

Mutual invadability near evolutionarily singular

2 strategies for multivariate traits, with special reference
to the strongly convergence stable case

4 Stefan A. H. Geritz · Johan A. J. Metz ·

Claus Rueffler

This paper is dedicated to Mats Gyllenberg for his 60th anniversary, as appreciation for his continual support of the adaptive dynamics cause.

6 Received: date / Revised: date

S. A. H. Geritz

Department of Mathematics and Statistics, 00014 University of Helsinki, Finland

J. A. J. Metz

Mathematical Institute & Institute of Biology, Leiden University, P.O. Box 9512, 2300RA
Leiden, Netherlands

2nd affiliation: Netherlands Centre for Biodiversity, Naturalis, P.O. Box 9517, 2300RA Leiden, Netherlands

3rd affiliation: Evolution and Ecology Program, International Institute of Applied Systems Analysis, A-2361 Laxenburg, Austria E-mail: j.a.j.metz@biology.leidenuniv.nl

C. Rueffler

Animal Ecology, Department of Ecology and Genetics, Uppsala University, Norbyvägen 18D,
75236 Uppsala, Sweden

E-mail: claus.rueffler@ebc.uu.se

Abstract Over the last two decades evolutionary branching has emerged as a possible mathematical paradigm for explaining the origination of phenotypic diversity. Although branching is well understood for one-dimensional trait spaces, a similarly detailed understanding for higher dimensional trait spaces is sadly lacking. This note aims at getting a research program of the ground leading to such an understanding. In particular, we show that, as long as the evolutionary trajectory stays within the reign of the local quadratic approximation of the fitness function, any initial small scale polymorphism around an attracting invadable evolutionarily singular strategy will evolve towards a dimorphism. That is, provided the trajectory does not pass the boundary of the domain of dimorphic coexistence and falls back to monomorphism (after which it moves again towards the singular strategy and from there on to a small scale polymorphism, etc.). To reach these results we analyze in some detail the behavior of the solutions of the coupled Lande-equations purportedly satisfied by the phenotypic clusters of a quasi- n -morphism, and give a precise characterisation of the local geometry of the set \mathcal{D} in trait space squared harbouring protected dimorphisms. Intriguingly, in higher dimensional trait spaces an attracting invadable ess needs not connect to \mathcal{D} . However, for the practically important subset of strongly attracting ess-es (i.e., ess-es that robustly locally attract the monomorphic evolutionary dynamics for all possible non-degenerate mutational or genetic covariance matrices) invadability implies that the ess does connect to \mathcal{D} , just as in 1-dimensional trait spaces. Another matter is that in principle there exists the possibility that the dimorphic evolutionary trajectory reverts to monomorphism still within the reign of the local

quadratic approximation for the invasion fitnesses. Such locally unsustainable
32 branching cannot occur in 1- and 2-dimensional trait spaces, but can do so in
higher dimensional ones. For the latter trait spaces we give a condition exclud-
34 ing locally unsustainable branching which is far stricter than the one of strong
convergence, yet holds good for a relevant collection of published models. It
36 remains an open problem whether locally unsustainable branching can occur
around general strongly attracting invadable ess-es

38 **Keywords** adaptive dynamics, evolutionary branching, multi-dimensional
trait space, mutual invadability, strong attractivity, local dimorphic divergence

40 **Mathematics Subject Classification (2000)** MSC 92D15, MSC 92D25

1 Introduction

42 Over the last two decades *evolutionary branching* has emerged as an important
concept for explaining the adaptive evolution of phenotypic diversity. Evolu-
44 tionary branching occurs at points in trait space (strategies) that initially
attract the evolutionary dynamics, but where selection changes from direc-
46 tional to disruptive once the population mean trait value comes sufficiently
close (Metz et al., 1996; Geritz et al., 1997, 1998; Rueffler et al., 2006; Dercole
48 and Rinaldi, 2008; Doebeli, 2011). (In line with tradition, this initial evolu-
tionary dynamics is assumed here to be (quasi-)monomorphic. We shall below
50 stick to this assumption, and refer to the attractors of this dynamics just as
evolutionary attractors, even though branching is coincident with their repul-
52 sion in the dimorphic realm.) As a result, at such points populations can split
into two or more phenotypic clusters. More specifically, evolutionary branch-

ing at a point \mathbf{x}^* in trait space requires that at least five requirements are fulfilled (Metz et al., 1996; Geritz et al., 1998; Doebeli, 2011): (i) The point \mathbf{x}^* has to be an attractor of the evolutionary dynamics. (ii) The point \mathbf{x}^* has to be locally invadable by mutants in at least one pair of opposite directions. (iii) In at least some of these directions nearby mutant phenotypes must be able to coexist in a protected dimorphism. (iv) There should be at least one such direction in which coexisting types experience divergent selection. (v) The coexistence cone emanating from $(\mathbf{x}^*, \mathbf{x}^*)$ should be sufficiently wide for the incipient branches to stay inside while they become visibly separated. In the clonal case branching is bound to occur if these conditions are fulfilled, while in the Mendelian case these conditions are necessary, but it depends on a lot more whether branching indeed occurs. In one-dimensional trait spaces conditions (i) to (v) are easy to check and it turns out that the former two imply the latter three (Metz et al., 1996; Geritz et al., 1997, 1998). However, in higher dimensional trait spaces this needs not to be the case. In particular, it is possible that requirement (i) and (ii) are fulfilled while (iii) is not (e.g. Doebeli, 2011, pp. 119), let alone (v).

In this paper we derive criteria for testing for (iii) and (v) within the reign of the local quadratic approximation for the invasion fitness function. It is known that in an n -dimensional trait space at most $n+1$ branches can coexist (Durinx et al., 2008). Hence, for scalar traits, branching can only be into two. Here we show that in higher dimensional trait spaces generically any polymorphism evolves in the direction of a dimorphism (or rather, quasi-dimorphism, as close to evolutionary attractors full mutation limitation fails,

78 so that h -morphisms get replaced by h concentrated clouds of trait values).
We can therefore confine ourselves to delimiting the set \mathcal{D} of trait pairs able
80 to coexist. As final step we derive conditions for further evolution to keep a
dimorphism in \mathcal{D} . The alternative is that the branching evolutionary trajectory
82 falls back to monomorphism, after which it may branch again, and so on.

A next question is whether there exist restricted model classes that can be
84 delimited in an intuitively natural manner and for which (iii) and/or (v) are
implied by (i) and (ii) like in the 1-dimensional case. To discuss this question
86 it is necessary to go a bit more deeply into the notion of attractor of the evo-
lutionary dynamics. In one-dimensional trait spaces, whether or not a point in
88 trait space is an attractor of the evolutionary dynamics is independent of the
mutational process. In trait spaces with more than one dimension, however,
90 the mutational input can affect the course of the evolutionary dynamics to the
extent that a particular point can be an attractor for one mutational variance-
92 covariance matrix but a repeller for another one. Leimar (2009) introduced
the following notions. First, a point \mathbf{x}^* is *absolutely convergence stable* when it
94 is an attractor of the evolutionary dynamics for any mutational process. Sec-
ond, a point \mathbf{x}^* is called *strongly convergence stable* when it is an attractor of
96 the evolutionary dynamics for any mutational process provided the mutational
step sizes are sufficiently small. Convergence stability in this case means that
98 \mathbf{x}^* is an asymptotically stable fixed point of the so-called *canonical equation*
of adaptive dynamics (Dieckmann and Law, 1996; Champagnat, 2003; Dur-
100 inx et al., 2008; Champagnat and Méléard, 2011; Collet et al., 2013; Metz
and de Kovel, 2013). Leimar (2009) furthermore established that for \mathbf{x}^* to be

robustly strongly convergence stable (below also referred to as *strongly attracting*) it is necessary and sufficient that the Jacobian matrix J of the selection gradient at \mathbf{x}^* is negative definite, i.e., that all eigenvalues of its symmetric part $\frac{1}{2}(J + J^T)$ are negative.

