

The adaptive dynamics of function-valued traits

Ulf Dieckmann^{a,b,*}, Mikko Heino^{a,c,d}, Kalle Parvinen^e

^a*Evolution and Ecology Program, International Institute for Applied Systems Analysis, Schlossplatz 1, A-2361 Laxenburg, Austria*

^b*Section Theoretical Biology, Institute of Biology, Leiden University, Kaiserstraat 63, 2311 GP Leiden, The Netherlands*

^c*Institute of Marine Research, P.O. Box 1870 Nordnes, N-5817 Bergen, Norway*

^d*Department of Biology, University of Bergen, P.O. Box 7800, N-5020 Bergen, Norway*

^e*Department of Mathematics, FIN-20014 University of Turku, Finland*

Received 28 April 2005; received in revised form 2 December 2005; accepted 2 December 2005

Available online 3 February 2006

Abstract

This study extends the framework of adaptive dynamics to function-valued traits. Such adaptive traits naturally arise in a great variety of settings: variable or heterogeneous environments, age-structured populations, phenotypic plasticity, patterns of growth and form, resource gradients, and in many other areas of evolutionary ecology. Adaptive dynamics theory allows analysing the long-term evolution of such traits under the density-dependent and frequency-dependent selection pressures resulting from feedback between evolving populations and their ecological environment. Starting from individual-based considerations, we derive equations describing the expected dynamics of a function-valued trait in asexually reproducing populations under mutation-limited evolution, thus generalizing the canonical equation of adaptive dynamics to function-valued traits. We explain in detail how to account for various kinds of evolutionary constraints on the adaptive dynamics of function-valued traits. To illustrate the utility of our approach, we present applications to two specific examples that address, respectively, the evolution of metabolic investment strategies along resource gradients, and the evolution of seasonal flowering schedules in temporally varying environments.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Function-valued traits; Adaptive dynamics; Canonical equation; Evolutionary constraints; Evolutionary branching

1. Introduction

Many adaptive features of biological organisms are best described as function-valued traits. Only scalar-valued adaptive traits can be captured by single variables; when many such variables are needed to quantify a trait, the trait is called vector-valued. More complex adaptive traits, however, require characterizing the variation of phenotypic components along a continuum: when given as a function of some other variable, such a trait is called function-valued. It is then the entire shape of the function that is subject to evolutionary change through mutation and selection, which is why such traits are also known as infinite-dimensional. Understanding the evolution of function-valued traits serves as an important step toward better

respecting the complexity of adaptive processes in evolutionary ecology.

For any function-valued trait, there is a quantity determining which specific component of the full function-valued trait is considered: we will refer to these quantities as the arguments of the function-valued trait. Examples of function-valued traits abound:

- In studies of phenotypic plasticity, the argument is given by environmental conditions, like temperature or salinity, and the function-valued trait is the reaction norm of an organism, which describes the phenotype that is expressed in response to a particular environmental condition.
- When considering the demography of physiologically structured populations, the argument may be given by age, weight, or size, while function-valued traits measure how, for instance, resource allocation to reproduction, or the rate of dispersal vary with age, weight, or size.

*Corresponding author. IIASA-EEP, A-2361 Laxenburg, Austria.

Tel.: +43 2236 807 386; fax: +43 2236 71313.

E-mail address: dieckmann@iiasa.ac.at (U. Dieckmann).

- In spatial ecology, the argument will often describe spatial distance: function-valued traits then measure the amount of competition or of dispersal occurring over a continuum of relevant distances.
- In resource utilization theory, the argument will frequently characterize types of resource—corresponding to gradients of size, quality, or of any other relevant continuous property in which resources could differ. A function-valued trait may then describe the harvesting rate or resource utilization spectrum of a consumer.
- In morphological evolution, function-valued traits can be employed to describe physical shapes or tissue composition. In the first case, the argument could be, e.g. the axial coordinate of a leaf, and the function-valued trait the leaf's width at each of these axial coordinates. In the second case, the argument could be, e.g. the radius measured from the center of a stem, and the function-valued trait could characterize the density of woody tissue along this radius.
- In the study of iterated pairwise games, the argument could be an opponent's last investment toward a focal individual, and the function-valued trait could characterize the level of investment with which the focal individual reciprocates.

In all these diverse cases, the considered adaptive trait is represented most naturally and accurately by a function of the argument. It is also self-evident from these illustrations that function-valued traits will rarely be selectively neutral: we must expect that some function shapes will amount to improved adaptation, so that variant phenotypes expressing such shapes will be able to invade and replace inferior resident types. In general, therefore, evolution will proceed to vary the shape of these functions until no selective advantage can be gained any more through such variation. In this study, we present a general framework for analysing function-valued adaptive dynamics.

In devising this framework, we build on two lines of preceding work. First, a comprehensive approach to describing the evolution of function-valued traits based on the methods of quantitative genetics has been developed by Mark Kirkpatrick, Richard Gomulkiewicz, and coworkers (Kirkpatrick and Heckman, 1989a; Kingsolver et al., 2001; and subsequent references in this paragraph). Two types of function-valued traits that have received particular attention in the context of this approach are growth trajectories, where the argument of the function-valued trait is age and the trait itself measures expected body size (Kirkpatrick, 1988, 1993; Kirkpatrick and Lofsvold, 1989, 1992; Kirkpatrick et al., 1990), and reaction norms, where the argument measures an environmental condition and the function-valued trait characterizes the phenotypes expressed in response to these conditions (Gomulkiewicz and Kirkpatrick, 1992). The mathematical structures underlying the evolution of function-valued traits in quantitative genetics have been elucidated by Gomulkiewicz and Beder (1996) and by

Beder and Gomulkiewicz (1998); these authors also derived results that facilitate analysis of several interesting classes of fitness functions used in quantitative genetics. Furthermore, the practical relevance of function-valued adaptation for livestock breeding has been pointed out (Kirkpatrick et al., 1994; Kirkpatrick, 1997).

Until now, analyses of function-valued evolution have been focused on evolution under frequency-independent selection. The pioneering studies highlighted in the preceding paragraph therefore did not yet concentrate attention on evolutionary processes in which the success of a variant phenotype depends on which other phenotypes are currently resident in the evolving population. Such dependence occurs when the ecological environment experienced by variant phenotypes is affected by the frequency of resident phenotypes. While the environment in breeding experiments can sometimes be so tightly controlled that frequency dependence is avoided, this is rarely the case for natural evolution: whenever individuals are interacting in ways affected by their adaptive traits, and the growth of populations is limited, density- and frequency-dependent selection arise generically. The framework of adaptive dynamics has been designed to study the course and outcome of long-term evolution under such conditions (Metz et al., 1992, 1996a; Dieckmann 1994, 1997; Dieckmann and Law, 1996; Geritz et al., 1997, 1998; Meszéna et al., 2001). By assuming asexual inheritance and mutation-limited evolution, adaptive dynamics is trading genetic detail for ecological realism: the effects of a wide range of density-dependent and frequency-dependent selection mechanisms, both in fecundity and in survival, can thus be included into the analysis of long-term evolution. In this way models of adaptive dynamics allow investigating evolutionary processes that are driven by realistic ecological conditions. A strong point of this framework is the intimate link it establishes between evolutionary dynamics and the underlying population dynamics. This results in fitness measures being rigorously derived for a particular ecological setting, instead of being just postulated.

Here we provide a synthesis of the two recent lines of evolutionary research described above. In other words, we introduce evolutionary dynamics that are particularly suited to study the adaptation of function-valued traits when this adaptation is driven by density-dependent and frequency-dependent selection. The structure of this article is as follows. Section 2 is setting the stage by cautioning against pitfalls that may be overlooked when modeling the evolution of function-valued traits. Section 3 presents three alternative models for the adaptive dynamics of function-valued traits: these models establish a close link between individual-based ecologies and the expected evolutionary trajectories resulting from processes of mutation and selection. Section 4 characterizes types of outcomes of function-valued evolution. Section 5 highlights the important role of constraints for evolving function-valued traits, and explains how to account for such constraints in

practice. Sections 6 and 7 provide two worked examples that apply the framework introduced here. These examples focus, in turn, on the evolution of metabolic investment strategies and on the evolution of seasonal flowering schedules.

2. Pitfalls in modeling function-valued adaptive dynamics

Paying too little attention to proper methods for analysing function-valued evolution is risky for at least three reasons:

- In the absence of an adequate framework, interesting questions of evolutionary ecology that critically require investigating the evolutionary dynamics of function-valued traits under frequency-dependent selection will remain unstudied. Overcoming such a state of affairs is the purpose of this article.
- Particular attention must be devoted to constraints on function-valued evolution. Without proper methods for, and experience with, accounting for such constraints, essential aspects of function-valued adaptation are likely to be misrepresented. We will elaborate on this caveat briefly below, and more extensively in Section 5.
- It is tempting to reduce, for simplicity's sake, the evolution of function-valued traits to that of vector-valued traits, by using a low number of parameters to approximate a specific function-valued trait. While such a simplification may often seem desirable at first sight, the danger of drawing spurious and misleading conclusions is considerable, as we will detail in the remainder of this section.

The significance of the second caveat derives from the fact that the higher the dimension of an adaptive trait, the more important becomes the mutational variance–covariance underlying its evolution. Such variance–covariance can be interpreted as imposing evolutionary constraints on the adaptation of vector-valued and function-valued traits: mutational variation in certain components of such traits may either be absent (variance constraints) and/or inevitably linked to variation in other components (covariance constraints). For function-valued traits, such constraints are most naturally expressed, derived, and analysed in terms of variance–covariance functions, as we will explain in Section 5.

Fig. 1 illustrates the third caveat. Let us assume that we wish to describe the evolution of a function-valued trait $x(a)$ along a continuum of argument values $0 \leq a \leq 1$. (Detailed descriptions of the underlying example, its ecological motivation, and the resulting selection pressures are given in Section 6; here we are only interested in the evolution of this trait as a means to highlight some general qualitative insights.) Function-valued traits are sometimes called “infinite-dimensional” since their accurate quantification requires specifying the value of $x(a)$ at infinitely many argument values a across the considered continuum.

It is therefore common to try and reduce this dimensionality. Based on some a priori intuition about expected evolutionary outcomes, we might decide, for example, to parameterize this particular function-valued trait in either of three ways: (i) by the amplitude c and parameter p of an exponential function, $x(a) = cE_p(a)$ with $E_p(a) = e^{pa}$; (ii) by the amplitude c , the mean m , and the standard deviation d of a normal function, $x(a) = cN_{m,d}(a)$; or (iii) by the amplitude c and parameters d and k of a sine function $x(a) = cS_{k,d}(a)$ with $S_{k,d}(a) = \sin[2\pi k(a - d)] + 1$. With each of these three choices we have reduced the infinite-dimensional trait $x(a)$ to a merely two- or three-dimensional vector-valued trait. This sounds attractive—yet, each choice demonstrates a different aspect of why such reduction can be misleading. Figs. 1a–d show evolutionary dynamics resulting from the three alternative parameterizations (initial conditions are shown as dotted curves, transients as thin curves, and evolutionary outcomes as thick curves):

- Fig. 1a reveals that the exponential parameterization is grossly inadequate since it is incapable of even qualitatively capturing the actual evolutionary outcome. It is important to realize, however, that this inference cannot be drawn from the vector-valued evolution as shown in the figure, but only from our knowledge of the outcome of the actual function-valued evolution as derived later in this article (shown in gray).
- Fig. 1b indicates that also the normal parameterization is of limited utility, since, although it does provide a rough approximation of the actual evolutionary outcome (compare the thick curve with the gray curve), it misses out on important qualitative features, like the significantly asymmetric shape of the actual solution and the fact that the actual solution vanishes at particular argument values.
- A particularly important and insidious discrepancy between parameterized and function-valued evolution is highlighted in Figs. 1c and d. While here the chosen parameterization would allow, in principle, to approximate the actual solution reasonably well (Fig. 1c), the parameterized evolutionary dynamics easily get trapped in spurious evolutionary attractors (Fig. 1d, where the initial condition is varied only very slightly compared to Fig. 1c). Low-dimensional parameterizations lead to spurious outcomes, since these parameterizations are prone to block so-called “extra-dimensional bypasses” (Conrad 1990). Such blockage is an undesired effect often imposed by low-dimensional dynamics and is illustrated schematically in Fig. 1e.

