Measure Dynamics on a One-Dimensional Continuous Trait Space: Theoretical Foundations for Adaptive Dynamics*

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

rcressma@wlu.ca

Josef Hofbauer
Department of Mathematics
University College London
London WC1E 6BT, U.K.

jhofb@math.ucl.ac.uk

Department of Mathematics
University of Vienna
1090 Vienna, Austria

josef.hofbauer@univie.ac.at

August, 2004

Abstract. The measure dynamics approach to modelling single-species coevolution with a one-dimensional trait space is developed and compared to more traditional methods of adaptive dynamics and the Maximum Principle. It is assumed that individual fitness results from pairwise interactions together with a background fitness that depends only on total population size. When fitness functions are quadratic in the real variables parameterizing the one-dimensional traits of interacting individuals, the following results are derived. It is shown that among monomorphisms (i.e. measures supported on a single trait value), the CSS (Continuously Stable Strategy) characterize those that are Lyapunov stable and attract all initial measures supported in an interval containing this trait value. In the cases where adap-

---

tive dynamics predicts evolutionary branching, convergence to a dimorphism is established. Extensions of these results to general fitness functions and/or multi-dimensional trait space are discussed.

Keywords: Adaptive dynamics, CSS, evolutionary branching, replicator equation, entropy, mean fitness, local superiority, strategy dominance, measure dynamics, weak topology

Running Head: Measure Dynamics on a Continuous Trait Space

1 Introduction

Interest in adaptive dynamics as a means to examine stability of coevolutionary systems has grown exponentially over the past decade (see Abrams (2001) and the references therein). Cornerstones for this theory are the stability conditions (e.g. continuously stable strategy, convergence stability) developed for the mean strategy dynamics of a single species with a one-dimensional continuous trait space. We briefly summarize this approach in Section 2.1 for the special case when individual fitness is given by two-variable quadratic functions defined on the trait space through pairwise interactions. As coevolution also involves a density dynamics on the total population size, we include a background fitness (that is strategy independent and decreasing with respect to density) to limit population growth. This has the effect that stability of the coevolutionary system is completely determined by the strategy dynamics. Here, adaptive dynamics predicts stability of a monomorphic equilibrium (i.e. one where all individuals in the population are using the same strategy) if, for all other monomorphisms that are small perturbations of this equilibrium, trait substitution through nearby mutations is only successful when this substitution moves the population closer to the equilibrium.

There is a general consensus among practitioners of adaptive dynamics (e.g. Abrams and Matsuda, 1997) that the assumptions underlying this approach (e.g. maintenance of monomorphisms through trait substitution and the suppression of population size effects) limit its applicability, especially as the theory progresses to analyzing non equilibrium limiting behavior. One alternative approach is to consider stability for only those coevolutionary systems where the distribution of strategies has finite support (i.e. there are only finitely many different individual strategies used by the population during the course of evolution), probably close to the monomorphic equilibrium. This approach, which in some sense ignores the possibility of continual though rare mutation, is closely related to the Maximum Principle promoted by Vincent and co-workers (Cohen et al., 1999; Vincent et al., 1996)[1] as summarized in Section 2.2. Both the adaptive dynamics and the Maximum Principle predictions for coevolution ultimately rest on the stability analysis of finite-dimensional, deterministic systems of differential equations that approximate the infinite-dimensional stochastic process of individual births, deaths, interactions, etc. The theory relies on the fact that the approximation is accurate as the number of individuals gets large.

It is always easier to criticize existing theories than to develop an alter-

---

[1]The literature here calls this the ESS maximum Principle. As the term ESS has several possibly different connotations, we prefer to either drop this qualification altogether or to replace it with the more neutral game-theoretic term of strict NE (Nash equilibrium).

native. The deterministic alternative we prefer is dynamic stability in the space of measures, an extension of the concepts developed for strategy distributions to models that include density dependence. This infinite-dimensional dynamical system has the advantage that it includes models of populations with an arbitrary number of individual strategies. Dynamics on strategy distributions (and not just the mean) with continuous strategy spaces have also been considered (Bomze, 1990, 1991; Oechssler and Riedel, 2001, 2002; Eshel and Sansone, 2003) where quadratic interaction terms are quite commonly used. In contrast to adaptive dynamics where monomorphic populations are invaded by rare mutants, this literature considers the evolution of distributions close to the monomorphic equilibrium distribution. This approach also has the potential to predict convergence to non monomorphic equilibria as well as the evolution of non equilibrium behavior.

For these reasons (see also the final section), we consider this dynamics (with the addition of background fitness) to better model the coevolutionary process. The main purpose of this paper is then to develop the measure dynamics method. To this end, Sections 3 and 4 generate convergence and stability conditions for the measure dynamics in a general setting. These results give exact conditions in Section 5 with our assumption of quadratic pairwise interactions and background fitness which are then compared to those of adaptive dynamics and the Maximum Principle. In particular, the Summary of Section 5.1 can be interpreted as providing a measure dynamics foundation for the theory of adaptive dynamics. Section 6 extends these methods to other fitness functions on a one-dimensional trait space. Further extensions to multi-dimensional trait space and to general fitness functions are discussed in the final section, emphasizing the added analytic problems that arise in these circumstances.

2   The Quadratic Pairwise Interaction Model

Suppose individuals in our single species use strategies that are parameterized by a single real variable x belonging to a closed and bounded interval $S = [\alpha, \beta]$. All the dynamics considered in this paper are based on the fitness of an individual with strategy (or trait) x. We take the game-theoretic approach (Maynard Smith, 1982; Eshel, 1983; Eshel et al., 1997) that assumes individual fitness results from payoffs received during pairwise interactions. For the quadratic pairwise interaction model, the payoff of an individual using strategy x against one using strategy y is

$$\pi(x, y) = ax^2 + bxy + cy^2 + dx + ey + f$$

4

where $x, y \in [\alpha, \beta]$.[2] Through these interactions, fitness of an individual using x is then taken to be the expected payoff this individual obtains in a random pairwise interaction with another individual in the population. Population size is assumed sufficiently large that stochastic effects resulting from a finite population can be ignored.[3]

Fitness in coevolutionary models also depends on population density N. To avoid some mathematical complications, we assume payoff is positive for all strategy pairs when the population size N is zero (i.e. no Allee effect) and negative when N is sufficiently large. The simplest way to accomplish this mathematically is to add an appropriate density term to the individual payoff function that is independent of the strategy pair (i.e. a "background" fitness term) and decreasing in density (Cressman, 1992). Specifically, we take

$$\pi(x, y, N) = ax^2 + bxy + cy^2 + dx + ey + f(N) \tag{1}$$

where $f(N)$ is a decreasing function of N with derivative $f'(N)$ bounded away from 0 and $f(0)$ chosen to make $\pi(x, y, 0) > 0$ for all $x, y \in [\alpha, \beta]$.

In the remainder of this section, we briefly describe the approaches of adaptive dynamics (Section 2.1) and the Maximum Principle (Section 2.2) as they apply to the stability analysis of monomorphic populations when payoff is given by (1).

2.1 Adaptive dynamics

The adaptive dynamics approach (Hofbauer and Sigmund, 1990) to stability of a monomorphism is based on a concept, introduced by Eshel and coworkers (e.g. Eshel, 1983; Eshel et al., 1997) for models without explicit density dependence, that has come to be known as *convergence stable* (Christiansen, 1991; Taylor, 1989). A monomorphism x^* is convergence stable if every y sufficiently close (but not equal) to x^* has a neighborhood $U(y)$ such that the fitness of any $x \in U(y)$ when playing against y should be greater than that of y against y if and only if x is closer to x^* than y.

With density dependent adaptive dynamics (Metz et al., 1996; Marrow et al., 1996; Dieckmann and Law, 1996), these fitnesses are calculated when population size is at its equilibrium value for the monomorphism x. We first find the equilibrium density $N(x_0)$ for x_0. That is, we solve $\pi(x_0, x_0, N(x_0)) = 0$

---

[2] Unless otherwise stated, our variables x, y, x^* etc are all assumed to belong to a closed and bounded interval $[\alpha, \beta]$. When these variables are restricted to a finite set (i.e. there are a finite number of traits), $\pi(x, y)$ is then an entry in the payoff matrix (Maynard Smith, 1982).

[3] Also ignored is the fact an individual does not interact with himself.

for $N(x_0)$ to obtain[4]

$$N(x_0) = f^{-1}\left(-(ax_0^2 + bx_0^2 + cx_0^2 + dx_0 + ex_0)\right).$$

We assume $x^* \in (\alpha, \beta)$ (i.e. in the interior of the trait space S). So x^* is convergence stable if and only if there exists an $\varepsilon > 0$ such that for all $0 < |y - x^*| < \varepsilon$ there is a $\delta > 0$ (which is usually taken less than ε and dependent on y) such that, for all $| x - y | < \delta$,

$$\pi(x, y, N(y)) > \pi(y, y, N(y)) = 0 \tag{2}$$

if and only if $0 < |x - x^*| < |y - x^*|$.

