On Hamilton's rule, the evolution of behavior rules and the wizardry of control theory

Laurent Lehmann*

Abstract

This paper formalizes selection on a quantitative trait affecting the evolution of behavior rules through which individuals act and react with their surroundings. Combining Hamilton's marginal rule for selection on scalar traits and concepts from optimal control theory, a necessary first-order conditions for the evolutionary stability of the trait in group-structured population are derived. The model, which is of intermediate level of complexity, fills a gap between the formalization of selection on evolving traits that are directly conceived as actions (no phenotypic plasticity) and selection on evolving traits that are conceived as strategies or function valued actions (complete phenotypic plasticity). By conceptualizing individuals as open deterministic dynamical systems expressing incomplete phenotypic plasticity, the model captures selection on a large class of phenotypic expression mechanisms, including the evolution of learning and preferences under life-history trade-offs. As an illustration of the results, a first-order condition for the evolutionary stability of behavior response rules from the social evolution literature is re-derived, strenghthened, and generalized. All results of the paper also generalize directly to selection on mutildimenstional quantitative traits affecting behavior rule evolution, thereby covering neural and gene network evolution.

Keywords: Selection; social evolution; control theory; Hamilton's rule; adaptive dynamics; behavior evolution; life-history evolution.

1 Introduction

This paper is about formalizing selection on quantitative traits and exploring the mixing of two pieces of magic. On one side, there is the magic of Hamilton's marginal rule, which describes the direction of selection on a quantitative trait in a group (or familly) structured population and provides a necessary first-order condition for evolutionary stability (Rousset, 2004; Van Cleve, 2015 for reviews). The magic here, actualy mathematics, is that the problem of solving the possibly gigantic system of equations describing the distribution of genetics states among locally interacting individuals is reduced to the much simpler task of computing the probability that two individuals from the same group share a common ancestor under a neutral evolutionary process (all individuals bear the same traits). On the other side, we have the magic of the control theory approach to the calculus of variations (Liberzon, 2011; Weber, 2011 for a reviews), which provides necessary first-order conditions for optimization problems involving objective functions that depend on whole trajectories of dynamical systems, e.g., life-time reproductive sucess depends on growth and resource allocation scheduling. The manifold magic here is that the problem of dynamic optimization and solving multidimensional partial differential equations capturing selection on reactions norms in an evolutionary biology context, is broken down to the much simpler task of static optimization and computing lower dimensional dynamical systems that are evaluated under a neutral process (often simply ordinary differential equations).

Both these approaches have opened the door to many applications in evolutionary biology as is illustrated by the vast literatures on social evolution and life-history theories (e.g. Stearns, 1992; Frank, 1998). Wizards have

*Department of Ecology & Evolution, University of Lausanne, Switzerland. Email: Laurent.Lehmann@unil.ch

also blended Hamilton's rule and control theory. In particular, Day and Taylor (1997, 1998, 2000) and Wild (2011) investigated selection on so-called open loop traits under limited genetic mixing. These describe phenotypic plastic expression as a function of time, for instance from birth to death, and represent the standard formalization of reaction norm evolution in life-history theory for panmictic populations. Here, control theory has actually long been applied under the heading of Prontryagin's maximum principle (e.g., León, 1976; Iwasa and Roughgarden, 1984; Perrin, 1992; Irie and Iwasa, 2005). But plastic phenotypes evolve to vary not only as a function of time but also as a function of any fitness relevant state variables determining the physiology, morphology or behavior of an individual, which may themselves depend on phenotypic expression. Such closed loop traits can be thought of as a contingency plans or strategies, since they specify a conditional trait expression rule according to fitness relevant conditions and close the output-input feedback loop between phenotypic expression and individual states. For this mode of trait control too, admixtures between Hamilton's rule and control theory have been explored (Avila et al., 2021), and the evolution of closed loop traits are often studied using dynamic programming and the spellbounding Hamilton-Jacobi-Bellman equation (e.g., Houston and McNamara, 1999; Mangel et al., 1988; Ewald et al., 2007; Nakamura and Ohtsuki, 2016).

In the bulk of these formalizations, however, open and closed loop traits display a form of complete plasticity, in the sense that given values of the independent variable(s), time and/or state(s), the phenotype is quantitative and can evolve freely (subject to physiological constraints). As such the genetically evolving trait is of large dimensions if the independent variables can take many values and evolving traits are typically of infinite dimensions when the range of the dependent variables is continuous. Yet, one can conceive situations of incomplete phenotypic plasticity, where an individual is conceptualized as an open system acting and reacting with its surroundings, but the number of genetic evolving traits affecting phenotypic scheduling is small. Examples actually abound in the literature and include models for the evolution of reactive strategies, behavior response rules, learning rules, or neural and gene networks with a few number of nodes (e.g., Ezoe and Iwasa, 1997; McNamara et al., 1999; Akçay and Van Cleve, 2009; Killingback and Doebeli, 2002; Taylor and Day, 2004; McNamara et al., 2004; André and Day, 2007; Wakano and Miura, 2014; Dridi and Lehmann, 2015; Kobayashi et al., 2019; Alon, 2020). In each of these cases, there is a dynamical system underlying phenotypic scheduling, which, in turn, affects survival and reproduction, and this dynamical system is affected by one or a few number of genetically evolving traits. To a first approximation, most analytical models for reaction norm evolution when interactions occur between individuals are actually of this type.

The goal of this paper is to show how stantard concepts from control theory, originally devised for situations of complete plasticity where the magic is most effective, are actually also usefull to analyze the evolution of behavioral interactions in group-structured populations under incomplete plasticity. While this topic was partly explored in Avila et al. (2021, their "Constant control result"), some of their results are here generalized. Moreover, life-history scenarios where the analysis turns out be completely analytically tractable are fullly worked out. Finally, a more detailed formalization of an individual interacting with its surroundings is provided, which allows to make explicit connections with the extensive literature on incomplete plasticity and social evolution. The rest of this paper is organized as follows. (i) Hamilton's marginal rule for scalar traits is recalled and a formalisation of an individual as an open system is provided. (ii) Control theory concepts are used to derive necessary first-order conditions for the evolutionary stability of a scalar trait affecting any feature of the dynamic behavior of interacting individuals. (iii) The generic first-order condition for selection on a trait affecting a behavior response rule in group-structued population (Akçay and Van Cleve, 2012, eq. 5, Akçay and Van Cleve, 2014, eq. 8) is re-derived, strenghtened, and generalized. (iv) Limitations and generalities of the model are discussed.

2 Model

2.1 Main assumptions

2.1.1 The evolutionary setting and Hamilton's marginal rule

Consider a population of homogeneous individuals (no class structure) that can be subdivided into a large number of groups of fixed size n and were the population is censused at discrete demographic time steps, i.e., the standard social evolution model setting as reviewed in Michod (1982) and Rousset (2004). Interactions may occur among individuals within groups and/or between groups before reproduction. These interaction are assumed to be affected by an evolving quantitive genetic trait u, which either belongs to the whole set $\mathbb{R}$ of real numbers or a subset $\mathcal{U} = [u_{\min}, u_{\max}]$ thereoff (i.e., $u \in \mathcal{U} \subset \mathbb{R}$, for instance when u is a probability, then $u_{\min} = 0$ and $u_{\max} = 1$). The interactions are also assumed to be such that a focal (or representative) individual from the population has its survival and reproduction affected by individuals in up to three distinct roles *sensu* Grafen (2006, p. 544) in relation to the focal. Namely, the focal individual itself with trait $u_\bullet$, its group members with average trait $u_\circ$, and individuals at large from the population with average trait u. Individuals in each role are homogeneous and are thus different types of actor on the focal individual's fitness $w(u_\bullet, u_\circ, u)$, which gives its expected number of successful offspring produced over one life-cycle iteration (including self through survival) when expressing trait value $u_\bullet$, when group neighbours express average trait value $u_\circ$, and individuals from the population at large express average trait value u (formally, $w : \mathcal{U}^3 \to \mathbb{R}_+$, which is assumed differentiable and demographic consistency due to density-dependent regulation requires that when the population is phenotypically monomorphic for trait value u, fitness is equal to one: $w(u, u, u) = 1$).

Let us further denote by $r(u)$ the neutral relatedness between two randomly sampled individuals from the same group when the population is phenotypically monomorphic for trait value u. Thus, $r(u)$ is the probability that in a neutral process (where all individuals are alike with trait u) two homologous genes of these individuals coalesce in a common ancestor (e.g., Michod and Hamilton, 1980; Frank, 1998; Roze and Rousset, 2003; Rousset, 2004; Lehmann and Rousset, 2014; Van Cleve, 2015). Then, a necessary first-order condition for trait value $u^* \in [u_{\min}, u_{\max}]$ to be uninvadable (all mutant deviations go extinct) is that

$$\begin{aligned} u^* &= u_{\min} && \text{only if} \quad s(u^*) \leq 0, \\ u_{\min} < u^* &< u_{\max} && \text{only if} \quad s(u^*) = 0, \\ u^* &= u_{\max} && \text{only if} \quad s(u^*) \geq 0, \end{aligned} \tag{1}$$

where

$$s(u) = \left.\frac{\partial w(u_\bullet, u_\circ, u)}{\partial u_\bullet}\right|_{u_\bullet = u_\circ = u} + r(u) \left.\frac{\partial w(u_\bullet, u_\circ u)}{\partial u_\circ}\right|_{u_\bullet = u_\circ = u} \tag{2}$$

is the inclusive fitness effect from changing trait value in a population monomorphic for u. This is the sum of the marginal effect of the focal individual varying its trait expression level on its own fitness plus the marginal effect of an average group neigbour on the focal's fitness weighted by neutral relatedness.

This condition for evolutionary stability (eqs. 1–2) is a central result of social evolution theory that has been derived in many different ways, some more precise and detailed than others, from the population genetics of allele frequency change and fixation probabilities, to the Price equation and evolutionary invasion analysis, and to quantitative genetics (see, e.g., Taylor and Frank, 1996; Frank, 1998; Rousset and Billiard, 2000; Roze and Rousset, 2003; Rousset, 2004; Taylor et al., 2007; Akçay and Van Cleve, 2012; Wakano et al., 2012; Lehmann and Rousset, 2014; Mullon et al., 2016; Van Cleve, 2015; Mullon and Lehmann, 2019; Van Cleve, 2020; Avila et al., 2021). All these derivations fully agree with each other and this stubbornly robust result will be taken as the evolutionary backbone and starting point of the present analysis. Eqs. (1)–(2) also imply that an increase in trait value brought by mutation is favored, to the first-order in a resident population at u, when $s(u) > 0$, i.e., Hamilton's marginal rule holds and a trait value u satisfying $s(u) = 0$ will be called a singular trait value. Nothing in Hamilton's marginal rule implies that the evolving trait is a particular behavioral action and this is true for Hamilton's (1964) initial

formulation of the inclusive fitness effect. It could be any quantitative genetic infinitesimal effect affecting more or less remotely the behavior of interacting individuals under the model's demographic assumptions. To bring this upfront, we need to first detail behavioral interactions and then describe how such interactions affect fitness.

2.1.2 The individual-surroundings interface

Each individual is conceptualized as being an open system exchanging energy, matter, and/or information with its local surroundings possibly at each time point $t \in [0, T]$ of the span T of behavioral interaction occurring within one demographic time step (life-cycle iteration). This behavioral time span T can be thought of as the time available for behavioral interactions. If behavior occurs on a fast time scale relative to reproduction one can take $T \to \infty$, i.e., a very large number of interactions occur before reproduction and the separation of times scales between demography and behavior is infinite (a frequently endorsed assumption). An individual as an open system implies that its behavior can be predicted from the knowledge of a set of internal states and a set of external inputs. This perspective is enshrined in animal behavior theory (McFarland and Sibly, 1975; McFarland and Houston, 1981; Enquist and Ghirlanda, 2005) and follows more generally from control theory (or system's theory e.g., Arbib, 1987; Sontag, 1998; Athans and Falb, 2007; Weber, 2011). As such, the focal individual (section 2.1.1) can be characterized by three phenotypic attributes at each time point $t \in [0, T]$ of the behavioral interactions. First, an action $a_\bullet(t) \in \mathcal{A}$ expressed at time t, which is the fundamental phenotypic unit by which the individual interacts with its surrounding, e.g., a motor pattern, a signal, a transfer of resources to another individual, and where $\mathcal{A}$ denotes the set of actions. Second, an internal state $x_\bullet(t) \in \mathcal{X}$, which is a description of an individual's internal phenotypic features such as those pertaining to physiology, morphology, or neurology e.g., fat reserve, brain and gonad size, memory and beliefs about locations of food items or family members, and where $\mathcal{X}$ is the set of (internal) states. Third, an external input $\phi_\bullet(t) \in \mathbf{\Phi}$, which is any more or less noisy private or public signal about other's actions or any other cue or material received from the surroundings, e.g., assimilated food, a signal, a perception of the actions of others, where $\mathbf{\Phi}$ denotes the set of inputs. While this state-space representation of behavior is in principle general (Haykin, 1999), we here make two specific sets of assumptions to obtain a tractable model. First, for simplicity of presentation of the main concepts, we start by assuming that actions and states are one-dimensional (i.e., $a_\bullet(t) \in \mathbb{R}$ and $x_\bullet(t) \in \mathbb{R}$, and the relevant multidimensional case will be discussed in section 4). Second, we assume a deterministic model and that the behavior of the focal at time $t \in [0, T]$ can be described by the following system of ordinary differential equations:

$$\begin{aligned} a_\bullet(t) &= d(x_\bullet(t), u_\bullet) \\ \frac{\mathrm{d}x_\bullet(t)}{\mathrm{d}t} &= \tilde{g}(a_\bullet(t), x_\bullet(t), \phi_\bullet(t), u_\bullet) \\ \phi_\bullet(t) &= h(a_\circ(t), a(t), u_\bullet). \end{aligned} \tag{3}$$

The first line says that the action of the focal individual depends on a decision rule mapping its internal state to action and that this decision rule is parameterized by the individual's quantitative trait $u_\bullet$ (formally $d : \mathbb{R} \times U \to \mathbb{R}$). The second line says that the change in the state of the focal at time t depends on its action $a_\bullet(t)$ and state $x_\bullet(t)$, as well as on the input $\phi_\bullet(t)$ received from the environment and that the mapping $\tilde{g}$ relating these quantities to behavior also depends on the evolving trait (formally $\tilde{g} : \mathbb{R} \times \mathbb{R} \times \mathbf{\Phi} \times U \to \mathbb{R}$). Finally, the last line says that the input $\phi_\bullet(t)$ to the focal is a mapping depending on the average action $a_\circ(t)$ of group neighbour, the average action $a(t)$ in the population at large, and the focal's genetically determined trait (formally $h : \mathbb{R}^2 \times U \to \mathbf{\Phi}$). Here and throughout all functions are assumed continuously differentiable.

