Stability of the replicator equation with continuous strategy space

Ross Cressman*

Department of Mathematics, Wilfrid Laurier University, Waterloo, Ontario Canada N2L 3C5

Received 1 April 2004; received in revised form 1 January 2005; accepted 1 March 2005
Available online 17 May 2005

Abstract

We extend previous work that analyzes the stability of evolutionary dynamics on probability distributions over continuous strategy spaces. The stability concept considered is that of “neighborhood” convergence to a rest point (i.e. an equilibrium distribution over the strategy space) under the dynamics in the weak topology for all initial distributions whose support is close to that of the rest point. Stability criteria involving strategy domination and neighborhood superiority are developed for monomorphic rest points (i.e. the equilibrium distribution is supported on a single strategy) and for distributions that have finite support.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Convergence; Coevolution; Replicator dynamics; Monomorphism; Weak topology

JEL classification: C72; C70

1. Introduction

Evolutionary dynamics for continuous strategy spaces have received considerable attention recently both among theoretical biologists who are interested in the coevolution of species traits and among economists who concentrate instead on predicting rational behavior of individuals whose payoffs are given through game interactions. Most
theoretical research on stability for coevolutionary models (e.g. Abrams, 2001; Doebeli and Dieckmann, 2000; Marrow et al., 1996 and the references therein) make the simplifying assumption that each species is monomorphic (or homogeneous) and remains so during the course of evolution (i.e. all individuals present in a given species exhibit the same behavior) and that the population is at its equilibrium size for the current monomorphism. This leads to the stability analysis of what are known as adaptive or strategy dynamics. Although there has been much less research in this area from the economic or game-theoretic perspective (e.g. Bomze, 1990, 1991; Seymour, 2000; Oechssler and Riedel, 2001, 2002), this latter literature typically considers the full dynamical system where aggregate behavior is described by a distribution on the space $S$ of individual strategy choices and assume individual payoffs are defined in terms of a function $f$ on $S \times S$.

In this paper, we follow the latter approach applied to a symmetric game with a continuous strategy space. These references in the economic literature spend a great deal of time developing the evolutionary dynamics on the set of probability distributions (e.g. the replicator equation), proving its solutions are well-defined, and relating its properties to static equilibrium conditions (that generalize those for the case where there are a finite set of strategies that may be used by the population as in the matrix games of Section 2.1). We benefit from their work by briefly summarizing this development at the beginning of Section 2 and devoting the remainder of the paper to analyzing the stability of equilibrium distributions (i.e. rest points) for the replicator equation.

Immediate issues that arise in this analysis are what constitutes closeness and/or convergence for probability distributions and for what initial distributions we expect this convergence. The main problem is that there are several ways to define these topological concepts that generalize the accepted approach when there are a finite set of strategies. Moreover, as clearly demonstrated by Oechssler and Riedel (2002), conclusions concerning stability depend critically on which definitions are taken. These issues are clarified in Section 2.1 by referring to well-known stability results for matrix games and, in the process, motivate our stability concept (see Definition 1 at the beginning of Section 3).

Our main goal is to derive conditions on $f$ that predict stability. Section 3 completely characterizes (Proposition 1 and Theorem 1) stability of monomorphic rest points (i.e. probability distributions concentrated at a single strategy) and relates these results to the coevolutionary literature for monomorphic populations. Section 4 extends these results to probability distributions with finite support, giving sufficient conditions for stability (Theorems 2 and 3). Several examples are developed to illustrate the theorems. Examples 1a, 1b and 1c in Sections 2, 3 and 4 respectively apply the theory to a model based on Cournot competition with $n$ identical firms that have a constrained maximum output. These latter two examples illustrate respectively situations where the evolutionary dynamics predicts all firms will have the same production level and where some firms produce at maximum output while others produce nothing. Examples 2 and 3 are based on functions $f$ taken from the adaptive dynamics literature, demonstrating in turn Theorem 1

1Notable exceptions are Vincent et al. (1996) and Cressman and Garay (2003) where polymorphic effects and the effects of varying population size are also taken into account. See also Example 2 in Section 3.
in a multi-dimensional continuous strategy space setting and Theorem 3 for dimorphic rest points. Section 5 concludes.

2. The model

Evolutionary games with an arbitrary strategy space and their corresponding replicator dynamics have been developed by a number of researchers over the past fifteen years. Here we briefly summarize this development as it applies to our model. The theory is most advanced for symmetric evolutionary games (Bomze, 1990, 1991; Oechssler and Riedel, 2001) with a one-dimensional strategy space (Oechssler and Riedel, 2002; Eshel and Sansone, 2003). Our games will continue to be symmetric but the continuous strategy set need not be one-dimensional.

In general, we assume that all individuals play a strategy in the same set \( S \) and the population state is given by a probability measure \( P \) with respect to a measure space \( (S, B) \). If \( B \subset B \), \( P(B) \) is interpreted as the proportion of individuals in the population who are using strategies in the set \( B \). For a given \( s \in S, \delta_s \) denotes the Dirac delta measure that assigns unit mass to \{s\}. Let \( \pi(s, P) = \pi(\delta_s, P) \) denote the expected payoff to an individual using strategy \( s \) when the population is in state \( P \). The mean payoff to a random individual in the population with state \( P \) is then \( \pi(P, P) = \int_S \pi(\delta_s, P) P(ds) \).

For us, \( S \) will be a nonempty compact subset of \( \mathbb{R}^n \) and \( B \) will be the Borel subsets of \( S \) (i.e. the \( \sigma \)-algebra of the Borel sets of \( \mathbb{R}^n \) intersected with \( S \) and so \( P \) is a Borel measure). Let \( A(S) \) denote the set of Borel probability measures with respect to \( (S, B) \). Since \( P \) is a Borel measure, there is a unique closed subset of \( S \), called the support of \( P \), such that the measure of its complement is 0 but every open set that intersects it has positive measure (Royden, 1988). We will be most interested in the situation where the payoff function \( \pi(s, P) \) is given through a continuous real-valued function \( f : S \times S \rightarrow \mathbb{R} \) by \( \pi(s, P) = \int_S f(s, y) P(dy) \). Unless otherwise stated, we will assume the existence of such an \( f \) for the remainder of the paper. In particular, standard symmetric evolutionary games with a finite trait space (see Section 2.1) that assume random pairwise interactions may be put in this form. An important example with a one-dimensional continuous strategy space (that will also be used in Examples 1b and 1c later on in the paper) is the following model based on Cournot competition with a linear demand function.

**Example 1a.** Suppose all firms producing a particular product have identical cost functions. Let \( C(x) \) be the cost of producing \( x \) units. Suppose there are a large, effectively infinite, number of such firms in the population and that \( n \) are drawn at random. If demand is linear, the payoff function (i.e. profit) of a firm producing \( x \) units (i.e. playing strategy \( x \)) can be normalized to the form \( x(1 - x - Y) - C(x) \) where \( Y \) is the total production of the

---

2See however Seymour (2000) and Heifetz et al. (2003) for evolutionary treatments of asymmetric games with continuous strategy spaces.

3In fact, Bomze and Pötscher (1989) argue that the existence of such an \( f \) means the evolutionary game can be interpreted as being based on pairwise interactions. It is only the form of \( \pi(s, P) \) that is important to us, not whether players are competing pairwise.
other $n - 1$ firms. Here, we assume that each firm plays a strategy $y \in S = [0, 1/n]$ to ensure that the price is nonnegative (see the following remark). With $P$ the output distribution of the current population, the expected value of $Y$ is $(n - 1) \int_S y P(dy)$. Thus the expected payoff $\pi(x, P)$ of a firm producing $x$ units can be rewritten in the form $\pi(x, P) = \int_S [x(1 - x - (n - 1)y) - C(x)]P(dy)$. That is, we can take

$$f(x, y) = x(1 - x - (n - 1)y) - C(x).$$  \hspace{1cm} (1)

**Remark 1.** The expected payoff for a general Cournot competition model with nonlinear demand function and/or unrestricted strategy space $S$ cannot always be represented through an $f(x, y)$, in which case the model is beyond the scope of this paper. In particular, linearity is useful since it implies that the other $n - 1$ firms can be viewed as being identical opponents, all choosing the same output $y = Y/(n - 1)$. Furthermore, if total production is too large (i.e. $x + Y > 1$), price will be zero and so payoffs of the form $x \max \{1 - x - Y, 0\} - C(x)$ are then more realistic. Unfortunately, unless there is a duopoly (i.e. $n = 2$), this payoff can no longer be represented through an $f(x, y)$ (such as $f(x, y)$ given in Eq. (1)) since it is no longer linear in the output $y$ of each of the other firms.