Fig. 1f finally shows the function-valued adaptive dynamics, as derived in Section 6, which is free from all these confounding impediments. Notice that the three types of pitfalls in parameterizing function-valued traits illustrated here can be hard to detect in practice: without analysing the function-valued evolution itself, it is

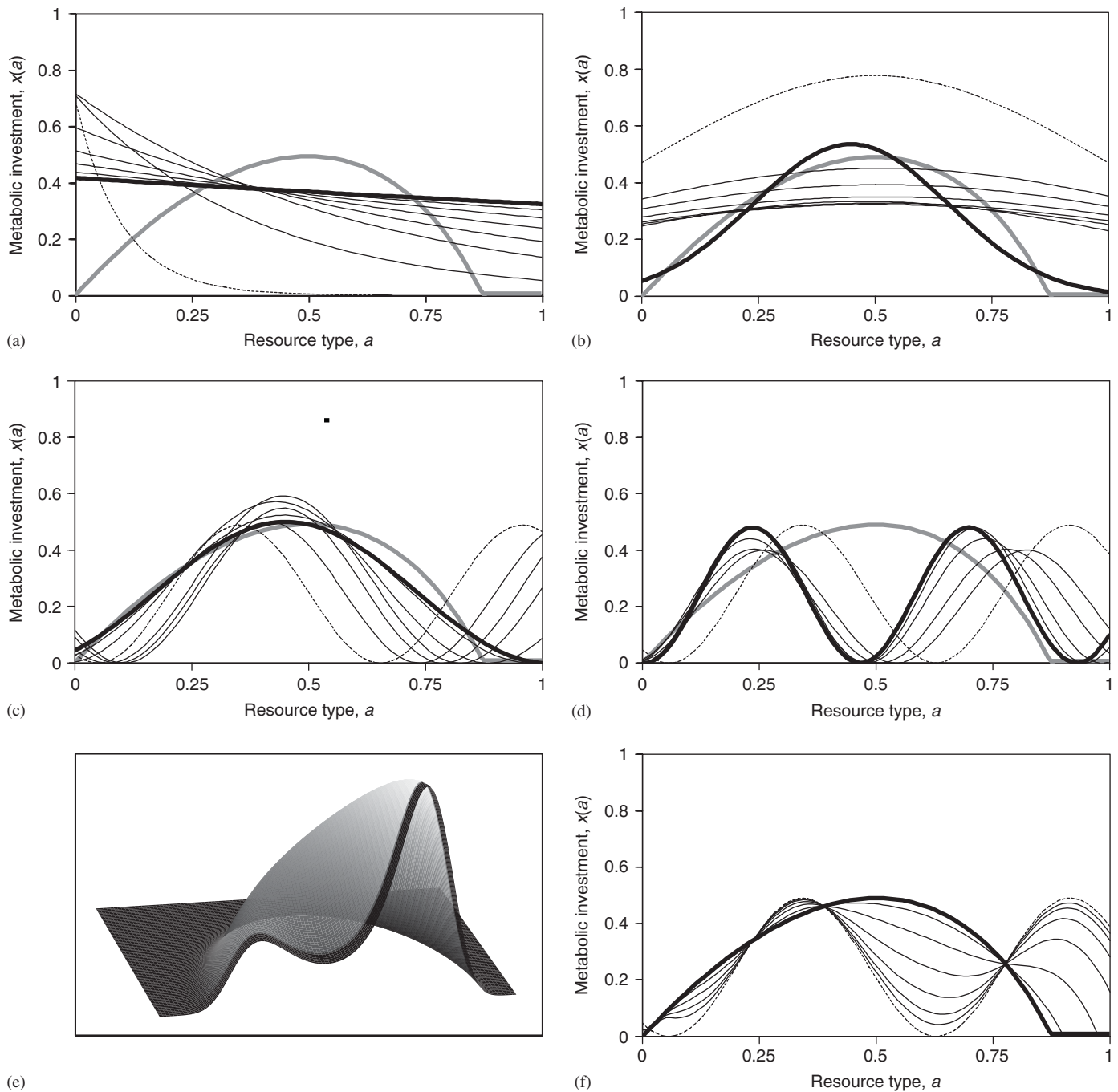


Fig. 1. Pitfalls in parametrizing function-valued traits. Panels (a)–(d) illustrate the evolutionary dynamics and outcomes resulting from different low-dimensional approximations of a function-valued trait. Dynamics are determined by the canonical equation of vector-valued adaptive dynamics (Dieckmann and Law, 1996; with mutational variances equal and covariances absent). Initial trait values are shown as thin dotted curves, intermediate values as thin continuous curves, and final values as thick continuous curves. Parametrizations: (a) exponential function, (b) normal function, (c) and (d) sine function (shown for two different initial trait values). Panel (e) illustrates an extra-dimensional bypass (Conrad, 1990). On the one-dimensional fitness landscape (black curve), two local maxima exist. By contrast, on the two-dimensional fitness landscape (gray surface) these points are connected by a monotonic ridge, resulting in a single local maximum. Panel (f) shows the dynamics and outcome resulting from function-valued evolution (as derived in Section 6). For easier comparison, this outcome is also indicated by continuous gray curves in panels (a)–(d).

impossible to decide which one of the alternative evolutionary predictions provided in Figs. 1a–d can actually be trusted. Only by investigating function-valued evolution directly, such problems and ambiguities can be overcome.

The considerations above lead us to realize that, while function-valued traits may seem complex at first sight, they compare favorably with seemingly simpler representations, by offering a more natural and less treacherous platform for many interesting studies in evolutionary ecology.

3. Models of function-valued adaptive dynamics

We now consider models for describing the evolution of a function-valued trait $x(a)$. We refer to the specific form or shape which the function x takes in an individual as that individual's trait value. The variable a belongs to what we call the trait's argument space or domain, while the function x itself belongs to the trait space of the considered population. In this article, we focus on the dynamics of a single function-valued trait with a one-dimensional argument; generalizations to the joint evolution of several traits and to multi-dimensional arguments are readily made.

To describe the ecological dynamics underlying function-valued evolution we assume, very generally, that the per capita birth and death rates of individuals depend on their trait value x , as well as on the phenotypic composition p of the population as a whole, $b_x(p)$ and $d_x(p)$. The variable p contains information about abundance and trait values of all individuals present in the population; for mathematical details see Appendix A. In addition to the selection pressures originating from ecological interactions, evolutionary processes require a mechanism for generating phenotypic variation on which natural selection can operate. In asexual organisms such variation is created through mutation. With the mutation probability μ_x a birth event in an individual with trait value x gives rise to mutant offspring with trait value x' , whereas the offspring faithfully inherits the parental trait value x with probability $1 - \mu_x$. The new trait value x' is chosen according to a mutation distribution $M(x', x)$. Natural selection arises from the resulting ecology, in particular from the dependence of the per capita growth rate, $f_x(p) = b_x(p) - d_x(p)$, on the function-valued traits of individuals. Unless the function-valued trait x is selectively neutral, or an evolutionarily stable trait has already been reached, a population with a resident trait value x is bound to be invaded and replaced by mutants with varied trait values x' .

Although individual-based models like the one specified here are healthily close to the ecological processes that underlie evolutionary change, directly following the resultant stochastic dynamics of function-valued traits through individual-based simulations is a fairly complex undertaking. Our goal in this article is therefore to provide a more tractable deterministic model describing the expected course of function-valued evolution. For this purpose, we make three simplifying assumptions:

- We consider populations of sufficient size for the demographic stochasticity of total population size to be negligible. (Under typical circumstances, this means that the resulting approximations are not suited for describing populations of less than about 100 individuals.)
- We adhere to the standard assumption that the adaptive process unfolds on an evolutionary time-scale that is longer than that of ecological change. This means that mutations which are both viable and advantageous are

supposed not to occur too frequently. (Exceptions to this assumption do exist in rapidly mutating organisms like pathogens.)

- We concentrate on processes of gradual adaptation, in which average mutational steps are small. (This focus seems relatively safe, since “hopeful monsters” arising from large mutational jumps are usually not viable in higher organisms.)

Even with these simplifying assumptions, stochastic aspects of evolutionary change remain important. This is for two reasons. First, mutations are inherently stochastic, and thus have to be treated as such. Second, the demographic stochasticity of mutants can never be safely ignored, since mutants enter a population at population size 1. Consequently, the formal link between the individual-based ecology with mutations specified above, which is fast and stochastic, and a description of the expected evolutionary process, which is slow and deterministic, has to be carefully constructed. To this end we consider a hierarchy of three evolutionary models that have previously been devised for studying scalar-valued and vector-valued evolution (Dieckmann, 1994; Dieckmann et al., 1995; Dieckmann and Law, 1996):

- *Polymorphic stochastic model, PSM.* This model provides a full, individual-based description of eco-evolutionary dynamics as specified above. It can be formulated as a function-valued master equation (Dieckmann, 1994), and accounts for the stochastic mutation–selection process in polymorphic populations, i.e. in populations that simultaneously comprise individuals with different trait values.
- *Monomorphic stochastic model, MSM.* The second model retains the stochasticity arising from the occurrence and demography of mutants, but assumes that the ecological and the evolutionary time-scale are sufficiently separated. The evolutionary process then is mutation-limited: selection usually has enough time to take effect before a new viable and advantageous mutant enters the population. Consequently, variation in the distribution of trait values is small enough for an assumption of monomorphism to provide a good approximation (Dieckmann and Law, 1996). The MSM describes the evolutionary process as a directed random walk in trait space. Stochastic steps occur when a resident trait value is replaced by a successfully invading advantageous mutant, $x \rightarrow x'$; a series of such substitutions is called a trait substitution sequence (Metz et al., 1992).
- *Monomorphic deterministic model, MDM.* This is a deterministic approximation to the MSM above, based on the assumption of mutational steps being relatively small (Dieckmann and Law, 1996). This model is given in terms of a differential equation describing a trait's expected evolutionary path, and is the main target of this article.

The sequence PSM → MSM → MDM, then, provides a route for deriving the adaptive dynamics of function-valued traits. Details of the derivation are given in Appendix A. One of the essential ingredients in this derivation is the concept of invasion fitness f : after a resident population with trait value x has attained its ecological attractor, it is the average initial per capita growth rate $f(x', x)$ of a new mutant with trait value x' arising in the resident population that determines whether or not the mutant may invade the resident population (Metz et al., 1992; Dieckmann, 1994; Rand et al., 1994; Ferrière and Gatto, 1995; Dieckmann and Law, 1996). When the resident population's ecological attractor is a fixed point, a model's invasion fitness is easily obtained from the underlying vital rates as $f(x', x) = b(x', x) - d(x', x)$ with $b(x', x) = b_x(p)$, $d(x', x) = d_x(p)$, where $p = \bar{n}_x \Delta_x$ denotes the phenotypic distribution of the resident population with trait value x and equilibrium population size \bar{n}_x . While a mutant with $f(x', x) < 0$ cannot invade, a mutant with $f(x', x) > 0$ can (owing to the high risk of accidental extinction at low mutant population size, successful invasions usually happen only after many unsuccessful trials). For mutations of small effect, $x' \approx x$, the condition $f(x', x) > 0$ does not only imply that the mutant can invade, but also that it generically will take over the resident population (Dieckmann, 1994; Geritz et al., 2002).

Appendix A shows that the MDM, and thus the expected adaptive dynamics of function-valued traits, is governed by the following equation:

$$\frac{d}{dt}x(a) = \frac{1}{2}\mu_x\bar{n}_x \int \sigma_x^2(a', a)g_x(a')da'. \quad (1)$$

Eq. (1) is referred to as the canonical equation of function-valued adaptive dynamics (or canonical equation for short). It applies when mutations are rare, small, and symmetrically distributed. Notice that the second and third of these assumptions can easily be relaxed (see the end of Appendix A). Notice also that the factor $\frac{1}{2}$ in the equation above simply reflects the fact that, whenever $g_x \neq 0$, always only exactly one-half of all possible small mutational steps $x \rightarrow x'$ are selectively advantageous. In Eq. (1), the equilibrium population size of the resident population with trait value x is denoted by \bar{n}_x , and σ_x^2 is the variance-covariance function of the mutation distribution M at trait value x ,

$$\sigma_x^2(a', a) = \int [x'(a') - x(a)][x'(a) - x(a)]M(x', x)dx', \quad (2)$$

where the integration extends over all feasible trait values x' . The function g_x is the selection gradient and is obtained as the functional derivative of f at trait value x (Kirkpatrick and Lofsvold, 1992; Gomulkiewicz and Beder, 1996),

$$g_x(a) = \lim_{\Delta x \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \frac{f(x + \varepsilon \Delta x, x) - f(x, x)}{\varepsilon}. \quad (3a)$$

This selection gradient is itself simply a function of a : for each a it describes the strength and direction of selection on

$x(a)$, by probing the sensitivity of invasion fitness f at a . This is accomplished by considering the fitness consequences of perturbations $\varepsilon \Delta x$ around the resident trait value x , as two limits are approached: first, the amplitude ε of these perturbations is sent to 0 (this is the same construction as is familiar from the definition of any scalar derivative), and, second, the perturbations are made increasingly localized around a . The latter is achieved by letting Δx converge to δ_a , the Dirac delta function peaked at location a ; see Appendix B for a gentle introduction to delta functions. Notice that residents are always neutral with regard to invading their own population, $f(x, x) = 0$ for all x , which simplifies the ratio in Eq. (3a) to $f(x + \varepsilon \Delta x, x)/\varepsilon$. We can rewrite Eq. (3a) as

$$g_x(a) = \left. \frac{\partial}{\partial \varepsilon} f(x + \varepsilon \delta_a, x) \right|_{\varepsilon=0}, \quad (3b)$$

yielding a simpler relation that is valuable in applications. When using this second expression for determining selection gradients, we just need to keep in mind that, in accordance with the definition in Eq. (3a), the epsilon derivative is to be taken before the delta function property is exploited in any calculation of expressions like that on the right-hand side of Eq. (3b). Finally, whenever the invasion fitness is given by an integral, $f(x', x) = \int_a^{a+} I(x'(a), x(a), a) da$, a third expression for calculating selection gradients is readily obtained from Eq. (3b),

$$g_x(a) = \left. \frac{\partial}{\partial y} I(y, x(a), a) \right|_{y=x(a)}. \quad (3c)$$

Since integral expressions for invasion fitness occur very frequently in applications, the simple relation in Eq. (3c), free from any consideration of delta functions, turns out to be of great practical utility (Parvinen et al., 2006).