Eq. (3) specifies how the focal acts and reacts with its surroundings and thus defines its *behavior rule*, namely the collection of functions d, $\tilde{g}$, and h together with initial conditions for eq. (3). In order to determine the resulting sequence of actions, we also need to specify the action and state dynamics of group members and individuals at large from the population. Since we assumed that individuals are homogeneous and thus differ only by way of expressing different genetic traits, their behavior rules can be taken to be characterized by the same functions d, $\tilde{g}$ and h. Thus the action of an average group neigbour and average individual from the population are, respectively,

given by $a_\circ(t) = d(x_\circ(t), u_\circ)$ and $a(t) = d(x(t), u)$. From this and eq. (3), we see that the action and change in state of the focal individual is determined once we know the collections

$$\boldsymbol{u}_\bullet = (u_\bullet, u_\circ, u) \qquad \boldsymbol{x}_\bullet(t) = (x_\bullet(t), x_\circ(t), x(t)) \tag{4}$$

of traits and states, respectively, of all possible actors on the focal's state dynamics. Thus, it is useful to introduce the function $g : \mathcal{U}^3 \times \mathbb{R}^3 \to \mathbb{R}$ defined as

$$g(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t)) = \tilde{g}(d(x_\bullet(t), u_\bullet), x_\bullet(t), h(d(x_\circ(t), u_\circ), d(x(t), u), u_\bullet), u_\bullet), \tag{5}$$

which simply concatenates the different dependencies given in eq. (3) and brings upfront the dependency of the focal's state dynamics on the collection of traits $\boldsymbol{u}_\bullet$ and states $\boldsymbol{x}_\bullet(t)$ of each actor on the focal's behavior rule. Thereby, we can write the focal state dynamics as

$$\frac{\mathrm{d}x_\bullet(t)}{\mathrm{d}t} = g(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t)), \tag{6}$$

and since all individuals are homogeneous we can express the rate of change of the state of an average neighbour of the focal and an average individual simply by permuting the arguments of the function g to obtain

$$\frac{\mathrm{d}x_\circ(t)}{\mathrm{d}t} = g(\boldsymbol{u}_\circ, \boldsymbol{x}_\circ(t)) \quad \text{and} \quad \frac{\mathrm{d}x(t)}{\mathrm{d}t} = g(\boldsymbol{u}, \boldsymbol{x}(t)), \tag{7}$$

where the vectors

$$\begin{aligned} \boldsymbol{u}_\circ &= (u_\circ, u_\mathrm{n}, u), & \boldsymbol{x}_\circ(t) &= (x_\circ(t), x_\mathrm{n}(t), x(t)) \\ \boldsymbol{u} &= (u, u, u), & \boldsymbol{x}(t) &= (x(t), x(t), x(t)) \end{aligned} \tag{8}$$

collect the (average) traits and states of actors on the state variables of an average neighbour of the focal individual (first line), and on an average individual in the population (second line), respectively (here and throughout all vectors are defined by default as being column vectors). These actors are thus second-order level actors on the focal recipient since they affect the state variables of actors affecting the focal's state dynamics. Note that the subscripts of the trait vectors ($\boldsymbol{u}_\bullet$, $\boldsymbol{u}_\circ$ and $\boldsymbol{u}$) and state vectors ($\boldsymbol{x}_\bullet(t)$, $\boldsymbol{x}_\circ(t)$ and $\boldsymbol{x}(t)$) emphasise the individual (actor) from who's perspective the second-order actors' control and variables are collected. Accordingly, since groups are of size n, the vectors in eq. (8) contain the elements

$$u_\mathrm{n} = \frac{1}{n-1}u_\bullet + \left(\frac{n-2}{n-1}\right)u_\circ, \qquad x_\mathrm{n}(t) = \frac{1}{n-1}x_\bullet(t) + \left(\frac{n-2}{n-1}\right)x_\circ(t), \tag{9}$$

which are, for an average neighbour of the focal, the control and state expressions of average neighbours viewed as actors on the focal individual.

Equations (3)-(7) together fully determine how an individuals acts and reacts with its surroudings. While tracking the dynamics of three state variables (eqs. 6–7) may appear complicated, it is much simpler than tracking the state of all individuals in a group separately plus those in the population, which would require as many equations as there are individuals in the group plus at least the average in the population, and this is the approach often taken in the game theory (e.g., Basar and Olsder, 1999; Weber, 2011) or the evolutionary literatures (e.g., Day and Taylor, 2000). The useful reduction in complexity embodied in eqs. (6)–(7) obtains because using only the average phenotypes and states variable of the different actors on the focal approximates possibly more complicated relationships between genetic states and phenotypic feature to the first-order. This is sufficient to evaluate the selection gradient $s(u)$, since to the first-order expectations of functions depending on phenotypes can be replaced by functions of expected phenotypes [see Rousset (2004, p. 95) for general social evolution considerations and Avila et al. (2021) for application to state space models].

2.1.3 Fitness

Having specified how individuals interact, let's now specify how the focal's fitness $w(u_\bullet, u_\circ, u)$ is determined as a result of these interactions and consider two types of fitness functions. First, the discounted fitness

$$w(u_\bullet, u_\circ, u) = \int_0^T l_\bullet(t)\tilde{w}(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))\,\mathrm{d}t + l_\bullet(T)\sigma(\boldsymbol{x}_\bullet(T)), \tag{10}$$

where $l_\bullet(t)$ is the probability that the focal individual survives until time t that satisfies

$$\frac{\mathrm{d}l_\bullet(t)}{\mathrm{d}t} = -\mu(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))l_\bullet(t) \text{ with initial condition } l_\bullet(t) = 1 \tag{11}$$

with $\mu(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))$ being the instanteneous death rate of the focal, which depends on traits and states (we also assume that $l_\bullet(t) \to 0$ as $t \to \infty$, i.e., death eventually occurs). The term $\tilde{w}(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))$ is the rate of increase of fitness due to effective reproduction at time t (or the contribution to effective reproduction of the states at time t). The term $l_\bullet(T)\sigma(\boldsymbol{x}_\bullet(T))$ is the so-called scrap value, which gives the contribution to fitness from effective reproduction $\sigma(\boldsymbol{x}_\bullet(T))$ at the final time $t = T$ (formally $\tilde{w} : \mathcal{U}^3 \times \mathbb{R}^3 \to \mathbb{R}_+$ and $\sigma : \mathbb{R}^3 \to \mathbb{R}_+$). Fitness is actually mediated by actions, since this is the medium by which individuals interact with their surroundings and so the vital rates, $\tilde{w}(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))$ and $\mu(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))$, should all depend on actions. But since actions in turn depend on states (recall eq. 3), this justifies to write the vital rates directly as function of the states.

The discounted fitness eq. (10) covers at least three classes of demographic situations. (a) Interaction between familly members, like betweeen siblings, in a familly structured populations (e.g, Michod, 1982; Roze and Rousset, 2004), where individuals can die during the ontogenic period. Depending on the model formulation $\tilde{w}(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))$ may be zero and all fitness contribution is captured by the scrap value. (b) Interactions among group members in a geographically structured population. With limited dispersal, the fitness component $\tilde{w}(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))$ can depends in a non-linear way on various vital rates of group members and individuals from the population at large (e.g, Roze and Rousset, 2003, eq. 35, Akçay and Van Cleve, 2012, eq. A12, Van Cleve, 2015, eq. 38, Mullon et al., 2016, eq. box 1a), which themselves may depend on integrals depending on the traits of the individuals in interaction. Such situations are taken into account in the above formalization either by (i) defining state variables whose integrated values represent the integral, and are covered by the scrap value $\sigma(\boldsymbol{x}_\bullet(T))$ in eq. (12) or (ii) by noting that to the first-order, functions of integrals can be replaced by integrals of first-order Taylor series of fitness and hence the $\tilde{w}(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))$ fitness component in eq. (12) may be evaluated as a first-order Taylor expansion of fitness in its vital rates (e.g., Van Cleve, 2015, eq. 39, Mullon et al., 2016, eq. A60-A61). (c) Finally, eq. (10) covers a case that is not fully apparent under the main assumptions (section 2.1.1) yet is useful to mention. This is the case of the standard life-history evolution context in a panmictic age-structured population in which case the vital rates $\tilde{w}$ and μ depend only on the state of the focal $x_\bullet(t)$ and its own trait $u_\bullet$ (e.g., Stearns, 1992; Perrin, 1992) and possibly the trait u of the population at large to capture density-dependent regulation or on $(u, x_\circ(t))$ to capture interactions among individuals of the same age (e.g., Day and Taylor, 1997). In this cases, relatedness is zero relatedness ($r(u) = 0$), $\tilde{w}(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))$ is simply the effective fecundity of the focal individual at age t, and $\tilde{w}(\boldsymbol{x}_\bullet(T))$ is usually considered nill.

Second, we consider the average fitness

$$w(u_\bullet, u_\circ, u) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \tilde{w}(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))\,\mathrm{d}t, \tag{12}$$

which is the average of the instanteneous contribution $\tilde{w}(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))$ to fitness. Since $T \to \infty$, the interactions during behavioral time are assumed to occur on a faster time scale than reproduction. Another way of interpreting this biologically, is to say that the behavioral dynamics is assumed to be able to reach some equilibrium before reproduction occurs. The average fitness eq. (10) can be applied to demographic situations (a)-(b) described above, but does not cover the standard life-history evolution context [demographic situation (c) above], since survival is implicitly assumed to be equal to one.

2.2 Selection on behavior rules

2.2.1 The Hamiltonian and the co-state variables

Having specified the assumptions about behavioral interactions (eq. (3) and eq. (6)) and fitness (eq. (10) and eq. (12)), we now turn to work out a meaningfull representations of the selection gradient $s(u)$ using concepts from

optimal control theory (Bryson and Ho, 1975; Athans and Falb, 2007; Sydsaeter et al., 2008; Weber, 2011). For this, let's use introduce the so-called Hamiltonian function defined here as

$$H\Big(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t), l_\bullet(t), \boldsymbol{\lambda}(t), v(t)\Big) = \Big(\tilde{w}(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t)) - \mu(\boldsymbol{u}_\bullet(t), \boldsymbol{x}_\bullet(t))\, v(t)\Big) l_\bullet(t) + g(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))\lambda_\bullet(t) + g(\boldsymbol{u}_\circ, \boldsymbol{x}_\circ(t))\lambda_\circ(t), \tag{13}$$

where $\boldsymbol{\lambda}(t) = (\lambda_\bullet(t), \lambda_\circ(t))$. Extending a classical interpretation (Dorfman, 1969, p. 822), the Hamiltonian (formally the function $H : U^3 \times \mathbb{R}^3 \times \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$) can be regarded as the contribution to the focal individual's fitness of the expression of own and others' current actions, holding the expression of all future actions constant and thus at their behavioral level determined by the trait u of the resident population. The Hamiltonian is thus the sum of the focal's current rate of fitness increase $l_\bullet(t)\tilde{w}(t, \boldsymbol{u}_\bullet(t), \boldsymbol{x}_\bullet(t))$ plus the changes in its survival $[-\mu(\boldsymbol{u}_\bullet(t), \boldsymbol{x}_\bullet(t))l_\bullet(t)]$ and fitness relevant states $[g(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))$ and $g(\boldsymbol{u}_\circ, \boldsymbol{x}_\circ(t))]$, resulting from current actions, were each such change is weighted by its marginal effect on the focal's expected remaining fitness. These marginal effects on fitness, $v(t)$, $\lambda_\bullet(t)$ and $\lambda_\circ(t)$, are collectively referred to as the costates and where $v(t)$ is actually Fisher's (1930) reproductive value (change in remaining fitness at t resulting from change in survival until t) and $\boldsymbol{\lambda}(t) = (\lambda_\bullet(t), \lambda_\circ(t))$ will be called the shadow values of the actions (change in remaining fitness stemming from change in states at t). All costates are evaluated in a monomorphic resident population where all individual bear trait value u (see Box 1 and Avila et al., 2021 for more formalities and interpretations about how the costates connect to reproductive value).