When $S$ is finite, the replicator dynamics Eq. (3) in Section 2.1 below is a biological model of the evolution of behavioral frequencies in a single species when there is asexual reproduction and payoffs are identified with reproductive success (Taylor and Jonker, 1978). It can also be justified as the dynamics that emerges when rational individuals base their strategy choices on imitative behavior (Schlag, 1998). For general $S$, this dynamics takes the form

$$\frac{dP}{dt}(B) = \int_B (\pi(\delta_s, P) - \pi(P, P))P(ds)$$  \hspace{1cm} (2)

It is a payoff monotone dynamics in that the relative rate of increase is greater for those sets of strategies $B$ that have a higher average payoff (i.e. $\frac{\dot{P}}{P}(B)/P(B) > \frac{\dot{P}}{P}(B')/P(B')$) if the average payoff to an individual in $B$, $\frac{1}{P(B)} \int_B (\pi(\delta_s, P) - \pi(P, P))P(ds)$, is higher than that in $B'$. For instance, in the Cournot competition model of Example 1a, the proportion of those firms in the current population which have higher expected payoffs when randomly matched in a group of $n$ firms will increase relative to those with lower expected payoff.

It has been shown (Oechssler and Riedel, 2001) that there is a unique solution $P(t)$ that satisfies the dynamics Eq. (2) for all positive $t$ given any initial probability measure $P(0)$ with compact support when $f$ is continuous.\(^5\) Furthermore, the support of $P(t)$ is the same

\(^4\)Specifically, if an individual observes one randomly chosen player per unit time and switches to the observed strategy (if his own payoff is less than the observed) with probability proportional to the payoff difference, then the replicator dynamics results.

\(^5\)This statement is actually valid for much wider classes of payoff functions $f$ and strategy spaces $S$. For instance, it is only necessary that $f$ be bounded (and measurable) on $S$. For a Borel subset $B$ of $S$, $\frac{dP}{dt}(B)$ at time $t$ is defined to be $\lim_{h \to 0} P((\frac{1}{h}t + P(0)) h^{-1})(B)$ with respect to the variational norm $\| P - Q \| = 2\sup_{B \in S} | P(B) - Q(B) |$ on $\Lambda(S)$. 


as $P(0)$ for all $t \geq 0$. In fact, the measure $P(t)$ is absolutely continuous with respect to $P(0)$ (i.e. if $P(t)(B) > 0$ then $P(0)(B) > 0$ for all $B \subseteq \mathcal{B}$) and conversely (Bomze, 1990).

We assume the replicator dynamics Eq. (2) describes how the population state evolves (i.e. its solution trajectory $P(t)$ in $\mathcal{A}(S)$ is the population state at time $t$). A population state $P^*$ is a rest point of Eq. (2) if and only if $\pi(\delta_i, P^*) - \pi(P^*, P^*) = 0$ for all $s \in \text{supp}(P^*)$. The main purpose of this paper is to analyze the dynamic stability of such a state $P^*$ when $P^*$ has finite support. Heuristically, dynamic stability of $P^*$ refers to the question of whether $P(t)$ stays close and/or converges to $P^*$ if the initial $P(0)$ is chosen appropriately in $\mathcal{A}(S)$. Most of our results are concerned with convergence although Lyapunov stability is also considered such as in Remark 3 (see Section 3 below). From Oechssler and Riedel (2002), it is clear that the answers to the stability question depend critically on the concept of closeness of probability measures (i.e. on the topology used for the space of Borel probability measures), especially when the strategy space is not a discrete subset of $\mathbb{R}^n$. We will return to this issue after the following section that begins with a finite strategy space.

2.1. Matrix games

Standard matrix games emerge when there are a finite number of possible strategies individuals may use. If there are $m$ such strategies, they are often thought of as “pure strategies” and then identified with the unit coordinate vectors $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ in $\mathbb{R}^m$ that have 1 in the $i$th component and 0 everywhere else. In our notation, $P(0)$ then has finite support contained in $\{e_1, \ldots, e_m\} = S$ where $S$ is now a subset of $\mathbb{R}^m$. Then $P(t) = \sum_{i=1}^m p_i(t) \delta_{e_i}$ where $p_i(t)$ is the proportion (or frequency) of individuals in the population using strategy $e_i$ at time $t$. From Eq. (2), the replicator equation becomes

$$
\dot{p}_i = p_i \left( \sum_{k=1}^m f(e_i, e_k) p_k - \sum_{j,k=1}^m f(e_j, e_k) p_j p_k \right).
$$

This dynamics can be rewritten in matrix form where $A$ is the $m \times m$ payoff matrix with entries $A_{ij} = f(e_i, e_j)$ as

$$
\dot{p}_i = p_i (e_i - p) : Ap.
$$

Here $u : Av$ is the standard inner product of column vectors $u$ and $Av$ in $\mathbb{R}^m$ (i.e. $u : Av = \sum_{j,k=1}^m u_j A_{jk} v_k$) and $p$ is the frequency vector in the $m - 1$ dimensional simplex $\Delta^m = \{(p_1, \ldots, p_m) | \sum p_i = 1, p_i \geq 0 \}$. Since there is a 1–1 correspondence between $\mathcal{A}(S)$ and $\Delta^m$, dynamic stability of $P^*$ becomes the stability of $p^* \in \Delta^m$ with respect to the dynamical system Eq. (3) on $\Delta^m$. It is well known that a “matrix” ESS is a locally asymptotically stable rest point of Eq. (3) and that the converse is not true for all matrix games when there are more than two pure strategies (i.e. $m > 2$).

\[\text{Since the term ESS is overused in the literature, it may have several meanings for some readers. We restrict its use in this paper to that of an evolutionarily stable strategy } p^* \in \Delta^m \text{ of an } m \times m \text{ payoff matrix } A \text{ (hence a “matrix” ESS) as developed by Maynard Smith (1982) where there is one universally accepted meaning. Such a } p^* \text{ is defined to be a Nash equilibrium (i.e. it satisfies } p : Ap^* \leq p^* : Ap^* \text{ for all } p \in \Delta^m \text{) that also fulfills the stability condition } p : Ap^* < p^* : Ap^* \text{ whenever } p : Ap^* = p^* : Ap^* \text{ and } p \neq p^*.\]
The matrix ESS concept may be developed through conditions of dynamic stability of a strategy $p^*$ when invaded by a mutant strategy $q$. This involves the extension of the above “pure-strategy” matrix game to the so-called “mixed strategy” model with a continuous strategy space. Following Hines (1980a), for a matrix game with $m$ pure strategies and $m \times m$ payoff matrix $A$, an individual is now allowed to play a mixed strategy $q \in \Delta^m$ where $q_i$ is then interpreted as the probability this individual will play strategy $e_i$ in a given contest. Thus $S$ becomes all of $\Delta^m$ and $P$ is a Borel probability measure on this continuous strategy space. Let $\tilde{p} \in \Delta^m$ be the mean strategy $\int_{\Delta^m} qP(dq)$ of the population state $P$. In this mixed strategy model, it is assumed that $\pi(e_i, P) = \sum_k A_{ik} \tilde{p}_k$ and that $\pi(q, P) = \sum_i q_i \pi(e_i, P)$. In particular, individual payoffs depend linearly on both the mean population state and on the components of the individual’s mean strategy.

To analyze stability for general payoffs, it is instructive to consider it first in this mixed strategy model when $p^*$ is invaded by a mutant strategy $q$. To this end, suppose that $p^*$ is a matrix ESS with payoff matrix $A$ and $q$ is some other mixed strategy sufficiently close to $p^*$. If $\text{supp}(P(0)) = \{p^*, q\}$ then the mixed strategy replicator dynamics that results from Eq. (2) is

$$
\frac{dP}{dt}(\{p^*\}) = P(\{p^*\})(p^* - \tilde{p}) : A \tilde{p}
$$

where $\tilde{p} = P(\{p^*\})p^* + P(\{q\})q$. Since $\tilde{p}$ is on the line segment from $q$ to $p^*$ (in particular, $\tilde{p}$ is closer to $p^*$ than $q$ is) and $p^*$ is a matrix ESS, $\frac{dP}{dt}(\{p^*\}) > 0$ for all $0 < P(\{p^*\}) < 1$. Thus $p^*$ is globally asymptotically stable.\(^7\)

In fact, $p^*$ is a matrix ESS if and only if $p^*$ is globally asymptotically stable for all these two-strategy dynamics with $\text{supp}(P(0)) = \{p^*, q\}$ and $q$ sufficiently close (but not equal) to $p^*$ (Hines, 1980b; Thomas, 1984; Bomze, 1990; Cressman, 1990). In particular, all such $P(t)$ converge to $\delta_{p^*}$. It is the generalization of this result to non-matrix symmetric games that is a key to our analysis in the following section.

We will also need to refer to the general mixed-strategy matrix game model with arbitrary $P(0)$. Here, the replicator dynamics Eq. (2) becomes

$$
\frac{dP}{dt}(B) = \int_B (q - \tilde{p}) : A \tilde{p}P(dq).
$$

Akin (1982) shows that the evolution of the strategy distribution $P(t)$ is completely determined by the initial distribution $P(0)$ and the evolution of the mean strategy $\tilde{p}(t)$. In particular, $\tilde{p}$ converges to a matrix ESS $p^*$ whenever $p^*$ is in the convex hull of $\text{supp}(P(0))$ and $\text{supp}(P(0))$ is sufficiently close to $p^*$.

