^+} \Phi(t\hat{x} + \hat{y}), \qquad \hat{x}, \hat{y} \in \hat{X}, p \geq 0. \tag{31}$$

By the above Fact and the continuity properties of Φ and $d\Phi$, for $0 < p < 1$ the right derivative in (31) is actually a two-sided derivative and it depends continuously on $\hat{y}$. Therefore, by (30), the right-hand side of (27) is equal to the expression

$$\left.\frac{d}{dt}\right|_{t=p} \Phi(tx + (1-p)y) - \left.\frac{d}{dt}\right|_{t=1-p} \Phi(px + ty),$$

which by the chain rule is equal to the derivative on the left-hand side of (27). Hence, that equality holds. Another corollary of (31) is the identity

[13] Note that, in the directional derivative $d\Phi$, the direction is specified by the first argument.

$$\int_0^t d\Phi(\hat{x}, p\hat{x} + \hat{y})\, dp = \Phi(t\hat{x} + \hat{y}) - \Phi(\hat{y}), \qquad \hat{x}, \hat{y} \in \hat{X}, t > 0,$$

which, used twice, gives

$$\int_0^{\lambda t} \big(d\Phi(\hat{z}, p\hat{z} + \lambda t\hat{x} + \hat{y}) + d\Phi(\hat{x}, p\hat{x} + \hat{y})\big)\, dp = \Phi(\lambda t\hat{z} + \lambda t\hat{x} + \hat{y}) - \Phi(\lambda t\hat{x} + \hat{y})$$

$$+\Phi(\lambda t\hat{x} + \hat{y}) - \Phi(\hat{y}) = \Phi(\lambda t\hat{z} + \lambda t\hat{x} + \hat{y}) - \Phi(\hat{y}), \qquad \hat{x}, \hat{y}, \hat{z} \in \hat{X}, \lambda, t > 0.$$

Dividing the right- and left-hand sides by t and taking the limit $t \to 0^+$ gives the identity

$$\lambda\, d\Phi(\hat{z}, \hat{y}) + \lambda\, d\Phi(\hat{x}, \hat{y}) = d\Phi(\lambda\hat{z} + \lambda\hat{x}, \hat{y}), \qquad \hat{x}, \hat{y}, \hat{z} \in \hat{X}, \lambda > 0.$$

The identify proves that $d\Phi$ (hence, g) is linear in the first argument. ∎

9 Symmetric multilinear games

Symmetric multilinear games are the N-player generalization of the two-player games considered in Section 6. The strategy space X is the unit simplex in a Euclidean space and $g: X^N \to \mathbb{R}$ is linear in each of the N arguments.

As Proposition 1 shows, stability in a symmetric multilinear game g is equivalent to stability in the population game $\bar{g}$ defined by (3). Requiring the same for *evolutionary stability* (which for a population game is given by Definition 1) yields the following natural definition.
A strategy y in a symmetric multilinear game g is said to be an ESS if, for every $x \neq y$, for sufficiently small $\epsilon > 0$ the strategy $x_\epsilon = \epsilon x + (1 - \epsilon)y$ satisfies

$$g(y, x_\epsilon, \dots, x_\epsilon) > g(x, x_\epsilon, \dots, x_\epsilon). \tag{32}$$

Strategy y is said to be an ESS with uniform invasion barrier if it satisfies the stronger condition that, for sufficiently small $\epsilon > 0$ (which cannot vary with x), inequality (32) holds for all $x \neq y$. Note that for the existence of a uniform invasion barrier it suffices that the last condition holds for *some* $0 < \epsilon < 1$, since this automatically implies the same for all smaller ϵ.

An equivalent definition of ESS is given by a generalization of condition (iv) in Proposition 3 (Broom et al. 1997; see also the proof of Lemma 3 below).

Lemma 1. A strategy y in a symmetric multilinear game g is an ESS if and only if, for every $x \neq y$, at least one of the N terms in the sum on the left-hand side of (7) is not zero, and the first such term is negative. In particular, an ESS is necessarily an equilibrium strategy (since the first term in (7) must be nonpositive).