Remark. More precisely, Leimar (2009) established that it is sufficient that J is negative definite, and necessary that it is negative semi-definite. In the borderline case any strong convergence is non-robust against arbitrary small perturbations of the model.

Negative definiteness is a strong requirement. However, it appears that most published models that describe the evolutionary dynamics of a multivariate trait by means of the adaptive dynamics approximation fulfill this criterion (Leimar, 2001; Vukics et al., 2003; Ackermann and Doebeli, 2004; Beltman and Metz, 2005; Ito and Shimada, 2007; Ravigné et al., 2009; Doebeli and Ispolatov, 2010; Svardal et al., 2011, 2014). It is therefore of some relevance to know whether for this special but apparently regularly occurring case a similar dependency exists as for one-dimensional trait spaces. We show that this is indeed to a certain extent the case: when the symmetric part of the Jacobian matrix of the fitness gradient is negative definite, condition (ii) implies conditions (iii) and (iv), so diversification at least will get started. However, it is not yet clear whether in these cases also (v) is implied. It thus remains an open problem whether for clonal reproducers strong convergence guarantees that an incipient diversification will culminate in more extended branching. The best we could do for the present is give some stronger conditions guaranteeing that such is the case.

126 2 Technical context

We start our treatment with a short methodological introduction. We are interested in the evolutionary dynamics of a population in which individuals are characterized by n quantitative traits. Thus, each individual is described by a trait vector $\mathbf{x} = (x_1, \dots, x_n)^T$. We follow the dynamics of the traits over evolutionary time as it results from repeated mutant substitutions. Specifically, we consider a simplified mutational process in a clonal population in which rare mutations of small effect change the trait values from \mathbf{x} to $\mathbf{y} = \mathbf{x} + \Delta_{\mathbf{x}}$ and where mutations can occur in all directions in trait space. The evolutionary dynamics can then be determined by following a series of mutation-substitution events in which the trait vector \mathbf{x} of the resident population changes over time. The fundamental tool to predict this dynamics are the invasion fitnesses $s(\mathbf{y}; \mathbf{x})$, which are defined as the expected long-term exponential growth rate of an infinitesimally small mutant subpopulation with trait vector \mathbf{y} in an environment in which all relevant components such as prey, pathogen and predator densities are determined by the resident population with trait vector \mathbf{x} (Metz et al., 1992; Metz, 2008)[the latter revised as (Metz, 2014)].

In the limit of rare mutation events and small unbiased mutational steps the evolutionary dynamics can be described by

$$\frac{d\mathbf{x}}{dt} = n_e(\mathbf{x})\theta\Sigma(\mathbf{x})\mathbf{g}(\mathbf{x}) \quad (1)$$

(Dieckmann and Law, 1996; Durinx et al., 2008; Metz and de Kovel, 2013; Metz and Jansen, in prep). Here, $n_e(\mathbf{x})$ is the effective population size as in population genetics, θ the mutation probability per birth event and Σ the

148 n -dimensional mutational variance-covariance matrix summarising the distri-
 149 bution of mutations supposed to be symmetric around the resident type \mathbf{x} .
 150 Finally, $\mathbf{g}(\mathbf{x})$ denotes the n -dimensional selection gradient with entries

$$g_i(\mathbf{x}) := \left. \frac{\partial s(\mathbf{y}; \mathbf{x})}{\partial y_i} \right|_{\mathbf{y}=\mathbf{x}}. \quad (2)$$

A point \mathbf{x}^* where $\mathbf{g}(\mathbf{x}^*) = 0$ is referred to as an *evolutionarily singular strategy*
 152 (*ess*). At such points the evolutionary dynamics described by Eq. (1) comes
 to a halt.

154 When mutation limitation fails, as is necessarily the case close to evolution-
 arily singular points, one can fall back on *Lande's equation* from quantitative
 155 genetics (modified to take account of the changes in the fitness landscape
 resulting from the trait evolution) which is similar to Eq. (1), except that
 158 the term $n_e(\mathbf{x})\theta\Sigma$ is replaced by the covariance matrix of the standing ge-
 netic variation, which we, with some slight abuse of notation, shall also denote
 160 as Σ (Lande, 1979, 1982). The usual additional assumption is that Σ is con-
 stant, interpreted as approximation for the case of relatively small evolutionary
 162 change (c.f. Figure 1). We will use this approximation when considering the
 initial divergence of the evolutionary branches. Of course, both the canonical
 164 and Lande's equation fail really close to the singular point when the spread-
 ing unimodal trait distribution is becoming multimodal. However, when the
 166 modes have grown sufficiently far apart their movement can initially again
 be modeled by a set of coupled Lande equations, that is, till the strength of
 168 directional selection in each branch increases to a level where the consumption
 of standing variation gets too large relative to its mutational replenishment.

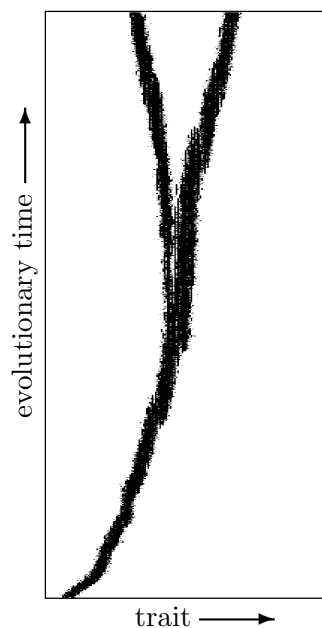


Fig. 1 Output of an individual based model showing a branching trajectory without full mutation limitation. At equally spaced times a dot was drawn for each individual present at that time. Note that the standing variation of the monomorphic population and of the branches remain roughly constant except during the widening of the distribution that precedes the splitting of the branches.

170 **Remark.** Three features distinguish the canonical and Lande's equation.
172 Firstly, the stress of the canonical equation is on the change of the fitness landscape that inevitably follows in the wake of trait evolution (the part Fisher
(1958) suppressed in the mathematical formulation of his fundamental theorem (p. 37), although he was obviously well aware of it (p. 45–49)), whereas
174 in Lande's equation, as standardly encountered, such changes are neglected.
176 Secondly, where the canonical equation is formulated in terms of the hypothetical underlying variation generating mechanism, Lande's equation uses
178 the empirically accessible standing genetic variation. Both differences make

Lande's equation more useful for concretely describing less extensive evolutionary changes, and the canonical equation more appropriate for the theoretical consideration of larger scale changes. Thirdly, the canonical equation has been rigorously underpinned (Dieckmann and Law, 1996; Champagnat, 2003; Tran, 2006; Durinx et al., 2008; Méléard and Tran, 2009; Champagnat and Méléard, 2011; Collet et al., 2013; Metz and de Kovel, 2013; Metz and Jansen, in prep), albeit using a biologically seemingly unrealistic limit procedure (but see the arguments of Metz and de Kovel (2013) about its domain of validity as an approximation), whereas the heuristic underpinning of Lande's equation so far has not been subjected to such a rigorous treatment.

For a matrix M we shall use $M > 0$ (< 0 , ≥ 0 , ≤ 0) to indicate that it is *positive* (*negative*, *positive semi-*, *negative semi-*) definite, i.e., $\mathbf{x}^T M \mathbf{x} > 0$ (< 0 , ≥ 0 , ≤ 0) for all $\mathbf{x} \neq 0$. In the case of non-symmetric M this means that the various kinds of definiteness are not so much properties of the full M as of its symmetric part $\frac{1}{2}(M + M^T)$, without involvement of its antisymmetric part $\frac{1}{2}(M - M^T)$.

The singular point is invadable by nearby mutants if the Hessian matrix H of the invasion fitness evaluated at \mathbf{x}^* , with entries

$$h_{ij} := \left. \frac{\partial^2 s(\mathbf{y}; \mathbf{x}^*)}{\partial y_i \partial y_j} \right|_{\mathbf{y}=\mathbf{x}^*}, \quad (3)$$

is not negative semi-definite and only if it is not negative definite, or, equivalently, if its dominant eigenvalue λ_1 is positive and only if it is non-negative. Note that the Hessian matrix is necessarily symmetric: $h_{ij} = h_{ji}$. If $\lambda_1 > 0$, then \mathbf{x}^* is not a local maximum of the fitness landscape but either a minimum or a saddle point and nearby mutants \mathbf{y} that correspond to a higher point

202 on this landscape are able to invade the population. In this case, selection is disruptive in at least some directions in trait space.