3. Convergence and stability of monomorphic populations

As stated earlier, we will mostly be concerned with convergence to a rest point $P^* \in \Delta(S)$ of Eq. (2) with finite support. This convergence requires a topological notion of

\(^7\)This is a slight abuse of notation that should not cause the readers undue confusion. Formally, in terms of the measure $P$, $P(\{p^*\})$ monotonically increases to 1 (as $P(\{q\})$ decreases to 0).
closeness on the set of probability measures. There are several topologies on $\Delta(S)$ that are all equivalent to the Euclidean topology on $\Delta^m$ when $S$ is a finite set with $m$ strategies (Bomze, 1990, 1991; Oechssler and Riedel, 2002). One is the strong topology based on the variational norm that is used to assert the existence and uniqueness of solutions to Eq. (2) in Section 2. However, following Oechssler and Riedel (2002) and Eshel and Sansone (2003), we will analyze convergence in the weak topology whereby $P(t)$ converges to $P^*$ if and only if

$$\lim_{t \to \infty} \int_S g(x)P(t)(dx) = \int_S g(x)P^*(dx)$$

for all continuous functions $g: S \to \mathbf{R}$.

One reason for this choice of topology is that two monomorphic populations $\delta_x$ and $\delta_y$ are then close to each other if and only if $x$ and $y$ are close in Euclidean distance (agreeing with the idea of closeness in the adaptive dynamics approach). In contrast, any $P$ close to $\delta_x$ in the strong topology requires $P(\{x\})$ to be close to 1. More generally, we feel populations described by $P$ with small variance whose mean strategy is near $x$ should be considered close to the monomorphism $\delta_x$. This is another intuitive property that distinguishes the weak from the strong topology.

By default, every monomorphism $\delta_{s^*}$ is a rest point of Eq. (2). A basis of open neighborhoods of $\delta_{s^*}$ in the weak topology are sets of the form

$$\{P \in \Delta(S) : P(U_\epsilon(s^*)) > 1 - \epsilon\}$$

where $U_\epsilon(s^*) = \{s \in S : |s - s^*| < \epsilon\}$ is the open ball of radius centered at $s^*$ intersected with $S$. That is, $P \in \Delta(S)$ is close to $\delta_{s^*}$ if most of its weight is close to $s^*$. However, such a $P$ may include elements $s$ in its support that are far away from $s^*$. To rule out this possibility and at the same time adapt the convergence results for the matrix game model of Section 2.1, we seek $s^* \in S$ that have a “neighborhood attracting” property for Eq. (2) by restricting $P(0)$ to have support sufficiently close to $s^*$. This is equivalent to considering the game on the reduced strategy space given by this support. Convergence and stability of other dynamic processes have been considered for such restricted games. For instance, Moulin (1984) investigates a local Cournot tatonnement adjustment process given by the sequence of local best replies. The following definition is then analogous to the concepts of local dominance solvable and of a locally stable Cournot tatonnement as discussed by Moulin (1984).

Our proofs of convergence to $P^*$ also require the “positive mass assumption” included in Definition 1; namely, that the initial $P$ place positive weight on each element in the support of $P^*$. This condition has previously proved useful for the analysis of the mixed strategy matrix model of Section 2.1 (e.g. Hines, 1980a) as well as for stability of monomorphisms under Eq. (2) with more general payoff functions (e.g. Oechssler and Riedel, 2002).

---

8For their one-dimensional strategy space $S$ and monomorphic $P^*$, Eshel and Sansone (2003) use a variant of the weak topology that they call the maximal shift topology on $\Delta(S)$. Convergence to such a $P^*$ under this variant is closely related to the concept of neighborhood attracting introduced in our Definition 1.
Definition 1. The probability measure $P^* \in \Delta(S)$ with finite support is neighborhood attracting (with respect to Eq. (2)) if, for all initial $P$ with positive weight on each $s^* \in \text{supp}(P^*)$ and with support sufficiently close to the support of $P^*$, $P(t)$ converges to $P^*$ in the weak topology.

Generalizing the convergence results for the matrix game model of Section 2.1 when a matrix ESS $p^*$ is invaded by a mutant strategy $q$, we require $P(t)$ converge to $\delta_{s^*}$ in the weak topology if $\text{supp}(P(0)) = \{s^*, s\}$ for all $s$ sufficiently close (but not equal) to $s^*$. In biological terms, the monomorphic population $\delta_{s^*}$ resists invasion by any mutant strategy $s$ sufficiently close to $s^*$. From Section 2, $\pi(\delta_{s^*}, P) = f(s^*, s^*)P(\{s^*\}) + f(s^*, s)P(\{s\})$, $\pi(\delta_s, P) = f(s, s^*)P(\{s^*\}) + f(s, s)P(\{s\})$ and $\pi(P, P) = P(\{s^*\})\pi(\delta_{s^*}, P) + P(\{s\})\pi(\delta_s, P)$. Thus

$$\frac{dP}{dt}(\{s^*\}) = P(\{s^*\})P(\{s\})[(f(s^*, s^*) - f(s, s^*))P(\{s^*\}) + (f(s^*, s) - f(s, s))P(\{s\})]$$

and so $\frac{dP}{dt}(\{s^*\}) > 0$ for all $0 < P(\{s^*\}) < 1$ if and only if

$$f(s^*, s^*) \geq f(s, s^*) \quad \text{and} \quad f(s^*, s) \geq f(s, s)$$

(5) with strict inequality in at least one of these for all $s$ sufficiently close (but not equal) to $s^*$. These inequalities state that, for all $s$ sufficiently close (but not equal) to $s^*$, $s^*$ weakly dominates $s$ in the two-strategy game between $s^*$ and $s$ with payoff matrix $^9$

$$A = \begin{bmatrix} f(s^*, s^*) & f(s^*, s) \\ f(s, s^*) & f(s, s) \end{bmatrix}.$$

The above discussion proves the result summarized in the following statement.

Proposition 1. Suppose $s^*$ weakly dominates $s$ in the two-strategy game between $s^*$ and $s$ for all $s$ sufficiently close (but not equal) to $s^*$. Then, $\delta_{s^*}$ is neighborhood attracting with respect to the invariant set of initial $P$ whose support is $\{s^*, s\}^{10}$. The converse is also true.

Proposition 1 is the analogue of the matrix ESS concept developed as a strategy that resists invasion by a mutant. As noted at the end of Section 2.1, a matrix ESS $p^*$ is also related to convergence in the mixed-strategy matrix game model. The corresponding question here becomes how the conditions of Proposition 1 relate to convergence properties of the measure-theoretic replicator equation Eq. (2) when $P(0)$ does not have finite support. The remainder of this section mostly considers this question for a monomorphic population $s^*$ in the interior of the strategy space $S$ (see however Remark 3 for results on local attracting at boundary $s^*$ and a discussion of other aspects of dynamic stability).

Weak domination is no longer sufficient in general (although it remains a necessary condition since the converse of Proposition 1 must still hold). To see this, consider $S = [0, 1]$ and $f(x, y) = -xy$. Then $s^* = 0$ weakly dominates every other $s \in S$. On the other hand, if $\text{supp}(P) = \{-\epsilon, 0, \epsilon\}$ and $P(\{-\epsilon\}) = P(\{\epsilon\})$, then $P$ is a rest point of the replicator equation since $\pi(\delta_{s^*}, P) = 0$ for all $s \in \text{supp}(P)$ (e.g. $\pi(\delta_{s^*}, P) = P(\{-\epsilon\})f(\epsilon, -\epsilon) +$)

$^9$In contrast to Eq. (5), $s^*$ is a matrix ESS for this two-strategy game if and only if the two conditions $f(s^*, s^*) \geq f(s, s^*)$ and, if $f(s^*, s^*) = f(s, s^*)$, then $f(s^*, s) \geq f(s, s)$ hold (see Section 2.1). These conditions only guarantee $\frac{dP}{dt}(\{s^\}) > 0$ if $P(\{s^\})$ is sufficiently close (but not equal) to 1 and so $s^*$ need not be neighborhood attracting.

$^{10}$This is a slight abuse of the neighborhood attracting concept of Definition 1 since these $P$ have only two points in their support.
\[ P(0)f(\epsilon,0) + P(\{\epsilon\})f(\epsilon,\epsilon) = P(\{\epsilon\})(\epsilon^2 - \epsilon^2) = 0. \] Thus \( \delta_0 \) is not neighborhood attracting. At best \( \delta_0 \) may be neutrally (i.e. Lyapunov) stable; although, in this mixed strategy model the population mean strategy does converge to 0 (see Eq. (4)) since this example can be reformulated as a mixed-strategy matrix game model as in Section 2.1 by the bilinearity of \( f \) in \( x \) and \( y \).