Eqs. (1)–(3) are readily interpreted. The first equation describes the expected evolutionary change in the function-valued trait $x(a)$ and comes in two parts. The first three factors on the right-hand side, including the mutation probability and the average population size, are all positive; therefore, they just scale the rate of evolutionary change. By contrast, the integral on the right-hand side can be either positive or negative, and hence determines the direction evolution takes. According to Eqs. (3), the selection gradient carries information about whether an increase in the current trait value x at argument value a is advantageous ($g_x(a) > 0$) or deleterious ($g_x(a) < 0$). In the former case, $x(a)$ would, at first sight, always be expected to increase, and to decrease in the latter case. This, however, does not yet account for effects of mutational covariance. According to Eq. (2), the function $\sigma_x^2(a', a)$ describes the average mutational side effects that an altered trait value at argument value a has on trait values at all other argument values $a' \neq a$. If, for example, a selective advantage is obtained by increasing x at argument value a , but this increase is inevitably linked to deleterious changes at other argument values a' , then $x(a)$ might, in fact,

decrease. Generally, for each a in Eq. (1), the integral balances the directional effect of $g_x(a)$ at a against directional effects of $g_x(a')$ at all other argument values a' , with the weighting of effects provided by $\sigma_x^2(a', a)$. Evolution thus proceeds in the direction of the net effect.

4. Outcomes of function-valued adaptive dynamics

Outcomes of function-valued evolution can be of different types. According to Eq. (1), three conditions can result in the expected evolutionary change $dx(a)/dt$ to vanish for all argument values a :

- *Selection-induced equilibria.* For the first type of equilibrium, it simply is the selection gradient that vanishes for all argument values a , $g_x(a) = 0$. The location of such equilibria in trait space is thus determined by selective forces alone. It must be noted, however, that, in contrast to their location, the asymptotic stability of these evolutionary equilibria under Eq. (1), determining whether they are attracting or repelling, may critically depend on subtle features of the mutation process (see Marrow et al., 1996 and Leimar, 2001 for examples that demonstrate this fact for two-dimensional vector-valued traits). Explicit criteria for this so-called convergence stability are provided below.
- *Covariance-induced equilibria.* Another important type of equilibrium can result from constraints imposed by the mutation process. If the variance–covariance function is singular, i.e. if $\int \sigma_x^2(a', a) g_x(a') da'$ vanishes for all argument values a and for some function $g_x \neq 0$, then a selection gradient that is proportional to g_x is said to lie in the null space of σ_x^2 , and cannot cause any evolutionary change. In such cases, the restricted availability of mutants around an adaptive trait value x causes positive and negative contributions to the integral in Eq. (1) to cancel: even though further adaptation would be possible without such covariances, evolutionary change is ground to a halt by mutational side effects.
- *Extinction-induced equilibria.* A third type of equilibrium arises when the evolutionary dynamics in Eq. (1) reaches viability boundaries in its trait space. Such boundaries are given by the condition $\bar{n}_x = 0$. In this case, the trait value x results in a non-viable population, and—short of any individuals that could produce mutants—evolutionary change evidently cannot proceed.

Mixtures of the first two fundamental types of evolutionary equilibrium can readily occur: this happens when evolution causes the selection gradient to vanish in parts of argument space, while evolution in the other parts is impeded by covariance constraints (Figs. 2d and e show an example).

Selection-induced evolutionary equilibria are also known as evolutionarily singular strategies or evolutionary singu-

larities (Metz et al., 1996a; Geritz et al., 1998), and are denoted by x^* below. Notice that evolutionary equilibria need not be asymptotically stable under the canonical equation of adaptive dynamics, Eq. (1): if they are attracting, they are called convergence stable (Christiansen, 1991). If a convergence stable evolutionary equilibrium is also locally evolutionarily stable, i.e. if it is immune to invasion by neighboring variants (Hamilton, 1967; Maynard Smith and Price, 1973; Maynard Smith, 1982), it is referred to as a continuously stable strategy (Eshel, 1983). Convergence stability, however, by no means implies local evolutionary stability or vice versa (Eshel and Motro, 1981; Eshel, 1983; Taylor, 1989; Christiansen, 1991; Takada and Kigami, 1991): evolutionary processes can therefore converge to equilibria at which small mutational steps in different directions generate viable invaders that can coexist with the original population at the evolutionary equilibrium. Such equilibria are called evolutionary branching points (Metz et al., 1996a; Geritz et al., 1997, 1998); their importance for understanding adaptation in function-valued traits will be highlighted in Section 7.

To formulate analytic conditions for the convergence stability and local evolutionary stability of an evolutionarily singular function-valued trait x^* , it is helpful first to construct the four second functional derivatives of invasion fitness $f(x', x)$,

$$h_{mm,x^*}(a', a) = \frac{\partial^2}{\partial \varepsilon' \partial \varepsilon} f(x^* + \varepsilon' \delta_{a'} + \varepsilon \delta_a, x^*) \Big|_{\varepsilon'=\varepsilon=0}, \quad (4a)$$

$$h_{mr,x^*}(a', a) = \frac{\partial^2}{\partial \varepsilon' \partial \varepsilon} f(x^* + \varepsilon' \delta_{a'}, x^* + \varepsilon \delta_a) \Big|_{\varepsilon'=\varepsilon=0}, \quad (4b)$$

$$h_{rm,x^*}(a', a) = \frac{\partial^2}{\partial \varepsilon' \partial \varepsilon} f(x^* + \varepsilon \delta_a, x^* + \varepsilon' \delta_{a'}) \Big|_{\varepsilon'=\varepsilon=0}, \quad (4c)$$

$$h_{rr,x^*}(a', a) = \frac{\partial^2}{\partial \varepsilon' \partial \varepsilon} f(x^*, x^* + \varepsilon' \delta_{a'} + \varepsilon \delta_a) \Big|_{\varepsilon'=\varepsilon=0}. \quad (4d)$$

The function h_{mm} is known as the mutant Hessian, h_{rr} as the resident Hessian, while h_{mr} and h_{rm} are referred to as mixed Hessians (note that $h_{mm}^T = h_{mm}$, $h_{rr}^T = h_{rr}$, and $h_{mr}^T = h_{rm}$). The condition for the local evolutionary stability of an evolutionarily singular function-valued trait x^* is that invasion fitness $f(x', x^*)$ possesses a local maximum at $x' = x^*$, which means that the mutant Hessian h_{mm,x^*} is negative definite. To evaluate the convergence stability of x^* , we construct the Jacobian of the right-hand side of Eq. (1) and thus conclude that x^* is convergence stable if and only if the dominant eigenvalue of $\sigma_{x^*}^2(h_{mm,x^*} + h_{mr,x^*})$ possesses a negative real part. From this we see that the convergence stability of x^* may be affected by the variance–covariance function $\sigma_{x^*}^2$. To identify situations in which changes in $\sigma_{x^*}^2$ cannot cause changes in convergence stability, Leimar introduced the very helpful notion of strong convergence stability (Leimar, 2001; Leimar, in press). Following his definition

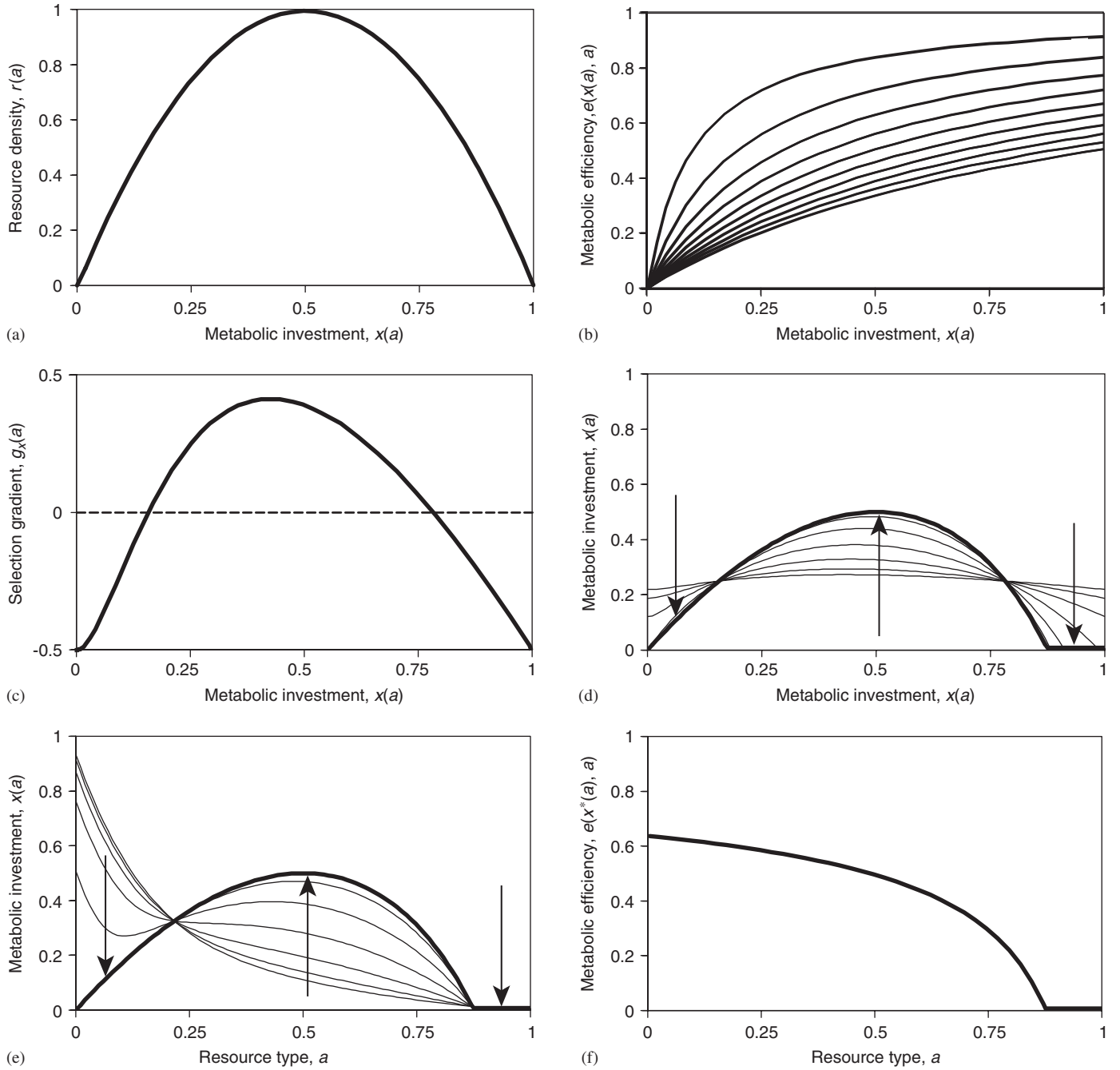


Fig. 2. Evolution of metabolic investment strategies. (a) Density $r(a)$ of different resource types a . (b) Relation between metabolic investment $x(a)$ and metabolic efficiency $e(x(a), a)$ for different resource types $a = 0.1, 0.2, \dots, 1.0$ (top to bottom). (c) Selection gradient $g_x(a)$ for the uniform trait $x(a) = \frac{1}{4}$. (d) and (e) Dynamics and outcomes of the evolution of metabolic investment $x(a)$, for two different initial trait values. Initial and intermediate trait values are shown as thin curves, and final values as thick curves. Arrows indicate the effect of the selection gradient. (f) Metabolic efficiencies $e(x^*(a), a)$ resulting from the evolutionary outcome x^* for different resource types a . Parameters: $c = \frac{1}{2}$ and $r(a) = 4a(1 - a)$.

and reasoning, we define an evolutionarily singular function-valued trait x^* as strongly convergence stable if and only if the symmetric component of $h_{mm,x^*} + h_{mr,x^*}$ is negative definite. Using the relation $h_{mm} + h_{mr} + h_{rm} + h_{rr} = 0$, which follows directly from $f(x, x) = 0$ for all x , we obtain $\frac{1}{2}(h_{mm,x^*} + h_{mr,x^*} + h_{mr,x^*} + h_{rr,x^*}) = \frac{1}{2}(h_{mm,x^*} - h_{rr,x^*})$ for the symmetric component of $h_{mm,x^*} + h_{mr,x^*}$. This means that we may equivalently define the strong convergence stability of x^* by the negative

definiteness of the difference between the mutant and resident Hessians, $h_{mm,x^*} - h_{rr,x^*}$.

5. Constraints on function-valued adaptive dynamics

The canonical equation of function-valued adaptive dynamics, Eq. (1), captures the interplay between the selection gradient and the mutational variance–covariance function of an adapting population. The selection gradient

is determined by the underlying ecology and its trait dependence, while the variance–covariance function reflects how adaptive constraints restrict the course of evolution. In this section we describe and formalize the later relation. As we will see, retaining the natural infinite-dimensional structure of a function-valued trait simplifies such reasoning.

Adaptive constraints on function-valued traits can take various forms and may originate from many causes. Let us look at three examples. First is a dispersal kernel, describing the probability $x(a)$ of a dispersing individual or propagule to settle after having covered a certain distance a . Here, the evolution of trait values x is limited by two separate constraints: first, by (i) $\int x(a)da = 1$, because x is a probability density and the disperser's probability of eventual settlement equals 1; and second, by (ii) $x(a) \geq 0$ for all a , since negative probability densities are not meaningful. Second is a distribution $x(a)$ of foraging times that an herbivore spends across a continuum of plant types a in the course of a day. In this case evolution again ought to respect the constraint $x(a) \geq 0$ for all a . An additional limitation is given by (iii) $\int x(a)da \leq 24\text{ h}$ —the total daily foraging time cannot exceed a day's length. Third is the growth trajectory of an individual, given by its body size $x(a)$ at age a . Once more, we have $x(a) \geq 0$ for all a , but we also need to fix the initial condition before growth starts: (iv) $x(0) = x_0$, where x_0 denotes the body size at birth.