204 We introduce the following notation:

$$\begin{aligned} C_{00} &:= \frac{1}{2} \frac{\partial^2 s(\mathbf{y}; \mathbf{x})}{\partial \mathbf{y}^2} \Big|_{\mathbf{y}=\mathbf{x}=\mathbf{x}^*}, & C_{10} &:= \frac{1}{2} \frac{\partial^2 s(\mathbf{y}; \mathbf{x})}{\partial \mathbf{x} \partial \mathbf{y}} \Big|_{\mathbf{y}=\mathbf{x}=\mathbf{x}^*}, \\ C_{01} &:= \frac{1}{2} \frac{\partial^2 s(\mathbf{y}; \mathbf{x})}{\partial \mathbf{y} \partial \mathbf{x}} \Big|_{\mathbf{y}=\mathbf{x}=\mathbf{x}^*}, & C_{11} &:= \frac{1}{2} \frac{\partial^2 s(\mathbf{y}; \mathbf{x})}{\partial \mathbf{x}^2} \Big|_{\mathbf{y}=\mathbf{x}=\mathbf{x}^*}. \end{aligned} \quad (4)$$

Thus, $2C_{00} = H$ and $C_{01} = C_{10}^T$.

206 Leimar (2009) showed that a singular point is robustly asymptotically stable for any variance-covariance matrix Σ (of the mutational or standing variation) if the Jacobian matrix $J = 2(C_{00} + C_{01})$ of the selection gradient at 208 \mathbf{x}^* is negative definite, while Σ has an essential influence when J is not negative semi-definite. From $s(\mathbf{x}; \mathbf{x}) = 0$ applied to the second order term in its 210 expansion around $\mathbf{x} = \mathbf{x}^*$ (in both positions) it follows that

$$C_{00} + C_{01} + C_{10} + C_{11} = 0, \quad (5)$$

212 which is equivalent to

$$C_{00} + \frac{C_{01} + C_{10}}{2} = -\frac{C_{01} + C_{10}}{2} - C_{11} \quad (6)$$

and therefore

$$\begin{aligned} (J + J^T)/2 < 0 &\Leftrightarrow C_{00} + \frac{C_{01} + C_{10}}{2} < 0 \Leftrightarrow C_{00} - \frac{C_{00} + C_{11}}{2} < 0 \\ &\Leftrightarrow \frac{C_{00} - C_{11}}{2} < 0 \Leftrightarrow C_{11} - C_{00} > 0. \end{aligned} \quad (7)$$

214 Thus, the condition for robust strong convergence stability *sensu* Leimar (2009) can be rephrased as $C_{11} - C_{00} > 0$.

216 3 Coexistence on the ecological time scale

For diversification to get of the ground it is necessary that close to the ess at
 218 least two phenotypes can coexist. Under certain smoothness conditions the co-
 existence of similar strategies is necessarily of the protected type, i.e., each phe-
 220 notype can invade into the other one (Geritz, unpublished), see also (Geritz,
 2005; Dercole and Geritz, submitted). Therefore we start with investigating
 222 the conditions for mutual invadability near an ess. The starting point is the
 Taylor approximation of the invasion fitness function

$$s(\mathbf{x}^* + \mathbf{v}; \mathbf{x}^* + \mathbf{u}) = \mathbf{v}^T \mathbf{C}_{00} \mathbf{v} + 2\mathbf{v}^T \mathbf{C}_{01} \mathbf{u} + \mathbf{u}^T \mathbf{C}_{11} \mathbf{u} + \text{h.o.t.} \quad (8)$$

224 *To diminish verbiage we shall phrase our arguments as if the reign of the*
quadratic approximation of s extends forever, as is the case when we look at
 226 *the geometry on the scale of the mutational steps.*

The conditions for mutual invadability are

$$\mathbf{u}_1^T \mathbf{C}_{00} \mathbf{u}_1 + 2\mathbf{u}_1^T \mathbf{C}_{01} \mathbf{u}_2 + \mathbf{u}_2^T \mathbf{C}_{11} \mathbf{u}_2 > 0 \quad (9a)$$

$$\mathbf{u}_2^T \mathbf{C}_{00} \mathbf{u}_2 + 2\mathbf{u}_2^T \mathbf{C}_{01} \mathbf{u}_1 + \mathbf{u}_1^T \mathbf{C}_{11} \mathbf{u}_1 > 0. \quad (9b)$$

228 To render these inequalities in a better interpretable form we introduce $\mathbf{m} :=$
 $\frac{1}{2}(\mathbf{u}_1 + \mathbf{u}_2)$, the mean of the two trait vectors, and $\mathbf{d} := \frac{1}{2}(\mathbf{u}_1 - \mathbf{u}_2)$, half their
 230 difference, so that $\mathbf{u}_1 = \mathbf{m} + \mathbf{d}$ and $\mathbf{u}_2 = \mathbf{m} - \mathbf{d}$ (Figure 2). Substitution of
 these expressions in Eq. (9) and perusing Eq. (5) gives

$$-\mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{11}) \mathbf{d} < 2\mathbf{d}^T (\mathbf{C}_{10} + \mathbf{C}_{11}) \mathbf{m} < \mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{11}) \mathbf{d}, \quad (10)$$

232 or, equivalently,

$$-\mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{11}) \mathbf{d} < 2\mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{01}) \mathbf{m} < \mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{11}) \mathbf{d}, \quad (11)$$

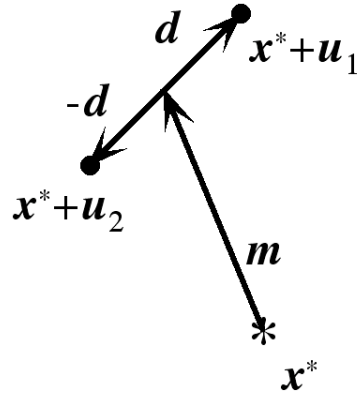


Fig. 2 Geometrical interpretation of the vectors m and d .

or, equivalently,

$$\mathbf{d}^T \mathbf{C}_{01} \mathbf{d} < \mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{01}) \mathbf{m} < -\mathbf{d}^T \mathbf{C}_{01} \mathbf{d}. \quad (12)$$

Eq. (11), (12) imply that there exist mutually invadable trait pairs near an evolutionarily singular point if and only if there exist vectors \mathbf{d} such that $\mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{11}) \mathbf{d} > 0$, or equivalently $-\mathbf{d}^T \mathbf{C}_{01} \mathbf{d} > 0$. (For the if direction take $\mathbf{m} = 0$.) These inequalities can hold good if and only if $\mathbf{C}_{00} + \mathbf{C}_{11}$ has at least one positive eigenvalue, or in other words, is not negative definite.

Remark. It may seem that we are a bit sloppy here as in a deterministic model a type may also invade when its invasion fitness is zero, except that it takes very long to do so. However, our deterministic models are only large system size limits of individual-based models. If the invasion fitness is zero, in the limit the probability that such a type invades, i.e., from a single individual its numbers grow to the order of magnitude of the system size, goes to zero. So in practice one can neglect this possibility, so that what in a strict mathematical sense is only a sufficient condition becomes an effectively necessary and sufficient one.

As a next step we take a closer look at the width of the $2n$ -dimensional set of coexisting pairs \mathcal{D} , characterised by Eq. (11), as this determines the ease with which the adaptive dynamics will step from the monomorphic to a polymorphic condition. We deliberately use the word 'width' since \mathcal{D} is scale-invariant, that is, $\alpha\mathcal{D} = \mathcal{D}$ for all $\alpha \in \mathbb{R}$. So what matters are the directions in \mathbb{R}^{2n} that correspond to coexistence.