Unlike in the special case $N = 2$ (Proposition 3), in a general multilinear game not every ESS has a uniform invasion barrier. It is easy to see that a sufficient condition for the existence of a uniform invasion barrier is that the ESS is locally superior, and this condition is in fact also necessary (Bomze and Weibull 1995, Theorem 3; Lemma 2 below). This raises the question of how stability (in the sense of Definition 4) compares with the two nonequivalent notions of ESS. As the following theorem shows, it is equivalent to neither of them, and instead occupies an intermediate position: weaker than one and stronger than the other. The two ESS conditions are also comparable with the stronger stability conditions derived from $\bar{p}$-stability (see Section 4.2). In fact, two of the latter turn out to be equivalent to ESS with uniform invasion barrier.

Theorem 5. In a symmetric multilinear game, with $N \geq 2$ players, the following implications and equivalences among the possible properties of a strategy hold:

> ESS ⇐ stable ⇐ ESS with uniform invasion barrier ⇔ locally superior ⇔ dependently-stable ⇔ independently-stable ⇐ symmetrically-stable.

Each of the three implications is actually an equivalence in the special case of two-player games, but not in general.

The proof of Theorem 5 uses the following two lemmas, which hold for every game g as in the theorem. The first lemma uses the following terminology. A strategy y in g is *conditionally locally superior* if it has a neighborhood where inequality (6) holds for every strategy $x \neq y$ that satisfies the reverse of inequality (1).

Lemma 2. For any $0 < p < 1$, the following properties of an equilibrium strategy y are equivalent, and imply that y is stable:

(i) Local superiority
(ii) Conditional local superiority
(iii) $\bar{p}$-stability with $\bar{p} = (p_1, p_2, \dots, p_N)$ given by (15)
(iv) $\bar{p}$-stability with $\bar{p} = (p_1, p_2, \dots, p_N)$ given by (16)
(v) ESS with uniform invasion barrier.

Proof. The implication (i) ⇒ (iii) is trivial: inequality (1) (from the equilibrium condition) and inequality (6) together give

$$(1-p)\big(g(x, y, \dots, y) - g(y, y, \dots, y)\big) + p\big(g(x, x, \dots, x) - g(y, x, \dots, x)\big) < 0.$$

Clearly, if the first term on the left-hand side of the last inequality is nonnegative, then the second term must be negative. This proves that (iii) ⇒ (ii).

To prove that (ii) ⇒ (i), assume that this implication does not hold: strategy y is not locally superior but it is conditionally locally superior. The assumption implies that there is a sequence $(x_k)_{k \geq 1}$ of strategies that converges to y such that for all k

$$g(x_k, x_k, \dots, x_k) - g(y, x_k, \dots, x_k) \geq 0 \tag{33}$$

and

$$g(x_k, y, \dots, y) - g(y, y, \dots, y) < 0.$$

The last inequality means that, when all the other players use y, strategy x_k is not a best response. Therefore, the latter can be presented as

$$x_k = \alpha_k z_k + (1 - \alpha_k) w_k, \tag{34}$$

where $0 < \alpha_k \leq 1$, z_k is a strategy whose support includes only pure strategies that are not best responses when everyone else uses the equilibrium strategy y, and w_k is a strategy that is a best response, i.e.,

$$g(w_k, y, \dots, y) - g(y, y, \dots, y) = 0. \tag{35}$$

Since there are only finitely many pure strategies, there is some $\delta > 0$ such that for all k

$$g(z_k, y, \dots, y) - g(y, y, \dots, y) < -2\delta. \tag{36}$$

By (33), (34), (35) and (36), for all k

$$\big(g(x_k, x_k, \dots, x_k) - g(x_k, y, \dots, y)\big) - \big(g(y, x_k, \dots, x_k) - g(y, y, \dots, y)\big) > 2\delta\alpha_k.$$