Adaptive constraints on function-valued traits can be classified according to two dichotomies: constraints can be given either by equalities (i and iv) or by inequalities (ii and iii), and such constraints may apply either locally (ii and iv) or globally (i and iii) in argument space. Equality constraints are sometimes also referred to as holonomic, and inequality constraints as non-holonomic. Equality and inequality constraints both limit the range of available trait values, but differ fundamentally in their way of doing so:

- Global equality constraints can always be expressed in the form $c(x) = 0$, where c is a suitable functional that maps trait values x to numbers. For such constraints, any location x in trait space has neighboring trait values that are prohibited by the constraint. A global equality constraint thus restricts adaptive dynamics to a constraint manifold in trait space: for a trait value lying on this manifold, only those neighbors that also lie on it correspond to permissible mutants. Local equality constraints affect trait values only at specific argument values and can formally be expressed as a special case of global equality constraints: this is demonstrated by the equivalence of $c(x) = \int x(a)\delta(a - a_0)da - x_0 = 0$ with $x(a_0) = x_0$. Also this local type of equality constraint thus implies that any trait value is surrounded by excluded trait values.
- The situation is quite different for inequality constraints: these only affect certain regions of trait space where trait values are getting close to a constraint manifold. Global inequality constraints take the form $c(x) \geq 0$. For trait

values x that are not yet in the vicinity of the constraint manifold given by $c(x) = 0$, all neighboring trait values are allowed. As long as mutational effects are mostly small, excluded mutants (situated on the far side of the manifold) are only likely to arise once trait values get close enough to the constraint manifold. Such exclusion thus unfolds as a gradual process: the closer an evolving population gets to the constraint manifold, the more potential mutants will be excluded, until eventually no feasible mutants remain that would take the population across the manifold. The same conclusion applies to local inequality constraints, for which a condition $x(a) \leq c_x^+(a)$ or $x(a) \geq c_x^-(a)$ holds for all or some values a in argument space. Notice that the functions c^\pm may vary with x .

The canonical equation of function-valued adaptive dynamics, Eq. (1), needs no modification for incorporating adaptive constraints. Instead, constraints like those described above require suitable variance–covariance functions to be utilized. How is this achieved? Again, the answer depends on the type of constraint. In particular, equality constraints require singular variance–covariance functions, and also the reverse is true: singular variance–covariance functions imply equality constraints. The null space of a singular variance–covariance function contains all directions of change that, if applied to a function-valued trait, would violate the associated equality constraints. For illustration, let us consider typical equality constraints of global and local type:

- *Global equality constraints* can take many forms. One of the most important example conserves the normalization, $\int x(a)da = 1$, of a probability density x . The simplest variance–covariance function respecting this constraint is

$$\sigma_x^2(a', a) = \sigma^2[\delta(a' - a) - 1/L], \quad (5a)$$

where σ^2 is the unconstrained variance and L is the length of argument space. This choice implies, first, that the variance at all argument values a is the same, since the amplitude of the delta function is uniform along the diagonal $a' = a$. Second, we immediately see from the constancy of $1/L$ that an increase or decrease of $x(a)$ at a particular argument value a is compensated uniformly across all other argument values a' , thus ensuring that the trait's normalization stays intact. A third feature is that, with this choice, there will be no cohesion of trait values $x(a)$ between adjacent argument values a , since the delta function's variance is zero. Taken literally, this is clearly unrealistic. We thus ought to think of the delta ridge along $a' = a$ in Eq. (5a)—or, for that matter, of the delta ridge in the simplest covariance-free function $\sigma_x^2(a', a) = \sigma^2\delta(a' - a)$ —as the idealization of a narrow ridge. Notice that, even though Eq. (5a) describes mutation processes that do not preserve the continuity of function-valued traits, this continuity is retained when considering the expected adaptive dynamics of function-valued traits in

Eq. (1). Again, this subtlety vanishes for variance–covariance functions with a narrow ridge instead of a delta ridge. An alternative normalization-preserving variance–covariance function that overcomes the somewhat unrealistic third feature, and may thus be of particular importance in many applications, is derived in Appendix C. The evolutionary implications of such a function are explored in Heino et al. (submitted). Other global equality constraints are dealt with by a variant of the method described for global inequality constraints; see below.

- *Local equality constraints* prevent the trait value at a particular argument value a_0 from changing, $x(a_0) = x_0$. Consequently, all corresponding mutational variances and covariances have to be zero,

$$\sigma_x^2(a', a_0) = \sigma_x^2(a_0, a) = 0 \quad \text{for all } x, a, \text{ and } a'. \quad (5b)$$

How gradually or abruptly such required troughs in the two-dimensional variance–covariance landscape are attained is at a modeler's discretion: if a local equality constraint at a_0 should affect trait evolution at adjacent argument values, variance–covariance functions should be continuous around the trough's bottom at 0. Local equality constraints can of course also be defined for continuous ranges of argument values, $a_0 \in A_0$. While the implications for the corresponding variance–covariance functions are obvious, such continuous ranges will rarely be of relevance in practice, since the interior of A_0 can then just as well be excluded from the evolutionary model.

Inequality constraints imply that the mutational variance–covariance function changes gradually across a boundary layer adjacent to the constraint manifold: sufficiently far away from the manifold the variance–covariance function is essentially unaffected, while very close to a smooth constraint manifold exactly one-half of all potential mutants is excluded. In some studies, this gradual transition might be of interest in its own right; in most cases, however, the complex details of how the variance–covariance function varies close to the constraint manifold do not matter at all, and using a simpler, approximate representation is then advisable. Such approximations are motivated by the fact that the canonical equation of adaptive dynamics is derived for small mutational steps (Section 3 and Appendix A). Accordingly, the boundary layer induced by inequality constraints will be very narrow whenever the canonical equation offers a valid description. This is why it will mostly be convenient to assume the layer's thickness to be zero, by utilizing approximate variance–covariance functions:

- *Global inequality constraints*, $c(x) \geq 0$, can be represented by linearly transforming an unconstrained variance–covariance function $U_x(a', a)$,

$$\sigma_x^2(a', a) = \iint \tilde{P}_x(a', a'') U_x(a'', a''') \tilde{P}_x(a''', a) da'' da''', \quad (5c)$$

with $\tilde{P}_x(a', a) = \delta(a' - a) - H(-c(x))H(-\int g_x(a'')\tilde{N}_x(a'')da'')P_x(a', a)$, $P_x(a', a) = \tilde{N}_x(a')\tilde{N}_x(a)$, $\tilde{N}_x(a) = N_x(a)/\sqrt{\int N_x^2(a')da'}$, and $N_x(a) = (\partial/\partial\epsilon)c(x + \epsilon\delta_a)|_{\epsilon=0}$. By its construction as a gradient, the function N_x is orthogonal to $c(x) = 0$ at x , which means that the function $\tilde{N}_x = N_x/|N_x|$ is the normalized normal on the constraint manifold at x . Accordingly, the transformation $P_x = P_x^T = \tilde{N}_x\tilde{N}_x^T$ projects any function onto \tilde{N}_x . Since the Heaviside function H equals 0 for negative arguments and is 1 otherwise, the transformation $\tilde{P}_x = \tilde{P}_x^T = \delta - H(-c(x))H(-g_x^T\tilde{N}_x)P_x$ removes the normal component of any function if the constraint manifold is reached, $c(x) = 0$, and further evolution along the selection gradient g_x would penetrate it, $g_x^T\tilde{N}_x \leq 0$. As long as these conditions are not met together, we simply have $\tilde{P}_x = \delta$, where the Dirac delta function δ represents the identity transformation. The transformation \tilde{P}_x in Eq. (5c), $\sigma_x^2 = \tilde{P}_x U_x \tilde{P}_x$, thus removes the normal component of the unconstrained variance–covariance function U_x if and only if the adaptive dynamics would violate the global inequality constraint $c(x) \geq 0$. This ensures that the adaptive dynamics slides along the constraint manifold, without ever penetrating it. Notice that all global equality constraints, $c(x) = 0$, can be dealt with by deleting the Heaviside function $H(-c(x))$ from the definition of \tilde{P}_x .

- *Local inequality constraints*, $x(a) \geq c_x^-(a)$ and $x(a) \leq c_x^+(a)$, can also be represented by multiplying an unconstrained variance–covariance function $U_x(a', a)$ with a correction term,

$$\sigma_x^2(a', a) = U_x(a', a) \times \prod_{a_0=a', a} [1 - H(s(x(a_0) - c_x^s(a_0))) H(sg_x(a_0))]. \quad (5d)$$

Notice that the second term in the square bracket is 1 at all argument values a_0 at which a local inequality constraint has been reached, $x(a_0) = c_x^\pm(a_0)$, and at which, moreover, pursuit of the selection gradient would violate these constraints, $\pm g_x(a_0) > 0$. Consequently, the correction term in Eq. (5d) annuls all variances and covariances that would breach a local inequality constraint. Constraints on the monotonicity of function-valued traits can be expressed as local inequality constraints on the derivative $(d/da)x(a)$.

The considerations above show how to account for equality and inequality constraints by choosing suitable variance–covariance functions. We note in passing that the constrained variance–covariance functions in Eqs. (5a)–(5d) preserve, as they must, both the symmetry and the nonnegative definiteness of the underlying unconstrained functions.

When constructing specific evolutionary models, one should take care not to confound evolutionary constraints,

which strictly exclude certain trait values, with fitness costs, which merely imply that certain trait values possess low fitness. A useful guideline is to restrict the use of constraints as those described above to immutable mathematical or physical requirements, and to employ costs to account for developmental, physiological, or morphological limitations.

6. First example: evolution of metabolic investment strategies

In this and the following section we illustrate our approach by investigating two specific eco-evolutionary problems. In the first example, we find evolutionarily singular metabolic investment strategies for consumers that are confronted with the opportunity of harvesting a range of different resource types. Almost any utilization of resources requires prior physiological or morphological investments in handling and processing. In particular, the chemical diversity of resource types necessitates a variety of adequate biochemical pathways for metabolizing these resources efficiently, and these pathways are expected to be costly. It is therefore important to address the question how a consumer should allocate its metabolic effort on building and maintaining the morphological and physiological machinery needed for accessing, decomposing, and utilizing the different types of resource it encounters in its diet.

We consider a function-valued trait $x(a)$ that describes the metabolic effort invested across different resource types a . Obviously, this effort cannot be negative, implying local inequality constraints $x(a) \geq 0$ for all a . Furthermore, the total effort $E(x) = \int x(a) da$ cannot become too large, since otherwise an individual's energy balance becomes negative. Below, we reflect this limitation through a fitness cost.

Types of resources are characterized by their digestibility a , standardized to take values between 0 and 1; low values of a correspond to high digestibility. The less digestible a type of resource, the more effort is needed to achieve a given level of efficiency in processing it. If no effort is invested on a resource type, then it is metabolized with efficiency zero, and consuming the resource can provide no gain. It is natural to assume that efficiency increases with increasing effort, while investing too much effort on a particular resource type leads to diminishing returns. Thus, metabolic efficiency is described by a positive function $e(a, x(a))$ that is decreasing in its first and increasing in its second argument (with an asymptotic value of 1). Such a function is given, for instance, by $e(a, x(a)) = x(a)/[x(a) + a]$.

To keep the example as simple as possible, the abundance of resource types is assumed to remain constant under utilization, so that the abundance of the different types a of resource is described by the resource density $r(a)$. Such a situation arises if ecological factors other than availability of the considered resource are limiting the abundance of the consumer population. Examples include

rare species within the grazing guild of savannas, or species that rely on an exogenous food supply, like trichopteran larvae filter-feeding on food particles drifting in a river. Other scenarios with more complicated ecological feedbacks can be analysed similarly (Heino et al., submitted).

The gain from a resource of type a is the product of its abundance and of the efficiency with which it is metabolized, $r(a)e(a, x(a))$. Assuming that there is no diet selection, the total intake of resources is then given by the integral of this product over the resource axis $G(x) = \int r(a)e(a, x(a)) da$. We assume that the costs incurred by a particular strategy x for the metabolic effort are given by $cE(x)$, with the positive constant of proportionality c scaling the expenses of metabolic investment. High metabolic efforts thus lead to higher gains, but at the same time are limited by the costs an individual can sustain.

To study the adaptive dynamics of metabolic investment, we now embed the model for the allocation of metabolic effort into a simple population dynamics. Specifically, we assume the birth rate of individuals to depend on their net gain, $b(x', x) = G(x') - cE(x')$. The death rate regulates population size n_x and is of logistic type with carrying capacity K , $d(x', x) = n_x/K$. At population dynamical equilibrium, the resident's population growth rate goes to zero, $b(x, x) = d(x, x)$; the equilibrium population size is thus given by $\bar{n}_x = \max(0, K[G(x) - cE(x)])$. For the invasion fitness we hence obtain

$$f(x', x) = [G(x') - cE(x')] - [G(x) - cE(x)]. \quad (6a)$$

The result shows that evolution in this example obeys an optimization principle: the optimal strategy x^* maximizes $G(x^*) - cE(x^*)$, and thus \bar{n}_{x^*} . This means that the calculus of variations can be used to determine x^* (Parvinen et al., 2006). Here we instead proceed with a dynamic mode of analysis that allows us to assess transient evolution as well as evolutionary outcomes.