As a warming up we first consider the one-dimensional case. There the condition of mutual invadability reduces to

$$-(c_{00} + c_{11})d < (c_{00} - c_{11})m < (c_{00} + c_{11})d \quad (13)$$

with $m = \frac{1}{2}(u_1 + u_2)$, $d = \frac{1}{2}(u_1 - u_2)$. The simplification relative to Eq. (11) derives from the fact that in the scalar case $c_{ij} = c_{ji}$ so that $2(c_{00} + c_{01}) = c_{00} + c_{01} - c_{10} - c_{11} = c_{00} - c_{11}$. In a mutual invadability plot as depicted in Figure 3, $\sqrt{2}d$ equals the distance of (u_1, u_2) to the diagonal, and $\sqrt{2}m$ the distance to the anti-diagonal through x^* . For singular points that are both attracting and invadable the coexistence cone always has a width of more than 90 degrees, and is symmetric around the diagonal as well as the anti-diagonal. As in the one-dimensional case, generally the pair (\mathbf{m}, \mathbf{m}) can be interpreted as the orthogonal projection of $(\mathbf{u}_1, \mathbf{u}_2)$ on the linear manifold given by $\mathbf{u}_1 = \mathbf{u}_2$, and $(\mathbf{d}, -\mathbf{d})$ as the difference of $(\mathbf{u}_1, \mathbf{u}_2)$ and that projection. In a similar vein, the symmetry of \mathcal{D} around the diagonal extends to symmetry in the \mathbf{d} directions around $\mathbf{d} = 0$ and symmetry over the anti-diagonal extends to symmetry in the \mathbf{m} directions around $\mathbf{m} = 0$.

The one-dimensional case also can be found embedded in the n -dimensional case in the form of pairs $(\mathbf{u}_1, \mathbf{u}_2)$ for which the line through \mathbf{u}_1 and \mathbf{u}_2 passes

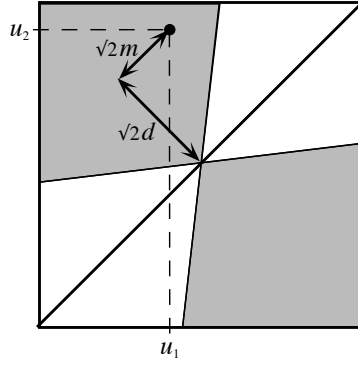


Fig. 3 How the quantities d and m relate to the coexistence region (grey) in the mutual invadability plot of u_1 and u_2 .

through the origin (situated at \mathbf{x}^*), so that we can write $\mathbf{u}_1 = u_1 \mathbf{r}$ and $\mathbf{u}_2 = u_2 \mathbf{r}$, giving $c_{ij} = \mathbf{r}^T \mathbf{C}_{ij} \mathbf{r}$.

Even when $\mathbf{C}_{00} + \mathbf{C}_{11}$ has only one positive eigenvalue κ_1 with eigenvector \mathbf{k}_1 , there is no need for \mathbf{d} to be aligned with \mathbf{k}_1 for a pair $(\mathbf{u}_1, \mathbf{u}_2)$ to lie in \mathcal{D} . If we express \mathbf{d} in a basis of eigenvectors, normalised such that $\mathbf{k}_i^T \mathbf{k}_i = 1$, all that is needed is $-\sum_2^n \kappa_i d_i^2 < \kappa_1 d_1^2$. We shall refer to \mathbf{d} such that $\mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{11}) \mathbf{d} > 0$ as 'allowable'. Shifting both members of a pair $(\mathbf{u}_1, \mathbf{u}_2)$ in any direction orthogonal to $\mathbf{d}(\mathbf{C}_{00} + \mathbf{C}_{01})$ does not affect either \mathbf{d} or $\mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{01}) \mathbf{m}$. Hence, to construct the set of all \mathbf{m} that go with a certain allowable \mathbf{d} we can start from $\mathbf{m} = \alpha \mathbf{h}$ with $\mathbf{h}^T := 2\mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{01})$ and α delimited by $-\mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{11}) \mathbf{d} < \mathbf{h}^T \mathbf{h} \alpha < \mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{11}) \mathbf{d}$, and add to these any \mathbf{m}' orthogonal to \mathbf{h} .

The upshot is that even when $\mathbf{C}_{00} + \mathbf{C}_{11}$ has only one positive eigenvalue, there are such a good amount of mutually invadable pairs that the step from mono- to dimorphism will occur rather sooner than later.

When \mathbf{x}^* is strongly attracting $\mathbf{C}_{11} - \mathbf{C}_{00} > 0$. Hence, when $\mathbf{r}^T \mathbf{C}_{00} \mathbf{r} > 0$ for some vectors \mathbf{r} , also $\mathbf{r}^T \mathbf{C}_{11} \mathbf{r} > 0$ and hence $\mathbf{r}^T (\mathbf{C}_{00} + \mathbf{C}_{11}) \mathbf{r} > 0$. Therefore, for

286 a strongly attracting ess invadability implies the existence of a multitude of
close by mutually invadable pairs of trait vectors, of which we will see in the
288 next section that they undergo disruptive selection, i.e., selective pressures on
each member of the pair that drive them further apart.

290 **4 Coexistence on the evolutionary time scale**

At a branching point the trait vectors representative for each of the incipient
292 branches are subject to disruptive selection, letting them grow apart at least
initially. We will follow this movement only within the reign of the quadratic
294 approximation of the invasion fitness function close to the ess, and will do so
under the assumption that the movement is adequately represented by coupled
296 Lande equations. (Note that with a quadratic approximation we do not mean
a Taylor approximation, as the latter is only applicable when the number of
298 coexisting trait vectors equals the dimension of the trait space (n) plus one,
see below.)

300 *We shall below again phrase our arguments as if the reign of the quadratic
approximation of s extends forever. Moreover, we without further ado proceed
302 on the assumption that coexistence results derived for the case of full muta-
tion limitation extend to any well separated quasi-monomorphic clusters that
304 replace the single phenotypes when there is less than strict mutation limita-
tion. Lastly, we will adapt the coordinate system so as to transform Σ into the
306 identity matrix.*

Other than perhaps expected from the scalar case, in the multivariate case
308 there is the possibility for $h > 2$ phenotypes to coexist near an ess. (The

reason why we have not gone into this potential complication in the previous
 310 section will become clear further on.) If the demographic parameters of the
 individuals under consideration depend smoothly on their phenotype then s
 312 will depend smoothly on the mutant trait vector \mathbf{v} (Ferrière and Gatto, 1995).
 Such smoothness cannot be assumed for the dependence on $(\mathbf{u}_1, \dots, \mathbf{u}_h)$ since
 314 the environment created by the residents $(\mathbf{x}^* + \mathbf{u}_1, \dots, \mathbf{x}^* + \mathbf{u}_h)$ is determined
 by the attractor of their community dynamics. Assuming smoothness in the
 316 resident phenotype for monomorphisms is pretty harmless, as at least for simple
 community attractors this is guaranteed away from community dynamical
 318 bifurcation points by some form of the inverse function theorem. This argument
 extends to polymorphisms, but not necessarily to the boundary of the
 320 region in \mathbb{R}^{hn} harbouring h -morphisms, as these are characterised by the occurrence
 of a bifurcation. In particular at corners of that boundary, like the
 322 point $(\mathbf{x}^*, \dots, \mathbf{x}^*) \in \mathbb{R}^{hn}$, differentiability can fail. Hence, we may expect the
 dependence of s on $(\mathbf{u}_1, \dots, \mathbf{u}_h)$ to have at best directional derivatives, but
 324 generally not to have a full derivative. An argument, in terms of the local
 geometry of the community dynamics, why directional derivatives can still be
 326 expected to exist can be found in (Durinx et al., 2008).