The intuition here is that mutations near y will only be successful if they are between y and the monomorphism, thereby driving the population to x^*. From (2), we consider the difference

$$
\begin{aligned}
&\pi(x, y, N(y)) - \pi(y, y, N(y)) \\
=\ & ax^2 + bxy + cy^2 + dx + ey + f(N(y)) - \left(ay^2 + by^2 + cy^2 + dy + ey + f(N(y))\right) \\
=\ & a(x^2 - y^2) + b(x - y)y + d(x - y) \\
=\ & (x - y)[a(x + y) + by + d].
\end{aligned}
$$

If $2ax^* + bx^* + d > 0$, then $\pi(x, y, N(y)) - \pi(y, y, N(y)) < 0$ if $x^* < x < y$ and y is sufficiently close to x^* so that $a(x + y) + by + d > 0$. That is, x^* is not convergence stable. By a similar argument with $2ax^* + bx^* + d < 0$, we have that a necessary condition for x^* to be convergence stable is

$$2ax^* + bx^* + d = 0.$$

That is, as a function of x, $\pi(x, x^*, N(x^*))$ has a critical point when $x = x^*$.

Furthermore, if $2a + b = 0$, then $d = 0$ and so

$$\pi(x, y, N(y)) - \pi(y, y, N(y)) = (x - y)[a(x + y) - 2ay] = a(x - y)^2.$$

Thus $2a + b \neq 0$ if x^* is convergence stable. This implies the dominating term in $\pi(x, y, N(y)) - \pi(y, y, N(y))$ is $(2a + b)(x - y)(y - x^*)$ and so x^* is convergence stable if and only if

$$
\begin{aligned}
2ax^* + bx^* + d\ &= 0 \\
2a + b\ &< 0.
\end{aligned}
$$

---

[4]Since $f(N)$ is strictly decreasing (and so has an inverse) with 0 in its range, $N(x_0)$ is well-defined and positive. For these same reasons, $N(x_0)$ is globally asymptotically stable for the density dynamics (cf. (4) below) when the population is monomorphic at x_0.

These conditions for convergence stability can be rewritten in their more traditional form (e.g. Marrow et al., 1996) as

$$\frac{\partial}{\partial x}\pi(x, x^*, N(x^*))|_{x=x^*} = 0$$

$$\frac{\partial^2}{\partial x^2}\pi(x, y, N(y))|_{x=y=x^*} + \frac{\partial^2}{\partial x \partial y}\pi(x, y, N(y))|_{x=y=x^*} < 0.$$

If there are terms in $\pi(x, y)$ of higher order than quadratic, then x^* may be convergence stable even if the last inequality is not strict (in which case these higher order terms need to be considered).

Adaptive dynamics is concerned with the evolution of the mean strategy of the population. Under the assumption that the ecological time scale (i.e. the time scale for changes in population size) is much faster than the evolutionary time scale on which the mean strategy evolves, adaptive dynamics eliminates the ecological effect by asserting that the coevolutionary system then tracks equilibrium population size (see also the discussion at the beginning of this section). The canonical equation for the mean strategy evolution near a monomorphic x^* is then

$$\frac{dy}{dt} = k(y)\frac{\partial}{\partial x}\pi(x, y, N(y))|_{x=y} \qquad (3)$$

where $k(y)$ is a positive function that is related to the evolutionary time scale and to equilibrium size. For our quadratic payoff model, we have $\frac{dy}{dt} = k(y)(2ay + by + d) = k(y)(2a + b)(y - x^*)$. We see y is asymptotically stable for the canonical equation if and only if $y = x^*$ where x^* is convergence stable.

2.2 The Maximum Principle

To simplify notation somewhat, we can shift the monomorphism $x^* = \frac{-d}{2a+b}$ to 0 (and so $x^* = 0 \in (\alpha, \beta)$) by replacing x and y with $x - \frac{d}{2a+b}$ and $y - \frac{d}{2a+b}$ respectively. This has the effect of eliminating the dx term in (1) so we now have[5]

$$\pi(x, y, N) = ax^2 + bxy + cy^2 + ey + f(N).$$

Vincent and coworkers (see Cohen et al., 1999 and the references therein) take a different approach to model dynamics stability in coevolutionary systems. Following Vincent et al. (1996), the strategy $x^* = 0$ (for them,

---

[5]This change of variables does shift e and $f(N)$ by constants but these have no effect on the mathematical analysis.

7

a coalition of one) is evolutionarily stable for the equilibrium size $N(x^*)$ if, for all choices of finitely many mutant strategies $\{x_2, ..., x_r\}$, the state $(N(x^*), 0, ..., 0)$ is asymptotically stable for the population dynamics

$$\dot{n}_i = n_i F_i(n_1, ..., n_r) \tag{4}$$

where n_i is the size of that part of the population where individuals use strategy x_i (here x_1 is identified with x^*) and $F_i(n_1, ..., n_r)$ is the expected fitness of strategy x_i when the population state is $(n_1, ..., n_r)$.

When applied to our model with quadratic payoff functions and random pairwise interactions that occur once per unit time for each individual, these fitnesses behave additively to yield

$$F_i(n_1, ..., n_r) = \sum_{j=1}^{m} n_j \frac{\pi(x_i, x_j, N)}{N}$$

where $N = \sum n_j$.

To check stability, the $r-$dimensional system is linearized at $(N(x^*), 0, ..., 0)$. This has the form of an upper triangular $r \times r$ matrix with diagonal entries

$$N(x^*)\frac{\partial F_1}{\partial n_1}, F_2, ..., F_r$$

where all these functions and partial derivatives are evaluated at $(N(x^*), 0, ..., 0)$. For $i > 1$, $F_i = \pi(x_i, x^*, N(x^*)) = ax_i^2$ and $\frac{\partial F_1}{\partial n_1} = \frac{\partial \pi(x^*, x^*, N)}{\partial N} = f'(N(x^*)) < 0$. Thus $x^* = 0$ is evolutionarily stable if $a < 0$ and unstable if $a > 0$.

Although the case $a = 0$ is quite important since it forms the basis of models where fitness is linear in the individual's choice of strategy (i.e. when $\pi(x, y, N)$ is linear in x), in our context we disregard this possibility as degenerate and so conclude that $x^* = 0$ is evolutionarily stable[6] according to Vincent and coworkers if and only if

$$a < 0.$$

The Maximum Principle is then equivalent to asserting that their "fitness generating function", $\pi(x, x^*, N(x^*))$, has a strict maximum at $x = x^* = 0$ as a function of x.

This condition seems to have no immediate connection to that of convergence stability. However, in the adaptive dynamics approach, it is often assumed (Metz et al., 1996; Marrow et al., 1996) no mutant strategies x

---

[6] As mentioned in the Introduction, we prefer to designate this condition as stating x^* is a strict NE.

can invade x^* (i.e. none have higher fitness than x^* when the population is monomorphic at x^*). This is equivalent (assuming $a \neq 0$) to $a < 0$. In fact, the condition $a < 0$ was already assumed by Eshel (1983) when he combined convergence stability with it to define a continuously stable strategy (CSS) x^* to be one that satisfies the two conditions, $a < 0, 2a + b < 0$.[7]

On the other hand, it should be noted that adaptive dynamics is also quite interested in the convergence stable situation with $a > 0$ since they view this as an instance of sympatric speciation (Doebeli and Dieckmann, 2000) and/or evolutionary branching (Metz et al., 1996; Geritz et al., 1997; Geritz et al., 1998). Furthermore, Vincent et al. (1993) (see also Cohen et al., 1999) have developed a mean strategy dynamics through their population dynamics model above that leads back to the canonical equation. Nevertheless, it is clear that there are discrepancies between these two approaches to modeling monomorphic stability in coevolutionary systems.

3   Measure Dynamics

The coevolutionary dynamics we consider is a generalization of the population dynamics (4) to the space of distributions of the population over the continuous trait space $S = [\alpha, \beta]$. Specifically, let μ be a finite measure defined on the $\sigma-$algebra $\mathcal{B}$ of Borel subsets of S. When the population is in state μ, the measure $\mu(B)$ for any $B \in \mathcal{B}$ is interpreted as the size of that part of the population that is using strategies in B.[8] Then $\mu(S)$ is the total population size which we assume to be positive. The fitness of an individual using strategy $x \in S$ is then its expected payoff in a random pairwise interaction (plus the background fitness); namely,

$$\pi(x, \mu) = \frac{1}{\mu(S)} \int_S \pi(x, y, \mu(S))\mu(dy) \tag{5}$$

For our quadratic payoff functions, we obtain

$$\pi(x, \mu) = \frac{1}{\mu(S)} \int_S (ax^2 + bxy + cy^2 + ey)\mu(dy) + f(\mu(S)).$$

The measure dynamics becomes

---

[7]This again requires quadratic fitness functions or else higher order terms may need to be examined in critical cases.