As $k \to \infty$, the two expressions in parentheses tend to zero, since $x_k \to y$. Therefore, $\alpha_k \to 0$, which by (34) implies that $w_k \to y$. Since y is conditionally locally superior and (35) holds for all k, for almost all k (that is, all $k > K$, for some integer K)

$$g(w_k, w_k, \dots, w_k) - g(y, w_k, \dots, w_k) \leq 0.$$

Therefore, for almost all k

$$\begin{aligned}
&\sum_{j=2}^{N} \frac{B_{j-1,N-1}(\alpha_k)}{\alpha_k} \big(g(w_k, \underbrace{z_k, \dots, z_k}_{j-1 \text{ times}}, \underbrace{w_k, \dots, w_k}_{N-j \text{ times}}) - g(y, \underbrace{z_k, \dots, z_k}_{j-1 \text{ times}}, \underbrace{w_k, \dots, w_k}_{N-j \text{ times}})\big) \\
&= \frac{1}{\alpha_k}\Big(\big(g(w_k, x_k, \dots, x_k) - g(y, x_k, \dots, x_k)\big) - (1-\alpha_k)^{N-1}\big(g(w_k, w_k, \dots, w_k) - g(y, w_k, \dots, w_k)\big)\Big) \\
&\geq \frac{1}{\alpha_k}\big(g(w_k, x_k, \dots, x_k) - g(y, x_k, \dots, x_k)\big).
\end{aligned}$$

The sum on the left-hand side tends to zero as $k \to \infty$, since $w_k \to y$. Therefore, for almost all k the expression on the right-hand side is less than δ, so that

$$g(w_k, x_k, \dots, x_k) - g(y, x_k, \dots, x_k) < \alpha_k \delta. \tag{37}$$

On the other hand, by (36) and since $x_k \to y$, for almost all k

$$\begin{aligned}
\alpha_k \Big(\big(g(y, y, \dots, y) - g(z_k, y, \dots, y)\big) + \big(g(z_k, y, \dots, y) - g(z_k, x_k, \dots, x_k)\big) \\
+ \big(g(w_k, x_k, \dots, x_k) - g(w_k, y, \dots, y)\big)\Big) > \alpha_k \delta.
\end{aligned}$$

By (34) and (35), the left-hand side is equal to $g(w_k, x_k, \dots, x_k) - g(x_k, x_k, \dots, x_k)$, which by (33) is less than or equal to

$$g(w_k, x_k, \dots, x_k) - g(y, x_k, \dots, x_k).$$

This contradicts (37). The contradiction proves that (ii) ⇒ (i).

To prove that (i) ⇒ (iv), assume that y is locally superior, and thus has a *convex* neighborhood U where (6) holds for every strategy $x \neq y$. By the convexity of U and the linearity of g in the first argument, for every $x \in U \setminus \{y\}$

$$g(x, x_p, \dots, x_p) - g(y, x_p, \dots, x_p) < 0, \tag{38}$$

where $x_p = px + (1 - p)y$. By the second equality in (11), inequality (38) is equivalent to (14), with $(p_1, p_2, \dots, p_N)$ given by (16). Thus, y has property (iv).

Clearly, the above arguments also apply with p replaced by any other number in $(0,1)$. Integration over this interval therefore gives that, for every $x \in U \setminus \{y\}$, (14) holds also with $p_1, p_2, \dots, p_N$ given (not by (16) but) by the left-hand side of the corresponding equality in (12). The equalities in (12) therefore prove that the locally superior strategy y is stable.

The proof of the reverse implication, (iv) ⇒ (i), is rather similar. As shown above, y has property (iv) if and only if it has a neighborhood U such that (38) holds for all strategies $x \neq y$ in U, or equivalently (6) holds for all $x \neq y$ in the set

$$U_p = \{ pz + (1 - p)y \mid z \in U \}.$$

In this case, y is locally superior, since U_p is also a neighborhood of y. Indeed, for *any* neighborhood U of any strategy y, $\{U_\epsilon\}_{0<\epsilon<1}$ is a base for the neighborhood system of y (see Bomze and Pötscher 1989, Lemma 42; Bomze 1991, Lemma 6).

The special case $U = X$ of the last topological fact gives the equivalence (i) ⇔ (v). A strategy y has a neighborhood where (6) holds for all $x \neq y$ if and only if it has such a neighborhood of the form X_ϵ, for some $0 < \epsilon < 1$. ■

Lemma 3. For a probability vector $\bar{p} = (p_1, p_2, \dots, p_N)$ with $p_N > 0$, every $\bar{p}$-stable strategy y is an ESS.