The Taylor expandability of s in \mathbf{v} gives

$$s(\mathbf{x}^* + \mathbf{v}; \mathbf{x}^* + \mathbf{u}_1, \dots, \mathbf{x}^* + \mathbf{u}_h) = a + \mathbf{b}^T \mathbf{v} + \mathbf{v}^T \mathbf{C}_{00} \mathbf{v}, \quad (14)$$

328 with a and \mathbf{b} functions of $(\mathbf{u}_1, \dots, \mathbf{u}_h)$, which we take to be second and first order
 respectively (on the strength of the existence of the directional derivatives).

330 The explicit expression for the quadratic term is found from the ecological con-

sistency condition

$$s(\mathbf{x}^* + \mathbf{v}; \mathbf{x}^*, \dots, \mathbf{x}^*) = s(\mathbf{x}^* + \mathbf{v}; \mathbf{x}^*). \quad (15)$$

332 One first result from the other ecological consistency conditions

$$s(\mathbf{x}^* + \mathbf{u}_i; \mathbf{x}^* + \mathbf{u}_1, \dots, \mathbf{x}^* + \mathbf{u}_h) = 0, \quad \text{for } i = 1, \dots, h, \quad (16)$$

is that close to \mathbf{x}^* generically at most $n+1$ phenotypes can coexist, as otherwise
 334 the number of equations for a and the components of \mathbf{b} exceeds the number
 of unknowns, a result going back to Christiansen and Loeschke (1987). When
 336 the number of coexisting phenotypes equals $n + 1$, Eqs. (15) and (16) fully
 determine s . When the number of phenotypes is less than $n + 1$ this is no
 338 longer the case and it becomes necessary to proceed through the harrowing
 procedure of calculating s from first principles. Luckily, there are still some
 340 results to be derived in a more lazy manner.

Our primary interest at this point is not s itself, but the selection gradients

$$\mathbf{g}_i^T(\mathbf{u}_1, \dots, \mathbf{u}_h) := \left. \frac{\partial s(\mathbf{v}; \mathbf{u}_1, \dots, \mathbf{u}_h)}{\partial \mathbf{v}} \right|_{\mathbf{v}=\mathbf{u}_i} = \mathbf{b}^T(\mathbf{u}_1, \dots, \mathbf{u}_h) + 2\mathbf{C}_{00}\mathbf{u}_i. \quad (17)$$

342 The form of Eq. (17) suggests defining $\mathbf{m} := h^{-1}(\mathbf{u}_1 + \dots + \mathbf{u}_h)$ and $\mathbf{d}_i :=$
 $\mathbf{u}_i - \mathbf{m}$, which when substituted in the Lande equations yields

$$\frac{d\mathbf{d}_i}{dt} = 2\mathbf{C}_{00}\mathbf{d}_i. \quad (18)$$

344 Hence, all \mathbf{d}_i will in the long run align in a direction parallel to the eigenvector
 \mathbf{z} corresponding to the dominant eigenvalue λ_1 of $\mathbf{H} = 2\mathbf{C}_{00}$, which we assume
 346 to be unique and positive. As a result all \mathbf{u}_i will get to lie at any given large
 time close to a single line $\{\mathbf{m} + \zeta\mathbf{z} | \zeta \in \mathbb{R}\}$. Restricted to such a line s becomes

348 a quadratic function of ζ . In combination with the consistency relation Eq.
 (16) this implies the following result:

350 **Proposition.** *Generically, expanding polymorphisms around ess-es initially
 evolve towards becoming dimorphisms.*

352 For dimorphisms, under the assumption that the community dynamics
 converges to an equilibrium point

$$s(\mathbf{v}; \mathbf{u}_1, \mathbf{u}_2) = \mathbf{m}^T \mathbf{C}_{11} \mathbf{m} - \mathbf{d}^T \mathbf{C}_{00} \mathbf{d} + \mathbf{v}^T \mathbf{C}_{00} \mathbf{v} \\ + 2 \left(\mathbf{m}^T \mathbf{C}_{10} \mathbf{v} - \frac{\mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{01}) \mathbf{m}}{\mathbf{d}^T \mathbf{C}_{01} \mathbf{d}} \mathbf{d}^T \mathbf{C}_{10} (\mathbf{v} - \mathbf{m}) \right) \quad (19)$$

354 with $\mathbf{d} = \mathbf{d}_1 = -\mathbf{d}_2$ (Durinx et al., 2008), which in the univariate case (for
 which $n + 1 = 2$) reduces to

$$s(v; u_1, u_2) = c_{00}(v - u_1)(v - u_2) \quad (20)$$

356 To see what can be deduced from Eq. (19) about the longer term coexis-
 tence of the diverging branches we calculate the selection gradients

$$\mathbf{g}_i(\mathbf{u}_1, \mathbf{u}_2) = 2 \left(\mathbf{C}_{01} \mathbf{m} - \frac{\mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{01}) \mathbf{m}}{\mathbf{d}^T \mathbf{C}_{01} \mathbf{d}} \mathbf{C}_{01} \mathbf{d} \right) + 2 \mathbf{C}_{00} \mathbf{u}_i, \quad (21)$$

358 which in the univariate case reduce to

$$g_1(u_1, u_2) = c_{00}(u_1 - u_2), \quad g_2(u_1, u_2) = c_{00}(u_2 - u_1). \quad (22)$$

Therefore the answer for the univariate case is easy. Since $dd/dt = 2c_{00}d$ and
 360 $dm/dt = 0$, the dimorphism generated at a branching point will just expand

over evolutionary time. In the multivariate case we get

$$\frac{d\mathbf{d}}{dt} = 2\mathbf{C}_{00}\mathbf{d}, \quad (23a)$$

$$\begin{aligned} \frac{d\mathbf{m}}{dt} &= 2\left(\mathbf{C}_{01}\mathbf{m} - \frac{\mathbf{d}^T(\mathbf{C}_{00} + \mathbf{C}_{01})\mathbf{m}}{\mathbf{d}^T\mathbf{C}_{01}\mathbf{d}}\mathbf{C}_{01}\mathbf{d}\right) + 2\mathbf{C}_{00}\mathbf{m} \\ &= 2\left(\text{id} - \frac{1}{\mathbf{d}^T\mathbf{C}_{01}\mathbf{d}}\mathbf{C}_{01}\mathbf{d}\mathbf{d}^T\right)(\mathbf{C}_{00} + \mathbf{C}_{01})\mathbf{m}, \end{aligned} \quad (23b)$$

362 with id the identity matrix.

Given the simple form of the mutual invadability results for strongly at-
 364 tracting ess-es on the community dynamical time scale, the obvious next
 step seems to be to look under what conditions the vector field specified by
 366 Eq. (23) points towards the interior of \mathcal{D} at points on its boundary. This
 amounts to seeing whether the scalar functions $\mathbf{d}^T\mathbf{C}_{01}\mathbf{d} - \mathbf{d}^T(\mathbf{C}_{00} + \mathbf{C}_{01})\mathbf{m}$
 368 and $\mathbf{d}^T\mathbf{C}_{01}\mathbf{d} + \mathbf{d}^T(\mathbf{C}_{00} + \mathbf{C}_{01})\mathbf{m}$ are bound to increase from their zero values.
 However, the expressions for the time derivatives of these functions, although
 370 simple looking, do not give any clear clues.

The next step is again to look at the large time behaviour of the solutions
 372 of Eq. (23).

In the case of (23a),

$$\mathbf{d}(t) \approx ce^{\lambda_1 t}\mathbf{z}. \quad (24)$$

374 From Eq. (12) and (24) it follows that in the longer run branching can only
 persist if

$$2\mathbf{z}^T\mathbf{C}_{01}\mathbf{z} = -\mathbf{z}^T(\mathbf{C}_{00} + \mathbf{C}_{11})\mathbf{z} < 0. \quad (25)$$

376 (as is the case for strongly attracting ess-es). As we are interested only in cases
 with non-empty coexistence cone *we proceed on the assumption that Inequality*
 378 *(25) holds good.*

From Eq. (23b) it moreover follows that

$$\mathbf{d}^T \frac{d\mathbf{m}}{dt} = 0. \quad (26)$$

380 Hence, in the long run \mathbf{m} either stays bounded, and therefore becomes negligible relative to \mathbf{d} , or becomes orthogonal to \mathbf{z} .