[8]Again, population size is assumed to be large enough that population densities may be used in place of numbers of individuals.

$$\frac{d\mu}{dt}(B) = \int_B \pi(x, \mu)\mu(dx). \tag{6}$$

This assumes each individual engages on average in one random pairwise interaction per unit time and reproduces offspring identical to itself (e.g. asexual reproduction) at a net rate equal to its expected fitness (5). If death rates are included in the model, they are assumed to be the same for all individuals and incorporated into the net rate through the background fitness function. Then (6) is the deterministic limit of the stochastic process as population size gets large.

The first question that arises is whether there are solutions to this measure dynamics. There are if μ has finite support[9] at time 0 (i.e. if $\mu = \sum_{i=1}^{r} n_i \delta_{x_i}$ where δ_{x_i} is the Dirac delta measure with support $\{x_i\}$ and $\delta_{x_i}(\{x_i\}) = 1$). Then $\mu(S) = \sum n_j = N$ and $\pi(x_i, \mu) = \frac{1}{N} \sum n_j \pi(x_i, x_j, N)$. The dynamics (6) is then the same as (4) in Section 2.2.

But we are more interested in the case where μ does not have finite (or discrete) support, perhaps given through a continuous density function. To show there are solutions to (6) in the general case, define the measure P as

$$P(B) = \mu(B)/\mu(S).$$

This is a probability measure (i.e. $P(S) = 1$) and we can rewrite individual fitness of strategy $x \in S$ as

$$\pi(x, P, \mu(S)) = \int_S (ax^2 + bxy + cy^2 + ey)P(dy) + f(\mu(S))$$

and population mean fitness as

$$\pi(P, P, \mu(S)) = \int_S \pi(x, P, \mu(S))P(dx).$$

A straightforward calculation using the quotient rule from calculus implies the measure dynamics for the probability space is

$$\frac{dP}{dt}(B) = \int_B \left(\pi(x, P, \mu(S)) - \pi(P, P, \mu(S)) \right) P(dx) \tag{7}$$

---

[9]The support $\text{supp}\,\mu$ of a measure μ is the closed set of those x for which every open neighbourhood of x has positive measure.

Since $\pi(x, P, \mu(S)) - \pi(P, P, \mu(S))$ does not depend on $\mu(S)$, we can ignore the background fitness $f(\mu(S))$ in (7) and take the fitness function to have the form $\pi(x, y, \mu(S)) = \pi(x, y)$.

Dynamics of the form (7) with $\pi(x, y) : S \times S \to R$ continuous have been shown (e.g. Bomze, 1991; Oechssler and Riedel, 2001)[10] to have solutions P_t for all $t \geq 0$ for any given initial condition where the derivative on the left-hand side is taken with respect to the variational norm.

Furthermore, evolution of the total population size satisfies $\frac{d\mu}{dt}(S) = (\int_S \pi(x, P_t, \mu(S)) P_t(dx)) \mu(S)$. This is a one-dimensional non-autonomous dynamics with continuous vector field and so has a unique solution for every initial condition. Also, since $\pi(x, P_t, 0) > 0$ and $\pi(x, P_t, \mu(S)) < 0$ for all P_t if $\mu(S)$ is sufficiently large, the solution is bounded. Moreover, if P_t evolves to P^*, then μ_t converges to $N^* P^*$ where N^* is the unique positive population size for which $\int_S \pi(x, P^*, N^*) P^*(dx) = 0$. That is, for convergence and stability of the measure dynamics (6), we can restrict attention to analyzing these same properties for (7) instead.

4 The Dynamics on the Space of Probability Measures

Our primary aim in the next section is the complete characterization of the convergence and stability properties of the probability dynamics (7) for all quadratic payoff functions and $S = [\alpha, \beta]$. However, many of our results that lead to this characterization in Section 5 are true for more general classes of payoff functions and other trait spaces S. These general results are collected in the present section. They rely on two Lyapunov functions, the relative (or cross) entropy and the mean payoff that are developed in Sections 4.1 and 4.2 respectively.

For the sake of concreteness, we assume S is a compact metric space and $\pi(x, y)$ is a continuous payoff function on $S \times S$. The *replicator equation* for a continuous trait space (Bomze, 1990, 1991; Oechssler and Riedel, 2001, 2002; Eshel and Sansone, 2003; Cressman, 2004) is then the measure dynamics

$$\frac{dP}{dt}(B) = \int_B (\pi(x, P) - \pi(P, P)) P(dx) \tag{8}$$

---

[10]In fact, they show that it is enough to assume that $\pi(x, y)$ is bounded and Borel measurable.

on the set of probability measures $\Delta(S)$ on the Borel $\sigma-$algebra. This again has a unique solution P_t for all initial $P_0 \in \Delta(S)$.

An important issue is the topology to be used on $\Delta(S)$. We feel the *weak topology* captures best the essence of convergence in coevolutionary systems (Oechssler and Riedel, 2002; Eshel and Sansone, 2003). One reason for this is that continuous population strategy distributions cannot be close to a monomorphism in the alternative strong topology, which is equivalent to the variational norm (Bomze, 1990, 1991; Oechssler and Riedel, 2001). Convergence of P_t to P in the weak topology (also called the weak* topology in functional analysis) means that $\int_S g(x)P_t(dx) \to \int_S g(x)P(dx)$ for all continuous functions $g : S \to \mathbf{R}$. Since S is a compact metric space, $\Delta(S)$ is compact and metrizable (Oechssler and Riedel, 2002). In particular, the set of limit points of the trajectory P_t is nonempty even if it does not converge.

The topology will mostly be applied to neighborhoods of monomorphic and dimorphic P^*. For a probability measure P^* with finite support $\{x_1, \ldots, x_m\}$, we can take a basis of $\varepsilon-$neighborhoods in the weak topology to be of the form

$$\{Q \in \Delta(S) : |Q(B_\varepsilon(x_i)) - P^*(\{x_i\})| < \varepsilon \quad \forall i = 1, \ldots, m\}$$

where $B_\varepsilon(x)$ is the open ball of radius ε centered at x. Thus, a continuous population strategy distribution is close to such a P^* if the proportion of the population near each x_i is approximately the proportion $P^*(\{x_i\})$ for P^*. Furthermore, unlike the strong topology where two monomorphisms are never close to each other, δ_{x_1} and δ_{x_2} are within ε of each other if and only if the Euclidean distance between these points is less than ε. In the following, all topological notions are taken for this weak topology, unless otherwise stated.

4.1 Local Superiority

When $S = \{x_1, ..., x_m\}$ is finite, the space of probability measures $\Delta(S)$ is the set $\{q = (q_1, ..., q_m) | \sum_{j=1}^m q_j = 1\}$ of probability vectors where q_i is the proportion of the population using the ith strategy. The probability measure dynamics is then the standard replicator game dynamics (Hofbauer and Sigmund, 1998) with $m \times m$ payoff matrix whose entries are $\pi(x_i, x_j)$. A standard way to prove the local asymptotic stability of a strategy p^* is by showing that it is a "matrix-ESS"[11] (i.e. a strategy which, if all individuals

---

[11] As mentioned earlier, the term ESS is overused in the literature and so may have several meanings for some readers. On the other hand, for games with a finite trait space, there is one universally accepted meaning originating with Maynard Smith (1982) as an evolutionarily stable strategy of the $m \times m$ payoff matrix A (hence a matrix-ESS).

adopt it, cannot be successfully invaded by any other rare mutant strategy) The proof uses the equivalent characterization of a matrix-ESS as a strategy p^* for which

$$\pi(p^*, q) > \pi(q, q) \tag{9}$$

for all q close to p^*. Here closeness is meant either on each ray connecting p^* with another strategy p, or simply in a Euclidean neighborhood.