Box 1. Shadow values and reproductive value as marginal values. In order to understand the biologically useful interpretation of the costates consider the discounted fitness case (eq. 10) and let's introduce the *future value reproductive value* defined as

$$\tilde{v}(t, \boldsymbol{x}(t), l(t)) = \int_t^T l(\tau)\tilde{w}(\boldsymbol{u}, \boldsymbol{x}(\tau))\,\mathrm{d}\tau + l(T)\sigma(\boldsymbol{x}(T)).$$

This is the expected remaining fitness from time t onwards of a individual when the current state variables of the individuals is $\boldsymbol{x}(t)$ and all variable are evaluated at the trait value u of the resident population (compare to eq. 10). By contrast, Fisher' s reproductive value, noted $v(t)$, is the expected remaining effective fecundity from time t onwards conditional on the individual surviving until time t (Fisher, 1930; Goodman, 1982) and it is therefore given as the *current value reproductive value*:

$$v(t) = \frac{\tilde{v}(t, \boldsymbol{x}(t), l(t))}{l(t)} = \int_t^T \exp\left(-\int_t^\tau \mu(\boldsymbol{u}, \boldsymbol{x}(h))\,\mathrm{d}h\right)\tilde{w}(\boldsymbol{u}, \boldsymbol{x}(\tau))\,\mathrm{d}\tau + \exp\left(-\int_t^T \mu(\boldsymbol{u}, \boldsymbol{x}(h))\,\mathrm{d}h\right)\sigma(\boldsymbol{x}(T)).$$

The second equality was reached by noting that $l(\tau)/l(t) = \exp\left(-\int_t^\tau \mu(\boldsymbol{u}, \boldsymbol{x}(h))\,\mathrm{d}h\right)$ is the probability of survival from time t to time $\tau \geq t$, which brings upfront that reproductive value $v(t)$ at t is independent of the value of survival $l(t) = \exp\left(-\int_0^t \mu(\boldsymbol{u}, \boldsymbol{x}(h))\,\mathrm{d}h\right)$ until t. Hence, it follows from this and the first equality in eq. (14) that

$$v(t) = \frac{\partial \tilde{v}(t, \boldsymbol{x}(t), l(t))}{\partial l(t)}.$$

That is, Fisher's reproductive value is the marginal change of the future value reproductive value with respect to change in survival (see also León, 1976; Perrin, 1992). Such interpretation of the costates variables as marginal changes in future value reproductive value holds generally, since one also have

$$\lambda_\bullet(t) = \frac{\partial \tilde{v}(t, \boldsymbol{x}_\bullet(t), l(t))}{\partial x_\bullet(t)} \qquad \lambda_\circ(t) = \frac{\partial \tilde{v}(t, \boldsymbol{x}_\bullet(t), l(t))}{\partial x_\circ(t)},$$

where all partial derivatives are evaluated at u and $\boldsymbol{x}_\bullet(t) = \boldsymbol{x}(t) = (x(t), x(t), x(t))$ (see Avila et al., 2021, Appendix B.1 for a proof that costates are generally marginal changes in $\tilde{v}(t, \boldsymbol{x}(t), l(t))$ regardless of the mode of trait control in an evolutionary context). Finally, note that demographic consistency entails that

$$w(u, u, u) = \tilde{v}(0, \boldsymbol{x}(0), l(0)) = v(0) = 1,$$

and that under an infinite horizon case ($T \to \infty$) for constant reproduction and mortality, say $\tilde{w}(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t)) = \tilde{w}(\boldsymbol{u}, \boldsymbol{x}_\bullet)$ and $\mu(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t)) = \mu(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))$, one has $v(t) = \tilde{w}(\boldsymbol{u}, \boldsymbol{x}_\bullet)/\mu(\boldsymbol{u}, \boldsymbol{x}_\bullet(t)) = 1$.

The Hamiltonian, among other things, is so defined as to generate the rate of changes of the costates, which are obtained by differentiation: $\mathrm{d}v(t)/\,\mathrm{d}t = -\partial H/\partial l_\bullet(t)$, $\mathrm{d}\lambda_\bullet(t)/\,\mathrm{d}t = -\partial H/\partial x_\bullet(t)$, and $\mathrm{d}\lambda_\circ(t)/\,\mathrm{d}t = -\partial H/\partial x_\circ(t)$, where all partial derivatives are evaluated at $\boldsymbol{u}_\bullet = \boldsymbol{u} = (u, u, u)$ and $\boldsymbol{x}_\bullet(t) = \boldsymbol{x}(t) = (x(t), x(t), x(t))$. From this and using eq. (9) in eq. (13), the costates are fully determined by the following system of ordinary differential equations

(ODE's):

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = g(\boldsymbol{u}, \boldsymbol{x}(t)) \qquad \text{with i.c. } x(0) = x_0, \tag{14a}$$

$$\frac{\mathrm{d}l(t)}{\mathrm{d}t} = -\mu(\boldsymbol{u}, \boldsymbol{x}(t))l(t) \qquad \text{with i.c. } l(0) = 1 \tag{14b}$$

$$-\frac{\mathrm{d}\boldsymbol{\lambda}(t)}{\mathrm{d}t} = \mathbf{b}(\boldsymbol{x}(t), l(t), v(t)) + \mathbf{A}(\boldsymbol{x}(t))\boldsymbol{\lambda}(t) \qquad \text{with b.c. } \boldsymbol{\lambda}(\zeta) = \boldsymbol{\lambda}_\zeta \tag{14c}$$

$$-\frac{\mathrm{d}v(t)}{\mathrm{d}t} = \tilde{w}(\boldsymbol{u}, \boldsymbol{x}(t)) - \mu(\boldsymbol{u}, \boldsymbol{x}(t))\, v(t) \qquad \text{with b.c. } v(\zeta) = v_\zeta, \tag{14d}$$

where

$$\mathbf{b}(\boldsymbol{x}(t), l(t), v(t)) = l(t) \begin{pmatrix} \frac{\partial \tilde{w}(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)} - \frac{\partial \mu(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)} v(t) \\ \frac{\partial \tilde{w}(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)} - \frac{\partial \mu(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)} v(t) \end{pmatrix}_{\boldsymbol{x}_\bullet(t) = \boldsymbol{x}(t)} \tag{15}$$

is a vector of marginal changes in fitness stemming from changes in state and

$$\mathbf{A}(\boldsymbol{x}(t)) = \begin{pmatrix} \frac{\partial g(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)} & \frac{\partial g(\boldsymbol{u}_\circ, \boldsymbol{x}_\circ(t))}{\partial x_\bullet(t)} \\ \frac{\partial g(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)} & \frac{\partial g(\boldsymbol{u}_\circ, \boldsymbol{x}_\circ(t))}{\partial x_\circ(t)} \end{pmatrix}_{\boldsymbol{x}_\bullet(t) = \boldsymbol{x}(t)} = \begin{pmatrix} \frac{\partial g(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)} & \frac{\partial g(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)} \frac{1}{n-1} \\ \frac{\partial g(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)} & \frac{\partial g(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)} + \frac{\partial g(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)} \left(\frac{n-2}{n-1}\right) \end{pmatrix}_{\boldsymbol{x}_\bullet(t) = \boldsymbol{x}(t)} \tag{16}$$

is a matrix of change in state dynamics where the second representation follows from using eq. (9), which allows to express $\mathbf{A}(\boldsymbol{x}(t))$ only in terms of the changes in state of the focal individual (both these representations of $\mathbf{A}(\boldsymbol{x}(t))$ can be usefull). In system (14), "i.c" stands for initial condition, i.e., the value of the states at time $t = 0$ and "b.c." stands for boundary condition, i.e., the value of the costates at some arbitrary time $t = \zeta$, which is due to the fact that, depending on the situation, the boundary conditions of the costates are initial or final values (more on this below). While eq. (14a) will generally be non-linear, eq. (14c) is a linear non-homogeneous ODE with varying coefficients and representations for its solution are discussed in Appendix A.

The ODE system (14) is centerpiece to the analyzes of selection and given initial condition for the states, boundary conditions for the costates, and trait value u, its solution is unique and so it can be thought to be parameterized by u. While the initial condition of the states, x_0 is a parameter, the boundary conditions of the costates are determined by the specificities of the situation at hand. No general recipe for determining the boundary conditions exists (e.g., Aseev and Kryazhimskiy, 2008; Sydsaeter et al., 2008), yet for the finite horizon case the boundary conditions are straighforwardly obtained from their definitions as marginal effects of changes in states on remaining fitness evaluated in a monomorphic population at u: $v(T) = \sigma(\boldsymbol{x}(T))$ and $\boldsymbol{\lambda}(T) = (l(T)\, \partial\sigma\left(\boldsymbol{x}_\bullet(T)\right)/\partial x_\bullet(T), l(T)\, \partial\sigma\left(\boldsymbol{x}_\bullet(T)\right)/\partial x_\circ(T))|_{\boldsymbol{x}_\bullet(T) = \boldsymbol{x}(T)}$ (see Box 1). Difficulties for determining the boundary condition arise for discounted fitness under the infinite horizon case (Sydsaeter et al., 2008; Aseev and Kryazhimskiy, 2008; Weber, 2011), where it may be felt that the costates go asymptotically to zero but this is not generally the case. For instance, for constant survival and reproduction, the reproductive value remains the same throuhgout lifespan and the reproductive value has always the constraint that its initial value is one (see Box 1). But in the presence of *dominating discounts*, where the death rate suppresses the growth rate of the gradient of effective fecundity, then the shawdow values should go to zero $\lim_{T\to\infty} \boldsymbol{\lambda}(T) = (0, 0)$, which fixes their boundary conditions. This is proved by Aseev and Kryazhimskiy (2008) for the case of constant discounts, but external mortality can usually be taken as constant, so there is alway a background of constant discounts to future effective fecundity. It is actually plausible that the majority of biological situations of interest actually do satisfy the dominating discounts assumption, at least by assuming high enough external mortality, one may systematically be able to enforce it (the work of Aseev and Kryazhimskiy, 2008 should be consulted for the exact mathematical specification of dominating discounts condition, which is beyond the scope of this paper). With this assumption in mind we are lead to our first result (proved in Appendix B.2.1) about the representation of the selection gradient.

Result 1: Discounted fitness with finite and infinite horizon. *Suppose individual fitness is given by eq.* (10) *with state dynamics by eqs.* (6)-(7)*, then the inclusive fitness effect for condition* (1) *can be represented as*

$$s(u) = \int_0^T s(t, u)\, \mathrm{d}t, \tag{17}$$

where

$$s(t,u)=\left[\frac{\partial H\Big(\boldsymbol{u}_\bullet,\boldsymbol{x}_\bullet(t),l_\bullet(t),\boldsymbol{\lambda}(t),v(t)\Big)}{\partial u_\bullet}+r(u)\frac{\partial H\Big(\boldsymbol{u}_\bullet,\boldsymbol{x}_\bullet(t),l_\bullet(t),\boldsymbol{\lambda}(t),v(t)\Big)}{\partial u_\circ}\right]_{\substack{\boldsymbol{u}_\bullet=\mathbf{u}\\ \boldsymbol{x}_\bullet(t)=\boldsymbol{x}(t)}}, \tag{18}$$

which, owing to eq. (13)*, can also be written as*

$$\begin{aligned}s(t,u)=\Bigg[&\left(\frac{\partial\tilde{w}(\boldsymbol{u}_\bullet,\boldsymbol{x}(t))}{\partial u_\bullet}-\frac{\partial\mu(\boldsymbol{u}_\bullet,\boldsymbol{x}(t))}{\partial u_\bullet}v(t)\right)l(t)+\frac{\partial g(\boldsymbol{u}_\bullet,\boldsymbol{x}(t))}{\partial u_\bullet}\lambda_\bullet(t)+\frac{\partial g(\boldsymbol{u}_\circ,\boldsymbol{x}(t))}{\partial u_\bullet}\lambda_\circ(t)\\&+r(u)\left(\left(\frac{\partial\tilde{w}(\boldsymbol{u}_\bullet,\boldsymbol{x}(t))}{\partial u_\circ}-\frac{\partial\mu(\boldsymbol{u}_\bullet,\boldsymbol{x}(t))}{\partial u_\circ}v(t)\right)l(t)+\frac{\partial g(\boldsymbol{u}_\bullet,\boldsymbol{x}(t))}{\partial u_\circ}\lambda_\bullet(t)+\frac{\partial g(\boldsymbol{u}_\circ,\boldsymbol{x}(t))}{\partial u_\circ}\lambda_\circ(t)\right)\Bigg]_{\boldsymbol{u}_\bullet=\mathbf{u}}.\end{aligned} \tag{19}$$

Therein, the states and costates satisfy eq. (14) *with b.c. of the costates given by* $v(T)=\sigma(\boldsymbol{x}(T))$ *and* $\boldsymbol{\lambda}(T)=\left(l(T)\,\partial\sigma\left(\boldsymbol{x}_\bullet(T)\right)/\partial x_\bullet(T),l(T)\,\partial\sigma\left(\boldsymbol{x}_\bullet(T)\right)/\partial x_\circ(T)\right)|_{\boldsymbol{x}_\bullet(T)=\boldsymbol{x}(T)}$ *for a finite horizon. For an infinite horizon, eq.* (17) *holds with the assumption that* $\lim_{T\to\infty}\boldsymbol{\lambda}(T)=(0,0)$*, i.e., dominating discounts obtains.*

From a biological perspective, this result says that selection on the evolving trait is the sum over all times, of its inclusive marginal effect on current and remaining reproduction, where the latter effect is mediated by how the change in trait value changes own and others' state expression, weighted by the marginal effect of changing these states on expected remaining reproduction. From, a computational perspective, it shows that the main technical difficulty to evaluate condition (1) is to solve the ODE system (14) and integrate $s(t,u)$ over time. This is generally a much simpler task than attempting to solve the problem by "brute force", i.e. solving eqs. (6)–(7), substituting into fitness (eq. (10) or eq. (12)). Hence, there are gains to use the control theory concepts for both obtaining a biolological interpretation of the selection pressure on a trait under incomplete plasticitiy, and for its concrete computation [result 1 generalizes to the infinite horizon case and to an explicit behavioral rule context the Constant control result of Avila et al. (2021, p. 11)].