382 To simplify the coming formulas we normalise \mathbf{z} such that $\mathbf{z}^T \mathbf{C}_{01} \mathbf{z} = -1$.
Substituting Eq. (24) in the differential equation for \mathbf{m} then gives

$$\frac{d\mathbf{m}}{dt} \approx 2(\text{id} + \mathbf{C}_{01} \mathbf{z} \mathbf{z}^T)(\mathbf{C}_{00} + \mathbf{C}_{01})\mathbf{m}, \quad (27)$$

384 In view of Eq. (26) the matrix $2(\text{id} + \mathbf{C}_{01} \mathbf{z} \mathbf{z}^T)(\mathbf{C}_{00} + \mathbf{C}_{01})$ has an eigenvalue 0,
and the eigenvectors corresponding to the other eigenvalues are orthogonal to
386 \mathbf{z} . Denote the largest eigenvalue of $2(\text{id} + \mathbf{C}_{01} \mathbf{z} \mathbf{z}^T)(\mathbf{C}_{00} + \mathbf{C}_{01})$ with eigenvector
 \mathbf{w} in the latter class as μ_1 . (We assume here that this eigenvalue is real. The
388 extension of the argument to a pair of complex eigenvalues is immediate but
tedious.) For the inequalities (11) to stay fulfilled

$$e^{-2\lambda_1 t} \mathbf{d}^T (\mathbf{C}_{00} + \mathbf{C}_{01})\mathbf{m} \asymp e^{(\mu_1 - \lambda_1)t} \mathbf{z}^T (\mathbf{C}_{00} + \mathbf{C}_{01})\mathbf{w} = e^{(\mu_1 - \lambda_1)t} \mathbf{z}^T \mathbf{C}_{01} \mathbf{w} \quad (28)$$

390 (\asymp : is asymptotically proportional to) should not grow out of bounds. More-
over, when the expressions in (28) stay bounded, for sufficiently small initial
392 \mathbf{m} the inequalities (11) stay fulfilled.

A sufficient condition for the expressions in (28) to stay bounded is that
394 $\lambda_1 > \mu_1$. This condition is also necessary when $\mathbf{z}^T \mathbf{C}_{01} \mathbf{w} \neq 0$. The condition
 $\mathbf{z}^T \mathbf{C}_{01} \mathbf{w} = 0$ together with the earlier found relations is equivalent to $2\mathbf{C}_{00} \mathbf{z} =$
396 $\lambda_1 \mathbf{z}$, $2(\mathbf{C}_{00} + \mathbf{C}_{01})\mathbf{w} = \mu_1 \mathbf{w}$, $\mathbf{z}^T \mathbf{w} = 0$. Although this of course depends on the
considered model family, the fulfillment of these three conditions together in

398 general is highly non-generic. Hence, generally the conditions

$$\mathbf{z}^T \mathbf{C}_{01} \mathbf{z} \leq 0 \quad \& \quad \lambda_1 > \mu_1 \quad (29)$$

are generically necessary and sufficient to make that for a sufficiently small
400 initial value of \mathbf{m} the two branches remain coexistent at least within the reign
of the local quadratic approximation of the invasion fitness function.

402 **Remark.** The above considerations also apply when dealing with more than
one evolving species. However, in the one-species case considered in this paper,
404 it is possible to make the stronger argument that $2(\mathbf{C}_{00} + \mathbf{C}_{01}) = \mathbf{J}$. As it
only makes sense to consider branching at attracting singular points, \mathbf{J} may
406 be supposed to have only eigenvalues with negative real parts. Hence when
 $\mathbf{z}^T \mathbf{C}_{01} \mathbf{w} = 0$, anyway $\lambda_1 > \mu_1$.

408 So far we have been unable to find an example of an invadable strongly
attracting ess-es that fails to satisfy (29), but neither have we been able to
410 prove that such ess-es do not exist. *So we flag the question whether strong
attraction and invadability together imply locally sustainable branching as open
412 problem.* The next section describes the results in this direction that we could
obtain under various additional assumptions.

414 5 Special cases

In this section we consider a number of special cases for which we could
416 get more information about the possible occurrence of locally unsustainable
branching.

418 We start with the case of 2-dimensional trait spaces. As a first step we
 observe that for such trait spaces we can without loss of generality assume
 420 that

$$C_{00} = \frac{1}{2} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad C_{01} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (30)$$

with $a < 0$ to guarantee the existence of a coexistence cone. (The form (30) can
 422 be reached by choosing the normalised eigenvectors of C_{00} as new orthogonal
 coordinate system. The resulting transformation of the matrices C_{ij} does not
 424 affect λ_1 or μ_1 .) This then gives

$$z = \begin{pmatrix} 1/\sqrt{-a} \\ 0 \end{pmatrix}. \quad (31)$$

μ_1 is the only nonzero eigenvalue of

$$\begin{aligned} & \left[\text{id} + \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} -a^{-1} & 0 \\ 0 & 0 \end{pmatrix} \right] \left[\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} + 2 \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right] \\ & = \begin{pmatrix} 0 & 0 \\ -a^{-1}c\lambda_1 & \lambda_2 + 2a^{-1}(ad - bc) \end{pmatrix}, \end{aligned} \quad (32)$$

426 $\lambda_2 + 2a^{-1}(ad - bc)$. Hence, the expansion of \mathbf{d} dominates, and the expanding
 branches can stay in the coexistence cone if

$$\lambda_1 - \lambda_2 - 2a^{-1}(ad - bc) > 0, \quad (33)$$

428 and only if (33) holds good with $>$ replaced, by \geq .

As it turns out (33) is implied by the requirement that the ess \mathbf{x}^* attracts
 430 for the chosen mutational or genetic covariance matrix. (Remember, λ_1 , λ_2 , a ,
 b , c and d where obtained from the original matrices C_{00} and C_{01} by a change
 432 of basis that transformed Σ into id .) The attractivity of \mathbf{x}^* is determined by

$J = 2(C_{00} + C_{01})$. If and only if the eigenvalues of J have non-positive real part

434 \mathbf{x}^* attracts. This is robustly the case if and only if

$$\text{trace}(J) = \lambda_1 + \lambda_2 + 2a + 2d < 0 \quad \& \quad \det(J) = (\lambda_1 + 2a)(\lambda_2 + 2d) - 4bc > 0 \quad (34)$$

With the help of the functions `Reduce` and `FindInstance` in *Mathematica*

436 (Wolfram Research, Inc.) we found that the Inequalities (34) imply (33).

Hence, locally unsustainable branching cannot occur in 2-dimensional trait

438 spaces.

By following a similar procedure we found instances of locally unsustainable
440 branching in three dimensions (see Appendix). However, when we concentrated
on strongly attracting ess-es *Mathematica* failed to resolve the issue.

442 As we have so far not been able to clarify whether in general strong at-
tractivity guarantees $\lambda_1 > \mu_1$, we went for potentially useful more stringent
444 conditions.

Proposition. *Assume that a coordinate system of the trait space exists such*
446 *that both C_{00} and C_{01} are diagonal matrices with diagonal entries p_{ii} and q_{ii} ,*
respectively. Furthermore, assume C_{00} has a unique largest positive diagonal
448 *entry equal to p_{11} and $p_{ii} + q_{ii} < 0$ for all i . Then $\mu_1 < 0$.*

Proof. Normalize the eigenvector \mathbf{z} corresponding to the dominant eigenvalue
of C_{00} such that $\mathbf{z}^T C_{01} \mathbf{z} = -1$. Then it is easy to see that $M := C_{01} \mathbf{z} \mathbf{z}^T$ has
 $m_{11} = -1$ and zeros elsewhere. Hence, $K := (\text{id} + C_{01} \mathbf{z} \mathbf{z}^T)(C_{00} + C_{01})$ is a
diagonal matrix with $k_{11} = 0$ and $k_{ii} = p_{ii} + q_{ii}$ for $i > 1$. \square

Since by assumption $\lambda_1 > 0$, the conditions of this proposition imply $\lambda_1 >$
450 μ_1 . They are fulfilled e.g. in the Lotka-Volterra models studied by Ackermann
and Doebeli (2004), Doebeli and Ispolatov (2010) and Svardal et al. (2014).