Whereas these various versions of closeness are all equivalent in finite games, there are many different versions for games with an infinite trait space S (see also the Summary in Section 5.1). The weaker the topology (or more general nearness concept) on $\Delta(S)$, the stronger the corresponding version of "ESS". Although we feel strongly that the results of Section 4 are compelling reasons to claim that the generalization of (9) to infinite trait spaces with respect to the weak topology deserves the ESS designation, we have used the phrase "locally superior in the weak topology" instead for this concept in the following definition to avoid confusion. Also, the notion of local superiority (Weibull, 1995) is now well-established for the case of a finite trait space S as an alternative phrase to denote a matrix-ESS.[12]

Definition 1 $P^* \in \Delta(S)$ *is a* locally superior strategy *(in the weak topology) if, for all $Q \neq P^*$ sufficiently close to P^*,*

$$\pi(P^*, Q) > \pi(Q, Q) \tag{10}$$

We say P^* is globally superior *if this inequality is true for all $Q \neq P^*$.*

Our first main result given in the following theorem uses the concept of *cross entropy*,[13] as developed by Bomze (1991) for probability measure dynamics. Indeed, suppose P is absolutely continuous with respect to Q and its corresponding Radon-Nikodym derivative $\phi = \frac{dP}{dQ}$ is bounded (i.e., there is a $C > 0$ such that $P(A) \leq CQ(A)$ for all Borel sets $A \subset S$). Then the cross entropy

$$L(Q) := K_{Q:P} = \int_S \log \frac{dP}{dQ} P(dx) = \int_S \phi \log \phi \, Q(dx)$$

is defined, nonnegative and finite. Furthermore, Lemma 2 in Bomze (1991) shows that $L(Q_t)$ is defined along the orbit of Q, and its time derivative satisfies

$$\frac{d}{dt} L(Q_t) = -\pi(P, Q_t) + \pi(Q_t, Q_t) \tag{11}$$

---

[12]In Oechssler and Riedel (2002), locally superior with respect to the weak topology is called "evolutionarily robust".

[13]In the finite case this cross entropy corresponds to the function $L(q) = \sum_i p_i^* \log \frac{p_i^*}{q_i}$ which like $\prod_i q_i^{p_i}$ is the well-known Lyapunov function near a matrix-ESS p^*.

13

for all $t \geq 0$. In particular, the cross entropy is decreasing if P is locally superior and Q_t is sufficiently close to P, a key fact in the proof of the following theorem.

Theorem 2 *If P^* is a locally superior strategy, which is Lyapunov stable, then for any initial Q sufficiently close to P^* with $\operatorname{supp} Q \supseteq \operatorname{supp} P^*$, $Q_t \to P^*$ as $t \to +\infty$. Moreover, if P^* is globally superior and Lyapunov stable then for any initial $Q \in \Delta(S)$ with $\operatorname{supp} Q \supseteq \operatorname{supp} P^*$, $Q_t \to P^*$ as $t \to +\infty$.*

Proof. Let U_1 be a compact neighborhood of P^* such that $\pi(P^*, P) - \pi(P, P) > 0$ holds for all $P \in U_1 \setminus \{P^*\}$. Since P^* is Lyapunov stable there is a neighborhood U_2 of P^* such that for all $Q \in U_2$ and $t \geq 0$ we have $Q_t \in U_1$.

Suppose now that P^* is not an $\omega-$limit point of such a Q. Then there is an open neighborhood U_3 of P^* with $Q_t \notin U_3$ for all $t \geq 0$. By compactness $\pi(P^*, P) - \pi(P, P) \geq 2c > 0$ for some $c > 0$ and all $P \in U_1 \setminus U_3$. By continuity, for all $\tilde{P}$ close enough to P^* (in the weak topology) we have

$$\pi(\tilde{P}, P) - \pi(P, P) \geq c > 0 \quad \forall P \in U_1 \setminus U_3. \tag{12}$$

Since $\operatorname{supp}(P^*) \subseteq \operatorname{supp}(Q)$, there is such a $\tilde{P}$ which is absolutely continuous with respect to Q and whose Radon-Nikodym derivative $\frac{d\tilde{P}}{dQ}$ is bounded.[14] By (11) and (12), the cross entropy $L(Q) := K_{Q:\tilde{P}}$ is defined and satisfies

$$\frac{d}{dt} L(Q_t) \leq -c < 0$$

along the solution Q_t for $t \geq 0$. Hence $L(Q_t) \to -\infty$, a contradiction to $L(Q) \geq 0$.

This shows that P^* is an $\omega-$limit point of Q. Since P^* is Lyapunov stable, it is the unique $\omega-$limit point of Q and hence $Q_t \to P^*$. Finally, if P^* is globally superior then take $U_1 = U_2 = \Delta(S)$. ∎

This result generalizes Theorem 3 of Oechssler and Riedel (2002) who proved it for monomorphisms $P^* = \delta_{x^*}$ and initial Q with $Q(\{x^*\}) > 0$ in place of our weaker assumption $x^* \in \operatorname{supp} Q$. It is an open problem whether the additional assumption of Lyapunov stability is really needed. When the trait space is finite, Lyapunov stability follows from local superiority

---

[14]Such a $\tilde{P}$ exists since the weak closure of the set of probability measures that are absolutely continuous with respect to Q and have bounded Radon-Nikodym derivative is the set of all probability measures whose support is contained in $\operatorname{supp} Q$. For example, for $s \in \operatorname{supp} Q$, and U_n the $\frac{1}{n}-$neighborhood of s, the uniform probability measures on $U_n \cap \operatorname{supp} Q$ converge to δ_s.

since the cross entropy $L(Q) = K_{Q:P^*} = \sum_i p_i^* \log \frac{p_i^*}{q_i}$ is then continuous in a neighborhood of P^*.[15]

It is essential $\operatorname{supp} Q \supseteq \operatorname{supp} P^*$ for the conclusions given in Theorem 2 to be valid. This is due to the fact that the replicator equation (8) shares the same property as (4) in that its support is invariant for all $t \geq 0$.[16] It is also important to recognize that local asymptotic stability of $(N(x^*), 0, ..., 0)$ in the dynamics (4) for any finite choice of strategies $\{x^*, x_2, ..., x_r\}$ does not imply the corresponding discrete measure converges weakly to $N(x^*)\delta_{x^*}$. At a minimum, weak convergence requires the asymptotic stability of (4) to be global for all finite choices of strategies that are sufficiently close to x^*.

Remark 1. The following observations are useful to identify locally/globally superior strategies. First, every locally superior strategy P^* is a Nash equilibrium (NE) (i.e. $\pi(P^*, P^*) \geq \pi(Q, P^*)$ for all $Q \in \Delta(S)$). Indeed, given $Q \in \Delta(S)$, for all nonnegative ε sufficiently close to zero

$$
\begin{aligned}
0 &\leq \pi(P^*, P^* + \varepsilon(Q - P^*)) - \pi(P^* + \varepsilon(Q - P^*), P^* + \varepsilon(Q - P^*)) \\
&= \varepsilon\pi(P^* - Q, P^* + \varepsilon(Q - P^*))
\end{aligned}
$$

Thus, for $\varepsilon \to 0$, we get $0 \leq \pi(P^* - Q, P^*)$.

Second, if the game is negative definite (i.e. $\pi(P - Q, P - Q) < 0$ for all $Q \neq P$) then there exists a globally superior strategy. To see this, let P^* be any NE. Then, for all $Q \neq P^*$

$$
\begin{aligned}
\pi(P^*, Q) - \pi(Q, Q) &= \pi(P^* - Q, P^*) + \pi(P^* - Q, Q - P^*) \\
&\geq \pi(P^* - Q, Q - P^*) > 0.
\end{aligned}
$$

Conversely, if P^* is locally superior (on each ray connecting P^* with another strategy Q) with full support then the game is negative definite.[17]

4.2 Potential Games

Consider now a symmetric payoff function $\pi : S \to S$, i.e. $\pi(x, y) = \pi(y, x)$, that is assumed to be continuous. Note that, for quadratic fitness func-

---

[15]Unfortunately, $L(Q)$ taken in the proof of Theorem 2 is not a continuous function in the weak topology for a continuous strategy space S and so Lyapunov stability is no longer automatic.

[16]In fact, Q_0 and Q_t are mutually absolutely continuous measures, as shown by Bomze (1991).

[17]If P^* is globally superior but does not have full support then the game is not necessarily negative definite, as already games with two strategies show. A game is negative definite if and only if the mean payoff function $P \mapsto \pi(P, P)$ is strictly concave on Δ. Our quadratic games are negative semi-definite if and only if $b \leq 0$.

tions and $S = [\alpha, \beta]$, the dynamics (7) is unchanged if we take the symmetric version $\pi(x, y) = ax^2 + bxy + ay^2$ as our payoff function. By common game-theoretic usage, games with symmetric payoff matrices are known as "potential" games.

By the symmetry of π, the expected payoff satisfies

$$\frac{d}{dt}\pi(P, P) = 2 \int_S \int_S \pi(x, y)[\pi(x, P) - \pi(P, P)]P(dx)P(dy)$$

$$= 2 \int_S (\pi(x, P) - \pi(P, P))^2 P(dx) \geq 0. \tag{13}$$

By the continuity of $\pi(x, P)$ in x, $\frac{d}{dt}\pi(P, P) = 0$ if and only if $\pi(x, P) = \pi(P, P)$ for all x in the support of P if and only if P is a rest point of (8).[18] Thus $\pi(P, P)$ is a strict Lyapunov function on $\Delta(S)$ in that it is strictly increasing under (8) unless at equilibrium.