Proof. Let $\bar{p}$ be a probability vector as above. For distinct strategies x and y and $0 < \epsilon < 1$,

$$\sum_{k=1}^{N} p_k (g(x_\epsilon, \underbrace{x_\epsilon, \dots, x_\epsilon}_{k-1 \text{ times}}, \underbrace{y, \dots, y}_{N-k \text{ times}}) - g(y, \underbrace{x_\epsilon, \dots, x_\epsilon}_{k-1 \text{ times}}, \underbrace{y, \dots, y}_{N-k \text{ times}}))$$

$$= \sum_{k=1}^{N} p_k \epsilon \sum_{j=1}^{k} B_{j-1,k-1}(\epsilon) (g(x, \underbrace{x, \dots, x}_{j-1 \text{ times}}, \underbrace{y, \dots, y}_{N-j \text{ times}}) - g(y, \underbrace{x, \dots, x}_{j-1 \text{ times}}, \underbrace{y, \dots, y}_{N-j \text{ times}}))$$

$$= \sum_{j=1}^{N} \left(\sum_{k=j}^{N} \binom{k-1}{j-1} (1-\epsilon)^{k-j} p_k \right) (g(x, \underbrace{x, \dots, x}_{j-1 \text{ times}}, \underbrace{y, \dots, y}_{N-j \text{ times}}) - g(y, \underbrace{x, \dots, x}_{j-1 \text{ times}}, \underbrace{y, \dots, y}_{N-j \text{ times}})) \epsilon^j .$$

The expression on the right-hand side is negative for sufficiently small $\epsilon > 0$ if and only if at least one of its N terms is not zero and the first such term (that is, the nonzero term ending with the smallest power of ϵ) is negative. Observe that the sign of each term is completely determined by the sign of the second expression in parentheses (the difference). The first term (the inner sum) is necessarily positive, since $p_N > 0$. This observation proves that if y is $\bar{p}$-stable, then the condition in Lemma 1 holds. Note, parenthetically, that in the special case $p_N = 1$ the observation also proves Lemma 1 itself. ■

Proof of Theorem 5. By Lemma 3, a strategy that has any of the seven properties in the theorem is an ESS, and hence (by Lemma 1) an equilibrium strategy. An immediate corollary of Lemma 2 is that, for an equilibrium strategy, the properties of local superiority, dependent- and independent stability, and ESS with uniform invasion barrier are all equivalent, and imply stability. The special case $p = 1/2$ of the same lemma (specifically, of the implication (iii) ⇒ (i)) shows that symmetric-stability implies local superiority.

With only two players ($N = 2$), there is no difference between stability and symmetric-stability, and thus the equivalence of all the properties in the theorem follows from the first part of the proof and Proposition 3. The counterexamples in Example 4 below (where $N = 4$) complete the proof. ∎

Example 4. A symmetric multilinear four-player game g is defined as follows. There are three pure strategies, so that the strategy space X consists of all probability vectors $x = (x_1, x_2, x_3)$ (with $x_1 + x_2 + x_3 = 1$). The payoff of a player using strategy x against opponents using strategies $y = (y_1, y_2, y_3)$, $z = (z_1, z_2, z_3)$ and $w = (w_1, w_2, w_3)$ is given by

$$g(x, y, z, w) = \sum_{i,j,k,l=1}^{3} g_{ijkl} x_i y_j z_k w_l.$$

It does not matter which of the other players uses which strategy, since the coefficients $(g_{ijkl})_{i,j,k,l=1}^{3}$ that define the game satisfy the symmetry condition $g_{ijkl} = g_{ij'k'l'}$, for all i and all triplets (j, k, l) and (j', k', l') that are permutations of one another. There are three versions of the game, with different coefficients, as detailed in the following table:

Coefficient	Version 1	Version 2	Version 3
g_{2211}	-2	-18	-4
g_{2221}	0	-16	-4
g_{3221}	4	4	0
g_{2331}	4	20	4
g_{2222}	3	-9	-3
g_{2332}	4	12	2
g_{3333}	-3	-15	-4
g_{2322}	4	4	0

Coefficients that are not listed in the table and cannot be deduced from it by using the symmetry condition are zero. In all three versions of the game, the strategy $y = (1,0,0)$ is an equilibrium strategy, since if all the other players use y, the payoff is zero regardless of the player's own strategy. However, the stability properties of y are different for the three versions.