Depending on the model, however it may remain a challenge to solve the ODE system (14) and we now describe a case where this can be markedly simplified. This is the case where the biological situation dictates that the initial state satisfies the resident equilibrium, i.e., $x_0=\hat{x}$ with $g(\boldsymbol{u},(\hat{x},\hat{x},\hat{x}))=0$. While this clearly excludes many situations pertaining to life-history evolution, it does include biologically relevant cases. For instance, it obtains when the initial condition of the states are inherited from the previous generation, which occurs in the case where knowledge, resources, or stationary behavior, is transferred between generations and this is typical of models for the evolution of learning or social interactions (e.g., Killingback and Doebeli, 2002; Taylor and Day, 2004; Aoki et al., 2012; Wakano and Miura, 2014; Kobayashi et al., 2019, see also section 3). To analyze this case, let $\bar{l}=\int_0^\infty l(t)\,\mathrm{d}t$ denote the average lifespan of an individual and $\bar{\boldsymbol{\lambda}}=(\bar{\lambda}_\bullet,\bar{\lambda}_\circ)=(\int_0^\infty\lambda_\bullet(t)\,\mathrm{d}t/\bar{l},\int_0^\infty\lambda_\circ(t)\,\mathrm{d}t/\bar{l})$ the average shadow values over lifespan. This leads us to the following result proved in Appendix B.2.2.

Result 2: Discounted fitness with infinite horizon and initial resident equilibrium state. *Suppose individual fitness is given by eq.* (10) *with state dynamics by eqs.* (6)-(7) *under an infinite horizon with dominating discounts, i.e.,* $T\to\infty$ *with* $\lim_{T\to\infty}\boldsymbol{\lambda}(T)=(0,0)$*, and that the initial state* $x_0=\hat{x}$ *satisfies* $g(\boldsymbol{u},\hat{\boldsymbol{x}})=0$ *where* $\hat{\boldsymbol{x}}=(\hat{x},\hat{x},\hat{x})$*. Then, the inclusive fitness for condition* (1) *can be written as*

$$s(u)=\bar{l}\left[\frac{\partial H\Big(\boldsymbol{u}_\bullet,\hat{\boldsymbol{x}},1,\bar{\boldsymbol{\lambda}},1\Big)}{\partial u_\bullet}+r(u)\frac{\partial H\Big(\boldsymbol{u}_\bullet,\hat{\boldsymbol{x}},1,\bar{\boldsymbol{\lambda}},1\Big)}{\partial u_\circ}\right]_{\boldsymbol{u}_\bullet=\mathbf{u}}, \tag{20}$$

or more explicitly as

$$\begin{aligned}s(u)\mu(\boldsymbol{u},\hat{\boldsymbol{x}})=&\left[\frac{\partial\tilde{w}(\boldsymbol{u}_\bullet,\hat{\boldsymbol{x}})}{\partial u_\bullet}-\frac{\partial\mu(\boldsymbol{u}_\bullet,\hat{\boldsymbol{x}})}{\partial u_\bullet}+\frac{\partial g(\boldsymbol{u}_\bullet,\hat{\boldsymbol{x}})}{\partial u_\bullet}\bar{\lambda}_\bullet+\frac{\partial g(\boldsymbol{u}_\circ,\hat{\boldsymbol{x}})}{\partial u_\bullet}\bar{\lambda}_\circ\right]_{\boldsymbol{u}_\bullet=\mathbf{u}}\\&+r(u)\left[\frac{\partial\tilde{w}(\boldsymbol{u}_\bullet,\hat{\boldsymbol{x}})}{\partial u_\circ}-\frac{\partial\mu(\boldsymbol{u}_\bullet,\hat{\boldsymbol{x}})}{\partial u_\circ}+\frac{\partial g(\boldsymbol{u}_\bullet,\hat{\boldsymbol{x}})}{\partial u_\circ}\bar{\lambda}_\bullet+\frac{\partial g(\boldsymbol{u}_\circ,\hat{\boldsymbol{x}})}{\partial u_\circ}\bar{\lambda}_\circ\right]_{\boldsymbol{u}_\bullet=\mathbf{u}},\end{aligned} \tag{21}$$

where $\bar{l}=1/\mu(\boldsymbol{u},\hat{\boldsymbol{x}})$ *and* $\bar{\boldsymbol{\lambda}}=\left(\mu(\boldsymbol{u},\hat{\boldsymbol{x}})\mathbf{I}-\mathbf{A}(\hat{\boldsymbol{x}})\right)^{-1}\mathbf{b}(\hat{\boldsymbol{x}},1,1)$ *with* $\mathbf{I}$ *being the identity matrix.*

From a biological perspective, this result says that the selection gradient on the trait is the sum of the inclusive effects on the stationary growth rate $(\tilde{w}(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}}) - \mu(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}}))$ of the focal and on the average (through lifespan) of remaining reproduction, where the latter effect is mediated by how the change in trait value changes own and others' state expression weighted by the marginal effect of these states changed on expected remaing reproduction averaged over lifespan. From, a computational perspective, there is no difficulty at all in evaluating $s(u)$, since it only requires matrix inversion. This result makes analytically tractable behavior rule evolution models when there is a trade-off between effects on fecundity and mortality. Let us finally turn to a closely related result proved in Appendix B.2.3, which is the representation of the selection gradient under average fitness (12).

Result 3: Average fitness. *Suppose individual fitness is given by eq.* (12) *with state dynamics by eqs.* (6)-(7) *and that given initial conditions* $x(0) = x_0$ *and* $\boldsymbol{\lambda}(0) = \boldsymbol{\lambda}_0$ *at* $t = 0$*, the dynamics* (14a) *and* (14c) *converge to the (hyperbolically stable) equilibrium* $\hat{\boldsymbol{x}} = \lim_{t\to\infty} \boldsymbol{x}(t)$ *and* $\hat{\boldsymbol{\lambda}} = \lim_{t\to\infty} \boldsymbol{\lambda}(t)$*. Then the inclusive fitness effect for condition* (1) *can be written as*

$$s(u) = \left[\frac{\partial H\left(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}}, 1, \hat{\boldsymbol{\lambda}}, 0\right)}{\partial u_\bullet} + r(u) \frac{\partial H\left(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}}, 1, \hat{\boldsymbol{\lambda}}, 0\right)}{\partial u_\circ} \right]_{\boldsymbol{u}_\bullet = \mathbf{u}}, \tag{22}$$

or more explicitly as

$$s(u) = \left[\frac{\partial \tilde{w}(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\bullet} + \frac{\partial g(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\bullet} \hat{\lambda}_\bullet + \frac{\partial g(\boldsymbol{u}_\circ, \hat{\boldsymbol{x}})}{\partial u_\bullet} \hat{\lambda}_\circ + r(u) \left(\frac{\partial \tilde{w}(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\circ} + \frac{\partial g(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\circ} \hat{\lambda}_\bullet + \frac{\partial g(\boldsymbol{u}_\circ, \hat{\boldsymbol{x}})}{\partial u_\circ} \hat{\lambda}_\circ \right) \right]_{\boldsymbol{u}_\bullet = \mathbf{u}}, \tag{23}$$

where $\hat{\boldsymbol{x}}$ *satisfies* $g(\boldsymbol{u}, \hat{\boldsymbol{x}}) = 0$ *and the costates* $\hat{\boldsymbol{\lambda}} = (\hat{\lambda}_\bullet, \hat{\lambda}_\circ)$ *satisfy* $\hat{\boldsymbol{\lambda}} = -\mathbf{A}(\hat{\boldsymbol{x}})^{-1}\mathbf{b}(\hat{\boldsymbol{x}}, 1, 0)$.

From a biological perspective, this result says that the selection gradient is now the inclusive marginal effect of the trait on current and future reproduction, where the latter effect is mediated by how the change in trait value changes own and others' future action expression weighted by the marginal effect of these actions on equilibrium future reproduction. This is so because the average fitness-(12) entails that only the equilibrium value of the state variable and shadow value will matter for fitness (a consequence of the ergodic theorem). From, a computational perspective, we are in the same straightfoward position as under result 2 to evaluate $s(u)$.

These results 1-3 can be seen from at least two perspectives. First, they allow to analyze new situations for the evolution of phenotypic scheduling under various social evolution and life-history contexts by providing a recipe to compute the selection gradient on the trait affecting phenotypic dynamics. Second, they provide a biological interpretation of selection on such traits using control theory concepts, which can be used to get insights about the selection pressure in previous evolutionary models under incomplete plasticity. In the next section, we turn to illustrate these points.

3 Application: behavioral response models

We now illustrate the results by applying them to the evolution of behavior response rules, in which one or a few scalar traits affect how individuals interact with each other under various type of games (e.g. McNamara et al., 1999; Killingback and Doebeli, 2002; Taylor and Day, 2004; Lehmann and Keller, 2006; André and Day, 2007; Akçay and Van Cleve, 2009; Akçay and Van Cleve, 2012). To make the connection to this literature, we make the main assumptions implied by the model formulation in this literature for the control theory framework. First, states correspond directly to actions; namely the decision rule is the identity function ($a_\bullet(t) = x_\bullet(t)$) and thus does depend on the evolving trait. Further, the input is simply the action of group neighbours ($\phi_\bullet(t) = a_\circ(t)$). Second, the transition rule describing change of action depend only on the actor's evolving trait; namely $g(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t)) = g(u_\bullet, \boldsymbol{x}_\bullet(t))$ and $g(\boldsymbol{u}_\circ, \boldsymbol{x}_\circ(t)) = g(u_\circ, \boldsymbol{x}_\circ(t))$. Third, reproduction and survival depend directly only on state variables and not on the evolving trait of the actor; namely

$$\tilde{w}(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t)) = \tilde{w}(u, \boldsymbol{x}_\bullet(t)) \text{ and } \mu(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t)) = \mu(u, \boldsymbol{x}_\bullet(t)), \tag{24}$$

where the dependence on the resident trait u in the vital rarte must remain owing to density-dependent regulation (i.e, the binding constraint that $w(u,u,u)=1$). Together, the second and third assumption imply that

$$\frac{g(\boldsymbol{u}_\circ,\boldsymbol{x}_\circ(t))}{\partial u_\bullet}=\frac{g(\boldsymbol{u}_\bullet,\boldsymbol{x}_\bullet(t))}{\partial u_\circ}=0,\ \frac{\partial\tilde{w}(\boldsymbol{u}_\bullet,\boldsymbol{x}_\bullet(t))}{\partial u_\bullet}=\frac{\partial\tilde{w}(\boldsymbol{u}_\bullet,\boldsymbol{x}_\bullet(t))}{\partial u_\circ}=0,\ \frac{\partial\mu(\boldsymbol{u}_\bullet,\boldsymbol{x}_\bullet(t))}{\partial u_\bullet}=\frac{\partial\mu(\boldsymbol{u}_\bullet,\boldsymbol{x}_\bullet(t))}{\partial u_\circ}=0. \quad (25)$$

This setting covers models for the evolution of so-called reactive strategies in the iterated prisonner's dilemma game or other repeated games (e.g. McNamara et al., 1999; Killingback and Doebeli, 2002; Taylor and Day, 2004; Lehmann and Keller, 2006; André and Day, 2007; Akçay and Van Cleve, 2014), certain models for the evolution of preferences in such games (e.g. Akçay and Van Cleve, 2009; Akçay and Van Cleve, 2012), as well as models for the evolution of learning in games (e.g. Dridi and Akçay, 2018; Leimar, 2021). This is so because all formalize the evolution of behavior rule paramaterized by one or a few quantitative genetic traits and it can further be usefull to note that a large class of (stochastic) learning dynamics can actually be described by differential equations fitting into the setting of the behavioral rule model eq. (3) (Fudenberg and Levine, 1998; Dridi and Lehmann, 2014).

3.1 Average fitness

Let us consider selection on a trait affecting the evolution of a behavioral response rule following assumptions (25) by using average fitness (eq. 12). This setting should cover the model analyzed in Akçay and Van Cleve (2012) (that generalizes Akçay and Van Cleve (2009) to interactions between relatives and n players) since this model is premised on the facts that (i) individuals interactions occur among groups of fixed size n in a geographically structured population; (ii) the dynamical system underlying individual actions has reached stationarity (Fig. 1 and eq. 12 of Akçay and Van Cleve, 2012); (iii) this dynamical system is influenced only by the evolving genetic trait of that individual (eq. 2 of Akçay and Van Cleve, 2009 and implied by the analysis in Akçay and Van Cleve, 2012); and (iv) the fitness/payoff depends only on state variables having reached equilibrium and not directly on the evolving traits (first equation on p. 260 of Akçay and Van Cleve, 2012). This is thus an ideal situation to check the consistency of the optimal control approach to behavior rule evolution, since we should be able to recover the first-order condition for uninvadability of Akçay and Van Cleve (2012, eq. 5).