To avoid this type of neutral stability, let us assume that, for all \( s \) sufficiently close (but not equal) to \( s^* \), \( s^* \) strongly dominates \( s \) in the two-strategy game between \( s^* \) and \( s \). That is, for all such \( s \),

\[
\begin{align*}
    f(s^*, s) &> f(s, s^*) \quad (s \neq s^*)
    \quad f(s^*, s) > f(s, s) \quad (s = s^*)
\end{align*}
\]

(6)

In particular, from the first inequality, \( s^* \) is locally a strict NE. For technical reasons (see Remark 2 below and the discussion that follows it), we will also assume that this domination is determined by the second order Taylor expansion of \( f(x, y) \), that \( f \) has continuous partial derivatives up to third order and that \( s^* \) is in the interior of \( S \). For example, when \( S \) is a subset of \( \mathbb{R} \), the Taylor expansion is

\[
f(s, y) = f(s^*, s^*) + f_1(s - s^*) + f_2(y - s^*)
+ \frac{1}{2} \left[ f_{11}(s - s^*)^2 + 2f_{12}(s - s^*)(y - s^*) + f_{22}(y - s^*)^2 \right] + \text{h.o.t.}
\]

(7)

where \( f_1 \) is the first order partial derivative of \( f(x, y) \) evaluated at \((s^*, s^*)\) with respect to the first variable, etc. Since \( s^* \) is in the interior of \( S \), \( f_1 = 0 \). By ignoring the higher order terms, the two inequalities in Ineq. (6) become

\[
\begin{align*}
    f_{11} &< 0 \quad \text{and} \quad f_{11} + 2f_{12} < 0
\end{align*}
\]

(8)

respectively.

With these assumptions, we then have the following result.

**Theorem 1.** Suppose \( s^* \) is in the interior of \( S \subset \mathbb{R}^n \) and that domination in the two-strategy game between \( s^* \) and \( s \) for all \( s \) sufficiently close (but not equal) to \( s^* \) is determined by the second order Taylor expansion of \( f(x, y) \) about \( x = y = s^* \). If \( s^* \) strongly dominates all such \( s \), then \( \hat{s} = s^* \) is neighborhood attracting. The converse is also true.

**Proof.** We restrict the proof here to the case of a one dimensional strategy space (i.e. \( S \) is a compact subset of \( \mathbb{R} \)). The general proof is in the Appendix A.

Assume \( s^* \) strongly dominates \( s \) for all \( s \) sufficiently close (but not equal) to \( s^* \). From Eq. (2),

\[
\begin{align*}
\frac{dP}{dt}(\{s^*\}) &= P(\{s^*\})(\pi(\delta_{s^*}, P) - \pi(P, P))
= P(\{s^*\}) \left( \int_S f(s^*, y)P(dy) - \int_S f(s, y)P(dy)P(ds) \right) \\
&= P(\{s^*\}) \int_S \int_S (f(s^*, y) - f(s, y))P(dy)P(ds).
\end{align*}
\]
From Eq. (7), \( f(s^*,y) - f(s,y) \) is given by
\[
-\frac{1}{2} f_{11}(s-s^*)^2 - f_{12}(s-s^*)(y-s^*) + \text{h.o.t.}
\]
\[
= -\frac{1}{2} \left[ f_{11} \left( (s-s^*)^2 - (s-s^*)(y-s^*) \right) + (f_{11} + 2f_{12})(s-s^*)(y-s^*) \right]
\]
\[+ \text{remainder}\]  \hspace{1cm} (9)

where the absolute value of the remainder is less than \( \frac{K}{6} \sum_{k=0}^{3} \left( \frac{3}{k} \right) (s-s^*)^k (y-s^*)^{3-k} \)
and \( K \) is the maximum absolute value of the third order partials of \( f \) evaluated at \( s, y \in \text{supp}(P) \).

Now, \( \int_S \int_S \left( (s-s^*)^2 - (s-s^*)(y-s^*) \right) P(dy)P(ds) = \int_S (s-\bar{s})^2 P(ds) \) \( \bar{s} = \int_S sP(ds) \) is the mean strategy of the population and \( \int_S \int_S ((s-s^*)(y-s^*)) P(dy)P(ds) = (\bar{s}-s^*)^2 \). Thus, from Eq. (9),
\[
\int_S \int_S (f(s^*,y) - f(s,y)) P(dy)P(ds)
\]
\[
= -\frac{1}{2} \left[ f_{11} \left( \int_S (s-\bar{s})^2 P(ds) \right) + (f_{11} + 2f_{12})(\bar{s}-s^*)^2 \right] + \text{h.o.t.}
\]
\[
\geq -\frac{1}{2} \left[ f_{11} \left( \int_S (s-\bar{s})^2 P(ds) \right) + (f_{11} + 2f_{12})(\bar{s}-s^*)^2 \right]
\]
\[
-\frac{K}{6} \sum_{k=0}^{3} \left( \frac{3}{k} \right) \int_S (s-s^*)^k P(ds) \int_S (y-s^*)^{3-k} P(dy)\right| \right| \geq \min \left\{ -\frac{1}{2} f_{11}, -\frac{1}{2} (f_{11} + 2f_{12}) \right\} \int_S (s-s^*)^2 P(ds)
\]
\[
-\frac{K}{6} \left| \sum_{k=0}^{3} \left( \frac{3}{k} \right) \int_S (s-s^*)^k P(ds) \right| \int_S (y-s^*)^{3-k} P(dy) \right| \right| \geq \min \left\{ -\frac{1}{2} f_{11}, -\frac{1}{2} (f_{11} + 2f_{12}) \right\} \int_S (s-s^*)^2 P(ds)
\]
Since \( |\int_S (s-s^*)^k P(ds) \int_S (y-s^*)^{3-k} P(dy)| \leq \epsilon \int_S (s-s^*)^2 P(ds) \) if \( \text{supp}(P) \) is within \( \epsilon \) of \( s^* \), \( \frac{dP}{ds} (\{s^*\}) > 0 \) with strict inequality unless \( \int_S (s-s^*)^2 P(ds) = 0 \) (i.e. unless \( P=\delta_{s^*} \)).

Thus, \( \lim_{t \to \infty} P(t)\{s^*\} = L \) exists and satisfies \( 0 < L \leq 1 \). Since \( \Lambda(S) \) is compact in the weak topology (Oechssler and Riedel, 2002), \( P(t) \) has some other \( \omega \)-limit point \( P^* \) if the trajectory does not converge to \( \delta_{s^*} \). Since \( P^*(\{s^*\}) \geq L \), \(^{11}\) this is impossible if \( L = 1 \). On the other hand, if \( L < 1 \), the above proof shows that \( \frac{dP}{ds} (\{s^*\}) > 0 \) if \( P \) is close to \( P^* \) (which contradicts \( L < 1 \)). Since all limit points are \( \delta_{s^*}, \delta_{s^*} \) is neighborhood attracting.

Conversely, suppose \( f_{11} > 0 \) or \( f_{11} + 2f_{12} > 0 \). Thus, \( s^* \) does not weakly dominate \( s \) in the two-strategy game between \( s^* \) and \( s \) for all \( s \) sufficiently close (but not equal) to \( s^* \). By Proposition 1, \( \delta_{s^*} \) is not neighborhood attracting. \( \Box \)

The above method of proof is related to that of Theorem 3 in Oechssler and Riedel (2002) who proved a similar statement when \( \delta_{s^*} \) is evolutionarily robust in their

\(^{11}\) It is straightforward to show that \( P^*(\{s^*\}) \geq L \). However, \( P^*(\{s^*\}) \) may be strictly greater than \( L \) since \( P(\{s^*\}) \) is not continuous in the weak topology.
terminology (see Remark 5 in the following section). This connection is explained further in Section 4 by introducing the concept of neighborhood superiority. In the special case when there are no terms higher than quadratic in the Taylor expansion of \( f(x,y) \), the proof of Theorem 1 immediately shows the following result on global convergence which is then applied to the Cournot competition model of Example 1.

**Corollary 1.** Suppose \( f(x,y) \) is a quadratic polynomial and \( s^* \) is in the interior of \( S \subset \mathbb{R}^n \). If \( s^* \) strongly dominates \( s \) in the two-strategy game between \( s^* \) and \( s \) for all \( s \) sufficiently close (but not equal) to \( s^* \), then \( \delta_{s^*} \) is globally attracting. Conversely, if \( \delta_{s^*} \) is neighborhood attracting, then \( s^* \) is a strict NE that weakly dominates all \( s \neq s^* \) in the two-strategy game between \( s^* \) and \( s \).