Claim. The equilibrium strategy $y = (1,0,0)$ is an ESS in all three versions of the game, but it is stable only in versions 2 and 3, independently-stable (equivalently, dependently-stable, locally superior, ESS with uniform invasion barrier) only in version 3, and symmetrically-stable in none of them.

The Claim has some significance beyond the present context. The fact that, in version 2 of the game, the ESS $(1,0,0)$ does not have a uniform invasion barrier and is not locally superior refutes two published results. A theorem of Crawford (1990), which is reproduced by Hammerstein and Selten (1994, Result 7), implies that every ESS in a symmetric multilinear game has a uniform invasion barrier. However, there is a known error in the proof of that theorem (Bomze and Pötscher 1993). Theorem 2 of Bukowski and Miekisz (2004) asserts that local superiority and the ESS condition are equivalent even for $N > 2$. However, these authors employ a definition of ESS that is different from that used here (and in other papers) in that it *requires* the existence of a uniform invasion barrier.

In view of Theorem 5, to prove the Claim it suffices to show that y is: (i) an ESS but not stable in version 1, (ii) stable but not independently-stable in version 2, and (iii) independently-stable but not symmetrically-stable in version 3.

In version 1 of g, (14) reads

$$-2p_2x_2^2 - 4p_3(x_1x_2^2 - x_2^2x_3 - x_2x_3^2) \\ - 3p_4(2x_1^2x_2^2 - 4x_1x_2^2x_3 - 4x_1x_2x_3^2 - x_2^4 - 4x_2^2x_3^2 + x_3^4 - 4x_2^3x_3) < 0.$$

Stability corresponds to $\bar{p} = (p_1, p_2, p_3, p_4) = (1/4,1/4,1/4,1/4)$, for which the above inequality can be simplified to

$$\frac{7}{16}x_2^2 < (x_2 - \frac{3}{8}(1 - x_1)^2)^2.$$

There are strategies $x = (x_1, x_2, x_3)$ arbitrarily close to but different from $(1,0,0)$ for which this inequality does not hold. For example, this is so whenever $x_2 = (3/8)(1 - x_1)^2 > 0$. This proves that the equilibrium strategy is not stable. To prove that it is nevertheless an ESS, consider (32), which in the present case can be simplified to

$$2x_2^2 < (2x_2 - \epsilon(1 - x_1)^2)^2.$$

For every (fixed) strategy $x = (x_1, x_2, x_3) \neq (1,0,0)$, this inequality holds for sufficiently small $\epsilon > 0$. Therefore, $(1,0,0)$ is an ESS.

In version 2 of the game, for $\bar{p} = (1/4,1/4,1/4,1/4)$ inequality (14) can be simplified to

$$-\frac{1}{80}x_2^2 < (x_2 - \frac{3}{8}(1 - x_1)^2)^2.$$

This inequality holds for all strategies x other than $(1,0,0)$, and therefore the latter is stable. However, it is not independently-stable, since for $\bar{p} = (1/8,3/8,3/8,1/8)$ inequality (14) can be simplified to

$$\frac{1}{10}x_2^2 < (x_2 - \frac{1}{4}(1 - x_1)^2)^2.$$

This inequality does not hold for strategies x with $x_2 = (1/4)(1 - x_1)^2 > 0$, which exist in every neighborhood of $(1,0,0)$.