The selection gradient is determined by applying Eq. (3c) to Eq. (6a),

$$g_x(a) = \frac{r(a)a}{(x(a) + a)^2} - c. \quad (6b)$$

Following Eqs. (1) and (5d), while otherwise assuming the simplest possible mutation structure—constant mutation probability, $\mu_x = \mu$, and absence of covariance, $\sigma_x^2(a', a) = \sigma^2 \delta(a' - a)$ —we obtain

$$\begin{aligned} \frac{d}{dt} x(a) &= \frac{1}{2} \mu \bar{n}_x \int \sigma^2 \delta(a' - a) C(x, a) C(x, a') g_x(a') da' \\ &= \frac{1}{2} \mu \sigma^2 \bar{n}_x \left[\frac{r(a)a}{(x(a) + a)^2} - c \right] C(x, a), \end{aligned} \quad (6c)$$

where the factor $C(x, a) = 1 - H(-x(a))H(-g_x(a))$ results from the local inequality constraints, $x(a) \geq 0$ for all a . Recall from above that H equals 0 for negative arguments and is 1 otherwise; consequently, $C(x, a)$ equals either 0 or 1, and thus $C^2(x, a) = C(x, a)$, which explains the result of the integral above.

According to Eq. (6c), the evolutionarily singular strategy x^* for the investment of metabolic effort, defined by $(d/dt)x^*(a) = 0$ for all a , is given by

$$x^*(a) = \begin{cases} \sqrt{(a/c)r(a)} - a & \text{if } r(a) > ac \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

This finding shows that metabolic effort will be invested only on resource types that are abundant enough to balance the costs of metabolic investment. The lower the digestibility of a resource type (large a), the higher is this investment threshold (since $r(a)$ must exceed ac).

Fig. 2 illustrates the results for this model for $c = \frac{1}{2}$ and for a quadratic resource density $r(a) = 4a(1 - a)$. Starting from different initial conditions, adaptive transients vary; yet, the evolutionarily singular strategy x^* is convergence stable and serves as a global evolutionary attractor. Notice that the evolutionarily singular metabolic effort is skewed: most metabolic effort is invested into resource types that are more easily digested than those that have the highest abundance. By contrast, only little effort needs to be invested into resource types that are easily digested: for these, a relatively high metabolic efficiency is achieved even with modest investment and, due to diminishing returns, higher investments would be wasteful. Beyond the threshold at $a = 1 - \frac{1}{4}c = \frac{7}{8}$, investment ceases completely.

7. Second example: evolution of seasonal flowering schedules

In our second example, we analyse the evolution of seasonal flowering schedules in plants that inhabit a temporally varying environment. Flowers opening at different parts of the season are facing different ecological conditions, for instance, in terms of the availability of pollinators or of the risk of being destroyed by herbivores. Assuming that a plant can produce only a limited number of flowers during the whole flowering season, it is expected that selection favors flowering schedules that track the seasonal pattern in expected pollination success. Here we introduce a simple model to study that hypothesis.

The flowering season in our model can either last throughout the year or for only part of it. We scale its duration to one, so that each moment in time during the flowering season is identified by a argument value a , ranging from $x \rightarrow x'$ to 1. We denote the amount of flowers produced at time a during the season by $x(a)$: the function-valued trait x thus characterizes a plant's temporal profile of flowering intensity, or seasonal flowering schedule. We assume that the total number of flowers a plant can produce during a season is limited by the amount of resources it has available, so that flowering schedules are restricted by the global equality constraint $\int x(a) da = \text{const.}$ For convenience, we normalize x such that $\int x(a) da = 1$.

Flowers produced at any time during the flowering season compete for pollinators and herbivore-free space; the effect of such competition is to decrease the probability

of setting seed. We describe this probability by $p_x(a) = e^{-n_x x(a)/K(a)}$, where n_x is population size, $n_x x(a)$ is the total amount of flowers open at time a , and $K(a)$ serves as a time-dependent carrying capacity. Accordingly, K describes seasonal factors affecting pollination success, such as the abundance of pollinators.

Traveling between flower heads decreases a pollinator's efficiency of foraging. With more flowers per flower head, foraging efficiency increases. Thus, pollinators are likely to find flower heads more attractive the more flowers they have. Among the flowering plants, this causes asymmetric competition for pollinators. A mutant individual which, at a given time of the season, sports a higher flowering intensity than the resident population possesses an advantage at that time. Since the total amount of flowers produced during the season is limited, gain at one time is necessarily coupled to loss at some other time. Specifically, we assume that asymmetric competition alters the pollination success at time a of a plant with flowering schedule x' competing with a plant with schedule x by a factor $c(e(a)) = 2/(1 + e^{-\alpha e(a)})$, where $e(a)$ denotes the mutant's excess flowering intensity compared to the resident. The exponent α determines the strength of asymmetric competition; for $\alpha = 0$ we obtain $c(e(a)) = 1$, corresponding to symmetric competition. For small excesses $e(a)$ the gain ($e(a) > 0$) or loss ($e(a) < 0$) in pollination success is proportional to $e(a)$; marginal gains and losses level off when differences in flowering intensity become large. The function $e(a) = [x'(a) - x(a)]/x(a)^\beta$, chosen below, allows measuring the mutant's excess flowering intensity along a range from absolute ($\beta = 0$) to relative ($\beta = 1$) differences.

The birth rate of a rare mutant individual with flowering schedule x' in a resident population with schedule x is given by $b(x', x) = \int x'(a) c([x'(a) - x(a)]/x(a)^\beta) p_x(a) da$, while the death rates of individuals are assumed to be constant, $d(x', x) = d$. Consequently, the invasion fitness for this ecological setting is given by

$$f(x', x) = \int x'(a) c([x'(a) - x(a)]/x(a)^\beta) \bar{p}_x(a) da - d \quad (8a)$$

with $\bar{p}_x(a) = e^{-\bar{n}_x x(a)/K(a)}$. The condition $f(x, x) = 0$ defines the equilibrium population size \bar{n}_x . From the invasion fitness we obtain the selection gradient according to Eq. (3c),

$$g_x(a') = \bar{p}_x(a') [1 + \frac{1}{2} \alpha x(a')^{1-\beta}]. \quad (8b)$$

Following Eq. (5a), the influence of the global equality constraint $\int x(a) da = 1$ is taken into account by using the variance-covariance function $\sigma_x^2(a', a) = \sigma^2[\delta(a' - a) - 1]$. We thus assume that any mutational increase (decrease) in flowering intensity at time a is balanced by a uniform decrease (increase) in flowering intensity at other times during the season. The mutation probability is assumed to be constant, $\mu_x = \mu$, and the local inequality constraints, $x(a) \geq 0$ for all a , are handled like in Section 6. The canonical equation of function-valued adaptive dynamics,

Eq. (1), thus becomes

$$\begin{aligned} \frac{d}{dt}x(a) &= \frac{1}{2}\mu\bar{n}_x \int \sigma^2[\delta(a' - a) - 1]C(x, a)C(x, a')g_x(a')da' \\ &= \frac{1}{2}\mu\sigma^2\bar{n}_x[\bar{p}_x(a)\left(1 + \frac{1}{2}\alpha x(a)^{1-\beta}\right) \\ &\quad - \int \bar{p}_x(a')\left(1 + \frac{1}{2}\alpha x(a')^{1-\beta}\right)C(x, a')da']C(x, a) \quad (8c) \end{aligned}$$

with $C(x, a) = 1 - H(-x(a))H(-g_x(a))$ as in Section 6. The evolutionarily singular strategy x^* for this example, satisfying $(d/dt)x^*(a) = 0$ for all a , cannot be obtained analytically. Numerical integration of Eq. (8c), however, is an efficient way of studying the adaptive dynamics leading to x^* . Fig. 3b illustrates how the seasonal flowering schedule evolves toward an evolutionarily singular schedule x^* that indeed tracks the temporal pattern, shown in Fig. 3a, of seasonal factors affecting pollination success. Integrating Eq. (8c) from various initial conditions demonstrates that all populations featuring a single flowering schedule converge toward the global evolutionary attractor x^* .

Evolutionary convergence toward such a monomorphic equilibrium, however, does not guarantee evolutionary stability. If the evolutionarily singular strategy x^* permits the invasion and coexistence of similar flowering schedules, evolution can give rise to a dimorphic population via evolutionary branching (Metz et al., 1992, 1996a; Geritz et al., 1997, 1998) and x^* is called an evolutionary branching point. The possibility of a population being driven by directional selection toward a point at which selection turns disruptive is a feature of many adaptive dynamics models that properly account for frequency-dependent selection (e.g. Doebeli and Dieckmann, 2000). To test whether x^* allows for evolutionary branching, and to investigate the potentially ensuing dimorphic evolution, we extend the monomorphic adaptive dynamics given by Eq. (8c) to the dimorphic case.

We denote the two resident types of a dimorphic population by $x_1(a)$ and $x_2(a)$, and their population sizes at ecological equilibrium by \bar{n}_{x_1} and \bar{n}_{x_2} , respectively. The degree of asymmetric competition $c(e(a))$ experienced by a rare mutant at any time a now depends on the population mean flowering intensity, $\bar{x} = (\bar{n}_{x_1}x_1 + \bar{n}_{x_2}x_2)/\bar{n}_{\bar{x}}$ with $\bar{n}_{\bar{x}} = \bar{n}_{x_1} + \bar{n}_{x_2}$, through $e(a) = [x'(a) - \bar{x}(a)]/\bar{x}(a)^\beta$. For the invasion fitness of a mutant x' in a resident population

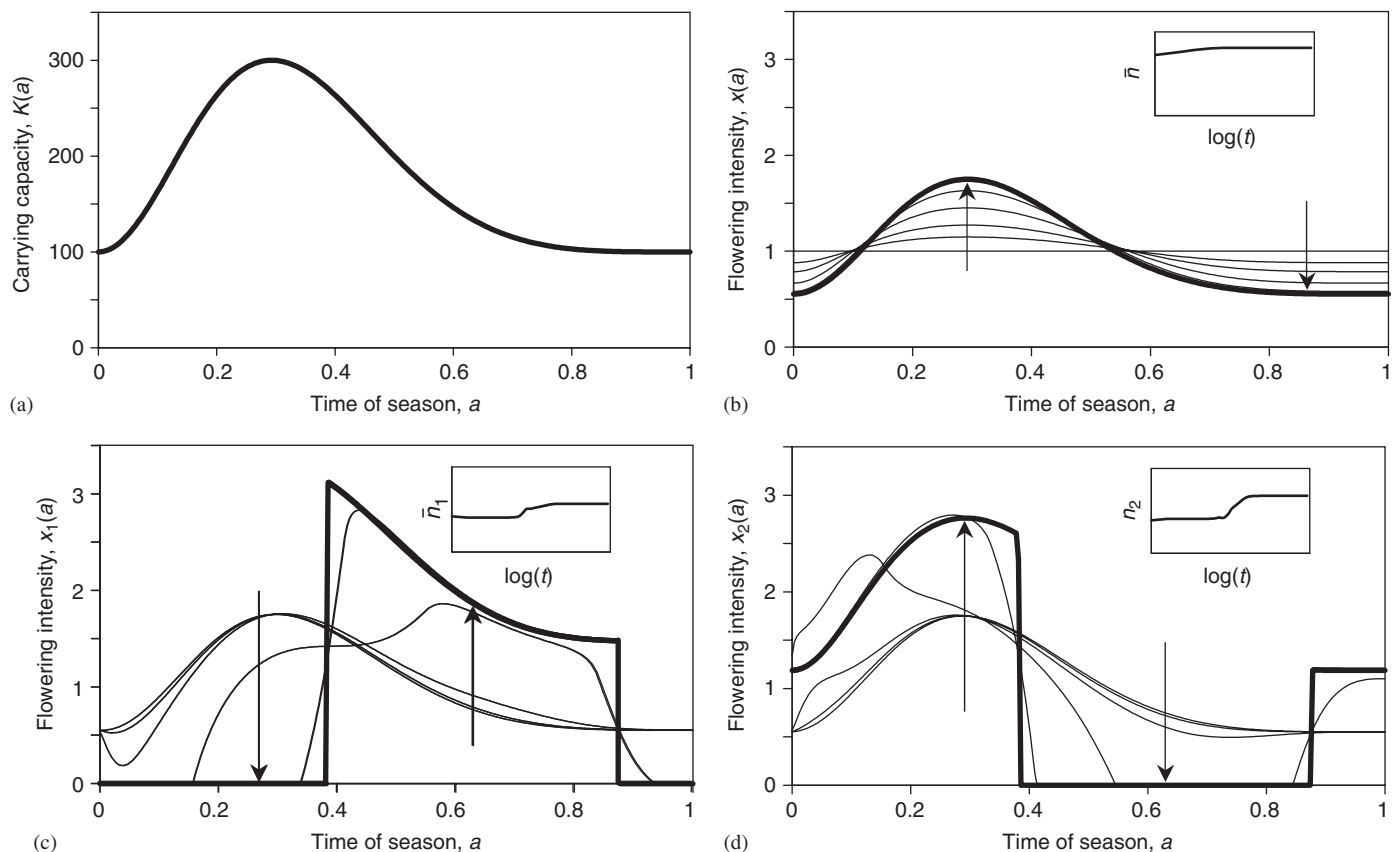


Fig. 3. Evolution of seasonal flowering schedules. (a) Carrying capacity $K(a)$ at different times a during the season. (b) Dynamics and outcome $x^*(a)$ of monomorphic evolution of flowering intensity $x(a)$. (c) and (d) Dynamics and outcome of dimorphic evolution of flowering intensities $x_1(a)$ and $x_2(a)$ started from the neighborhood of the monomorphic evolutionary outcome $x^*(a)$. Initial and intermediate trait values are shown as thin curves, and final values as thick curves. Insets show changes in equilibrium population sizes resulting from the depicted evolutionary change. Parameters: $\alpha = 1$, $\beta = 0.9$, $d = \frac{1}{2}$, and $K(a) = 100[2 + \sin(2\pi((a - 1)^2 - \frac{1}{4}))]$.