Since $\Delta(S)$ is compact in the weak topology and $\pi(P, P)$ is a continuous function, P^* will be Lyapunov stable if it is an isolated local maximizer of $\pi(P, P)$ with respect to the weak topology. For finite games, a strategy P^* is a local maximizer of $\pi(P, P)$ if and only if it is locally superior, see e.g. Hofbauer and Sigmund (1998). In general only the direction proved in the following lemma is true, as shown by the counterexample given in the Summary of Section 5.1.

Lemma 3 *If P^* is locally superior then it is an isolated local maximizer of the mean fitness function $\pi(P, P)$.*

Proof. Since P^* is a Nash equilibrium, $\pi(P^*, P^*) \geq \pi(Q, P^*) = \pi(P^*, Q) > \pi(Q, Q)$ for all Q sufficiently close to P^*. ∎

Combining the above discussion with Theorem 2 we get

Theorem 4 *If P^* is a locally superior strategy (with respect to the weak topology) in a potential game, then P^* is Lyapunov stable and for any initial Q sufficiently close to P^* with $\operatorname{supp} Q \supseteq \operatorname{supp} P^*$, $Q_t \to P^*$. If P^* is globally superior with $\operatorname{supp} Q \supseteq \operatorname{supp} P^*$, then $Q_t \to P^*$.*

This theorem was proved by Bomze (1990) in the special case where $\pi(x, y)$ depends only on x and by Oechssler and Riedel (2002) when $\pi(x, y)$ is symmetric and P^* is a monomorphism.

---

[18]This is the extension to continuous strategy spaces of one part of the Fundamental Theorem of Natural Selection that states mean fitness increases unless at equilibrium. If π is not symmetric, then $\frac{d}{dt}\pi(P, P) = \int_S \int_S [\pi(x, y) + \pi(y, x)] [\pi(x, P) - \pi(P, P)]P(dx)P(dy)$ may be negative.

5 Dynamic Stability for Quadratic Payoff Functions

Let us apply the general theory above to our quadratic payoff function $\pi(x, y) = ax^2 + bxy + ay^2$ where the interesting monomorphism is $x^* = 0$ and we assume $a \neq 0$. For this, the following formulas for $\pi(P, Q)$ etc. in terms of the mean $E(P)$ and the variance $Var(P)$ of a probability measure P are useful. It is even convenient to consider higher order moments: let $P_k = \int x^k P(dx)$ be the kth moment of P. Then $P_1 = E(P)$ and $P_2 = Var(P) + P_1^2$. We get $\pi(x, P) = ax^2 + bxP_1 + aP_2$,

$$\pi(P, Q) = aP_2 + bP_1Q_1 + aQ_2 =$$
$$= a(Var(P) + E(P)^2 + Var(Q) + E(Q)^2) + bE(P)E(Q)). \qquad (14)$$
$$\pi(P, P) = 2aVar(P) + (2a + b)E(P)^2, \qquad (15)$$
$$\pi(x, P) - \pi(P, P) = a(x^2 - P_2) + b(xP_1 - P_1^2), \qquad (16)$$
$$\pi(P - Q, P - Q) = b(E(P) - E(Q))^2. \qquad (17)$$

Our classification of the stability of δ_0 in Sections 5.1 and 5.2 is based first on whether a is negative or positive and then on subclasses depending on the value of b. This classification scheme is similar to that given by Metz et al. (1996) (see also Geritz et al., 1997; Diekmann, 2004) for the adaptive dynamics approach.

One reason for using this classification scheme is that the subspace of probability measures that are symmetric about 0 is invariant for our quadratic payoff functions, and on this subspace, the variance is increasing if $a > 0$ and decreasing if $a < 0$. To see this invariance, note that symmetry implies $P_1 = 0$ and so $\frac{dP}{dt}(B) = \frac{dP}{dt}(-B)$ from (16) and (8).

Furthermore, from (16) and

$$\dot{P}_k = \int x^k d\dot{P} = \int x^k \left[\pi(x, P) - \pi(P, P)\right] P(dx),$$

the differential equations for the moments are

$$\dot{P}_1 \quad = \quad aP_3 + (b - a)P_1P_2 - bP_1^3 \qquad (18)$$
$$\dot{P}_2 \quad = \quad aP_4 - aP_2^2 + bP_1P_3 - bP_1^2P_2 \qquad (19)$$
$$\dot{P}_3 \quad = \quad aP_5 - aP_2P_3 + bP_1P_4 - bP_1^2P_3 \qquad (20)$$
$$\cdots$$

Obviously, if the initial P is symmetric around 0, then so is P_t, hence the odd moments vanish and the variance satisfies $\dot{P}_2 = a(P_4 - P_2^2)$. Since $P_4 \geq P_2^2$ (with equality for point measures), variance increases if $a > 0$ and decreases if $a < 0$. In particular, this shows instability of δ_0 for $a > 0$.

5.1 Case 1: $a < 0$

This is the case where $x^* = 0$ is a strict NE, i.e. $\pi(0,0) > \pi(x,0)$ for all $x \neq 0$.

5.1.1 Case 1a: $a + b < 0$.

From (15), we see that $\pi(P,P) \leq 0$ with equality if and only if $P = P^* = \delta_0$. Thus P^* is the unique global maximizer of the mean fitness function (and there are no other local maximizers). Hence $P^* = \delta_0$ is Lyapunov stable by section 4.2. Furthermore, from (16), if $Q \neq P^*$, then

$$\pi(\delta_0, Q) - \pi(Q,Q) = -aVar(Q) - (a+b)E(Q)^2 > 0.$$

Thus $P^* = \delta_0$ is globally superior and by Theorem 2 it attracts all initial Q_0 that have $0 \in \operatorname{supp} Q_0$.

5.1.2 Case 1b: $2a + b < 0 \leq a + b$.

We still have $P^* = \delta_0$ as the unique global maximizer of mean fitness and so Lyapunov stable but it is no longer locally superior. However, the following theorem that uses an iterated domination argument between pure strategies shows P^* still attracts all initial Q_0 with full support.

Theorem 5 *Suppose $a < 0$ and $2a+b < 0 \leq a+b$. If the support of Q_0 is an interval that contains $x^* = 0$, then Q_t converges to δ_0 (in the weak topology).*

   Proof. Without loss of generality, assume $\operatorname{supp} Q \supseteq [0, \beta]$. Let $x_0 = -\frac{b\beta}{2a}$. Then $0 < x_0 < \beta$. Take $A = [x_0 + 3\varepsilon, \beta]$ and $B = [x_0 + \varepsilon, x_0 + 2\varepsilon]$ where $x_0 + 3\varepsilon < \beta$ and ε is positive.
   Then, another application of the quotient rule yields

$$\frac{d}{dt}\left(\frac{Q(B)}{Q(A)}\right) = \frac{1}{Q(A)^2} \int_S \left[\int_A \int_B \left(\pi(x,z) - \pi(y,z)\right) Q(dx)Q(dy) \right] Q(dz).$$
(21)

For $x \in B, y \in A, z \in S$, we have

$$\pi(x,z) - \pi(y,z) = ax^2 + bxz - ay^2 - byz = (x-y)[a(x+y) + bz] > -2a\varepsilon^2 > 0.$$
(22)

Thus, $\frac{d}{dt}(\frac{Q(B)}{Q(A)}) \geq -\frac{Q(B)}{Q(A)}2a\varepsilon^2 > 0$. Thus $\frac{d}{dt}(\frac{Q(B)}{Q(A)})$ grows to infinity with exponential order. In particular, $\lim Q_t(A)$ is 0 as t approaches infinity. Now

choose ε so that $x_0 + 3\varepsilon = (x_0 + \beta)/2$ (i.e. $\varepsilon = (\beta - x_0)/6$). Then $\lim Q_t([(x_0 + \beta)/2, \beta]$ is 0. Now suppose we have shown $\lim Q_t[\gamma, \beta] = 0$ for some $0 < \gamma < \beta$. We will iterate the above argument to show $\lim Q_t[\frac{x_0+\beta}{2}\gamma, \gamma] = 0$. Take $A = [-\beta\gamma/2a + 3\varepsilon, \gamma]$ and $B = [-\beta\gamma/2a + \varepsilon, -\beta\gamma/2a + 2\varepsilon]$ where $-\beta\gamma/2a + 3\varepsilon < \gamma$. Then

$$
\frac{d}{dt}\left(\frac{Q(B)}{Q(A)}\right) = \frac{1}{Q(A)^2}\int_{[\gamma,\beta]}[\int_A\int_B(\pi(x,z) - \pi(y,z))\,Q(dx)Q(dy)]Q(dz)
$$

$$
+ \quad \frac{1}{Q(A)^2}\int_{S-[\gamma,\beta]}[\int_A\int_B(\pi(x,z) - \pi(y,z))\,Q(dx)Q(dy)]Q(dz)
$$

$$
\geq \quad -(Q([\gamma,\beta]2K + Q(S - [\gamma,\beta])2a\varepsilon^2)\frac{Q(B)}{Q(A)}
$$

$$
\geq \quad -a\varepsilon^2\frac{Q(B)}{Q(A)}
$$

for t sufficiently large. Here $K = \max_{x,y\in S}|\pi(x,y)|$. In particular, $\lim Q_t(A) = 0$. Thus, by iteration, we can take $\gamma = [(x_0 + \beta)/2]^n\beta$ for all $n \in Z$ and so $\lim Q_t([\gamma, \beta]) = 0$ for all $\gamma > 0$.