**Example 1b.** Suppose a firm’s cost to produce \( x \geq 0 \) units in Example 1a is given by the increasing linear function \( C(x) = ax + b \). Also suppose ten firms are drawn at random (i.e. \( n = 10 \)) and so \( S = [0,1/10] \). From Eq. (1),

\[
f(x,y) = x(1 - x - 9y) - (ax + b)
\]

and so the payoff function is quadratic. There is a strict NE with \( s^* > 0 \) if \( a < 1 \) and it is given by \( s^* = \frac{1-a}{11} \). By Corollary 1, since \( f_{11} < 0 \) and \( f_{11} + 2f_{12} < 0 \), every initial \( P(0) \) with \( s^* \) in its support converges weakly to \( \delta_{s^*} \) with resultant payoff \( \left(\frac{1-a}{11}\right)^2 - b \). If \( a > 1 \) (i.e. if each firm’s marginal cost always exceeds its marginal revenue), \( s^* = 0 \) is a strict NE since \( f_1 < 0 \) (see Remark 3 below) and \( \delta_0 \) is now globally attracting.

**Remark 2.** The negativity conditions on \( f_{11} \) and \( f_{11} + 2f_{12} \) in Ineq. (8), that follow from the Taylor expansion of Ineq. (6) when \( S \) is one dimensional, have received considerable attention in the coevolutionary literature. The first is often called the ESS criterion for \( s^* \) when the payoff function is nonlinear in its first variable since it implies that a monomorphic population \( s^* \) cannot be invaded by a rare mutant strategy \( s \), paralleling the intuition developed by Maynard Smith (1982). However, as remarked earlier, the term ESS is potentially ambiguous here and so the phrase “local strict NE” is preferred.

The second negativity condition of Ineq. (8) is that \( s^* \) is a NIS (neighborhood invader strategy) (McKelvey and Apaloo, 1995; Apaloo, 1997). It is interesting to note that together these two inequalities imply that \( f_{11} + f_{12} < 0 \). This latter inequality combined with \( f_{11} < 0 \) provide the original criteria (Eshel, 1983) for stability of monomorphic populations called CSS (continuously stable strategy). On its own, the inequality \( f_{11} + f_{12} < 0 \) has come to be known as the condition for convergence stability (Christiansen, 1991) in the adaptive dynamics literature.

Situations where the Taylor expansion is not valid have also been considered. For example, Seymour (2000) has strengthened the strictness concept to that of a super-strict NE whereby, for some \( \varepsilon > 0 \), \( f(s^*,s^*) > f(s,s^*) + \varepsilon \) whenever \( s \neq s^* \). In particular, \( f \) is not continuous. However, if \( f \) is still continuous in its second variable, \( f(s^*,y) - f(s,y) > 0 \) for all \( (s,y) \) sufficiently close (but not equal) to \( (s^*, s^*) \) with \( s \) different than \( s^* \). The above method of proof then shows Theorem 1 remains valid in these circumstances, a result that also follows from Seymour’s analysis.

---

\(^{12}\text{Global means there is no restriction on how far elements of } \text{supp}(P) \text{ are away from } s^* \text{ as long as } P \text{ satisfies the positive mass assumption that } P(\{s^*\}) > 0.\)
Example 2. Let \( S \subset \mathbb{R}^n \) and consider the payoff function given by

\[
f(x,y) = 1 - e^{-\frac{(x-y)(x-y)}{2\sigma^2}} \frac{e^{-\frac{y^2}{2\sigma^2}}}{e^{-\frac{x^2}{2\sigma^2}}}. \tag{10}
\]

When \( n = 1 \), payoffs of this form are frequently used in the adaptive dynamics literature (e.g. Vincent et al., 1996; Doebeli and Dieckmann, 2000) to model strategy evolution of a monomorphic population when the strength of interactions between individuals is larger as traits become more similar (as reflected by the factor \( e^{-\frac{(x-y)^2}{2\sigma^2}} \) where \( \sigma \) is a positive parameter measuring the significance of this similarity). Here \( e^{-\frac{y^2}{2\sigma^2}} \) is interpreted as the carrying capacity of a monomorphic population with strategy \( x \). In particular, traits near zero have higher equilibrium size (perhaps through more efficient use of resources). Then \( e^{-\frac{y^2}{2\sigma^2}} \) is a normalization factor to ensure that monomorphic populations are at their equilibrium size (i.e. \( f(y,y) = 0 \)).

When \( n > 1 \), Eq. (10) models a situation where the effect between \( x \) and \( y \) of being similar traits is the same in every direction from the monomorphism \( y \) and the monomorphic carrying capacity depends only on the distance from \( 0 \). In general, these factors could be direction dependent (e.g. factors of the form \( e^{-\frac{(x-y)^2}{2\sigma^2}} \cdot C(x-y)/2\sigma^2 \) where \( C \) is an \( n \times n \) covariance matrix).

If \( x \) is in the interior of \( S \subset \mathbb{R}^n \) and \( \delta_x \) is neighborhood attracting, then the gradient of \( f(x,y) \) with respect to \( x \) must be zero. That is, from Eq. (10),

\[
0 = \left. \frac{\partial f(x,y)}{\partial x_i} \right|_{y=x} = e^{-\frac{(x-y)(x-y)}{2\sigma^2}} \frac{e^{-\frac{y^2}{2\sigma^2}}}{e^{-\frac{x^2}{2\sigma^2}}} \left( x_i - y_i \right) \frac{1}{\sigma^2} \frac{e^{-\frac{y^2}{2\sigma^2}}}{e^{-\frac{x^2}{2\sigma^2}}} \left( x_i - y_i \right) \frac{1}{\sigma^2}
\]

for all \( i \). Thus, the only possible stable monomorphism is \( \delta_0 \). Furthermore, from the Appendix A, domination is determined by the negative definiteness of the \( n \times n \) matrices with entries \( \left[ \frac{\partial^2 f}{\partial x_i \partial x_j} \right] \) and \( \left[ \frac{\partial^2 f}{\partial x_i \partial y_j} + \frac{\partial^2 f}{\partial x_j \partial y_i} + \frac{\partial^2 f}{\partial y_i \partial y_j} \right] \) evaluated at \( x = y = 0 \). From Eq. (10), these are \( \left( \frac{1}{\sigma^2} - 1 \right) I \) and \( -\left( 1 + \frac{1}{\sigma^2} \right) I \) respectively. By Theorem 1, \( \delta_0 \) is neighborhood attracting if and only if \( \sigma > 1 \).

Theorem 1 cannot be extended to initial \( P(0) \) that have most of their support near \( s^* \) or to cases where strong domination is not determined by second order terms of the Taylor expansion as the following two examples with one-dimensional strategy space demonstrate.

An example of the latter phenomenon is provided by Eshel and Sansone (2003) where

\[
f(x,y) = (x - y)^4 - 2x^4.
\]

It is straightforward to show that 0 strictly dominates all other \( s \in \mathbb{R} \). However, with \( \text{supp}(P(0)) = \{-s,0,s\} \) and \( P(0)(\{s\}) = P(0)(\{-s\}) \), we find

\[
\frac{dP}{dt}(\{0\}) = -P(\{0\})(1 - P(\{0\}))(5 - 6P(\{0\}))s^4 < 0
\]

if \( P(\{0\}) < \frac{5}{6} \). Thus \( P(t) \) does not converge to \( \delta_0 \) even though \( P(0) \) can be taken arbitrarily close to \( \delta_0 \) in the weak topology.

\[\text{Here we ignore the transitional case of } \sigma = 1 \text{(i.e. the case where the effects of similar traits have the same spread as the carrying capacity at 0).}\]
An example from Oechssler and Riedel (2002) shows it is important \( \text{supp}(P(0)) \) is close to \( s^* \). Let \( f(x,y) = -x^2 + x^2y^2 \). It is again straightforward to confirm that \( s^* = 0 \) strongly dominates \( s \) in the two-strategy game between \( s^* \) and \( s \) for all \( 0 < |s| < 1 \). However, \( \frac{dP}{dt}(\{s^*, s\}) < 0 \) if \( \text{supp}(P(0)) = \{s^*, s\} \) and \( P(\{s\}) s^4 - s^2 > 0 \), and so \( P(t) \) does not converge to \( \delta_s \) in the weak topology for all initial \( P(0) \) if \( |s| > 1 \). The problem here is that initial states can be close to \( \delta_s \) in the weak topology without the Hausdorff distance (see Section 4 below) between their compact supports being small.

This problem led Oechssler and Riedel (2002) to consider other static conditions, based on payoff comparisons, that potentially imply dynamic stability with respect to the weak topology. Their most promising static concept (see evolutionary robustness in Remark 5 of Section 4) is based on intuitive stability for initial \( P(0) \) that allow for “both a large change of strategic play by a small fraction of players as well as a small change of strategic play by a large fraction of the population” (which is an informal description of the weak topology). Unfortunately, they were unable to prove a general stability result using this concept which gives a main impetus for the approach adopted in this paper.

**Remark 3.** Most of the above analysis concerns convergence to interior monomorphisms. For one-dimensional strategy space with quadratic payoff functions (as in Examples 1b and 1c), Cressman and Hofbauer (2005) also showed Lyapunov stability of monomorphisms corresponding to strongly dominant strategies \( s^* \) (i.e. those satisfying to Ineq. (8)). Combined with Corollary 1, this shows global asymptotic stability (with respect to initial \( P(0) \) with \( s^* \) in their support) in this case. Furthermore, they were able to analyze what happens when no monomorphism is asymptotically stable (see Example 1c in the following section).