(x_1, x_2) we then have

$$f(x', x_1, x_2) = \int x'(a) c([x'(a) - \bar{x}(a)] / \bar{x}(a)^\beta) \bar{p}_{\bar{x}}(a) da - d. \quad (9a)$$

The equilibrium population sizes $(\bar{n}_{x_1}, \bar{n}_{x_2})$ can be obtained numerically from the two equations $f(x_1, x_1, x_2) = 0$ and $f(x_2, x_1, x_2) = 0$. With the two selection gradients

$$g_{x_i}(a) = \bar{p}_{\bar{x}}(a) c(e_i(a)) \times \left[1 + \frac{1}{2} \alpha_{x_i}(a) c(e_i(a)) e^{-\alpha_{x_i}(a)} \bar{x}(a)^{-\beta} \right], \quad (9b)$$

for $i = 1, 2$ and with $e_i(a) = [x'_i(a) - \bar{x}(a)] / \bar{x}(a)^\beta$, we obtain the canonical equation for the dimorphic flowering schedule model,

$$\frac{d}{dt} x_i(a) = \frac{1}{2} \mu \sigma^2 \bar{n}_{x_i} [g_{x_i}(a) - \int g_{x_i}(a') C(x_i, a') da'] C(x_i, a). \quad (9c)$$

As initial condition for the dimorphic adaptive dynamics in Eq. (9c) we chose a pair of small perturbations of the evolutionarily singular flowering schedule, by shifting x^* slightly to the left and to the right, see Figs. 3c and d. If x^* was evolutionarily stable, then the resulting dimorphic population would converge back to the monomorphic equilibrium x^* . However, as Figs. 3c and d demonstrate, the two strategies instead start to diverge from x^* and from each other, until a complete phenological segregation of flowering schedules is reached. This result is the first demonstration that evolutionary branching also naturally occurs in function-valued traits.

In nature, the initiation and cessation of flowering is bound to be less sharp than predicted by the simple model analysed here. Once localized covariances, as introduced in Appendix C, are incorporated, the vertical flanks in the presented dimorphic flowering schedules will naturally be smoothed out. Also the simplistic global equality constraint considered above would benefit from being replaced with a mechanistically motivated fitness cost. These interesting and ecologically relevant refinements exceed the scope of the introductory treatment here and will be investigated elsewhere (Parvinen et al., in preparation).

8. Discussion

We have introduced a comprehensive and flexible theoretical framework for studying the long-term evolution of function-valued traits. Three features of our approach may deserve being highlighted. First are the advantages of respecting the function-valued nature of many phenotypes of interest in evolutionary ecology. This prevents falling prey to the misleading results of oversimplified models (Section 2) and enables analyses of questions that would be difficult to address otherwise (Sections 6 and 7). Second is the need to deal with realistic feedback between ecological

and evolutionary dynamics. Our approach permits studying evolutionary problems involving any kind of density-dependent and/or frequency-dependent selection pressures (Section 3), which in turn opens up new perspectives on the richness of phenomena driven by the long-term evolution of function-valued traits (Section 7). The intimate connection between adaptive dynamics and population dynamics, which enables the derivation of invasion fitness, is particularly valuable here (Sections 3, 6, and 7; Appendix A). Third is the systematic treatment of evolutionary constraints, which proper analyses of function-valued evolution cannot do without (Sections 4–7). We have shown how such constraints are readily incorporated by choosing suitable variance–covariance functions for describing the availability of trait values. The general solutions offered here for working with the four most important classes of constraints, local and global, as well as based on equalities and inequalities, should considerably facilitate their treatment in specific models (Section 5).

The two worked examples in Sections 6 and 7 showcase the utility of our approach. While these examples are based on reasonably realistic ecological models, they were kept deliberately simple to serve as illustrations. Models with a higher degree of ecological complexity based on the framework introduced here have been devised to study the evolution of reaction norms in spatially heterogeneous environments (Ernande and Dieckmann, 2004), of reaction norms for age and size at maturation (Ernande et al., 2004), of seasonal flowering schedules (Parvinen et al., in preparation), and of resource utilization spectra under complex ecological conditions (Heino et al., submitted).

Our framework for the evolution of function-valued traits is based on adaptive dynamics theory and complements an earlier body of corresponding work (e.g. Kirkpatrick and Heckman, 1989a; Kingsolver et al., 2001) based on quantitative genetics theory. At first sight, there is a striking similarity between the canonical equation of adaptive dynamics (Dieckmann and Law, 1996) and equations based on Lande's model of quantitative genetics (Lande, 1976, 1979). In both types of model, evolutionary rates are proportional to selection gradients and to measures of genetic variability. This observation, in conjunction with a recognition of the multitude of closely related models (e.g. Brown and Vincent, 1987a, 1987b, 1992; Rosenzweig et al., 1987; Hofbauer and Sigmund, 1988, 1990; Vincent, 1990; Iwasa et al., 1991; Takada and Kigami, 1991; Taper and Case, 1992; Abrams et al., 1993; Marrow and Cannings, 1993; Vincent et al., 1993), provided the original motivation for referring to dynamics like those in Eq. (1) as “canonical” (Dieckmann and Law, 1996).

The similarity between the equations resulting from adaptive dynamics and quantitative genetics has misled some authors into suggesting that the corresponding models are not just superficially similar, but fundamentally equivalent (e.g. Waxman and Gavrillets, 2005). Since such uncritical lumping undermines a proper understanding of the complementary strengths and weaknesses of models

alternatively based on adaptive dynamics or quantitative genetics, it may be helpful, also in the present context of function-valued evolution, to recall the main differences between these models:

1. First are elementary distinctions in the described biological processes. As detailed in Section 3 and Appendix A, the canonical equation describes the expected course of an evolutionary random walk resulting from mutation and selection, through the successive invasion of advantageous mutants into essentially monomorphic resident populations. By using such adaptive dynamics models, we are thus looking at the expectation of a monomorphic stochastic process. By contrast, Lande's model and its later extension to function-valued traits describe how the mean of a population's distribution of quantitative traits evolves under the action of selection. By using such quantitative genetics models, we are thus looking at the mean of a polymorphic deterministic process.
2. These differences naturally go along with disparate key assumptions made in the derivation of the two types of model. The canonical equation describes the effects of selection when successful mutations are rare and small, while Lande's model describes the effects of selection when a population's trait distribution is normal and its genetic variance–covariance constant.
3. The fundamentally different nature of both models becomes most evident when these simplifying assumptions are relaxed. When larger mutational steps are considered in models of adaptive dynamics, correction terms are incorporated into the canonical equation (Dieckmann and Law, 1996; Appendix A). These terms contain higher derivatives of invasion fitness beyond the selection gradient and higher moments of the mutation distribution beyond its variance. When variable genetic variances are considered in models of quantitative genetics, new equations for the variance's dynamics need to be introduced and coupled with the equation for the mean's dynamics (e.g. Vincent et al., 1993; Turelli and Barton, 1994). Since quantitative genetics models describe the dynamics of distributions of quantitative traits, they are equivalent only to the entire hierarchy of moment equations, given by the dynamics of the distribution's mean, variance, skewness etc. Accordingly, the distribution's variance assumed to be constant in Lande's model in actual fact is a hidden state variable, whose changes are exactly prescribed by the distributional dynamics from which Lande's model for the change in the distribution's mean is derived. The simplicity of Lande's model is thus afforded by making the mental leap of treating this dynamic variable as a constant parameter, so as to truncate the otherwise intractable hierarchy of moment equations. It should be kept in mind that no moment hierarchy, and therefore also no such truncation, underlies the derivation of the canonical equation.

A proper appreciation of these details allows us to draw the following conclusions:

1. Adaptive dynamics and quantitative genetics models are both based on simplifying assumptions that are unlikely to be closely met in real processes of long-term evolution. At the simplest level of modeling, however, the two approaches result in equations that are very similar and therefore can often be used interchangeably. Also the solutions presented in Section 5 and Appendix C for the systematic construction of constrained mutational variance–covariance functions carry over to the variance–covariance functions of constrained genetic polymorphisms needed in quantitative genetics models.
2. Yet the similarity between the two approaches must not be exaggerated. Specifically, the essential equivalence of results breaks down (a) when the scaling of evolutionary rates with population abundances matters (in adaptive dynamics models, evolutionary rates are proportional to population size, while in quantitative genetics models they are not), (b) when the effects of frequency-dependent selection are to be described accurately (which in Lande's model requires variance–covariance corrections in the fitness function itself, causing deviations from the invasion fitness underpinning adaptive dynamics models), (c) when the dynamical emergence of dimorphisms through evolutionary branching has to be treated (which in adaptive dynamics models is easy, and can be done consistently with the described mutation–selection process, while Lande's model is not geared to this task), or, generally, (d) when the simplifying assumptions underlying the models are relaxed (which results in very different types of correction).

It may thus be conjectured that models of function-valued evolution based on quantitative genetics theory are better suited for describing short-term evolution, during which selection acts on substantial additive genetic variation within populations. In such settings, the variance–covariance functions characterizing population-level genetic variation, as well as the fitness functions characterizing selection pressures, may be assumed to stay reasonably constant over short periods of time. This makes models of this type especially relevant for describing selection experiments. By contrast, the models introduced here might be deemed better suited for describing long-term evolution, during which selection relies on infrequent evolutionary innovations, and fitness functions substantially change over time, owing to density and frequency dependence. Clearly, also mutation distributions may change in the course of long-term evolution, and such changes could be taken into account in models of function-valued adaptive dynamics. Yet, assuming the constancy of the variance–covariance function of mutation distributions (as is often done in adaptive dynamics models) is more justifiable than assuming the constancy of the variance–covariance function of population distributions (as is often

done in quantitative genetics models), since the latter is directly subject to selection. This holds in particular when the effects of mutational constraints approached by an evolutionary process are accounted for by the methods introduced in Section 5.

A valuable extension of the framework presented here stems from its linkage with the calculus of variations (e.g. Wan, 1995). Only in the absence of frequency-dependent selection can evolution be described by optimization principles (Mylius and Dieckmann, 1995; Metz et al., 1996b; Heino et al., 1998; Meszéna et al., 2001). In such special circumstances, the calculus of variations can be used directly to predict outcomes of function-valued evolution. Interestingly, however, methods from the calculus of variations can be harnessed more widely, even for applying the framework advanced here to ecological settings in which selection is frequency-dependent. The resulting novel methodological avenues complement the approach presented here and are explored in a separate study (Parvinen et al., 2006; concrete applications are also presented in Heino et al., submitted).

We suggest that the scope of exciting problems in evolutionary ecology that become accessible by studying the evolutionary dynamics of function-valued traits under frequency-dependent selection is enormous. In particular, the evolutionary branching of function-valued traits opens up fascinating opportunities for scrutinizing the interplay between individual-level plasticity (described by function-valued traits) and population-level diversity (described by polymorphisms thereof). It is therefore hoped that the methods introduced in this study make a valuable contribution to the increasingly powerful toolbox of theoretical evolutionary ecologists.

Acknowledgments

It is a pleasure to thank Mark Kirkpatrick, Hans Metz, and an anonymous reviewer for thoughtful comments on an earlier version of this article. The authors gratefully acknowledge financial support by the European Research Training Network *ModLife* (Modern Life-History Theory and its Application to the Management of Natural Resources, funded through the Human Potential Programme of the European Commission; UD, MH & KP); by the Austrian Science Fund (UD); and by the Austrian Federal Ministry of Education, Science, and Cultural Affairs (UD).

Appendix A. Derivation of the canonical equation

In this appendix we describe the steps by which the canonical equation of function-valued adaptive dynamics, Eq. (1), is derived. As outlined in Section 3, this derivation proceeds in three steps.

Polymorphic stochastic model, PSM. We start from an individual-based description of the ecology of an evolving

population (Dieckmann, 1994; Dieckmann et al., 1995). The phenotypic distribution p of a population of n individuals is given by $p = \sum_{i=1}^n \Delta_{x_i}$, where x_i are the function-valued traits of individual i , and Δ_{x_i} denotes a generalized delta function peaked at x_i (see Appendix B). We can think of $p(x)$ as a density distribution in the space of function-valued traits x , with one peak positioned at the trait value of each individual in the population. This implies $p(x) = 0$ unless x is represented in the population. Since $\int \Delta_{x_i}(x) dx = 1$ for any x_i , we also have $\int p(x) dx = n$. If $p(x) \neq 0$ for more than one x , the population is called polymorphic, otherwise it is referred to as being monomorphic. The birth and death rates of individual i are given by $b_{x_i}(p)$ and $d_{x_i}(p)$. Each birth by a parent with trait value x gives rise, with probability μ_x , to mutant offspring with a trait value $x' \neq x$, distributed according to $M(x', x)$. (Appendix C shows how to obtain such distributions from assumptions about the underlying mutational changes.) A master equation (e.g. van Kampen, 1981) describes the resultant birth–death–mutation process,

$$\frac{d}{dt} P(p) = \int [r(p, p') P(p') - r(p', p) P(p)] dp'.$$

The equation describes changes in the probability $P(p)$ for the population to be in state p . This probability increases with transitions from states $p' \neq p$ to p (first term) and decreases with transitions away from p (second term). A birth event causes a single generalized delta function, peaked at the trait value x of the new individual, to be added to p , $p \rightarrow p' = p + \Delta_x$, whereas a death event corresponds to subtracting such a generalized delta function from p , $p \rightarrow p' = p - \Delta_x$. The rate $r(p', p)$ for the transition $p \rightarrow p'$ is thus given by

$$r(p', p) = \int [r_x^+(p) \Delta(p + \Delta_x - p') + r_x^-(p) \Delta(p - \Delta_x - p')] dx.$$

Notice that subscripts of the generalized delta function equaling 0 are omitted. The terms $\Delta(p + \Delta_x - p')$ and $\Delta(p - \Delta_x - p')$ ensure that the transition rate r vanishes unless p' can be reached from p through a birth event (first term) or death event (second term). The death rate $r_x^-(p)$ at trait value x is given by multiplying the per capita death rate $d_x(p)$ with the abundance $p(x)$ of individuals at that trait value,

$$r_x^-(p) = d_x(p)p(x).$$

Similarly, the birth rate $r_x^+(p)$ at trait value x is given by

$$r_x^+(p) = (1 - \mu_x) b_x(p) p(x) + \int \mu_{x'} b_{x'}(p) p(x') M(x', x) dx',$$

with the first and second terms corresponding to births without and with mutation, respectively. The master equation above, together with its transition rates, describes so-called “generalized replicator dynamics” (Dieckmann 1994) and offers a generic formal platform for deriving

simplified descriptions of individual-based mutation–selection processes.