452 **6 Discussion**

The evolutionarily singular strategies of published eco-evolutionary models
454 with multivariate traits often turn out to be strongly attracting (i.e., robustly
convergence stable *sensu* Leimar (2009)). Under the assumption that the latter
456 is the case we established that for the initiation of evolutionary branching it
suffices that the ess is invadable. In several published studies this has been tac-
458 itly assumed, based on the hope that the classical results for one-dimensional
trait spaces extend unmodified to the multi-dimensional case. We thus proved
460 that this is indeed the case for the initiation of branching, but unfortunately
we were not able to prove that under the same conditions a similar statement
462 holds true for its continuation at least within the realm of a quadratic ex-
pansion of the invasion fitness function for the dimorphism. To arrive at these
464 results we analysed the geometry of mutual invadability around general mul-
tivariate ess-es. This gave the initial positive result. Next we derived criteria
466 for checking whether the initial mutual invadability extends to the expand-
ing dimorphisms that ensue from disruptive selection. However, here we could
468 not make a link with the negative definiteness of the Jacobian matrix of the
monomorphic selection gradient at the ess (the signature of strong attractivity)
470 due to the cross-derivatives in the Taylor expansion of the monomorphic inva-
sion fitness function turning up in the dimorphic selection gradients. Hence,

472 whether a branching point indeed spawns temporarily persisting branches on
the evolutionary time scale is a question which in general requires separate
474 investigation. The relations between the five conditions for branching, (i) at-
traction, (ii) invadability, (iii) nearby population dynamical coexistence, (iv)
476 nearby disruptive selection, (v) nearby evolutionary coexistence, given in the
introduction thus become: conditions (i) and (ii) are independent, (iii) is im-
478 plied by (i) and (ii) in the strongly convergent stable case (which includes
the case of one-dimensional trait spaces), but in general is independent, (ii)
480 and (iii) always imply (iv), finally (iii) is necessary for (v) but not sufficient,
except when the trait space is one- or two-dimensional. However, we were un-
482 able to resolve whether in the strongly convergent stable case (iii) implies (v).
Hence, the best we could do was give some appreciably stronger conditions un-
484 der which the latter implication holds good. Finally, and perhaps biologically
most relevantly, we established that, within the reign of the local quadratic
486 approximation of the fitness function, expanding polymorphisms around ess-es
in general initially evolve towards (quasi)-dimorphisms, or after a short while
488 fall back to (quasi)-monomorphism (which in case the ess attracts will again
lead to an expanding polymorphism in an ever ongoing cycle).

490 **Acknowledgements** The authors gratefully acknowledge the help of Mattias Siljestam
with the numerics (including a number of stochastic simulations the results of which did
492 not end up in the paper, but which were quite informative as they eventually suggested
that locally unsustainable branching cannot occur in two-dimensional trait spaces). This
494 work benefitted from the support from the "Chair Modélisation Mathématique et Biodiver-
sité of Veolia Environnement-Ecole Polytechnique-Museum National d'Histoire Naturelle-
496 Fondation X".

Appendix: 3-dimensional trait spaces

498 The case of 3-dimensional trait spaces proceeds analogous to that of 2-dimensional
trait spaces in Section 5. Let $\theta_i := \frac{1}{2}\lambda_i$ and $\kappa_i = \frac{1}{2}\mu_i$, and

$$\mathbf{C}_{00} := \begin{pmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & \theta_3 \end{pmatrix}, \quad \mathbf{C}_{01} := \begin{pmatrix} p & q & r \\ u & v & w \\ x & y & z \end{pmatrix},$$

500 with $p < 0$ to guarantee the local existence of a coexistence cone. The Jacobian
matrix of the selection gradient at \mathbf{x}^* then becomes

$$\mathbf{J} = 2 \begin{pmatrix} p + \theta_1 & q & r \\ u & v + \theta_2 & w \\ x & y & z + \theta_3 \end{pmatrix}.$$

502 For \mathbf{x}^* to attract all eigenvalues of \mathbf{J} should have negative real part. The
Routh-Hurwitz criteria tell that this is the case if and only if

$$(i) \ a_1 = -\text{trace}\left(\frac{1}{2}\mathbf{J}\right) > 0 \ \& \ (ii) \ a_3 = -\det\left(\frac{1}{2}\mathbf{J}\right) > 0 \ \& \ (iii) \ a_1 a_2 > a_3,$$

504 where a_1 to a_3 are the coefficients of the characteristic polynomial $\lambda^3 + a_1\lambda^2 +$
 $a_2\lambda + a_3$ of $\frac{1}{2}\mathbf{J}$. These inequalities evaluate to

$$(i) \ p + v + z + \theta_1 + \theta_2 + \theta_3 < 0$$

$$(ii) \ pvz + qwx + ruy - pwy - quz - rvx + vz\theta_1 - wy\theta_1 + pz\theta_2$$

$$- rx\theta_2 + pv\theta_3 - qu\theta_3 + z\theta_1\theta_2 + p\theta_2\theta_3 + v\theta_1\theta_3 + \theta_1\theta_2\theta_3 < 0$$

$$(iii) \ (p + \theta_1)^2(v + \theta_2) + (p + \theta_1)(v + \theta_2)^2 + (p + \theta_1)^2(z + \theta_3) + (p + \theta_1)(z + \theta_3)^2$$

$$+ (v + \theta_2)^2(z + \theta_3) + (v + \theta_2)(z + \theta_3)^2 + 2(p + \theta_1)(v + \theta_2)(z + \theta_3)$$

$$< qwx + ruy + qu(p + \theta_1 + v + \theta_2) + rx(p + \theta_1 + z + \theta_3) + wy(v + \theta_2 + z + \theta_3).$$

506 Criteria for the strong attraction of \mathbf{x}^* can be derived by applying the Routh-Hurwitz criteria to

$$\frac{1}{2} (\mathbf{J} + \mathbf{J}^T) = \begin{pmatrix} 2(p + \theta_1) & (q + u) & (r + x) \\ (q + u) & 2(v + \theta_2) & (w + y) \\ (r + x) & (w + y) & 2(z + \theta_3) \end{pmatrix},$$

508 which results in

$$(i) \quad p + v + z + \theta_1 + \theta_2 + \theta_3 < 0$$

$$(ii) \quad 4(p + \theta_1)(v + \theta_2)(z + \theta_3) + (q + u)(w + y)(r + x) \\ - (w + y)^2(p + \theta_1) - (r + x)^2(v + \theta_2) - (q + u)^2(z + \theta_3) < 0$$

$$(iii) \quad 4[2(p + \theta_1)(v + \theta_2)(z + \theta_3) + (p + \theta_1)^2(v + \theta_2) + (p + \theta_1)^2(z + \theta_3) \\ + (p + \theta_1)(v + \theta_2)^2 + (p + \theta_1)(z + \theta_3)^2 + (v + \theta_2)^2(z + \theta_3) + (v + \theta_2)(z + \theta_3)^2] \\ < (q + u)(w + y)(r + x) + (r + x)^2(p + \theta_1) + (q + u)^2(p + \theta_1) \\ + (w + y)^2(v + \theta_2) + (q + u)^2(v + \theta_2) + (w + y)^2(z + \theta_3) + (r + x)^2(z + \theta_3).$$

The conditions for locally sustainable branching are that the eigenvalues of

$$\left[\text{id} + \begin{pmatrix} p & q & r \\ u & v & w \\ x & y & z \end{pmatrix} \begin{pmatrix} -p^{-1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right] \begin{pmatrix} \theta_1 + p & q & r \\ u & \theta_2 + v & w \\ x & y & \theta_3 + z \end{pmatrix} - \theta_1 \text{id} \\ = \begin{pmatrix} -\theta_1 & 0 & 0 \\ -u\theta_1/p & \theta_2 - \theta_1 + v - uq/p & w - ur/p \\ -x\theta_1/p & y - xq/p & \theta_3 - \theta_1 + z - xr/p \end{pmatrix}$$