It is worth pointing out\(^{14}\) that the local strict NE condition, \( f_{11} < 0 \), for \( s^* \) is not sufficient for \( \delta_s \) to be neighborhood attracting by Theorem 1 (e.g. take \( f(x,y) = -x^2 + 4xy \) with \( s^* = 0 \)). This contrasts with the situation for finite strategy spaces where the local asymptotic stability of a (local) strict NE is one of the main results of the Folk Theorem of Evolutionary Game Theory (Hofbauer and Sigmund, 1998; Cressman, 2003).

On the other hand, strictness of a NE on the boundary of \( S \) may be sufficient. For instance, if \( S \) is the compact interval \([a,b]\), the endpoint \( s^* = a \) will be a strict NE if \( f_1 < 0 \). Then, from Eq. (7), \( \frac{1}{P(\{s^*\})} \frac{dP}{dt}(\{s^*\}) \equiv -f_2(\bar{s} - s^*) > 0 \) unless \( \bar{s} = s^* \) (i.e. unless \( P = \delta_{s^*} \)) by following the first steps of the proof of Theorem 1. The remainder of the proof can then be adapted to show \( \delta_{s^*} \) is neighborhood attracting. If \( S \) is multi-dimensional and \( s^* \) is a NE on the boundary of \( S \), the combination of being a strict NE as determined by the gradient of \( f \) together with the necessary conditions from Theorem 1 on the second order directional derivatives parallel to the boundary will again imply \( \delta_{s^*} \) is neighborhood attracting.

### 4. Convergence and stability of dimorphic populations

In contrast to the study of stability for monomorphic populations \( P^* = \delta_{s^*} \), very little research has been done that analyzes this question for a general rest point of the replicator

--

\(^{14}\)This is also pointed out in Oechssler and Riedel (2002) for the one dimensional case.
dynamics. Although the main results (Theorems 2 and 3 below) apply to all $P^*$ with finite support, we concentrate here on the most elementary extension; namely, when $P^*$ has two strategies $\{s^*, r^*\}$ in its support (hence, a dimorphism). To generalize Theorem 1, we want conditions for which $P(t)$ converges to $P^*$ in the weak topology if $P(s^*)P(r^*) > 0$ initially and $\text{supp}(P(0))$ is sufficiently close to $\{s^*, r^*\}$.\(^{15}\)

In particular, the dimorphism $P^*$ must be attracting for the replicator dynamics when it is invaded by a mutant strategy $s$ near $s^*$ or $r^*$. Conditions for this with respect to the resultant $3 \times 3$ payoff matrix

\[
A = \begin{bmatrix}
    f(s^*, s^*) & f(s^*, r^*) & f(s^*, s) \\
    f(r^*, s^*) & f(r^*, r^*) & f(r^*, s) \\
    f(s, s^*) & f(s, r^*) & f(s, s)
\end{bmatrix}
\]

are not so precise as for the monomorphisms of Section 2 (see Proposition 1 there). In this section, we consider two separate approaches (Sections 4.1 and 4.2 respectively) that guarantee $P^*$ is locally attracting. It should also be emphasized here that the two approaches will be developed for any $P^*$ with finite support and not only for the dimorphic (or monomorphic) model.

4.1. Neighborhood superiority

The first approach is based on the following definition.

**Definition 2.** The probability measure $P^* \in \Delta(S)$ is neighborhood superior if, for all other $P$ with support sufficiently close to the support of $P^*$, $\pi(P^*, P) > \pi(P, P)$.

**Remark 4.** Weibull (1995) defined the concept of locally superior for matrix game models and noted that a $p^* \in \Delta^m$ is locally superior if and only if $p^*$ is a matrix ESS. His definition (that $\pi(\bar{p}^*, p) > \pi(\bar{p}, \bar{p})$ for all $\bar{p}$ sufficiently close (but not equal) to $\bar{p}^*$) is equivalent to Definition 2 in the mixed strategy matrix model if we only consider those $P$ for which the mean strategy $\bar{p}$ is different from that of $P^*$ since $\pi(P, Q) = \pi(\bar{p}, \bar{q})$. For probability measures on a continuous strategy space, we have used the phrase neighborhood superior in place of locally superior to avoid possible confusion that “local” may be misinterpreted as being with respect to the topology on $\Delta(S)$ rather than with respect to the topology of the strategy space $S$.

Moreover, when Definition 2 is applied to a monomorphic $P^* = \delta_{s^*}$, $P^*$ is neighborhood superior if and only if $s^*$ strictly dominates $s$ in the two person game between $s$ and $s^*$ whenever $s$ is sufficiently close to $s^*$ and domination is determined by the second order Taylor expansion of $f(x, y)$. This is essentially what is proven in Theorem 1.

---

\(^{15}\)We take the Hausdorff distance (Gulick, 1992) as a measure of closeness between two compact sets. Since $\{s^*, r^*\}$ is contained in $\text{supp}(P(0))$, the Hausdorff distance between these sets is $\max_{s \in \text{supp}(P(0))} \min \{s - s^*, |s - r^*|\}$. For the monomorphisms of Section 3, the Hausdorff distance between the supports of $P(0)$ and $\delta_{s^*}$ is the distance from $s^*$ to the farthest point on the boundary of $\text{supp}(P(0))$, $\max_{s \in \text{supp}(P(0))} \{s - s^*\}$.\(\)
By taking supp($P$) = $\{s^*, r^*\}$, we find neighborhood superiority of the dimorphism $P^*$ implies that $f(s^*, r^*) > f(r^*, r^*)$ and $f(r^*, s^*) > f(s^*, s^*)$. Then $P^*$ must be the interior NE $(p_1^*, p_2^*) = (P(\{s^*\}), P(\{r^*\}))$ given by

$$p_1^* = \frac{f(s^*, r^*) - f(r^*, r^*)}{f(s^*, r^*) - f(r^*, r^*) + f(r^*, s^*) - f(s^*, s^*)}$$

$$p_2^* = \frac{f(r^*, s^*) - f(s^*, s^*)}{f(s^*, r^*) - f(r^*, r^*) + f(r^*, s^*) - f(s^*, s^*)}$$

(12)

that is a matrix ESS and globally asymptotically stable with respect to the replicator dynamics for the two-strategy game between $s^*$ and $r^*$. Also, if $P$ is of the form $(1 - \varepsilon) P^* + \varepsilon\delta_s$ for some $0 < \varepsilon < 1$, neighborhood superiority implies that $P^*$ is a locally quasi-strict NE for the three-strategy game with payoff matrix given by Eq. (11) for all $s$ near $s^*$ or $r^*$. 16

The following result connects neighborhood superiority to neighborhood attracting under the measure dynamics version Eq. (2) of the replicator equation.

**Theorem 2.** If $P^*$ is a neighborhood superior probability measure with finite support, then $P^*$ is neighborhood attracting.

**Proof.** We only provide the proof for the case where supp($P^*$) = $\{s^*, r^*\}$. The general proof follows analogously. Consider the function $V: \Delta(S) \to \mathbb{R}$ given by $V(P) = P(\{s^*\})p_1^*P(\{r^*\})p_2^*$. A straightforward calculation shows that, along any trajectory of Eq. (2), $\frac{d}{dt}V = \pi(P^*, P) - (P, P) > 0$ if $P \neq P^*$ and supp($P(0)$) is sufficiently close to supp($P$). Thus $V(P)$ is strictly increasing toward a limit $L$ less than or equal to its maximum value attained only when $P(\{s^*\}) = p_1^*$ and $P(\{r^*\}) = p_2^*$ (i.e. when $P = P^*$). Suppose $Q^*$ is an $\omega$-limit point of $P(t)$. Then $V(Q^*) \geq L > 0$. If $Q^* \neq P^*$, then $V(P^*) > V(Q^*)$ and $\frac{d}{dt}V > \delta > 0$ whenever $P(t)$ is sufficiently close to $Q^*$ in the weak topology (since $\pi(P^*, P) - \pi(P, P)$ is a continuous function of $P$ in the weak topology). This contradicts

$$\lim_{t \to \infty} V(P(t)) \leq V(Q^*).$$

Thus $P^*$ is the only $\omega$-limit point of $P(t)$ and the result follows from the compactness of $\Delta(S)$ in the weak topology (as in the proof of Theorem 1). $\Box$

**Example 1c.** Consider the Cournot competition model of Example 1a with ten firms drawn at random as in Example 1b. However, suppose that each firm’s cost function is now the increasing quadratic $C(x) = \frac{1}{2}x - 2x^2$ where production $x$ is assumed to be limited to the interval $S = [0, 1_{10}]$. From Eq. (1),

$$f(x, y) = x(1 - x - 9y) - \left(\frac{1}{2}x - 2x^2\right) = \frac{1}{2}x + x^2 - 9xy.$$
No interior $s \in (0, \frac{1}{10})$ is a local NE since this payoff function is an upward parabola in $x$. Furthermore, neither endpoint is a local NE since $f_1(0,0) > 0$ and $f_1 \left( \frac{1}{10}, \frac{1}{10} \right) < 0$ (see Remark 3). Thus, no monomorphism $\delta_s$ is neighborhood attracting.