Monomorphic stochastic model, MSM. If the time intervals between viable and advantageous mutations are long enough for evolution to be mutation-limited, $\mu_x \rightarrow 0$ for all x , the evolving population will be monomorphic at almost any moment in time. We can then consider transitions resulting from the emergence of single viable and advantageous mutant individuals in a monomorphic resident population that has attained its ecological equilibrium. Denoting trait values and population sizes by x and n for the resident and by x' and n' for the mutant, we can substitute the dimorphic density $p = n\Delta_x + n'\Delta_{x'}$ into the generalized replicator dynamics defined above to obtain a master equation for the probability $P(n, n')$ of jointly observing resident population size n and mutant population size n' . Assuming that the mutant is rare while the resident is sufficiently abundant to be described deterministically, this master equation is equivalent to the joint dynamics

$$\frac{d}{dt}n = [b_x(n\Delta_x) - d_x(n\Delta_x)]n,$$

for the resident population and

$$\frac{d}{dt}P(n') = b_{x'}(n\Delta_x)P(n' - 1) + d_{x'}(n\Delta_x)P(n' + 1)$$

for the mutant population, where $P(n')$ denotes the probability of observing mutant population size n' . When the resident is at its equilibrium population size \bar{n}_x , defined by $b_x(\bar{n}_x\Delta_x) = d_x(\bar{n}_x\Delta_x)$, the rare mutant thus follows a homogeneous and linear birth–death process. The probability of an emerging mutant to survive accidental extinction through demographic stochasticity is thus given by $s(x', x) = \max(0, f(x', x))/b(x', x)$ (e.g. [Athreya and Ney, 1972](#)), with $b(x', x) = b_{x'}(\bar{n}_x\Delta_x)$, $d(x', x) = d_{x'}(\bar{n}_x\Delta_x)$, and $f(x', x) = b(x', x) - d(x', x)$. Once mutants have grown beyond the range of low population sizes in which accidental extinction through demographic stochasticity is still likely, they are generically bound to go to fixation and thus to replace the former resident, provided that their trait value is sufficiently close to that of the resident, $x' \approx x$ ([Geritz et al. 2002](#)). Hence, the transition rate $r(x', x)$ for the trait substitution $x \rightarrow x'$ is given by multiplying (i) the distribution $\mu_x b(x, x)\bar{n}_x M(x', x)$ of arrival rates for mutants x' among residents x , with (ii) the probability $s(x', x)$ of mutant survival given arrival, and with (iii) the probability 1 of mutant fixation given survival,

$$r(x', x) = \mu_x b(x, x)\bar{n}_x M(x', x)s(x', x).$$

Accordingly, the master equation for the probability $P(x)$ of observing trait value x ,

$$\frac{d}{dt}P(x) = \int [r(x, x')P(x') - r(x', x)P(x)]dx',$$

describes the directed random walks in trait space that result from sequences of such trait substitutions.

Monomorphic deterministic model, MDM. If mutational steps are small, the average of many realizations of the evolutionary random walk model described above is closely approximated by

$$\frac{d}{dt}x(a) = \int [x'(a) - x(a)]r(x', x)dx'$$

(e.g. [van Kampen, 1981](#)). After inserting $r(x', x)$ as derived above, this yields

$$\frac{d}{dt}x(a) = \mu_x b(x, x)\bar{n}_x \int s(x', x)[x'(a) - x(a)]M(x', x)dx'.$$

By expanding $s(x', x) = \max(0, f(x', x))/b(x', x)$ around x to first order in x' , we obtain $s(x', x) = \max(0, (x' - x)b^{-1}(x, x)g_x)$ with g_x as defined in Eqs. (3). We can rewrite this by spelling out the multiplication of the two functions $x' - x$ and g_x and by extracting $b(x, x) \geq 0$, which gives $s(x', x) = b^{-1}(x, x) \max(0, \int [x'(a') - x(a')]g_x(a')da')$. This means that in the x' -integral above only half of the total x' -range contributes, while for the other half the corresponding integrand is 0. If we assume that $M(x', x)$ is symmetric, $M(x + \Delta x, x) = M(x - \Delta x, x)$ for all x and Δx , we obtain

$$\begin{aligned} \frac{d}{dt}x(a) &= \frac{1}{2}\mu_x \bar{n}_x \iint [x'(a') - x(a')] \\ &\quad \times [x'(a) - x(a)]M(x', x)dx'g_x(a')da'. \end{aligned}$$

Denoting the inner integral with $\sigma_x^2(a', a)$ according to Eq. (2), we recover the canonical equation of function-valued adaptive dynamics, Eq. (1).

Two of the simplifying assumptions used in the derivation above can be relaxed, should this be desirable for specific applications. First, when mutational steps $x \rightarrow x'$ are small but asymmetrically distributed, we have to replace $\sigma_x^2(a', a)$ in Eq. (1) with

$$\begin{aligned} \tilde{\sigma}_x^2(a', a) &= 2 \int [x'(a') - x(a')][x'(a) - x(a)] \\ &\quad \times H([x'(a') - x(a')]g_x(a'))M(x', x)dx', \end{aligned}$$

where the Heaviside function H equals 0 for negative arguments and is 1 otherwise. Notice that g_x breaks the symmetry of $\tilde{\sigma}_x^2$, so that the relation $\sigma_x^2(a', a) = \sigma_x^2(a, a')$ does not extend to $\tilde{\sigma}_x^2$. Second, when mutational steps $x \rightarrow x'$ are not small, introducing correction terms up to order $k > 1$ improves the accuracy of the canonical equation ([Dieckmann and Law 1996](#)),

$$\begin{aligned} \frac{d}{dt}x(a) &= \mu_x b(x, x)\bar{n}_x \int \max(0, \sum_{i=1}^k (x' - x)^i D_i(x)/i!) \\ &\quad \times [x'(a) - x(a)]M(x', x)dx'. \end{aligned}$$

Here the mutation distribution $M(x', x)$ may be symmetric or asymmetric, and $D_i(x)$ denotes the i th functional derivative of $f(x', x)/b(x', x)$ with respect to x' evaluated at x .

Appendix B. Delta functions

In this appendix we summarize the essential properties of delta functions, for readers unfamiliar with this tool. Delta functions can be defined for integers, real numbers, and functions; these are described in turn below.

Integers. It is simplest to start with the delta function for integers, also known as the Kronecker symbol $\delta_{ii'}$. This function depends on two integers, i and i' , and takes the value 1 if these are equal and the value 0 if not. Formally speaking, the Kronecker symbol can be defined by requiring the relation $\sum_i F_i \delta_{ii'} = F$ to hold for all vectors F . Notice that, for $F = 1$, this implies $\sum_i \delta_{ii'} = 1$.

Real numbers. The notion of a function of two arguments that differs from 0 only if the arguments coincide was generalized from integers to real numbers by Dirac (1926, 1958). He defined what is known today as the Dirac delta function by the relation $\int F(r') \delta(r' - r) dr' = F(r)$ for all continuous functions F . This identity is called the sifting property of the Dirac delta function: since the term $\delta(r' - r)$ has to be 0 except at $r' = r$ (where the delta function's argument vanishes), only at that point can the value of F contribute to the integral. By setting $F = 1$ one sees that the sifting property implies $\int \delta(r' - r) dr' = 1$. That integral can only equal 1—while $\delta(r' - r) = 0$ for $r' \neq r$ —if $\delta(0)$ is infinite. The shape of $\delta(r' - r)$ thus becomes clear: it is 0 everywhere, except at $r' = r$ where it possesses an infinitely high peak. This also implies that the Dirac delta function is symmetric, $\delta(r' - r) = \delta(r - r')$. Sometimes the Dirac delta is not defined algebraically by the sifting property as above, but analytically as the limit of a series of regular functions. Setting $h_\varepsilon(r)$ to $1/\varepsilon$ for $|r| < \varepsilon/2$ and to 0 elsewhere, we can write $\delta(r) = \lim_{\varepsilon \rightarrow 0} h_\varepsilon(r)$. Alternatively setting h_ε to a normal distribution with mean 0 and standard deviation ε has the same effect. For convenience, the location of a delta function's peak is often given as a subscript, $\delta(r' - r) = \delta_r(r')$.

Functions. The generalized delta function extends the idea of the Kronecker symbol and the Dirac delta function from the realm of integers and real numbers, respectively, to that of functions (Dieckmann, 1994, Dieckmann and Law, 2000). We can envisage the generalized delta function Δ as an infinitely narrow and infinitely high peak in the space of functions, with its maximum being located at a particular function x . Such a heuristic notion is made exact by defining the generalized delta function by the sifting property $\int F(x') \Delta(x' - x) dx' = F(x)$ for an arbitrary continuous functional F . Again, the location of Δ 's peak may be given as a subscript, $\Delta(x' - x) = \Delta_x(x')$. It is important to realize that here we do not have to be concerned at all with the potential intricacies of integrating over function spaces, since the algebraic definition of Δ amounts to a mere rewriting rule. Once combined according to this definition, neither functional integrations nor generalized delta functions remain in the end result. Or, in the words of Dirac (1958): “The use of delta functions thus does not involve any lack of rigour in the theory, but is merely a convenient

notation, enabling us to express in a concise form certain relations which we could, if necessary, rewrite in a form not involving delta functions, but only in a cumbersome way which would tend to obscure the argument.”

Appendix C. Normalization-preserving variance–covariance functions

In this appendix we derive a family of variance–covariance functions that preserve the normalization of a function-valued trait while also featuring localized covariances.

Mutation distribution. For the purpose of this derivation, we consider mutational changes Δx in a function-valued trait x with argument space $0 \leq a \leq 1$ that have only two actual degrees of freedom: the argument value a_0 at which the amount of mutational change is maximal, and the amplitude Δx_0 that scales this amount. We also assume that (a) the values of a_0 are distributed uniformly over the interval $0 \leq a \leq 1$, (b) the values of Δx_0 are distributed normally with mean 0 and standard deviation $\sigma_{\Delta x}$, (c) the mutational impact around a_0 attenuates according to a normal distribution with mean 0 and standard deviation σ_a , and (d) the overall effect of mutation leaves the normalization of the function-valued trait intact, $\int_0^1 x(a) da = 1$. From these assumptions, we obtain the mutation distribution $M(x + \Delta x, x)$ as

$$M(x + \Delta x, x) = \int_0^1 \int_0^1 U(a_0) N_{\sigma_{\Delta x}}(\Delta x_0) \Delta(\Delta x - \Delta x_0 m_{a_0, \sigma_a}) da_0 d\Delta x_0$$

with

$$U(a_0) = \begin{cases} 1 & 0 \leq a_0 \leq 1 \\ 0 & \text{otherwise,} \end{cases}$$

$$N_{\sigma_{\Delta x}}(\Delta x_0) = \exp(-\frac{1}{2} \Delta x_0^2 / \sigma_{\Delta x}^2) / (\sigma_{\Delta x} \sqrt{2\pi}),$$

$$m_{a_0, \sigma_a}(a) = \exp(-\frac{1}{2} (a - a_0)^2 / \sigma_a^2) - c(a_0, \sigma_a),$$

$$c(a_0, \sigma_a)$$

$$= \sigma_a \sqrt{\pi/2} [\operatorname{erf}(a_0 / (\sigma_a \sqrt{2})) + \operatorname{erf}((1 - a_0) / (\sigma_a \sqrt{2}))],$$

and with Δ denoting the generalized delta function (Appendix B). The offset c in the definition of m_{a_0, σ_a} guarantees that each possible mutation preserves the function-valued trait's normalization.