A similar argument on the interval $[\alpha, \gamma]$ with $\gamma < 0$ completes the proof. ∎

In game-theoretic terms, inequality (22) asserts that every $x \in B$ strictly dominates every $y \in A$. The proof is then essentially the iterated elimination of strictly dominated pure strategies. This technique is well-known for games with finite trait space (e.g. Samuelson and Zhang, 1992; Hofbauer and Weibull, 1996) but this seems to be one of the first instances where it is used in games with a continuum of pure strategies.[19]

The method of proof can extend the statement of the Theorem to measures that do not have full support as long as the "gap" between points in the support of Q_0 is not too great. This gap must decrease as we get closer to x^*. In particular, if one wants to approximate the measure dynamics with a discrete version similar to (4), then one needs the grid to become finer as we approach x^*. Otherwise, say if the grid is uniform, the most we can expect is that the support of Q_t will approach an interval containing x^* and that this interval will approach x^* as the number of points in the grid increases.

---

[19]Since the original version of this paper was submitted for publication, Frank Riedel (personal communication) pointed out Heifetz et al. (2003) used this method for their analysis of the evolution of preferences based on the replicator equation.

5.1.3 Case 1c: $2a + b = 0$.

$\pi(P, P) = 0$ for all $P = \delta_x$ with $x \in [\alpha, \beta]$. In this degenerate case the payoff function $\pi(x, y) = a(x - y)^2$ is translation invariant. Every δ_s is a strict NE and maximizer of $\pi(P, P)$.

5.1.4 Case 1d: $2a + b > 0$.

Here $P = \delta_0$ is a saddle point, and $P = \delta_\alpha$ and $P = \delta_\beta$ are the only local maximizers of $\pi(P, P)$. These endpoints are also locally superior with respect to those Q whose support is either $[\alpha, -\varepsilon]$ or $[\varepsilon, \beta]$ respectively. This gives us a bistable situation where some initial Q_0 close to δ_0 evolve to one monomorphism supported at one endpoint and some to the other. In fact, by continuity of $\pi(Q, Q)$, δ_β attracts those Q_0 with full support that have $Q_0([\alpha, 0])$ sufficiently small.

Summary. The four subcases of this section clarify the relevance of the CSS concept and the importance of the topology chosen for $\Delta(S)$.

First, Cases 1a and 1b combine to show that a CSS $x^* = 0$ in the interior of $[\alpha, \beta]$ (i.e. $a < 0$ and $2a + b < 0$) is Lyapunov stable and every initial Q with full support converges to δ_0 in the weak topology. Moreover, it is already clear from (15) that δ_0 is unstable if $2a + b > 0$. These results give a strong measure theoretic justification of the CSS concept that lies at the heart of adaptive dynamics.

It must be pointed out, however, that there is a significant difference between the basins of attraction of δ_0 that are CSS depending on the sign of $a+b$. If $a+b < 0$, δ_0 is known as a *good invader* (Kisdi and Meszéna, 1995) or a *neighborhood invader strategy* (NIS) (McKelvey and Apaloo, 1995; Apaloo, 1997). This latter condition can be used to prove convergence in Case 1a with a single domination argument (Cressman, 2004) (see also Eshel and Sansone, 2003) that avoids the entropy technique used in the proof of Theorem 2. To illustrate this difference, suppose Q_0 is a dimorphism with support $\{0, s\}$, with $0 < s \leq \beta$. Then $a + b < 0$ implies Q_t converges to δ_0 in the weak topology.

On the other hand, if $a+b > 0$, the dynamics (8) restricted to the support $\{0, s\}$ is bistable: with $q = Q_t(\{0\})$ and $1 - q = Q_t(\{s\})$ we get

$$\begin{aligned}
\dot{q} &= q(1 - q)[q(\pi(0, 0) - \pi(s, 0)) + (1 - q)(\pi(0, s) - \pi(s, s))] \\
&= q(1 - q)s^2[bq - (a + b)] \qquad (23)
\end{aligned}$$

Hence $\dot{q} < 0$ if $0 < q < (a + b)/b$. (Note that $0 < (a + b)/b < 1/2$ which means that δ_0 has the larger basin of attraction on this line than δ_s.) Thus,

$q \to 0$ if q is sufficiently small initially. This result also follows from (15) since $\pi(q\delta_0 + (1-q)\delta_s, q\delta_0 + (1-q)\delta_s) = s^2(bq^2 - 2(a+b)q + 2a+b)$ which is a quadratic function of q with minimum at $q = (a+b)/b$. Since $\pi(Q, Q)$ is increasing, $Q_t \to \delta_s$ if $0 < Q_0(\{0\}) < (a+b)/b$. Furthermore, by continuity of $\pi(Q, Q)$ in the weak topology, for ε sufficiently small, if $\operatorname{supp} Q_0 = [-\varepsilon, \varepsilon] \cup [s-\varepsilon, s+\varepsilon]$ and $Q_0([s-\varepsilon, s+\varepsilon]) < (a+b)/b$, then $Q_t([s-\varepsilon, s+\varepsilon]) \to 1$ as $t \to \infty$ (actually, Q_t converges weakly to $\delta_{s-\varepsilon}$ by the argument in the proof of Theorem 5).

The above analysis also shows that the convergence results of Theorem 4 need not be true if we only assume P^* is the unique global maximizer of the expected payoff $\pi(P, P)$.

Cases 1c and 1d illustrate the importance of the chosen topology for convergence and stability results. Specifically, δ_0 is locally superior with respect to the variational norm[20] if and only if $a < 0$. That is, even in these last two cases when mean fitness at δ_0 is less than that of any other monomorphism, δ_0 is locally superior in the strong topology (in fact, every δ_s in Case 1c is locally superior). Thus, δ_0 is Lyapunov stable in this strong topology, and further Q_t converges weakly to δ_0, if $Q_0(\{0\})$ is close to 1 as shown by Oechssler and Riedel (2001) (see also Bomze, 1990).

5.2   Case 2: $a > 0$

From (15), any local maximizer of $\pi(P, P)$ must have as large a variance as possible given $E(P^*) =: E$. Thus the support of P^* is contained in $\{\alpha, \beta\}$. In fact, $P_E^* = p_1^*\delta_\beta + p_2^*\delta_\alpha$ where $p_1^* = \frac{E-\alpha}{\beta-\alpha}$ and $p_2^* = \frac{\beta-E}{\beta-\alpha}$. Thus, we need to maximize

$$\operatorname{Var}(P_E^*) = 2a(p_1^*(E-\beta)^2 + p_2^*(E-\alpha)^2) + (2a+b)E^2$$

for $E \in [\alpha, \beta]$. This expression simplifies to

$$\operatorname{Var}(P_E^*) = -2a(E-\alpha)(E-\beta) + (2a+b)E^2 = bE^2 + 2a(\alpha+\beta)E - 2a\alpha\beta.$$

If $b < 0$, this is a downwards parabola that has a unique maximum at $E^* = -\frac{a}{b}(\alpha+\beta)$. Depending on whether E^* is outside or inside the interval S, the unique local maximizer is given by $P^* = \delta_\beta$ if $(a+b)\beta + a\alpha \geq 0$, by $P^* = \delta_\alpha$ if $(a+b)\alpha + a\beta \leq 0$ and by the above dimorphism if $(a+b)\beta + a\alpha < 0$ and $(a+b)\alpha + a\beta > 0$. Since P^* is a NE and the game is negative semidefinite

---

[20]This norm corresponds to the strong topology with respect to which local superiority is often called "strongly uninvadable" (e.g. Bomze, 1991).

by (17),

$$\begin{aligned}
\pi(P^*, Q) - \pi(Q, Q) &= \pi(P^* - Q, P^*) + \pi(P^* - Q, Q - P^*) \\
&\geq \pi(P^* - Q, Q - P^*) = -b(E(P^*) - E(Q))^2 \geq 0.
\end{aligned}$$

Furthermore, if $E(P^*) = E(Q)$, then $\pi(P^*, Q) - \pi(Q, Q) = a(Var(P^*) - Var(Q))$ from (14) and (15). Thus, $\pi(P^*, Q) - \pi(Q, Q) > 0$ unless P^* and Q have the same mean and variance. Since P^* is the unique probability measure that has the largest variance for a given mean, P^* is globally superior. By Theorem 4, P^* is Lyapunov stable and attracts every Q_0 whose support contains that of P^*.