Under these assumptions, the analysis of the model is covered by result 3. On substituting eq. (25) into eq. (23) shows that the selection gradient reduces to

$$s(u)=\frac{\partial g(u_\bullet,\hat{\boldsymbol{x}})}{\partial u_\bullet}\left(\hat{\lambda}_\bullet+r(u)\hat{\lambda}_\circ\right). \quad (26)$$

Selection thus depends on how the change in trait of the focal affects the change in its action dynamics, $\partial g(\boldsymbol{u}_\bullet,\hat{\boldsymbol{x}}/\partial u_\bullet$, weighted by $\hat{\lambda}_\bullet+r(u)\hat{\lambda}_\circ$, which can be interpreted as the stationary inclusive shadow value of the focal's actions. This is the sum of the effect $\hat{\lambda}_\bullet$ of the focal's action on its (stationary) fitness and the effect $\hat{\lambda}_\circ$ of these actions on the (stationary) fitness of an average neighbour weighted by relatedness $r(u)$ between the two individuals. A singular trait must satisfy $\partial g(u_\bullet,\hat{\boldsymbol{x}})/\partial u_\bullet$ and/or $\hat{\lambda}_\bullet+r(u)\hat{\lambda}_\circ=0$. Since the behavior response rule literature usually assumes that the function g is monotonic in its first argument, we focus on characterizing the condition $\hat{\lambda}_\bullet+r(u)\hat{\lambda}_\circ=0$.

Under assumptions (24)–(25) and from eq. (14c), the equilibrium shadow values satisfy

$$\begin{aligned}&\frac{\partial\tilde{w}(u,\boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)}+\hat{\lambda}_\bullet\frac{\partial g(u,\boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)}+\hat{\lambda}_\circ\frac{\partial g(u,\boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)}\frac{1}{n-1}\\&\frac{\partial\tilde{w}(u,\boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)}+\hat{\lambda}_\circ\frac{\partial g(u,\boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)}+\hat{\lambda}_\bullet\left[\frac{\partial g(u,\boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)}+\frac{\partial g(u,\boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)}\left(\frac{n-2}{n-1}\right)\right]=0.\end{aligned} \quad (27)$$

where recall that all derivatives, here and throughout, are evaluated at $\boldsymbol{x}_\bullet(t)=\hat{\boldsymbol{x}}$. Note that for the behavior response rule setting, the first column of matrix $\mathbf{A}(\boldsymbol{x}(t))$ (recall eq. (16)) describes the change in action of the focal individual when its own action and that of its neighbours are varied [first and and second entry, respectively, given by $\partial g(u,\boldsymbol{x}_\bullet(t))/\partial x_\bullet(t)$ and $\partial g(u,\boldsymbol{x}_\bullet(t))/\partial x_\circ(t)$ in eq. (27)]. The second colum of $\mathbf{A}(\boldsymbol{x}(t))$ describes the change of action of a neighbour of the focal when the focal and its neighbours change action. Matrix $\mathbf{A}(\boldsymbol{x}(t))$ can thus be interpreted

as describing the action reaction to action between interacting individuals and since the equilibrium shadow values satisfy $\hat{\boldsymbol{\lambda}} = (\hat{\lambda}_\bullet, \hat{\lambda}_\circ) = -\mathbf{A}(\hat{\boldsymbol{x}})^{-1}\mathbf{b}(\hat{\boldsymbol{x}}, 1, 0)$, the equilibrium action interdependence is captured by matrix $-\mathbf{A}(\hat{\boldsymbol{x}})^{-1}$. The ratio $\rho(u) = a_{12}/a_{11}$, where a_{ij} stands for entry ij of matrix $-\mathbf{A}(\hat{\boldsymbol{x}})^{-1}$, can thus be regarded as a behavioral response coefficient that tells us how much, in equilibrium, the action of a single neighbour of the focal is varied when the focal varies its action relative to the focal's own induced action variation. Algebraic rearrangements (easily done with a symbolic algebra system) then show that the response coefficient can be expressed as

$$\rho(u) = -\frac{\frac{1}{(n-1)}\frac{\partial g(u,\boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)}}{\frac{(n-2)}{(n-1)}\frac{\partial g(u,\boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)} + \frac{\partial g(u,\boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)}}. \tag{28}$$

Now solving for $(\hat{\lambda}_\bullet, \hat{\lambda}_\circ)$ in eq. (27), substituting into eq. (26), and rearranging shows that the condition for u to be singular can be written as

$$\underbrace{\frac{\partial \tilde{w}(u,\boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)}}_{-c} + \rho(u)\underbrace{\frac{\partial \tilde{w}(u,\boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)}}_{(n-1)b} + r(u)\left(\Big(1+\rho(u)(n-2)\Big)\underbrace{\frac{\partial \tilde{w}(u,\boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)}}_{(n-1)b} + \rho(u)(n-1)\underbrace{\frac{\partial \tilde{w}(u,\boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)}}_{-c}\right) = 0. \tag{29}$$

This first-order condition turns out to be conceptually exactly equivalent to eq. (5) of Akçay and Van Cleve (2012) [and eq. (8) of Akçay and Van Cleve, 2014], since we can write $\partial\tilde{w}(\boldsymbol{x}_\bullet(t))/\partial x_\circ(t)) = (n-1)b$, where b is conceptually equivalent to the effect of Akçay and Van Cleve (2012) of a single neighbor changing action on the focal's fitness and the response coefficient (28) is actually exactly eq. (13) of Akçay and Van Cleve (2012). This can be seen by noting that $\partial g(\mathbf{u}, \boldsymbol{x}_\bullet(t))/\partial x_\circ(t)$ is the effect of the whole set of group neigbours on the action dynamics of the focal individual, while $\partial^2 x_j/(\partial a_j \partial a_i)$ is the effect of a single neighbour on the action dynamics of the focal in Akçay and Van Cleve (2012) and so $\partial g(\mathbf{u}, \boldsymbol{x}_\bullet(t))/\partial x_\circ(t)/(n-1) = \partial^2 x_j/(\partial a_j \partial a_i)$ (the special notation $\partial^2 x_j/(\partial a_j \partial a_i)$ stems from the fact that Akçay and Van Cleve, 2012 endorse a gradient dynamics).

Since the response coefficient can be interpreted as the effect of the focal changing action on the action of a single neighbour (see also eq. 2 of Akçay and Van Cleve, 2012), this paves the way to an intuitive interpretation of eq. (29) by considering that $-c = \partial\tilde{w}(\boldsymbol{x}_\bullet(t))/\partial x_\bullet(t)$ is a fitness cost. Namely, eq. (29) says that the fitness cost to self of varying (infinitesimally) action expression can be recouped by two pathways. First, by the sum of the effect of each neigbour changing actions on the focal fitness weighted by the extent $\rho(u)$ to which any neighbour's action is varied when the focal varies its action. Second, by the net indirect effect on the fitness of any related neighbour, which consists of three distinct fitness effects: (i) the immediate effect of the focal varying action on the fitness of each related neighbour [effect of intensity $r(u)(n-1)b$]; (ii) the correlated effect resulting from the change of action of each neighbour of a focal's neighbour when the focal varies its action, which induces a change in fitness in any of the focal's neighbour related [effect of intensity $r(u)\rho(u)(n-2)(n-1)b$]; and (iii) the correlated effect of the change of action of each related neighbour of a focal's neighbour when the focal varies its action, which changes those neighbors fitness [effect of intensity $-r(u)\rho(u)(n-1)c$]. As emphasized by (Akçay and Van Cleve, 2014) this entails an interesting but neglected interaction between relatedness and behavioral responses, owing to the product $\rho(u)r(u)$ appearing in eq. (29).

Three other facts about the first-order condition eq. (29) are noteworthy. First, its original derivation by Akçay and Van Cleve (2012) is more direct and simpler as it uses the spatial average fitness. Namely, it consist of letting the state first go to equilibrium and use fitness at that state as $w(u_\bullet, u_\circ, u) = \tilde{w}(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})$ (which can be obtained from eq. (12) by invoking the ergodic theorem), and using implicit differentiation of the growth rates of actions at equilibrium with respect to trait values, which eventually leads to eq. (29). Second, eq. (29) is written at the fitness levels and the units of measurements therein are thus different than those of eq. (5) of Akçay and Van Cleve (2012), since they do not consider effective fecundity $\tilde{w}$, but actual fecundity. As such r in their eq. (5) is not relatedness but the scaled relatedness coefficient for an iteroparous population. This connects to the point of the last paragraph of section 2.1.3: given some specific life-fcycle assumptions express fitness in terms of vital rates and then use the chain rule to express fitness effects in terms of vital rates effect and then one would recover

eq. (5) of Akçay and Van Cleve (2012) with the same units of measurement; in other words eq. (29) and eq. (5) of Akçay and Van Cleve (2012) are scaled differently. Third, while preference evolution was emphasized in the original derivation of condition (29), it can as well be used to study the evolution of reactive strategies (e.g., Killingback and Doebeli, 2002; Taylor and Day, 2004; André and Day, 2007), which is a connection that is worked out in Akçay and Van Cleve (2014) and so will not be repeated here. Yet more generally, the first-order condition (29) applies to the evolution of traits affecting learning dynamics. In fact, it applies to any model that can be put under the form of the behavior rule dynamics eq. (3) [along with assumptions (24)–(25)] and so is robust to change in quite a large number of assumptions about behavioral dynamics. Hence, depending on the context, evolution my lead to very different values of the response coefficient $\rho(u)$ and it is not a conclusion that $\rho(u)$ always evolves to be positive (for instance, results on preference evolution under incomplete information (Alger et al., 2020) implies it goes to zero).

3.2 Discounted fitness

Let us now consider selection on a trait affecting the evolution of a behavioral response rule under assumptions (25) by using discounted fitness under an infinite time horizon (eq. 10 with $T \to \infty$). Then, the analysis is covered by result 2, and on substituting eqs. (24)–(25) into eq. (21) shows that the the selection gradient reduces to

$$s(u) = \frac{\partial g(u_\bullet, \hat{\boldsymbol{x}})}{\partial u_\bullet} \left(\bar{\lambda}_\bullet + r(u)\bar{\lambda}_\circ \right). \tag{30}$$

Hence, selection now depends on how the change in trait affect the change in the action dynamics of the focal individual, $\partial g(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})/\partial u_\bullet$ weighted by the inclusive average shadow value of these actions.

As for the previous model, let us focus on the condition where $\bar{\lambda}_\bullet + r(u)\bar{\lambda}_\circ = 0$ and in order to characterize it note that as before matrix $\mathbf{A}(\boldsymbol{x}(t))$ describes the action reaction to action between interacting individuals. Yet individuals may die as they age and for this model, the average action interdependence is captured by the matrix $(\mu(\hat{\boldsymbol{x}})\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1}$ since $\bar{\boldsymbol{\lambda}} = (\mu(\hat{\boldsymbol{x}})\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1}\mathbf{b}(\hat{\boldsymbol{x}}, 1, 1)$. As such the ratio $\tilde{\rho}(u) = a_{12}/a_{11}$, where a_{ij} now stands for entry ij of matrix $(\mu(\hat{\boldsymbol{x}})\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1}$, can be regarded as the behavioral response coefficient that tells us how much, on average throughout the whole span of interactions, the action of a single neighbour of the focal is varied when the focal varies its action relative to the focal own induced action variation. Algebraic rearrangements show that the response coefficient is now given by

$$\tilde{\rho}(u) = -\frac{\frac{1}{(n-1)}\frac{\partial g(u,\boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)}}{\frac{(n-2)}{(n-1)}\frac{\partial g(u,\boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)} + \frac{\partial g(u,\boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)} - \mu(\hat{\boldsymbol{x}})}, \tag{31}$$

and the difference with eq. (28) is thus only the $-\mu(\hat{\boldsymbol{x}})$ in the denominator. This will act to reduce the value of the response coefficient (owing to the minus sign in front of the ratio). This makes biological sense: the change of action by an individual may not feedback on future fitness through change of action by neighbours if it dies.

Using assumptions (24)–(25) in eqs. (15)–(16), computing $\bar{\boldsymbol{\lambda}} = (\mu(\hat{\boldsymbol{x}})\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1}\mathbf{b}(\hat{\boldsymbol{x}}, 1, 1)$, substituting into eq. (30), and rearranging shows that the condition for u to be singular can be written as

$$\frac{\partial h(u,\boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)} + \tilde{\rho}(u)\frac{\partial h(u,\boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)} + r(u)\left(\left(1 + \tilde{\rho}(u)(n-2)\right)\frac{\partial h(u,\boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)} + \tilde{\rho}(u)(n-1)\frac{\partial h(u,\boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)}\right) = 0, \tag{32}$$

where

$$h(u,\boldsymbol{x}_\bullet(t)) = \tilde{w}(u,\boldsymbol{x}_\bullet(t)) - \mu(u,\boldsymbol{x}_\bullet(t)) \tag{33}$$

is the stationary growth rate of an individual with trait value u. Equation (32) has exactly the same structural form as eq. (29), and so we don't need to repeat its interpretation since everything carries over *mutatis mutandis* from eq. (29). The difference, however, is that one can now study behavior response evolution analytically in the presence of survival-fecundity trade-off and thus embed such evolution in a more explicit life-hitory context. We can also conclude that the structural feature of the selection gradient derived by Akçay and Van Cleve (2012) is robust to changes in a number of demographic assumptions.

4 Discussion

This paper combined Hamilton's marginal rule for scalar quantitative traits and concepts from optimal control theory to derive the selection gradient on a trait affecting phenotypic scheduling of an individual interacting with its surroundings. The different first-order conditions (results 1-3) allow to study selection on behavior rules under incomplete plasticity for a reasonably large class of different scenarios for (i) evolution in homogeneous group (or familly) structured populations and (ii) life-history evolution under panmixia.