Variance–covariance function. From this specification of the full distribution M of mutational effects, we obtain the mutational variance–covariance function for the function-valued trait x according to Eq. (2),

$$\begin{aligned} \sigma_x^2(a, a') &= \int \Delta x(a) \Delta x(a') M(x + \Delta x, x) d\Delta x \\ &= \sigma_{\Delta x}^2 \int_0^1 m_{a_0, \sigma_a}(a) m_{a_0, \sigma_a}(a') da_0. \end{aligned}$$

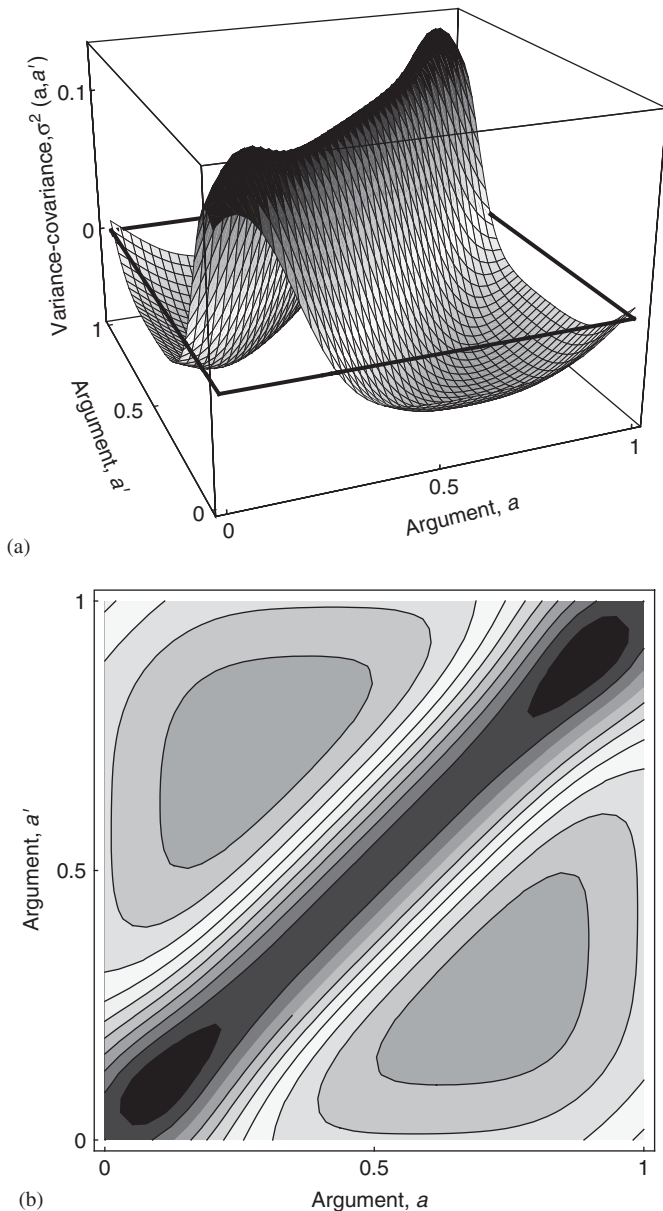


Fig. 4. Normalization-preserving variance-covariance function with localized covariances, as derived in Appendix C. (a) Surface plot. (b) Contour plot. Departures of the variance-covariance function $\sigma^2(a', a)$ from 0, occurring both in the positive or negative direction, are indicated by increasingly dark shades of gray, with white corresponding to $\sigma^2(a', a) = 0$. Parameters: $\sigma_{Ax} = 1$ and $\sigma_a = 0.1$.

(On the right-hand side, the mutational parameters σ_{Ax} and σ_a could, in general, be varied with x .) The normalization-preserving variance-covariance function provided by this construction is illustrated in Fig. 4.

References

- Abrams, P.A., Harada, Y., Matsuda, H., 1993. On the relationship between ESS and quantitative genetic models. *Evolution* 47, 982–985.
- Athreya, K.B., Ney, P.E., 1972. *Branching Processes*. Springer, Berlin, Germany.

- Beder, J.H., Gomulkiewicz, R., 1998. Computing the selection gradient and evolutionary response of an infinite-dimensional trait. *J. Math. Biol.* 36, 299–319.
- Brown, J.S., Vincent, T.L., 1987a. Coevolution as an evolutionary game. *Evolution* 41, 66–79.
- Brown, J.S., Vincent, T.L., 1987b. Predator–prey coevolution as an evolutionary game. In: Cohen, Y. (Ed.), *Applications of Control Theory in Ecology*. Lecture Notes in Biomathematics 73. Springer, Berlin, Germany, pp. 83–101.
- Brown, J.S., Vincent, T.L., 1992. Organization of predator–prey communities as an evolutionary game. *Evolution* 46, 1269–1283.
- Christiansen, F.B., 1991. On conditions for evolutionary stability for a continuously varying character. *Am. Nat.* 138, 37–50.
- Conrad, M., 1990. The geometry of evolution. *Biosystems* 24, 61–81.
- Dieckmann, U., 1994. *Coevolutionary Dynamics of Stochastic Replicator Systems*. Central Library of the Research Center Juelich, Juelich, Germany.
- Dieckmann, U., 1997. Can adaptive dynamics invade. *Trends Ecol. Evol.* 12, 128–131.
- Dieckmann, U., Law, R., 1996. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.* 34, 579–612.
- Dieckmann, U., Law, R., 2000. Relaxation projections and the method of moments. In: Dieckmann, U., Law, R., Metz, J.A.J. (Eds.), *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*. Cambridge University Press, Cambridge, UK, pp. 412–455.
- Dieckmann, U., Marrow, P., Law, R., 1995. Evolutionary cycling in predator–prey interactions: population dynamics and the Red Queen. *J. Theor. Biol.* 176, 91–102.
- Dirac, P.A.M., 1926. The physical interpretation of the quantum dynamics. *Proc. R. Soc. London A* 113, 621–641.
- Dirac, P.A.M., 1958. *The Principles of Quantum Dynamics*, fourth ed. Clarendon Press, Oxford, UK.
- Doebeli, M., Dieckmann, U., 2000. Evolutionary branching and sympatric speciation caused by different types of ecological interactions. *Am. Nat.* 156, S77–S101.
- Ernande, B., Dieckmann, U., 2004. The evolution of phenotypic plasticity in spatially structured environments: implications of intraspecific competition, plasticity costs and environmental characteristics. *J. Evol. Biol.* 17, 613–628.
- Ernande, B., Dieckmann, U., Heino, M., 2004. Adaptive changes in harvested populations: plasticity and evolution of age and size at maturation. *Proc. R. Soc. London Ser. B* 271, 415–423.
- Eshel, I., 1983. Evolutionary and continuous stability. *J. Theor. Biol.* 103, 99–111.
- Eshel, I., Motro, U., 1981. Kin selection and strong evolutionary stability of mutual help. *Theor. Popul. Biol.* 19, 420–433.
- Ferrière, R., Gatto, M., 1995. Lyapunov exponents and the mathematics of invasion in oscillatory or chaotic populations. *Theor. Popul. Biol.* 48, 126–171.
- Geritz, S.A.H., Metz, J.A.J., Kisdi, E., Meszéna, G., 1997. Dynamics of adaptation and evolutionary branching. *Phys. Rev. Lett.* 78, 2024–2027.
- Geritz, S.A.H., Kisdi, E., Meszéna, G., Metz, J.A.J., 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.* 12, 35–57.
- Geritz, S.A.H., Gyllenberg, M., Jacobs, F.J.A., Parvinen, K., 2002. Invasion dynamics and attractor inheritance. *J. Math. Biol.* 44, 548–560.
- Gomulkiewicz, R., Beder, J.H., 1996. The selection gradient of an infinite-dimensional trait. *SIAM J. Appl. Math.* 56, 509–523.
- Gomulkiewicz, R., Kirkpatrick, M., 1992. Quantitative genetics and the evolution of reaction norms. *Evolution* 46, 390–411.
- Hamilton, W.D., 1967. Extraordinary sex ratios. *Science* 156, 477–488.
- Heino, M., Dieckmann, U., Parvinen, K. (submitted). Evolution of foraging strategies on resource gradients.
- Heino, M., Metz, J.A.J., Kaitala, V., 1998. The enigma of frequency-dependence. *Trends Ecol. Evol.* 13, 367–370.

- Hofbauer, J., Sigmund, K., 1988. *Theory of Evolution and Dynamical Systems*. Cambridge University Press, New York, USA.
- Hofbauer, J., Sigmund, K., 1990. Adaptive dynamics and evolutionary stability. *Appl. Math. Lett.* 3, 75–79.
- Iwasa, I., Pomiankowski, A., Nee, S., 1991. The evolution of costly mate preferences. II. The 'handicap' principle. *Evolution* 45, 1431–1442.
- Kingsolver, J., Gomulkiewicz, R., Carter, P.A., 2001. Variation, selection, and evolution of function-valued traits. *Genetica* 112/113, 87–104.
- Kirkpatrick, M., 1988. The evolution of size in size-structured populations. In: Ebenman, B., Persson, L. (Eds.), *The Dynamics of Size-Structured Populations*. Springer, Berlin, Germany, pp. 13–28.
- Kirkpatrick, M., 1993. Evolution of size and growth in fisheries and other harvested natural populations. In: Stokes, K., McGlade, J.M., Law, R. (Eds.), *The Exploitation of Evolving Resources*. Lecture Notes in Biomathematics 99. Springer, Berlin, Germany, pp. 145–154.
- Kirkpatrick, M., 1997. Genetic improvement of livestock growth using infinite-dimensional analysis. *Anim. Biotechnol.* 8, 55–62.
- Kirkpatrick, M., Heckman, N., 1989a. A quantitative-genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *J. Math. Biol.* 27, 429–450.
- Kirkpatrick, M., Lofsvold, D., 1989b. The evolution of growth trajectories and other complex quantitative characters. *Genome* 31, 778–783.
- Kirkpatrick, M., Lofsvold, D., 1992. Measuring selection and constraint in the evolution of growth. *Evolution* 46, 954–971.
- Kirkpatrick, M., Lofsvold, D., Bulmer, M., 1990. Analysis of the inheritance, selection, and evolution of growth trajectories. *Genetics* 124, 979–993.
- Kirkpatrick, M., Hill, W.G., Thompson, R., 1994. Estimating the covariance structure of traits during growth and aging, illustrated with lactation in dairy cattle. *Genet. Res.* 64, 57–69.
- Lande, R., 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30, 314–334.
- Lande, R., 1979. Quantitative genetic analysis of multivariate evolution applied to brain:body size allometry. *Evolution* 33, 402–416.
- Leimar, O., 2001. Evolutionary change and Darwinian demons. *Selection* 2, 65–72.
- Leimar, O. (in press). Multidimensional convergence stability and the canonical adaptive dynamics. In: Dieckmann, U., Metz, J.A.J. (Eds.), *Elements of Adaptive Dynamics*. Cambridge University Press, Cambridge, UK.
- Marrow, P., Cannings, C., 1993. Evolutionary instability in predator–prey systems. *J. Theor. Biol.* 160, 135–150.
- Marrow, P., Dieckmann, U., Law, R., 1996. Evolutionary dynamics of predator–prey systems: an ecological perspective. *J. Math. Biol.* 34, 556–578.
- Maynard Smith, J., 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge, UK.
- Maynard Smith, J., Price, G., 1973. The logic of animal conflict. *Nature* 246, 15–18.
- Mészéna, G., Kisdi, É., Dieckmann, U., Geritz, S.A.H., Metz, J.A.J., 2001. Evolutionary optimisation models and matrix games in the unified perspective of adaptive dynamics. *Selection* 2, 193–210.
- Metz, J.A.J., Nisbet, R.M., Geritz, S.A.H., 1992. How should we define fitness for general ecological scenarios. *Trends Ecol. Evol.* 7, 198–202.
- Metz, J.A.J., Geritz, S.A.H., Mészéna, G., Jacobs, F.J.A., van Heerwaarden, J.S., 1996a. Adaptive dynamics, a geometrical study of the consequences of nearly faithful reproduction. In: van Strien, S.J., Verduyn Lunel, S.M. (Eds.), *Stochastic and Spatial Structures of Dynamical Systems*. North-Holland, Amsterdam, The Netherlands, pp. 183–231.
- Metz, J.A.J., Mylius, S.D., Dieckmann, O., 1996b. When does evolution optimize? On the relation between types of density dependence and evolutionarily stable life history parameters. IIASA Working Paper WP-96-004. International Institute for Applied Systems Analysis, Laxenburg, Austria.
- Mylius, S.D., Dieckmann, O., 1995. On evolutionarily stable life histories, optimization and the need to be specific about density dependence. *Oikos* 74, 218–224.
- Parvinen, K., Dieckmann, U., Heino, M., 2006. Function-valued adaptive dynamics and the calculus of variations. *J. Math. Biol.* 52, 1–26.
- Parvinen, K., Heino, M., Dieckmann, U. The evolution of seasonal flowering schedules, in preparation.
- Rand, D.A., Wilson, H.B., McGlade, J.M., 1994. Dynamics and evolution: evolutionarily stable attractors, invasion exponents and phenotype dynamics. *Philos. Trans. R. Soc. London Ser. B* 343, 261–283.
- Rosenzweig, M.L., Brown, J.S., Vincent, T.L., 1987. Red Queens and ESS: the coevolution of evolutionary rates. *Evol. Ecol.* 1, 59–94.
- Takada, T., Kigami, J., 1991. The dynamical attainability of ESS in evolutionary games. *J. Math. Biol.* 29, 513–529.
- Taper, M.L., Case, T.J., 1992. Models of character displacement and the theoretical robustness of taxon cycles. *Evolution* 46, 317–333.
- Taylor, P.D., 1989. Evolutionary stability in one-parameter models under weak selection. *Theor. Popul. Biol.* 36, 125–143.
- Turelli, M., Barton, N.H., 1994. Genetic and statistical analyses of strong selection on poly-genic traits—what, me normal? *Genetics* 138, 913–941.
- van Kampen, N.G., 1981. *Stochastic Processes in Physics and Chemistry*. North Holland, Amsterdam, the Netherlands.
- Vincent, T.L., 1990. Strategy dynamics and the ESS. In: Vincent, T.L., Mees, A.I., Jennings, L.S. (Eds.), *Dynamics of Complex Interconnected Biological Systems*. Birkhäuser, Basel, Switzerland, pp. 236–262.
- Vincent, T.L., Cohen, Y., Brown, J.S., 1993. Evolution via strategy dynamics. *Theor. Popul. Biol.* 44, 149–176.
- Wan, F.Y.M., 1995. *Introduction to the Calculus of Variations and Its Applications*. Chapman & Hall, New York, USA.
- Waxman, D., Gavrillets, S., 2005. Issues of terminology, gradient dynamics, and the ease of sympatric speciation in Adaptive Dynamics. *J. Evol. Biol.* 18, 1214–1219.