On the other hand, for the dimorphism $P^* = \frac{1}{3} \delta_0 + \frac{2}{3} \delta_{\frac{1}{10}}$,

$$\pi(P^*, P) - \pi(P, P) = \frac{1}{3} f(0, \tilde{P}) + \frac{2}{3} f \left( \frac{1}{10}, \tilde{P} \right) - \left( \frac{1}{2} \tilde{P} + \text{Var}(P) + \tilde{P}^2 - 9\tilde{P}^2 \right)$$

where $\tilde{P} = \int_S y P(dy)$ and $\text{Var}(P) = \int_S (y - \tilde{P})^2 P(dy)$ are the mean and variance of $P$ respectively. Since $\text{Var}(P) \leq \frac{\tilde{P}}{10} - \tilde{P}^2$ (with equality if and only if $\text{supp}(P) \subset \{0, \frac{1}{10}\}$),

$$\pi(P^*, P) - \pi(P, P) \geq \frac{2}{3} \left( \frac{1}{20} + \frac{1}{100} - \frac{9}{10} \tilde{P} \right) - \left( \frac{1}{2} \tilde{P} + \tilde{P}^2 - 9\tilde{P}^2 \right)$$

$$= \frac{9\tilde{P}^2}{5} - \frac{6}{25} \tilde{P} = 9 \left( \tilde{P} - \frac{1}{15} \right)^2 \geq 0$$

with equality if and only if $\tilde{P} = \frac{1}{15}$ and $\text{supp}(P) = \{0, \frac{1}{10}\}$ (i.e. if and only if $P = P^*$).

Thus, $P^*$ is globally superior and, by a straightforward extension of Theorem 2, is globally attracting. That is, the evolutionary approach to this Cournot competition example restricted to the strategy set $[0, 1/10]$ predicts two out of three firms in the large population will produce at the maximum output level of $\frac{1}{10}$ and the other firms will produce nothing. Furthermore, each firm will receive profit $\pi(P^*, P^*) = 0$ at these production levels.

**Remark 5.** Definition 2 and Theorem 2 are closely related to the concept of evolutionarily robust introduced by Oechssler and Riedel (2002) as a $P^* \in \Delta(S)$ for which $\pi(P^*, P) - \pi(P, P) > 0$ for all $P$ sufficiently close (but not equal) to $P^*$ in the weak topology. Their main goal was to show an evolutionarily robust $P^*$ attracts all $P$ sufficiently close in the weak topology whose support contains $\text{supp}(P^*)$ (see their Conjecture 1). However, they were only able to obtain partial results for monomorphic $P^*$ and report none for other $P^*$ with finite support. The main obstacle they encounter is how to control the payoff effects of strategies far from the support of $P^*$. From this perspective, the neighborhood attracting property of Definition 1 removes this obstacle by only looking at measures with support close to that of $P^*$, leading to the positive results of Theorems 2 and 3 in Sections 4.1 and 4.2 respectively.

### 4.2. Strategy domination method for dimorphic rest points

It is well known that local superiority (i.e. the matrix ESS condition as in Remark 4) is not a necessary condition for a dimorphism to be attracting in the matrix game model when it is invaded by a mutant strategy $s$. For instance, following Akin (1980), strategy domination may be used instead to eliminate the mutant strategy under the replicator dynamics.

We can no longer expect $s^*$ to dominate all strategies in $\text{supp}(P(0))$ as in Proposition 1 and Theorem 1 since it does not dominate $r^*$ in the two-strategy game between $s^*$ and $r^*$. Instead, if $\text{supp}(P^*) = \{s^*, r^*\}$, we look for dominance of nearby strategies that implies
\[ P(B) \text{ converges to 0 if } B \text{ is a compact set sufficiently close to and disjoint from } s^*. \] (and also for such sets \( C \) near \( r^* \)). Our two domination results are based on the following expression for the evolution of \( \frac{dP(B)}{dt} \), assuming initial positive mass on \( s^* \). From Eq. (2),

\[
\frac{d}{dt} \left( \frac{P(B)}{P(s^* \cup r^*)} \right) = \frac{1}{(P(s^*))^2} \int_B (\pi(s,P) - \pi(P,P))P(ds) - P(s^*) \pi(s^*,P) - \pi(P,P)P(B) \]

\[
= \frac{1}{(P(s^*))^2} \int_B (\pi(s,P) - \pi(s^*,P))P(ds) \]

\[
= \frac{1}{(P(s^*))^2} \int_S \int_B (f(s,y) - f(s^*,y))P(dy)P(ds). \]

First, suppose that, in the game with strategy set \( \text{supp}(P(0)) \), \( s^* \) strongly dominates \( s \) for all other \( s \) near \( s^* \) (i.e. \( f(s^*,y) > f(s,y) \) for all \( y \in \text{supp}(P(0)) \)) and \( r^* \) strongly dominates \( r \) for all other \( r \) near \( r^* \). Then, \( \frac{d}{dt} \left( \frac{P(B)}{P(s^*)} \right) < -\delta \frac{P(B)}{P(s^*)} \) where \( \delta = \min \{ f(s^*,y) - f(s,y) : s \in B, y \in \text{supp}(P(0)) \} \) is positive since \( B \) and \( \text{supp}(P(0)) \) are compact and \( f \) is continuous.

For the second result, suppose that \( s^* \) strongly dominates \( s \) for all other \( s \) near \( s^* \) in the two-strategy game between \( s^* \) and \( s \) and that \( s^* \) weakly dominates \( s \) for all other \( s \) near \( s^* \) in the three-strategy game \( \{ s^*,s,y \} \) with \( y \) near \( s^* \). Similarly, \( r^* \) strongly dominates \( r \) for all other \( r \) near \( r^* \) in the two-strategy game between \( r^* \) and \( r \) and that \( r^* \) weakly dominates \( r \) for all other \( r \) near \( r^* \) in the three-strategy game \( \{ r^*,r,y \} \) with \( y \) near \( s^* \). If \( \text{supp}(P(0)) \) is sufficiently close to \( \text{supp}(P^*) \) in Hausdorff distance, \( \frac{d}{dt} \left( \frac{P(B)}{P(s^*)} \right) < -\delta \frac{P(B)}{P(s^*)} \) since \( f(s,y) \leq f(s^*,y) \) for all \( y \) near \( r^* \). Now \( \delta \) is the positive lower bound for \( \frac{1}{(P(s^*))^2} \int_B (f(s^*,y) - f(s,y))P(dy)P(ds) \) found in the proof of Theorem 1.

In either case, \( \frac{d}{dt} \left( \frac{P(B)}{P(s^*)} \right) \) is monotone decreasing to zero and so \( P(B) \) must also converge to 0. Similarly, \( P(C) \) converges to 0 for all compact sets sufficiently close to and disjoint from \( r^* \). These results lead to the two parts of the following theorem.

**Theorem 3.** Suppose there are \( m \) strategies in the support of \( P^* \) and \( P^* \) is a globally asymptotically stable equilibrium in the corresponding \( m \)-strategy game.

(a) If, for all \( s^* \in \text{supp}(P^*) \), \( s^* \) strongly dominates every other nearby \( s \) in the game whose strategy set \( S' \) is sufficiently close to \( \text{supp}(P^*) \) in Hausdorff distance, then \( P^* \) is neighborhood attracting.

(b) Suppose, for all \( s^* \in \text{supp}(P^*) \) and nearby \( s \), that \( s^* \) strongly dominates in the two-strategy game between \( s^* \) and \( s \) and that \( s^* \) weakly dominates \( s \) in the three-strategy game \( \{ s^*,s,y \} \) with \( y \) near \( \text{supp}(P^*) \) \( \backslash \{ s^* \} \). Then \( P^* \) is neighborhood attracting.