5.2.1 Case 2a: $2a + b < 0$.

In this case we have the situation which is often referred to in the adaptive dynamics literature as *evolutionary branching* (Geritz et al, 1997). It is straightforward to show that P^* is then a dimorphism. By the above reasoning for $b < 0$, P^* is Lyapunov stable and attracts any initial $Q(0)$ whose support includes $\{\alpha, \beta\}$ since it is globally superior.

5.2.2 Case 2b: $2a + b \geq 0$ and $b < 0$.

P^* is still globally superior but could be either a dimorphism or one of the monomorphisms δ_β or δ_α.

5.2.3 Case 2c: $b > 0$.

In this final case, $f(E)$ is an upwards parabola which has δ_β and δ_α as the local maximizers of $\pi(P, P)$. This is again the bistable situation as in Case 1d.

5.3 Normal Distributions with $S = \mathbf{R}$

Following Oechssler and Riedel (2002) we consider special solutions Q_t of (8) with $S = \mathbf{R}$ that are normal distributions with mean $m(t)$ and variance $V(t)$, i.e.,

$$\frac{dQ_t}{dx} = \frac{1}{\sqrt{2\pi V(t)}} \, e^{-\frac{(x - m(t))^2}{2V(t)}} \tag{24}$$

From $\frac{d\dot{P}}{dP}(x) = \pi(x, P) - \pi(P, P)$ we obtain

$$-\frac{d}{dt}\left[\frac{(x - m(t))^2}{2V(t)} + \frac{1}{2}\log V(t)\right] = a(x^2 - m(t)^2) + bm(t)(x - m(t)) - aV(t)$$

22

which reduces to

$$\dot{V} = 2aV^2 \qquad \dot{m} = (2a + b)mV \qquad (25)$$

Hence (24) are solutions of (8) if and only if the mean and variance satisfy the differential equations (25).

This illustrates nicely the meaning of the two crucial parameters a and $2a+b$: the equation for the mean resembles the canonical equation of adaptive dynamics. Note that, for $a > 0$, the variance goes to infinity in finite time. Such a blow up is possible only for non–compact S.

6   Non-Quadratic Payoff Functions

The method developed in Section 4 and the proof of Theorem 5 can be used to analyze probability measure dynamics beyond the quadratic payoff functions considered in Section 5. For instance, the domination argument used in the proof of Theorem 5 remains valid if we restrict the support of Q_0 to be an interval sufficiently close to x^* so that the quadratic terms in the Taylor expansion of $\pi(x, y)$ about (x^*, x^*) are predominant. In particular, a CSS that is not an NIS continues to attract all such initial Q_0. Moreover, the one-time domination argument of Cressman (2004) (see the Summary in Section 5.1) shows the same result for arbitrary support sufficiently close to x^* when x^* is CSS and NIS. Thus, the CSS condition will continue to guarantee local convergence in the absence of the symmetry implied by only quadratic terms in the payoff function - what is lost is whether the CSS remains Lyapunov stable. Conversely, if x^* is not CSS then a reverse domination argument shows instability of δ_{x^*}.

In the other case of particular interest to adaptive dynamics (i.e. the evolutionary branching of Case 2a), a convergence stable x^* with $\pi_{xx}(x^*, x^*) > 0$ will be unstable in the measure dynamics whether $\pi(x, y)$ is quadratic or not. What is not so clear is what the measure dynamics will evolve to in this situation without the quadratic payoffs that imply a globally stable dimorphism emerges that is supported on the endpoints of the trait space S since mean fitness grows as variance increases. However, our results apply to nonquadratic payoff functions of the form $\pi(x, y) = \phi(x) + bxy + \phi(y)$ with $b < 0$ and ϕ an arbitrary smooth function. Since such games are negative semidefinite, if there is a unique maximizer P^* of the mean fitness then P^* is globally superior and hence attracts all initial Q with full support. If ϕ is a symmetric function (i.e. $\phi(-x) = \phi(x)$) and $\phi_{xx}(0) > 0$ then there is no monomorphic NE, and hence P^* is supported on at least two traits. As an example take $b = -5$ and $\phi(x) = 2x^2 - x^4$ on the trait space $S = \mathbf{R}$. If

23

supp Q_0 contains both maxima of ϕ (i.e. $\{-1, 1\}$), the theory from Section 4 proves Q_t converges to the Lyapunov stable dimorphism $P^* = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$ in the weak topology. This follows from the fact P^* is globally superior since

$$
\begin{aligned}
&\pi(P^*, Q) - \pi(Q, Q) \\
=\ &\int_S \int_S (\phi(x) - 5xy + \phi(y))P^*(dx)Q(dy) - \int_S \int_S (\phi(x) - 5xy + \phi(y))Q(dx)Q(dy) \\
=\ &\frac{1}{2}(\phi(1) + \phi(-1)) - \int_S \phi(y)Q(dy) + 5E(Q)^2 \\
>\ &0
\end{aligned}
$$

unless supp $Q = \{-1, 1\}$ and $E(Q) = 0$ (i.e. unless $Q = P^*$.)

7   Discussion

This paper is meant to introduce measure dynamics as a means to model coevolutionary systems and to compare this theory to other, more established, approaches such as adaptive dynamics. The comparison is most complete in the basic model of coevolution where stability of a monomorphism in a single species with a one-dimensional continuous trait space is analyzed. This necessarily entailed several simplifying assumptions that we would like to address in this concluding section.

We have already discussed to some extent the issue of symmetric payoff functions in Section 6. There is no doubt symmetry is an important tool in our development of the theory for a general setting in Section 4 (see especially Lemma 3 and its proof that shows a locally superior strategy is Lyapunov stable). On the other hand, the quadratic approximation to fitness functions about a monomorphism x^* provide this symmetry and can be used to describe local behavior of the dynamics of probability measures with support near x^*.

Dependence on total population size (i.e. density dependence) is included in our model by assuming a background fitness that is strategy independent. This has the effect that density dependence essentially disappears from the measure space dynamics of Section 3 (as well as the adaptive models summarized in Section 2). General density dependence is more difficult from a technical perspective. One attempt to avoid this problem is to assume, as is often done in the adaptive dynamics approach, that the population instantaneously tracks its equilibrium density for a given probability measure. This approach is usually justified by appealing to a dichotomy between the time scales for the population dynamics (ecological time scale) versus strategy evolution (evolutionary time scale). That is, it is assumed ecological changes are

much faster than evolutionary ones. The dynamics is then reduced to one on the "stationary density surface" (SDS) where the fitness $\pi(x, P)$ is assumed to be given by $\pi(x, P, N(P))$ where $N(P)$ is on the stationary density surface. We feel such an assumption places a severe restriction on the applicability of adaptive dynamics, especially as the theory considers non monomorphic limiting behavior. On the other hand, there is some evidence the SDS continues to play an important role when there is no separation of time scales, at least when there is a finite number of strategies as in Section 2.2. The SDS can then be used to separate the density effect from the local asymptotic stability analysis of the mean strategy dynamics for both monomorphisms and polymorphisms in single species (Cressman, 1992) or multiple species (Cressman and Garay, 2003) models of coevolution. Of particular relevance for us is the result for monomorphisms; namely, a monomorphism is locally asymptotically stable in the coevolutionary model of combined density and strategy evolution when there is a finite number of strategies if and only if the monomorphic density dynamics is asymptotically stable at equilibrium strategy and the induced strategy dynamics on the stationary density surface is asymptotically stable (no matter what the relative rates of ecology and evolution are). For our measure dynamics model with general density dependent payoffs, we conjecture a stationary density surface will continue to predict local behavior near a monomorphism but not for general equilibria P^* distributed over the trait space (see also our comments on higher dimensional trait space below).

Mutation has not been explicitly added to our model. Our perspective is that this effect is already included by considering arbitrary initial population distributions over the strategy space. In this sense, all potential mutations are already present in the system to start with and, if this "mutation" is successful, it will grow in relative size compared to other possible strategies. In particular, our formalism includes systems that are initially concentrated near a monomorphic equilibrium with a small subpopulation concentrated around a "mutant" strategy. There is no doubt this modelling of "mutation" is different than the mutation and trait substitution or the finite number of fixed mutants of Sections 2.1 and 2.2 respectively. On the other hand, our results also allow one to introduce rare mutations in the more traditional sense by letting the system evolve arbitrarily close to a stable equilibrium between mutation events. Theorems 2 and 4 then show the perturbed system will stay close to a locally superior equilibrium distribution P^* that is Lyapunov stable and eventually return to it.