As an application and check of the results, the seminal first-order condition for behavior response rule evolution in group structured populations of Akçay and Van Cleve (2012, eq. 5) was re-derived and generalized to broader behavioral and life-history contexts. This illustrates that the control theory approach provides an overarching framework to formalize selection on quantitative traits affecting phenotypic scheduling that feed back on survival and reproduction. Control theory concepts straddle the layers of plasticity and complexity, since they can be applied to models of complete plasticity ranging from standard open and closed loop traits to models of incomplete plasticity as considered here; and for each case it allows to blend in Hamilton's rule seemlessly to consider evolution in spatially structured populations (Avila et al., 2021).

An obvious limitation of the present formalization is that it applies only to homogeneous group-structured populations and it would thus relevant to think about developing it for heterogeneous group-structure and possibly isolation-by-distance. This is in principle doable since control theory can be applied to age-structured populations that are also class-structured (Avila et al., 2019). Another limitation is that the model is deterministic and thus does not consider stochastic effects on state dynamics. But such effects may not only impact state dynamics direclty but also indirectly through environmental stochasticity, which will impact environmentally dependent state variables. In order to include stochasticity one could replace the system of ordinary differential equations (ODE's) for the state dynamics by stochastic differential equations (SDE's), and then attempt to apply results from stochastic control theory (e.g. Basar and Olsder, 1999; Kamien and Schwartz, 2012; Fleming and Soner, 2006) to generalize the first-order conditions to the case where fitness involves expectation of integrals over lifespan. This opens avenues for future research.

For ease of presentation of the concepts, the two central phenotypic attributes of the model, the state x and evolving trait u, were assumed to be one dimensional. Is anything changed if these variables take values in some finite muldimensional space; namely, if state $x = (x_1, x_2, ...)$ is a finite vector with x_i being the i-th state variable and trait $u = (u_1, u_2, ..)$ is a finite vector with u_i being the i-th quantitative trait? For this case, the selection gradient on trait u_i is still given by the inclusive fitness effect (eq. 2). Further, all other equations of the paper are actually maintained functionally, one just need to carefully read all one-dimensional mappings as multidimensional ones and any product of scalars by the relevant dot product between vectors if the scalar has a vectorial counterpart. But no new concept is needed and therefore results 1-3 direclty generalize to finite multidimensional traits, states, and actions.

An interesting and perhaps biologically surprising fact that results by going from the one-dimensional to the finite multidimensional trait and state case is that this is sufficient to make the model universal in the sense that any behavior rule can in principle be modeled with it. The reason is that a finite muldimenstional state dynamics is sufficient to describe the dynamics of a finite recurrent neural network, and where the neural network weights can be taken as the evolving quantitative traits. In turn, finite recurrent neural network can implement any type of behavior, i.e., finite recurrent networks are computationaly equivalent to Turing machines (Siegelman and Sontag, 1995; Haykin, 1999). In other words, evolving neural networks allow to consider in principle the evolution of any behavior rule, and here one could let the level of plasiticity itself be evolving. This also opens avenues for future research.

It has long been stressed that constructing more explict mechanistic models for the evolution of phenotypic expression would be usefull (e.g. McNamara and Houston, 1999; Lotem, 2012; Fawcett et al., 2012; Akçay, 2020). Perhaps one reason this has not been systematically undertaken is the lack of identification by evolutionary biologists of a coherent framework that allows to analyze adapative models accross a wide spectrum of mechanistic

formulations. Control theory blended with adaptive dynamics provides just that and so my hope is that the present formalization and more generally the many insights from control theory (e.g., Bryson and Ho, 1975; Arbib, 1987; Sontag, 1998; Haykin, 1999; Dockner et al., 2000; Liberzon, 2011; Athans and Falb, 2007; Sydsaeter et al., 2008; Astrom and Murray, 2008; Weber, 2011) can be used as sources of inspirations to attempt to close the loop between mechanistic and functional aspects of phenotypic expression evolution.

Acknowledgments

I thank M. Cant for a discussion that greatly stimulated the writing of this paper and P. Avila for comments on the manuscript.

Appendix A Cauchy formula

We here present the general solution to the non-homogeneous linear system of ordinary differential equations of the form $\mathrm{d}\boldsymbol{\lambda}(t)/\,\mathrm{d}t = -\mathbf{b}(t) - \mathbf{A}(t)\boldsymbol{\lambda}(t)$ with b.c. $\boldsymbol{\lambda}(\zeta)$ at $t = \zeta$. This covers the shadow value dynamics eq. (14c) of the main text since for that case $\mathbf{b}(\tau) = \mathbf{b}(\boldsymbol{x}(\tau), l(\tau), v(\tau))$ and $\mathbf{A}(t) = \mathbf{A}(\boldsymbol{x}(t))$. According to the Cauchy formula, the following solution then holds for all $t \in [0, T]$:

$$\boldsymbol{\lambda}(t) = \Psi(t,\zeta)\boldsymbol{\lambda}(\zeta) - \int_{\zeta}^{t} \Psi(t,\tau)\mathbf{b}(\tau)\,\mathrm{d}\tau, \tag{A-1}$$

where $\Psi(t,\tau)$ is the so-called fundamental matrix satisifying for all $t \in [\tau, T]$ the initial value problem

$$\frac{\mathrm{d}\Psi(t,\tau)}{\mathrm{d}t} = -\mathbf{A}(t)\Psi(t,\tau) \quad \text{with i.c.} \quad \Psi(\tau,\tau) = \mathbf{I}, \tag{A-2}$$

where $\mathbf{I}$ is the identity matrix (e.g., Athans and Falb, 2007, pp. 127–130, Weber, 2011, pp. 69–72, Aseev and Kryazhimskiy, 2008, eq. 15). The ith column of $\Psi(t,\tau)$ is the solution $\boldsymbol{\lambda}(t)$ of the system $\mathrm{d}\boldsymbol{\lambda}(t)/\,\mathrm{d}t = -\mathbf{A}(t)\boldsymbol{\lambda}(t)$ when the initial condition $\boldsymbol{\lambda}(\tau)$ is the standard basis with unit value at entry i, zero otherwise. For the one dimensional case, this can be solved and the solution is

$$\Psi(t,\tau) = \exp\left(-\int_{\tau}^{t} A(h)\,\mathrm{d}h\right). \tag{A-3}$$

For the multidimensional case, there is no analytic representation for $\Psi(t,\tau)$ when the system has variable coefficients and various method have been developped to evaluate $\Psi(t,\tau)$ (Weber, 2011, p. 71). For the constant coeffcient case for all t, where $\mathbf{A}(t) = \mathbf{A}$, we get the standard matrix exponential form solution

$$\Psi(t,\tau) = e^{-(t-\tau)\mathbf{A}}, \tag{A-4}$$

whereby eq. (A-1) reduces to

$$\boldsymbol{\lambda}(t) = e^{-(t-\zeta)\mathbf{A}}\boldsymbol{\lambda}(\zeta) - \int_{\zeta}^{t} e^{-(t-\tau)\mathbf{A}}\mathbf{b}(\tau)\,\mathrm{d}\tau. \tag{A-5}$$

The above results can also be used to solve for the reproductive value dynamics $\mathrm{d}v(t)/\,\mathrm{d}t = -\tilde{w}(\boldsymbol{u}, \boldsymbol{x}(t)) + \mu(\boldsymbol{u}, \boldsymbol{x}(t))\,v(t)$ under the finite horizon case (recall eq. 14d), which, in terms of the above notation, is obtained by setting $\lambda(t) = v(t)$, $\mathbf{b}(\tau) = \tilde{w}(\boldsymbol{u}, \boldsymbol{x}(\tau))$, and $\mathbf{A}(t) = -\mu(\boldsymbol{u}, \boldsymbol{x}(t))$. Then taking $\zeta = T$ as the final condition with $v(T) = \sigma(\boldsymbol{x}(T))$ and recalling that $\int_{\tau}^{t} f(h)\,\mathrm{d}h = -\int_{t}^{\tau} f(h)\,\mathrm{d}h$, we see that by substituting eq. (A-3) into eq. (A-1), we get

$$v(t) = \exp\left(-\int_{t}^{T} \mu(\boldsymbol{u}, \boldsymbol{x}(h))\,\mathrm{d}h\right)\sigma(\boldsymbol{x}(T)) + \int_{t}^{T} \exp\left(-\int_{t}^{\tau} \mu(\boldsymbol{u}, \boldsymbol{x}(h))\,\mathrm{d}h\right)\tilde{w}(\boldsymbol{u}, \boldsymbol{x}(\tau))\,\mathrm{d}\tau, \tag{A-6}$$

which is exactly the expression for the reproductive value given in Box 1, as it should.

Appendix B Representation of the selection gradient

Appendix B.1 Generic expression

We here present the generic representation of the selection gradient $s(u)$ when fitness takes the form of eq. (10) or eq. (12), and the state dynamics satisfy eqs. (6)-(7). The representation of $s(u)$ for these cases is implicitly contained in the calculations of Avila et al. (2021, Appendix B.2.1), who derived an expression for the selection gradient in the more general context where the evolving trait itself is dynamic and thus infinite dimensional. But since this more general context involves more complicated mathematics and notations, the represention of $s(u)$ is here (re)-derived for clarity and completeness.

In order to derive the representation $s(u)$, note first that the survival probability $l_\bullet(t)$ can itself be thought of as a state variable and eq. (10) of the main text can then be written as

$$\begin{aligned} w(u_\bullet, u_\circ, u) &= \int_0^T f(\boldsymbol{u}_\bullet, \mathbf{y}_\bullet(t))\,\mathrm{d}t + \Phi(\mathbf{y}_\bullet(T)) \\ &= \int_0^T f(\boldsymbol{u}_\bullet, \mathbf{y}_\bullet(t))\,\mathrm{d}t + \Phi(\mathbf{y}_\bullet(T)) + \underbrace{\int_0^T \left[\boldsymbol{\Lambda}(t)\cdot\left(\boldsymbol{g}(\boldsymbol{u}_\bullet, \mathbf{y}_\bullet(t)) - \frac{\mathrm{d}\mathbf{y}_\bullet(t)}{\mathrm{d}t}\right)\right]\mathrm{d}t}_{=0}, \end{aligned} \tag{A-7}$$

where $f(\boldsymbol{u}_\bullet(t), \mathbf{y}_\bullet(t)) = l_\bullet(t)\tilde{w}(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))$, $\Phi(\mathbf{y}_\bullet(T)) = l_\bullet(T)\sigma(\boldsymbol{x}_\bullet(T))$, the vector $\mathbf{y}_\bullet(t) = (\boldsymbol{x}_\bullet(t), l_\bullet(t)) = (x_\bullet(t), x_\circ(t), l_\bullet(t))$ collects all the state variables, the vector $\boldsymbol{\Lambda}(t) = (\lambda_\bullet(t), \lambda_\circ(t), v(t))$ collects all costates, and the vector $\mathrm{d}\mathbf{y}_\bullet(t)/\,\mathrm{d}t = \boldsymbol{g}(\boldsymbol{u}_\bullet, \mathbf{y}_\bullet(t)) = (g(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t)), g(\boldsymbol{u}_\circ, \boldsymbol{x}_\circ(t)), -\mu(\boldsymbol{u}_\bullet(t), \boldsymbol{x}_\bullet(t))l_\bullet(t)$ collects all changes in states. The symbol $\cdot$ denotes a dot product between two vectors and integrating the last term in eq. (A-7) by parts yields

$$\begin{aligned} &-\int_0^T \boldsymbol{\Lambda}(t)\cdot\frac{\mathrm{d}\mathbf{y}_\bullet(t)}{\mathrm{d}t}\,\mathrm{d}t \\ &= \int_0^T \frac{\mathrm{d}\boldsymbol{\Lambda}(t)}{\mathrm{d}t}\cdot\mathbf{y}_\bullet(t)\,\mathrm{d}t - \boldsymbol{\Lambda}(T)\cdot\mathbf{y}_\bullet(T) + \boldsymbol{\lambda}(0)\cdot\mathbf{y}_\bullet(0), \end{aligned} \tag{A-8}$$

and hence eq. (A-7) becomes

$$\begin{aligned} w(u_\bullet, u_\circ, u) = &\int_0^T \left(f(\boldsymbol{u}_\bullet(t), \mathbf{y}_\bullet(t)) + \boldsymbol{\Lambda}(t)\cdot\boldsymbol{g}(\boldsymbol{u}_\bullet, \mathbf{y}_\bullet(t)) + \frac{\mathrm{d}\boldsymbol{\Lambda}(t)}{\mathrm{d}t}\cdot\mathbf{y}_\bullet(t)\right)\mathrm{d}t \\ &- \boldsymbol{\Lambda}(T)\cdot\mathbf{y}_\bullet(T) + \boldsymbol{\Lambda}(0)\cdot\mathbf{y}_\bullet(0) + \Phi(\mathbf{y}_\bullet(T)). \end{aligned} \tag{A-9}$$

Using

$$H\Big(\boldsymbol{u}_\bullet, \underbrace{\boldsymbol{x}_\bullet(t), l_\bullet(t)}_{\mathbf{y}_\bullet(t)}, \underbrace{\boldsymbol{\lambda}(t), v(t)}_{\boldsymbol{\Lambda}(t)}\Big) = f(\boldsymbol{u}_\bullet(t), \mathbf{y}_\bullet(t)) + \boldsymbol{g}(\boldsymbol{u}_\bullet, \mathbf{y}_\bullet(t))\cdot\boldsymbol{\Lambda}(t) \tag{A-10}$$

and

$$\Phi_{\mathrm{aug}}(T, \mathbf{y}_\bullet(T)) = \Phi(\mathbf{y}_\bullet(T)) - \boldsymbol{\Lambda}(T)\cdot\mathbf{y}_\bullet(T), \tag{A-11}$$

allows to write eq. (A-9) as

$$w(u_\bullet, u_\circ, u) = \int_0^T \left[H(\boldsymbol{u}_\bullet, \mathbf{y}_\bullet(t), \boldsymbol{\Lambda}(t)) + \frac{\mathrm{d}\boldsymbol{\Lambda}(t)}{\mathrm{d}t}\cdot\mathbf{y}_\bullet(t)\right]\mathrm{d}t + \Phi_{\mathrm{aug}}(T, \mathbf{y}_\bullet(T)) + \boldsymbol{\Lambda}(0)\cdot\mathbf{y}_\bullet(0). \tag{A-12}$$