510 have negative real part. (Rationale: The real part of the rightmost eigenvalue κ_1 of $\mathbf{K} := (\text{id} + \mathbf{C}_{01} \mathbf{z} \mathbf{z}^T)(\mathbf{C}_{00} + \mathbf{C}_{01})$ should be smaller than θ_1 . The eigenvalues

512 of $K - \theta_1 \text{id}$ equal $\kappa_i - \theta_1$. So the statement above is equivalent to $\text{Re}(\kappa_i - \theta_1) < 0$
 for all i .) This is the case if and only if the trace of

$$\begin{pmatrix} \theta_2 - \theta_1 + v - uq/p & w - ur/p \\ y - xq/p & \theta_3 - \theta_1 + z - xr/p \end{pmatrix}$$

514 is negative and the determinant is positive. This can be written as

$$p(2\theta_1 - \theta_2 - \theta_3) + qu + rx - pv - pz < 0$$

and

$$\begin{aligned} p(\theta_1 - \theta_2)(\theta_1 - \theta_3) + (rx - pz)(\theta_1 - \theta_2) + (qu - pv)(\theta_1 - \theta_3) \\ + p(vz - wy) + q(wx - uz) + r(uy - vx) < 0. \end{aligned}$$

516 Finding cases where \mathbf{x}^* attracts and the branching is either locally sus-
 tainable or not, using a mixture of inspired guesses with a little help from
 518 *Mathematica* (Wolfram Research, Inc.), turned out not to be too difficult.
 However, in the case where \mathbf{x}^* strongly attracts both *Mathematica* and we
 520 were unable to resolve the inequalities.

References

- 522 Ackermann, M., and M. Doebeli. 2004. Evolution of niche width and adaptive
diversification. *Evolution* 58:2599–2612.
- 524 Beltman, J. B., and J. A. J. Metz. 2005. Speciation: more likely through a
genetic or through a learned habitat preference? *Proceedings of the Royal*
526 *Society London B* 272:1455–1463.
- Champagnat, N. 2003. Convergence of adaptive dynamics n -morphic jump
528 processes to the canonical equation and degenerate diffusion approximation.
Preprint of the University of Nanterre (Paris 10) No. 03/7.
- 530 Champagnat, N., and S. Méléard. 2011. Polymorphic evolution sequence and
evolutionary branching. *Probability Theory and Related Fields* 151:45–94.
- 532 Christiansen, F. B., and V. Loeschke. 1987. Evolution and intraspecific com-
petition. III. One-locus theory for small additive gene effects and multidimensional resource qualities. *Theoretical Population Biology* 31:33–46.
- 534 Collet, P., S. Méléard, and J. A. J. Metz. 2013. A rigorous model study of the
adaptive dynamics of mendelian diploids. *Journal of Mathematical Biology*
536 67:569–607.
- 538 Dercole, F., and S. A. H. Geritz. submitted. Unfolding the resident invader
dynamics of similar strategies. *Journal of Theoretical Biology* .
- 540 Dercole, F., and S. Rinaldi. 2008. *Analysis of Evolutionary Processes: The
Adaptive Dynamics Approach and Its Applications*. Princeton University
542 Press, Princeton, NJ.
- Dieckmann, U., and R. Law. 1996. The dynamical theory of coevolution: A
544 derivation from stochastic ecological processes. *Journal of Mathematical*

Biology 34:579–612.

546 Doebeli, M. 2011. Adaptive Diversification, vol. 48 of *Monographs in Population Biology*. Princeton University Press.

548 Doebeli, M., and I. Ispolatov. 2010. Complexity and diversity. *Science* 328:494–497.

550 Durinx, M., G. Meszéna, and J. A. J. Metz. 2008. Adaptive dynamics for physiologically structured population models. *Journal of Mathematical Biology* 56:673–742.

552 Ferrière, R., and M. Gatto. 1995. Lyapunov exponents and the mathematics of invasion in oscillatory of chaotic populations. *Theoretical Population Biology* 48:126–171.

556 Fisher, R. A. 1958. *The genetical theory of natural selection*. 2nd ed. Dover, New York.

558 Geritz, S. A. H. 2005. Resident-invader dynamics and the coexistence of similar strategies. *Journal of Mathematical Biology* 50:67–82.

560 Geritz, S. A. H., É. Kisdi, G. Meszéna, and J. A. J. Metz. 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology* 12:35–57.

562 Geritz, S. A. H., J. A. J. Metz, É. Kisdi, and G. Meszéna. 1997. Dynamics of adaptation and evolutionary branching. *Physical Review Letters* 78:2024–2027.

566 Ito, H. C., and M. Shimada. 2007. Niche expansion: coupled evolutionary branching of niche position and width. *Evolutionary Ecology Research* 9:675–695.

- Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: Body size allometry. *Evolution* 33:402–416.
- . 1982. A quantitative genetic theory of life history evolution. *Ecology* 63:607–615.
- Leimar, O. 2001. Evolutionary change and darwinian demons. *Selection* 2:65–72.
- . 2009. Multidimensional convergence stability. *Evolutionary Ecology Research* 11:191–208.
- Méléard, S., and V. C. Tran. 2009. Trait substitution sequence process and canonical equation for age-structured populations. *Journal of Mathematical Biology* 58:881–921.
- Metz, J. A. J. 2008. Fitness. Pages 1599–1612 *in* S. Jørgensen and B. Fath, eds. *Evolutionary Ecology*. Vol. [2] of *Encyclopedia of Ecology*. Elsevier.
- Metz, J. A. J. 2014. Fitness. Reference Module in Earth Systems and Environmental Sciences, Elsevier, 2014. 11-Sep-14 doi: 10.1016/B978-0-12-409548-9.09361-1.
- Metz, J. A. J., and C. G. F. de Kovel. 2013. The canonical equation of adaptive dynamics for mendelian diploids and haplo-diploids. *Interfaces Focus* 3:20130025.
- 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. Pages 183–231 *in* S. van Strien and S. Verduyn Lunel, eds. *Stochastic and spatial structures of dynamical systems*, Proceedings of the Royal Dutch

Academy of Science. North Holland, Dordrecht, Netherlands; available at

594 <http://www.iiasa.ac.at/Research/ADN/Series.html>.

Metz, J. A. J., and V. A. A. Jansen. in prep. Adaptive dynamics for mendelian

596 genetics: relating the speed of evolution to the effective population size .

Metz, J. A. J., R. M. Nisbet, and S. A. H. Geritz. 1992. How should we define

598 ‘fitness’ for general ecological scenarios? *Trends in Ecology and Evolution*
7:198–202.

600 Ravigné, V., U. Dieckmann, and I. Olivieri. 2009. Live where you thrive: Joint

evolution of habitat choice and local adaptation facilitates specialization

602 and promotes diversity. *The American Naturalist* 174:E141–E169.

Rueffler, C., T. J. M. Van Dooren, O. Leimar, and P. A. Abrams. 2006. Disrup-

604 tive selection and then what? *Trends in Ecology and Evolution* 21:238–245.

Svardal, H., C. Rueffler, and M. Doebeli. 2014. Organismal complexity and

606 the potential for evolutionary diversification. *Evolution* 3248–3259.

Svardal, H., C. Rueffler, and J. Hermisson. 2011. Comparing environmental

608 and genetic variance as adaptive response to fluctuating selection. *Evolution*
65:2492–2513.

610 Tran, V. C. 2006. Modèles particuliers stochastiques pour des problèmes

d’évolution adaptative et pour l’approximation de solutions statistiques.

612 Ph.D. thesis. Université Paris X - Nanterre, 12., <http://tel.archives-ouvertes.fr/tel-00125100>.

614 Vukics, A., J. Asbóth, and G. Meszéna. 2003. Speciation in multidimensional

evolutionary space. *Physical Review E* 68:041903.

616 Wolfram Research, Inc. 2015. *Mathematica* 10.3. Champaign, Illinois.