The adaptive dynamics approach has also been extended to multi-dimensional strategy spaces (e.g. Meszéna et al., 2001) and/or multi-species models (e.g. Marrow et al., 1996). For a single-species monomorphic equilibrium x^*

25

in a compact trait space in $\mathbf{R}^n$ (such as a ball or hypercube), the measure dynamics approach applied here suggests the method of assuming the CSS conditions hold along any ray (as in Meszéna et al., 2001) from x^* is particularly relevant. Indeed, it is straightforward to see that the CSS conditions along any ray are necessary for stability of δ_0 (simply take initial measures supported on this ray). On the other hand, under the additional condition corresponding to NIS, convergence to x^* can be shown in the weak topology when initial distributions have support sufficiently close to x^* (this uses an argument similar to Cressman (2004)). A precise characterization of measure-theoretic stability is an open problem.

However, in other multi-dimensional or multi-species situations (e.g. non-monomorphic P^* or especially if there is non-equilibrium behavior of the adaptive dynamics such as cyclic behavior in a single-species model or in a two-species predator-prey system), we are quite skeptical that the conclusions from the measure dynamics approach will correspond to that from adaptive dynamics. There is in fact evidence from coevolutionary models based on finite trait spaces (e.g. Abrams and Matsuda, 1997) that questions the relevance of those adaptive dynamics approaches that assume population size(s) tracks its equilibrium value in these circumstances (Marrow et al, 1996; Dieckmann and Law, 1996). It is hoped that future results from the measure dynamics model will contribute to this ongoing debate. Models that exhibit non-equilibrium behavior also call into question the assumption that fitness functions are based on pairwise interactions. It seems more reasonable that in such situations, $\pi(x, \mu)$ should depend on the measure in a more complicated way than simply averaging the payoffs between individual interactions. In fact, such non-pairwise fitness functions were proposed from the outset of coevolutionary models (e.g. Roughgarden, 1979) where Lotka-Volterra type models were proposed with interaction coefficients given by Gaussian distributions depending on the separation of strategies from the mean.

We view this paper as a first, but crucial, step to compare the infinite-dimensional measure dynamics approach to modelling coevolutionary systems with more traditional finite-dimensional dynamical methods, thereby providing a solid theoretical foundation to predict long term behavior in coevolutionary systems. From this perspective, our results (see especially the Summary in Section 5.1 that the CSS plays a central role in understanding convergence and stability of the measure dynamics at a monomorphism) show the adaptive method is clearly a valid shortcut to analyzing measure-theoretic coevolutionary models when there are quadratic pairwise interactions. As discussed above, there are many obstacles to a general theory, but the end results should be equally rewarding.

References

[1] Abrams, P. (2001) Modelling the adaptive dynamics of traits involved in inter- and intraspecific interactions: an assessment of three methods. *Ecol. Lett.* **4**, 166-175.

[2] Abrams, P. and H. Matsuda (1997) Fitness minimization and dynamic instability as a consequence of predator-prey coevolution. *Evol. Ecol.* **11**, 1-20.

[3] Apaloo, J. (1997) Revisiting strategic models of evolution: The concept of neighborhood invader strategies, *Theoret. Pop. Biol.* **52**, 71-77.

[4] Bomze, I. (1990) Dynamical aspects of evolutionary stability. *Monatsh. Math.* **110**, 189-206.

[5] Bomze, I. (1991) Cross entropy minimization in uninvadable states of complex populations, *J. Math. Biol.* **30**, 73-87.

[6] Christiansen, F.B. (1991) On conditions for evolutionary stability for a continuously varying character. *Amer. Nat.* **138**, 37-50.

[7] Cohen, Y., T.L. Vincent and J.S. Brown (1999) A G-function approach to fitness minima, fitness maxima, evolutionarily stable strategies and adaptive landscapes. *Evol. Ecol. Res.* **1**, 923-942.

[8] Cressman, R. (1992) *The Stability Concept of Evolutionary Game Theory (A Dynamic Approach)*, Lecture Notes in Biomathematics **94**, Springer-Verlag, Heidelberg.

[9] Cressman, R. (2004) Dynamic stability of the replicator equation with continuous strategy space. IIASA Report IR-04-017.

[10] Cressman, R. and J. Garay (2003) Stability in N-species coevolutionary systems. *Theoret. Pop. Biol.* **64**, 519-533.

[11] Dieckmann, U. and R. Law (1996) The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.* **34**, 579-612.

[12] Diekmann, O. (2004) A beginners guide to adaptive dynamics. In *Mathematical Modelling of Population Dynamics* (R. Rudnicki, ed.), 47-86 Banach Center Publications, Volume 63, Institute of Mathematics, Polish Academy of Sciences.

[13] Doebeli, M. and U. Dieckmann (2000) Evolutionary branching and sympatric speciation caused by different types of ecological interactions. *Amer. Nat.* **156**, S77-S101.

[14] Eshel I. (1983) Evolutionary and continuous stability. *J. Theor. Biol.* **103**, 99-111.

[15] Eshel I., U. Motro and E. Sansone (1997) Continuous stability and evolutionary convergence. *J. Theor. Biol.* **185**, 333-343.

[16] Eshel, I and E. Sansone (2003) Evolutionary and dynamic stability in continuous population games. *J. Math. Biol.* **46**, 445-459.

[17] Geritz, S.A.H., J.A.J. Metz, E. Kisdi and G. Meszéna (1997) Dynamics of adaptation and evolutionary branching. *Phys. Rev. Lett.* **78**, 2024-2027.

[18] Geritz, S.A.H., E. Kisdi and G. Meszéna, J.A.J. Metz (1998) Evolutionarily stable strategies and the adaptive growth of the evolutionary tree. *Evol. Ecol.* **12**, 35-57.

[19] Heifetz, A., C. Shannon and Y. Spiegel (2003) What to maximize if you must. mimeo.

[20] Hofbauer, J. and K. Sigmund (1990) Adaptive dynamics and evolutionary stability. *Appl. Math. Lett.* **3**, 75-79.

[21] Hofbauer, J. and K. Sigmund (1998) *Evolutionary Games and Population dynamics*, Cambridge University Press.

[22] Hofbauer, J. and J. Weibull (1996) Evolutionary selection against dominated strategies. *J. Econ. Theory* **71**, 558-573.

[23] Kisdi, É and G. Meszéna (1995) Life histories with lottery competition in a stochastic environment: ESSs which do not prevail. *Theoret. Pop. Biol.* **47**, 191-211.

[24] Marrow, P., U. Dieckmann and R. Law (1996) Evolutionary dynamics of predator-prey systems: an ecological perspective. *J. Math. Biol.* **34**, 556-578.

[25] Maynard Smith, J. (1982) *Evolution and the Theory of Games*, Cambridge University Press.

[26] McKelvey R. and J. Apaloo (1995) The structure and evolution of competition-organized ecological communities, *Rocky Mountain J. Math.* **25**, 417-436.

[27] Meszéna G., É. Kisdi, U. Dieckmann, S.A.H. Geritz and J.A.J. Metz (2001) Evolutionary optimisation models and matrix games in the unified perspective of adaptive dynamics. *Selection* **2**, 193-210.

[28] Metz J.A.J., S.A.H. Geritz, G. Meszéna, F.J.A. Jacobs and J.S. van Heerwaarden (1996) Adaptive dynamics, a geometrical study of the consequences of nearly faithful reproduction. In *Stochastic and Spatial Structures of Dynamical Systems* (S.J. van Strien and S.M. Verduyn Lunel, eds), 183-231 North Holland, Amsterdam.

[29] Oechssler, J. and F. Riedel (2001) Evolutionary dynamics on infinite strategy spaces, *Econ. Theory* **17**, 141-162.

[30] Oechssler, J. and F. Riedel (2002) On the dynamics foundation of evolutionary stability in continuous models, *J. Econ. Theory* **107**, 223-252.

[31] Roughgarden J. (1979) *Theory of Population Genetics and Evolutionary Ecology: An Introduction,* Macmillan, New York.

[32] Samuelson, L. and J. Zhang (1992) Evolutionary stability in asymmetric games. *J. Econ. Theory* **59**, 363-391.

[33] Taylor, P.D. (1989) Evolutionary stability in one-parameter models under weak selection. *Theor. Popul. Biol.* **36**, 125-143.

[34] Vincent T.L., Y. Cohen and J.S. Brown (1993) Evolution via strategy dynamics. *Theoret. Popul. Biol.* **44**, 149-176.

[35] Vincent T.L., M.V. Van and B.S. Goh (1996) Ecological stability, evolutionary stability and the ESS maximum principle. *Evol. Ecol.* **10**, 567-591.

[36] Weibull, J. (1995) *Evolutionary Game Theory*, MIT Press, Cambridge, MA.