**Proof.** We only provide the proof of part (a) when \( P^* \) is a dimorphism with \( \text{supp}(P^*) = \{ s^*,r^* \} \). The proof of part (b) and for polymorphic \( P^* \) with \( m > 2 \) is then straightforward given the above estimates for \( \frac{d}{dt} \left( \frac{P(B)}{P(s^*)} \right) \).
This last expression is positive if and only if \( \frac{d}{dt} \left( \frac{P(U_x(r^*))}{P(U_x(s^*))} \right) \) is given by

\[
P(U_x(s^*)) \int_{U_x(r^* \cup s^*)} f(r, y)P(dy)P(dr) - P(U_x(r^*)) \int_{U_x(s^*)} f(s, y)P(dy)P(ds)
\]

\[
\geq P(U_x(r^*)) \left[ f(r^*, r^*)P(U_x(r^*)) + f(r^*, s^*)P(U_x(s^*)) \right]
\]

\[
- \left[ f(s^*, r^*)P(U_x(r^*)) + f(s^*, s^*)P(U_x(s^*)) \right]
\]

\[
= P(U_x(r^*)) \left[ (f(r^*, r^*) - f(s^*, r^*))P(U_x(r^*)) \right]
\]

\[
+ (f(r^*, s^*) - f(s^*, s^*)) (1 - P(U_x(r^*))) \right].
\]

This last expression is positive if and only if \( P(U_x(r^*)) < \frac{f(r^*, s^*) - f(s^*, s^*)}{f(s^*, r^*) - f(r^*, s^*) + f(r^*, r^*) - f(s^*, s^*)} \) \( \int_{0}^{1} \). Thus, as \( P(U_x(r^*)) + P(U_x(s^*)) \) approaches 1, \( P(U_x(r^*)) \) approaches \( \frac{f(r^*, s^*) - f(s^*, s^*)}{f(s^*, r^*) - f(r^*, s^*) + f(r^*, r^*) - f(s^*, s^*)} \). That is, \( P(t) \) converges in the weak topology to the dimorphism given by Eq. (12). □

**Example 3.** Suppose that \( S \subset R \) and that both \( r^* = -1 \) and \( s^* = 1 \) are neighborhood attracting according to Theorem 1. For instance, near either of these points, take \( f(x, y) \) to be the translations either of the quadratic payoff function of Example 1b or the exponential payoffs of Example 2 with \( \sigma > 1 \). Also, suppose the payoff effect mentioned in Example 2 for interactions between individuals with traits \( x \) and \( y \) disappears if \( |x - y| > 1 \). Then \( f(x, s^*) \) is constant if \( x \) is near \( r^* \) and \( f(x, r^*) \) is constant if \( x \) is near \( s^* \). If interactions between similar traits have negative effects on payoffs, then \( f(s^*, r^*) > f(r^*, r^*) \) and \( f(r^*, s^*) > f(s^*, s^*) \). Thus, by Theorem 3, the dimorphism \( P^* \) given by Eq. (12) is neighborhood attracting.

This example shows how a small mass introduced far away from a monomorphism \( \delta_x \) that is neighborhood attracting has the potential to destabilize the monomorphism with the resultant system evolving to a neighborhood attracting dimorphism \( P^* \). This again demonstrates the potential problems that arise if convergence to a monomorphism \( \delta_x \) in the weak topology is sought without restricting the support of initial \( P \) to be close to \( s^* \).

Sections 4.1 and 4.2 describe two different methods to approach convergence and stability questions of dimorphic populations (or, more generally, populations with finite support). In fact, other approaches may also be feasible since neither neighborhood superiority nor domination criteria completely characterize dimorphic \( P^* \) being attracting under invasion by a mutant strategy through the payoff matrix Eq. (11)\(^{17} \).

\(^{17}\)This lack of precision was also mentioned at the beginning of Section 4.
5. Conclusion

Predicting the behaviors of rational individuals involved in game interactions through analyzing the stability of rest points of evolutionary dynamics is a well-accepted approach (Weibull, 1995; Hofbauer and Sigmund, 1998; Cressman, 2003), especially when the game has a finite number of strategies. When there is a continuous strategy space $S$ and a rest point $P^*$ is given by a measure (i.e. a distribution over $S$), stability conditions for $P^*$ based on the game’s payoffs are not so well known, perhaps because there is no general consensus on what constitutes such stability. In this paper, we have taken the concept to mean that all distributions with support close to that of $P^*$ (and that place positive weight on each element in the support of $P^*$) converge to $P^*$ in the weak topology. This is the neighborhood attracting property of Definition 1.

From this perspective, Section 2 demonstrates that our concept is a clear extension of the finite strategy model. Sections 3 and 4 then develop the theory for a continuous strategy space. Here it is seen that conditions on $f(x,y)$ for stability in terms of strategy domination (Proposition 1 and Theorems 1 and 3) emerge naturally by considering finite strategy games contained in the continuous model. These domination conditions (e.g. Ineq. (8)) are relatively elementary inequalities to confirm for a given function $f$. On the other hand, the intuitive neighborhood superiority condition of Theorem 2 is not as easy to verify but does have the potential to predict rational behavior when the equilibrium does not have finite (or even discrete) support. For instance, the candidate stable equilibrium $P^*$ for the standard War of Attrition Game (Oechssler and Riedel, 2001; Cressman, 2003) with continuous strategy space as compact interval has the entire first half of this interval contained in its support. Oechssler and Riedel (2001) show that this $P^*$ is globally superior and are able to prove this implies convergence in the weak topology.

The techniques developed in this paper are more directly applicable to a rest point $P^*$ with finite support. They are particularly relevant for the stability analysis of monomorphic populations (Section 3), a topic closely connected to recent work on coevolutionary models among theoretical biologists (see Remark 2 especially). Furthermore, the interpretation of these techniques in terms of strategy domination (an approach more familiar to game theorists interested in rational behavior) has generalizations to the study of stability for dimorphic populations as shown in Section 4. These monomorphic and dimorphic results suggest that the emerging theory of evolutionary dynamics on continuous strategy spaces will continue the tradition of evolutionary game theory that fosters corresponding models and methods in predicting behavior both of human and of other species.

Acknowledgements

The author appreciates suggestions for improvements on the original version by Immanuel Bomze and Josef Hofbauer as well as by an Associate Editor and a referee. This research was completed while the author was a Fellow at the Collegium Budapest. He thanks both the Collegium and its staff for their hospitality and research support. Financial
assistance from the Natural Sciences and Engineering Research Council of Canada is also gratefully acknowledged.

Appendix A

Proof of the remainder of Theorem 1. From the assumptions that $s^*$ is in the interior of $S \subset \mathbb{R}^n$ and that domination in the two-strategy game between $s^*$ and $s$ for all $s$ sufficiently close (but not equal) to $s^*$ is determined by the second order Taylor expansion of $f(x,y)$ about $x = y = s^*$, we have that the inequalities in Ineq. (8) must hold where the second order partial derivatives are now in the direction from $s^*$ to $s$. The first inequality implies that

$$
\sum_{i,j} \frac{\partial^2 f(x,y)}{\partial x_i \partial x_j} (s_i - s^*_i) (s_j - s^*_j) < 0
$$

for all $s$ in a neighborhood of $s^*$ (here the second order partial derivatives are evaluated at $(s^*,s^*)$). That is, the $n \times n$ matrix with $ij$ entry $\frac{\partial^2 f}{\partial x_i \partial x_j}$ (i.e. the matrix $[\frac{\partial^2 f}{\partial x_i \partial x_j}]$) is negative definite. Similarly, from the second inequality in to Ineq. (8), the $n \times n$ matrix $[\frac{\partial^2 f}{\partial x_i \partial y_j} + \frac{\partial^2 f}{\partial y_i \partial x_j}]$ is also negative definite.\footnote{The use of negative definiteness in the proof is similar to the method used by Meszéna et al. (2001) (see also Bomze, 1990; Apaloo, 1997) where $\int_S (s_i - \overline{s}_i)(s_i - \overline{s}_j)P(ds)$ give entries of the covariance of the population mean strategy.}

From the second order Taylor expansion, $\int_S \int_S (f(s^*,y) - f(s,y))P(dy)P(ds)$ is now given by

$$
-\frac{1}{2} \left[ \int_S \int_S \sum_{i,j} f_{x_i x_j} ((s_i - s^*_i) (s_j - s^*_j) - (s_i - s^*_i) (y_j - s^*_j))P(dy)P(ds) 
+ \int_S \int_S \sum_{i,j} (f_{x_i x_j} + f_{x_j x_i}) (s_i - s^*_i) (y_j - s^*_j)P(dy)P(ds) \right]
$$

$$
= -\frac{1}{2} \left[ \int_S \sum_{i,j} f_{x_i x_j} (s_i - \overline{s}_i) (s_j - \overline{s}_j)P(ds) 
+ \sum_{i,j} (f_{x_i x_j} + f_{x_j x_i}) (\overline{s}_i - \overline{s}^*_i) (\overline{s}_j - \overline{s}^*_j) \right].
$$

Since $[f_{x_i x_j}]$ is negative definite, $\int_S \sum_{i,j} f_{x_i x_j} (s_i - \overline{s}_i) (s_j - \overline{s}_j)P(ds)$ is negative unless $s_i = \overline{s}_i$ for all $i$ whenever $s_i \in \text{supp}(P)$. Similarly, the negative definiteness of $[f_{x_i x_j} + f_{x_j x_i} + f_{x_i x_j}]$ implies $\sum_{i,j} (f_{x_i x_j} + f_{x_j x_i} + f_{x_i x_j}) (\overline{s}_i - \overline{s}^*_i) (\overline{s}_j - \overline{s}^*_j)$ is negative unless $\overline{s}_i = \overline{s}^*_i$ for all $i$. The remainder of the proof follows as in the main text. \qed
References