Taking the derivative of fitness (A-12) with respect to the focal's trait $u_\bullet$ yields in force of the chain rule and Leibniz's formula (Sydsaeter et al., 2008, p. 160) that

$$\frac{\partial w(u_\bullet, u_\circ, u)}{\partial u_\bullet} = \int_0^T \left(\frac{\partial H(\boldsymbol{u}_\bullet, \mathbf{y}_\bullet(t), \boldsymbol{\Lambda}(t))}{\partial u_\bullet} + \left(\nabla H(\boldsymbol{u}, \mathbf{y}_\bullet(t), \boldsymbol{\Lambda}(t)) + \frac{\mathrm{d}\boldsymbol{\Lambda}(t)}{\mathrm{d}t}\right)\cdot\frac{\partial\mathbf{y}_\bullet(t)}{\partial u_\bullet}\right)\mathrm{d}t + \nabla\Phi_{\mathrm{aug}}(T, \mathbf{y}_\bullet(T))\cdot\frac{\partial\mathbf{y}_\bullet(T)}{\partial u_\bullet}, \tag{A-13}$$

where the gradient operator ∇ is with respect to variable $\mathbf{y}_\bullet(t)$, e.g., $\nabla H = (\partial H/\partial x_\bullet(t), \partial H/\partial x_\circ(t), \partial H/\partial l_\bullet(t))$, $\partial \mathbf{y}_\bullet(T)/\partial u_\bullet = (\partial x_\bullet(T)/\partial u_\bullet, \partial x_\circ(T)/\partial u_\bullet, \partial l_\bullet(T)/\partial u_\bullet)$, and all derivatives, here and throughout, are evaluated at $\boldsymbol{u}_\bullet = \boldsymbol{u} = (u, u, u)$ and $\mathbf{y}_\bullet(t) = (\boldsymbol{x}(t), l(t)) = ((x(t), x(t), x(t)), l(t))$. Likewise, we have

$$\frac{\partial w(u_\bullet, u_\circ u)}{\partial u_\circ} = \int_0^T \left(\frac{\partial H(\boldsymbol{u}_\bullet, \mathbf{y}_\bullet(t), \boldsymbol{\Lambda}(t))}{\partial u_\circ} + \left(\nabla H(\boldsymbol{u}, \mathbf{y}_\bullet(t), \boldsymbol{\Lambda}(t)) + \frac{\mathrm{d}\boldsymbol{\Lambda}(t)}{\mathrm{d}t} \right) \cdot \frac{\partial \mathbf{y}_\bullet(t)}{\partial u_\circ} \right) \mathrm{d}t + \nabla \Phi_{\mathrm{aug}}(T, \mathbf{y}_\bullet(T)) \cdot \frac{\partial \mathbf{y}_\bullet(T)}{\partial u_\circ}, \tag{A-14}$$

where $\partial \mathbf{y}_\bullet(T)/\partial u_\circ = (\partial x_\bullet(T)/\partial u_\circ, \partial x_\circ(T)/\partial u_\circ, \partial l_\bullet(T)/\partial u_\circ)$.

Note that the derivatives of the term $\boldsymbol{\Lambda}(0) \cdot \mathbf{y}_\bullet(0)$ in eqs. (A-13)–(A-14) have vanished because the initial state variables are fixed $x_\bullet(0) = x_\circ(0) = x(0) = x_{\mathrm{init}}$. Note also that

$$\nabla H(\boldsymbol{u}, \mathbf{y}_\bullet(t), \boldsymbol{\Lambda}(t)) = \nabla f(\boldsymbol{u}, \mathbf{y}_\bullet(t))) + \nabla \boldsymbol{g}(\boldsymbol{u}, \mathbf{y}_\bullet(t)) \boldsymbol{\Lambda}(t) \tag{A-15}$$

where

$$\nabla \boldsymbol{g}(\boldsymbol{u}(t), \boldsymbol{x}(t)) = \begin{pmatrix} \frac{\partial g(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)} & \frac{\partial g(\boldsymbol{u}_\circ, \boldsymbol{x}_\circ(t))}{\partial x_\bullet(t)} & -\frac{\partial \mu(\boldsymbol{u}_\bullet(t), \boldsymbol{x}_\bullet(t))}{\partial x_\bullet(t)} l_\bullet(t) \\ \frac{\partial g(\boldsymbol{u}_\bullet, \boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)} & \frac{\partial g(\boldsymbol{u}_\circ, \boldsymbol{x}_\circ(t))}{\partial x_\circ(t)} & -\frac{\partial \mu(\boldsymbol{u}_\bullet(t), \boldsymbol{x}_\bullet(t))}{\partial x_\circ(t)} l_\bullet(t) \\ 0 & 0 & -\mu(\boldsymbol{u}_\bullet(t), \boldsymbol{x}_\bullet(t)) \end{pmatrix}_{\substack{\boldsymbol{u}_\bullet = \boldsymbol{u} \\ l_\bullet(t) = l(t) \\ \boldsymbol{x}_\bullet(t) = \boldsymbol{x}(t)}} \tag{A-16}$$

is a matrix, the gradient of vector $\boldsymbol{g}$ (hence $\nabla \boldsymbol{g}(\boldsymbol{u}, \mathbf{y}_\bullet(t)) \boldsymbol{\Lambda}(t)$ is the multiplication of a matrix by a vector). Finally,

$$\nabla \Phi_{\mathrm{aug}}(T, \mathbf{y}_\bullet(T)) = \nabla \Phi(\mathbf{y}_\bullet(T)) - \boldsymbol{\Lambda}(T) \tag{A-17}$$

where

$$\nabla \Phi(\mathbf{y}_\bullet(T)) = (l(T) \partial \sigma(\boldsymbol{x}_\bullet(T))/\partial x_\bullet(T), l(T) \partial \sigma(\boldsymbol{x}_\bullet(T)/\partial x_\circ(T), \sigma(\boldsymbol{x}(T)))|_{\boldsymbol{x}_\bullet(t) = \boldsymbol{x}(t)}. \tag{A-18}$$

Now if we set

$$-\frac{\mathrm{d}\boldsymbol{\Lambda}(t)}{\mathrm{d}t} = \nabla H(t, \boldsymbol{u}, \mathbf{y}_\bullet(t), \boldsymbol{\Lambda}(t)), \tag{A-19}$$

then the second term in the integral of eqs. (A-13)–(A-14) vanishes. Thus, on substituting eqs. (A-13)–(A-14) into the inclusive fitness effect (2) of the main text

$$s(u) = \left. \frac{\partial w(u_\bullet, u_\circ, u)}{\partial u_\bullet} \right|_{u_\bullet = u_\circ = u} + r(u) \left. \frac{\partial w(u_\bullet, u_\circ u)}{\partial u_\circ} \right|_{u_\bullet = u_\circ = u}, \tag{A-20}$$

and simplifying shows that the inclusive fitness effect with discounted fitness (10) can be written as

$$s(u) = \int_0^T s(t, u) \, \mathrm{d}t + F(T), \tag{A-21}$$

with

$$s(t, u) = \left[\frac{\partial H\Big(\boldsymbol{u}_\bullet, \boldsymbol{x}(t), l(t), \boldsymbol{\lambda}(t), v(t)\Big)}{\partial u_\bullet} + r(u) \frac{\partial H\Big(\boldsymbol{u}_\bullet, \boldsymbol{x}(t), l(t), \boldsymbol{\lambda}(t), v(t)\Big)}{\partial u_\circ} \right]_{\boldsymbol{u}_\bullet = \mathbf{u}} \tag{A-22}$$

and

$$F(T) = \nabla \Phi_{\mathrm{aug}}(T, \boldsymbol{x}(T)) \cdot \left[\frac{\partial \mathbf{y}_\bullet(T)}{\partial u_\bullet} + r(u) \frac{\partial \mathbf{y}_\bullet(T)}{\partial u_\circ} \right], \tag{A-23}$$

and since all state variables therein are evaluated at $\boldsymbol{u}_\bullet = \mathbf{u}$, they satisfy the dynamics

$$\begin{aligned} \frac{\mathrm{d}x(t)}{\mathrm{d}t} &= g(\boldsymbol{u}, \boldsymbol{x}(t)) \\ \frac{\mathrm{d}l(t)}{\mathrm{d}t} &= -\mu(\boldsymbol{u}, \boldsymbol{x}(t)) l(t). \end{aligned} \tag{A-24}$$

Exactly the same line of arguments can be applied to the integral appearing in the average fitness (12), with the difference that there is no state variable associated to survival and thus no costate. Carrying out the parallel calculations, then shows that the selection gradient (A-20) under average fitness can be represented as

$$s(u) = \lim_{T\to\infty} \frac{1}{T}\Big(\int_0^T \left[\frac{\partial H\Big(\boldsymbol{u}_\bullet, \boldsymbol{x}(t), 1, \boldsymbol{\lambda}(t), 0\Big)}{\partial u_\bullet} + r(u)\frac{\partial H\Big(\boldsymbol{u}_\bullet, \boldsymbol{x}(t), 1, \boldsymbol{\lambda}(t), 0\Big)}{\partial u_\circ}\right] \mathrm{d}t + F_{\mathrm{a}}(T)\Big) \tag{A-25}$$

with

$$F_{\mathrm{a}}(T) = -\boldsymbol{\Lambda}(T)\cdot\left[\frac{\partial \mathbf{y}_\bullet(T)}{\partial u_\bullet} + r(u)\frac{\partial \mathbf{y}_\bullet(T)}{\partial u_\circ}\right]. \tag{A-26}$$

Appendix B.2 The three results

Appendix B.2.1 Result 1

We here prove result 1. Let us first consider the finite horizon case ($T < \infty$). Then, the boundary conditions for the costates are $v(t) = \sigma(\boldsymbol{x}(T))$ and $\boldsymbol{\lambda}(T) = \nabla\sigma(\boldsymbol{x}_\bullet(T))$ because this satisfies the definition of the co-state as marginal changes of fitness at T (Box 1). These boundary conditions are also standard in control theory (Bryson and Ho, 1975; Dockner et al., 2000; Athans and Falb, 2007; Sydsaeter et al., 2008; Aseev and Kryazhimskiy, 2008; Weber, 2011) and imply that $\nabla\Phi_{\mathrm{aug}}(T, \boldsymbol{x}(T)) = (0, 0, 0)$. Hence, it follow from eq. (A-23) that $F(T) = 0$, whereby the first line of eq. (A-21) reduces to eq. (17) and since together eq. (A-19) and eq. (A-24) are equivalent to eq. (14), this proves the finite horizon case of result 1.

Let us now consider the infinite horizon case ($T \to \infty$). Owing to the assumptions of dominating discounts, i.e., $\lim_{T\to\infty} \boldsymbol{\lambda}(T) = (0, 0)$ and that $\lim_{t\to\infty} l_\bullet(t) = 0$, whereby $\lim_{t\to\infty} \partial l_\bullet(\infty)/\partial u_\bullet = \lim_{t\to\infty} \partial l_\bullet(\infty)/\partial u_\circ = 0$, we have

$$\begin{aligned}
\lim_{T\to\infty} \nabla\Phi(\mathbf{y}_\bullet(T)) &= (0, 0, \lim_{T\to\infty} \sigma(\boldsymbol{x}(T))) \\
\lim_{T\to\infty} \boldsymbol{\Lambda}(T) &= (0, 0, \lim_{T\to\infty} v(T)) \\
\lim_{T\to\infty} \partial\mathbf{y}_\bullet(T)/\partial u_\circ &= (\lim_{T\to\infty} \partial x_\bullet(T)/\partial u_\circ, \lim_{T\to\infty} \partial x_\circ(T)/\partial u_\circ, 0) \\
\lim_{T\to\infty} \partial\mathbf{y}_\bullet(T)/\partial u_\bullet &= (\lim_{T\to\infty} \partial x_\bullet(T)/\partial u_\bullet, \lim_{T\to\infty} \partial x_\circ(T)/\partial u_\bullet, 0).
\end{aligned} \tag{A-27}$$

Substituting into eq. (A-23) shows that $\lim_{T\to\infty} F(T) = 0$ (provided all entries remain bounded). Thereby eq. (A-21) reduces to eq. (17), which proves the infinite horizon part of result 1.