Finally, in version 3 of the game, for $\bar{p} = (1/8\,,3/8,3/8,1/8)$ inequality (14) can be simplified to

$$-x_3^4 < 3(4x_2 - (x_2 + x_3)^2)^2.$$

This inequality holds for all strategies x other than $(1,0,0)$. Therefore, by Lemma 2 (which implies that, if (iv) holds for $p = 1/2$, it holds for all $0 < p < 1$), $(1,0,0)$ is independently-stable. However, it is not symmetrically-stable. There are probability vectors $\bar{p}$ satisfying (17) for which (14) does not hold for some strategies x arbitrarily close to $(1,0,0)$. For examples, for $\bar{p} = (1/20,9/20,9/20,1/20)$, inequality (14) can be simplified to

$$24x_2^2 - \frac{1}{3}x_3^4 < (8x_2 - (1 - x_1)^2)^2.$$

For strategies x with $x_2 = (1/8)(1 - x_1)^2$, this inequality is equivalent to $(1 - x_1)^4 - 32(1 - x_1)^3 + 384(1 - x_1)^2 - 2048(1 - x_1) > 512$. Hence, it does not hold if x_1 is sufficiently close to 1. This completes the proof of the Claim.

Appendix A. Other notions of stability

Static and dynamic stability are not the only kinds of stability in strategic games considered in the game-theoretic literature. For completeness, some of the other categories are briefly reviewed below.

One kind of stability refers to the effects of perturbations of the players' *strategy spaces* (e.g., allowing only completely mixed strategies) or a combination of perturbations of the strategy spaces and of the strategies themselves. The requirement that a strategy profile in a strategic game is stable against these kinds of perturbations gives the notions of (trembling-hand) perfect equilibrium (Selten 1975), proper equilibrium (Myerson 1978), strict perfection (Okada 1981) and (strategic) stability and full stability (Kohlberg and Mertens 1986). Stability may also refer to the effects on a given equilibrium of perturbations of the *payoff functions*. Essentiality (Wu and Jiang 1962) and strong stability (Kojima et al. 1985) are examples of this kind of stability, which has interesting links with some of the other kinds. For example, in a multilinear game, every essential equilibrium is strictly perfect (van Damme 1991, Theorem 2.4.3), and in a symmetric $n \times n$ game, every regular ESS is essential (Selten 1983). Another example of a link between different kinds of stability is the finding that, in several classes of games, the (local) degree of an equilibrium (or of a connected component of equilibria) is equal to its index (Govindan and Wilson 1997; Demichelis and Germano 2000). The index of an equilibrium is related to its asymptotic stability or instability with respect to a large class of natural dynamics, which determine how strategies in the game change over time. The degree, by contrast, expresses a topological property of the equilibrium when viewed as a point in a manifold that includes the various equilibria of *different* games (Ritzberger 2002). The index (= degree) of an equilibrium is connected with stability also in that, in a nondegenerate bimatrix game, it determines whether the equilibrium can be made the unique equilibrium by extending the game: adding one or more pure strategies to one of the players (von Schemde and von Stengel 2008).

Whether any of these alternative notions can be linked with statics stability is yet to be determined.

References

Apaloo, J., 1997. Revisiting strategic models of evolution: The concept of neighborhood invader strategies. Theoretical Popul. Biol. 52, 71–77.

Bomze, I. M., 1991. Cross entropy minimization in uninvadable states of complex populations. J. Math. Biol. 30, 73–87.

Bomze, I. M., Pötscher, B. M., 1989. *Game Theoretical Foundations of Evolutionary Stability*. Springer, Berlin.

Bomze, I. M., Pötscher, B. M., 1993. "On the definition of an evolutionarily stable strategy in the playing the field model" by V. P. Crawford. J. Theoretical Biol. 161, 405.

Bomze, I. M., Weibull, J. W., 1995. Does neutral stability imply Lyapunov stability? Games Econ. Behav. 11, 173–192.

Broom, M., Cannings, C., Vickers, G. T., 1997. Multi-player matrix games. Bull. Math. Biol. 59, 931–952.

Bukowski, M., Miekisz, J., 2004. Evolutionary and asymptotic stability in symmetric multi-player games. Internat. J. Game Theory 33, 41–54.

Christiansen, F. B., 1991. On conditions for evolutionary stability for a continuously varying character. Am. Nat. 138, 37–50.

Crawford, V.P., 1990. On the definition of an evolutionarily stable strategy in the "playing the field" model. J. Theoretical Biol. 143, 269–273.

Demichelis, S., Germano, F., 2000. On the indices of zeros of Nash fields. J. Econ. Theory 94, 192–217. (An extended version: Discussion paper, CORE, Louvain-la-Neuve, 2000).

Dold, A., 1980. *Lectures on Algebraic Topology*, Second Edition. Springer-Verlag, Berlin.

Eshel, I., 1983. Evolutionary and continuous stability. J. Theoretical Biol. 103, 99–111.

Eshel, I., Motro, U., 1981. Kin selection and strong evolutionary stability of mutual help. Theoretical Popul. Biol. 19, 420–433.

Fudenberg, D., Tirole, J., 1995. *Game Theory*. MIT Press, Cambridge, MA.

Govindan, S., Wilson, R., 1997. Equivalence and invariance of the index and the degree of Nash equilibria. Games Econ. Behav. 21, 56–61.

Hammerstein, P., Selten, R., 1994. Game theory and evolutionary biology. In: Aumann, R. J., Hart, S. (Eds.), *Handbook of Game Theory with Economic Applications*, Vol. 2, pp. 929–993. Elsevier Science, Amsterdam.

Harsanyi, J., Selten, R., 1988. *A General Theory of Equilibrium Selection in Games*. MIT Press, Cambridge, MA.

Hofbauer, J., Sigmund, K., 1998. *Evolutionary Games and Population Dynamics.* Cambridge University Press, Cambridge, UK.

Kelley, J. L., 1955. *General Topology.* Springer, New York.

Kohlberg, E., Mertens, J.-F., 1986. On the strategic stability of equilibria. Econometrica 54, 1003–1037.

Kojima, M., Okada, A., Shindoh, S., 1985. Strongly stable equilibrium points of *n*-person noncooperative games. Math. Oper. Res. 10, 650–663.

Maynard Smith, J., 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge, UK.

Milchtaich, I., 2012. Comparative statics of altruism and spite. Games Econ. Behav. 75, 809–831.

Monderer, D., Shapley, L. S., 1996. Potential games. Games Econ. Behav. 14, 124–143.

Myerson, R. B., 1978. Refinements of the Nash Equilibrium Concept. Internat. J. Game Theory 7, 73–80.

Oechssler, J., Riedel, F., 2002. On the dynamic foundation of evolutionary stability in continuous models. J. Econ. Theory 107, 223–252.

Okada, A., 1981. On the stability of perfect equilibrium points. Internat. J. Game Theory 10, 67–73.

Ritzberger, K., 2002. *Foundations of Non-Cooperative Game Theory*. Oxford University Press, Oxford.

Samuelson, P. A., 1983. *Foundations of Economic Analysis*, Enlarged Edition. Harvard University Press, Cambridge, MA.

Selten, R., 1975. A reexamination of the perfectness concept for equilibrium points in extensive games. Internat. J. Game Theory 4, 25–55.

Selten, R., 1983. Evolutionary stability in extensive two-person games. Math. Social Sci. 5, 269–363.

Taylor, P. D., 1989. Evolutionary stability in one-parameter models under weak selection. Theoretical Popul. Biol. 36, 125–143.

van Damme, E., 1991. *Stability and Perfection of Nash Equilibria*, Second Edition. Springer-Verlag, Berlin.

von Schemde, A., von Stengel, B., 2008. Strategic characterization of the index of an equilibrium. In: Monien, B., Schroeder, U.-P. (Eds.), *Symposium on Algorithmic Game Theory (SAGT) 2008*, Lecture Notes in Computer Science 4997, pp. 242–254, Springer-Verlag, Berlin.

Weibull, J. W., 1995. *Evolutionary Game Theory.* MIT Press, Cambridge, MA.

Weissing, F. J., 1991. Evolutionary stability and dynamic stability in a class of evolutionary normal form games. In: Selten, R. (Ed.), *Game Equilibrium Models*, pp. 29–97. Springer-Verlag, Berlin.

Wu, W.-T., Jiang, J.-H., 1962. Essential equilibrium points of *n*-person non-cooperative games. Sci. Sinica 11, 1307–1322.