Appendix B.2.2 Result 2

We here prove result 2, which relies on the assumption that the initial condition of the state variable satisfies the resident equilibrium, i.e., $x_0 = \hat{x}$ satisfying $g(\boldsymbol{u}, (\hat{x}, \hat{x}, \hat{x})) = 0$. With this and owing to the fact that the selection coefficient $s(t, u)$ in eq. (A-22) is evaluated at the residend state $\boldsymbol{x}(t)$ and thus at $\boldsymbol{x}(t) = \hat{\boldsymbol{x}} = (\hat{x}, \hat{x}, \hat{x})$ for all $t \in [0, \infty)$, allows us to write eq. (A-22) as

$$\begin{aligned}
s(t, u) = \Bigg[& l(t)\left(\frac{\partial\tilde{w}(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\bullet} - \frac{\partial\mu(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\bullet} v(t)\right) + \frac{\partial g(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\bullet}\lambda_\bullet(t) + \frac{\partial g(\boldsymbol{u}_\circ, \hat{\boldsymbol{x}})}{\partial u_\bullet}\lambda_\circ(t) \\
& + r(u)\left(l(t)\left(\frac{\partial\tilde{w}(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\circ} - \frac{\partial\mu(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\circ} v(t)\right) + \frac{\partial g(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\circ}\lambda_\bullet(t) + \frac{\partial g(\boldsymbol{u}_\circ, \hat{\boldsymbol{x}})}{\partial u_\circ}\lambda_\circ(t)\right)\Bigg]_{\boldsymbol{u}_\bullet = \mathbf{u}}.
\end{aligned} \tag{A-28}$$

All vital rates and transition functions are here independent of time, since they are all evaluated at $\hat{x}$. Hence, substituting eq. (A-28) into eq. (A-21), the vital rates and transition functions can be factored out from the integral

and since an infinite time horizon with dominating discounts is assumed, we have $F(T) = 0$. Therefore

$$s(u) = \bar{l}\left[\frac{\partial \tilde{w}(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\bullet} - \frac{\partial \mu(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\bullet}\bar{v} + \frac{\partial g(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\bullet}\bar{\lambda}_\bullet + \frac{\partial g(\boldsymbol{u}_\circ, \hat{\boldsymbol{x}})}{\partial u_\bullet}\bar{\lambda}_\circ\right]_{\boldsymbol{u}_\bullet = \mathbf{u}}$$
$$+ r(u)\bar{l}\left[\frac{\partial \tilde{w}(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\circ} - \frac{\partial \mu(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\circ}\bar{v} + \frac{\partial g(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}})}{\partial u_\circ}\bar{\lambda}_\bullet + \frac{\partial g(\boldsymbol{u}_\circ, \hat{\boldsymbol{x}})}{\partial u_\circ}\bar{\lambda}_\circ\right]_{\boldsymbol{u}_\bullet = \mathbf{u}}, \quad \text{(A-29)}$$

where $\bar{l} = \int_0^\infty l(t)\,\mathrm{d}t$, $\bar{v} = \int_0^\infty v(t)l(t)\,\mathrm{d}t/\bar{l}$, $\bar{\boldsymbol{\lambda}} = (\bar{\lambda}_\bullet, \bar{\lambda}_\circ) = (\int_0^\infty \lambda_\bullet(t)\,\mathrm{d}t/\bar{l}, \int_0^\infty \lambda_\circ(t)\,\mathrm{d}t/\bar{l})$. Using the definition of the Hamiltonian function, this can be equivalently written as

$$s(u) = \bar{l}\left[\frac{\partial H\left(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}}, 1, \bar{\boldsymbol{\lambda}}, \bar{v}\right)}{\partial u_\bullet} + r(u)\frac{\partial H\left(\boldsymbol{u}_\bullet, \hat{\boldsymbol{x}}, 1, \bar{\boldsymbol{\lambda}}, \bar{v}\right)}{\partial u_\circ}\right]_{\boldsymbol{u}_\bullet = \mathbf{u}}, \quad \text{(A-30)}$$

which gives eq. (20) for $\bar{v} = 1$.

To complete the proof of result 2, it thus remains to prove that $\bar{l} = 1/\mu(\boldsymbol{u}, \hat{\boldsymbol{x}})$, $\bar{v} = 1$, and $\bar{\boldsymbol{\lambda}} = (\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1}\,\mathbf{b}(\hat{\boldsymbol{x}}, 1, 1)$. For this, let us first note that under the assumptions underlying result 2, which involves dominating discounts, eq. (14) reduces to

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = g(\boldsymbol{u}, \hat{\boldsymbol{x}}) = 0 \quad \text{(A-31a)}$$
$$\frac{\mathrm{d}l(t)}{\mathrm{d}t} = -\mu(\boldsymbol{u}, \hat{\boldsymbol{x}})l(t) \qquad \text{with i.c. } l(0) = 1 \quad \text{(A-31b)}$$
$$-\frac{\mathrm{d}\boldsymbol{\lambda}(t)}{\mathrm{d}t} = \mathbf{b}(\hat{\boldsymbol{x}}, l(t), v(t)) + \mathbf{A}(\hat{\boldsymbol{x}})\boldsymbol{\lambda}(t) \qquad \text{with b.c. } \lim_{t\to\infty}\boldsymbol{\lambda}(t) = (0, 0) \quad \text{(A-31c)}$$
$$-\frac{\mathrm{d}v(t)}{\mathrm{d}t} = \tilde{w}(\boldsymbol{u}, \hat{\boldsymbol{x}}) - \mu(\boldsymbol{u}, \hat{\boldsymbol{x}})\,v(t) \qquad \text{with i.c. } v(0) = 1, \quad \text{(A-31d)}$$

where for the boundary condition of the reproductive value we used the demographic consistency relation of Box 1. Straightfoward integration of eq. (A-31c), yields $\bar{l} = 1/\mu(\boldsymbol{u}, \hat{\boldsymbol{x}})$. And since survival and effective fecundity is constant, we have $v(0) = \tilde{w}(\boldsymbol{u}, \hat{\boldsymbol{x}})/\mu(\boldsymbol{u}, \hat{\boldsymbol{x}}) = 1$ (Box 1), and so $v(t) = 1$ for all t satisfies eq. (A-31d), and thus $\bar{v} = 1$. In order to prove the third equality, let us use the vectors

$$\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) = \Big(\partial\tilde{w}(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))/\partial x_\bullet(t), \tilde{w}(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))/\partial x_\circ(t), \partial\tilde{w}(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))/\partial x(t)\Big)_{\boldsymbol{x}_\bullet(t) = \hat{\boldsymbol{x}}} \quad \text{(A-32)}$$
$$\mathbf{b}_{\mu}(\hat{\boldsymbol{x}}) = \Big(\partial\mu(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))/\partial x_\bullet(t), \mu(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))/\partial x_\circ(t), \partial\mu(\boldsymbol{u}, \boldsymbol{x}_\bullet(t))/\partial x(t)\Big)_{\boldsymbol{x}_\bullet(t) = \hat{\boldsymbol{x}}},$$

whereby eq. (15) can be written

$$\mathbf{b}(\hat{\boldsymbol{x}}, l(t), v(t)) = \mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}})l(t) - \mathbf{b}_{\mu}(\hat{\boldsymbol{x}})l(t)v(t). \quad \text{(A-33)}$$

Substituting this into eq. (A-31) and intergrating over time yields

$$\int_0^\infty \frac{\mathrm{d}\boldsymbol{\lambda}(t)}{\mathrm{d}t}\,\mathrm{d}t = -\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}})\int_0^\infty l(t)\,\mathrm{d}t + \mathbf{b}_{\mu}(\hat{\boldsymbol{x}})\int_0^\infty l(t)v(t)\,\mathrm{d}t - \mathbf{A}(\hat{\boldsymbol{x}})\int_0^\infty \boldsymbol{\lambda}(t)\,\mathrm{d}t. \quad \text{(A-34)}$$

From the fundamental theorem of calculus $\int_0^\infty \frac{\mathrm{d}\boldsymbol{\lambda}(t)}{\mathrm{d}t}\,\mathrm{d}t = \boldsymbol{\lambda}(\infty) - \boldsymbol{\lambda}(0)$ and so we can write

$$\boldsymbol{\lambda}(\infty) - \boldsymbol{\lambda}(0) = \bar{l}\left(-\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) + \mathbf{b}_{\mu}(\hat{\boldsymbol{x}})\bar{v} - \mathbf{A}(\hat{\boldsymbol{x}})\bar{\boldsymbol{\lambda}}\right) = \bar{l}\left(-\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) + \mathbf{b}_{\mu}(\hat{\boldsymbol{x}}) - \mathbf{A}(\hat{\boldsymbol{x}})\bar{\boldsymbol{\lambda}}\right). \quad \text{(A-35)}$$

Since $\boldsymbol{\lambda}(\infty) = 0$ by assumption of dominating discounts, we need an expression for $\boldsymbol{\lambda}(0)$ to obtain $\bar{\boldsymbol{\lambda}}$. To than end,

insert eq. (A-33) into eq. (A-5), which, for $t = 0$ and $\zeta = T$ with $T = \infty$ and $\boldsymbol{\lambda}(\infty) = 0$, yields

$$\begin{aligned}
\boldsymbol{\lambda}(0) &= -\int_{\infty}^{0} e^{\mathbf{A}(\hat{\boldsymbol{x}})\tau}\mathbf{b}(\hat{\boldsymbol{x}}, l(\tau), v(\tau))\,\mathrm{d}\tau \\
&= \int_{0}^{\infty} e^{\mathbf{A}(\hat{\boldsymbol{x}})\tau}\mathbf{b}(\hat{\boldsymbol{x}}, l(\tau), v(\tau))\,\mathrm{d}\tau \\
&= \int_{0}^{\infty} e^{\mathbf{A}(\hat{\boldsymbol{x}})\tau} l(\tau)\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}})\,\mathrm{d}\tau - \int_{0}^{\infty} e^{\mathbf{A}(\hat{\boldsymbol{x}})\tau} l(\tau)v(\tau)\mathbf{b}_{\mu}(\hat{\boldsymbol{x}})\,\mathrm{d}\tau \\
&= \int_{0}^{\infty} e^{\mathbf{A}(\hat{\boldsymbol{x}})\tau} e^{-\mu\tau}\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}})\,\mathrm{d}\tau - \int_{0}^{\infty} e^{\mathbf{A}(\hat{\boldsymbol{x}})\tau} e^{-\mu\tau}\mathbf{b}_{\mu}(\hat{\boldsymbol{x}})\,\mathrm{d}\tau \\
&= \int_{0}^{\infty} e^{\mathbf{A}(\hat{\boldsymbol{x}})\tau} e^{-\mu\tau}\,\mathrm{d}\tau\,(\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) - \mathbf{b}_{\mu}(\hat{\boldsymbol{x}})) \\
&= (\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1}(\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) - \mathbf{b}_{\mu}(\hat{\boldsymbol{x}}))\,,
\end{aligned} \tag{A-36}$$

where the second inequality follows from interchanging the boundaries, the third from inserting eq. (A-33), the fourth from inserting for $v(t) = 1$ and $l(\tau) = e^{-\mu\tau}$ (and using the shorthand notation $\mu = \mu(\boldsymbol{u}, \hat{\boldsymbol{x}})$), the fifth from reanranging, and the sixth from noting that the fifth line defines a Laplace transform and so $\int_0^\infty e^{\mathbf{A}(\hat{\boldsymbol{x}})\tau}\mathbf{b}e^{-\mu\tau}\,\mathrm{d}\tau = (\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1}\mathbf{b}$ for a constant vector $\mathbf{b}$ (Athans and Falb, 2007, p. 140). Inserting this into eq. (A-35) and solving for $\bar{\boldsymbol{\lambda}}$ finally yields

$$\begin{aligned}
\bar{\boldsymbol{\lambda}} &= (1/\bar{l})\mathbf{A}(\hat{\boldsymbol{x}})^{-1}\left[\boldsymbol{\lambda}(0) - \bar{l}\,(\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) - \mathbf{b}_{\mu}(\hat{\boldsymbol{x}}))\right] \\
&= (1/\bar{l})\mathbf{A}(\hat{\boldsymbol{x}})^{-1}\left[(\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1}(\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) - \mathbf{b}_{\mu}(\hat{\boldsymbol{x}})) - \bar{l}\,(\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) - \mathbf{b}_{\mu}(\hat{\boldsymbol{x}}))\right] \\
&= (1/\bar{l})\mathbf{A}(\hat{\boldsymbol{x}})^{-1}\left[(\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1} - \bar{l}\right](\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) - \mathbf{b}_{\mu}(\hat{\boldsymbol{x}})) \\
&= \mathbf{A}(\hat{\boldsymbol{x}})^{-1}\left[(1/\bar{l})\,(\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1} - 1\right](\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) - \mathbf{b}_{\mu}(\hat{\boldsymbol{x}})) \\
&= \left[(1/\bar{l})\mathbf{A}(\hat{\boldsymbol{x}})^{-1}(\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1} - \mathbf{A}(\hat{\boldsymbol{x}})^{-1}\right](\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) - \mathbf{b}_{\mu}(\hat{\boldsymbol{x}}))\,.
\end{aligned} \tag{A-37}$$

In order to further simplify this expression, note that $\mathbf{I} = (\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))(\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1} = \mu\,(\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1} - \mathbf{A}(\hat{\boldsymbol{x}})\,(\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1}$. Hence, we have $(\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1} = (1/\mu)\left(\mathbf{I} + \mathbf{A}(\hat{\boldsymbol{x}})\,(\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1}\right)$ and on substituting into the last line of eq. (A-37) and using $(1/\bar{l}) = \mu$ produces

$$\begin{aligned}
\bar{\boldsymbol{\lambda}} &= \left[\mathbf{A}(\hat{\boldsymbol{x}})^{-1}\left(\mathbf{I} + \mathbf{A}(\hat{\boldsymbol{x}})\,(\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1}\right) - \mathbf{A}(\hat{\boldsymbol{x}})^{-1}\right](\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) - \mathbf{b}_{\mu}(\hat{\boldsymbol{x}})) \\
&= \left[\mathbf{A}(\hat{\boldsymbol{x}})^{-1} + (\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1} - \mathbf{A}(\hat{\boldsymbol{x}})^{-1}\right](\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) - \mathbf{b}_{\mu}(\hat{\boldsymbol{x}})) \\
&= (\mu\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1}(\mathbf{b}_{\mathrm{f}}(\hat{\boldsymbol{x}}) - \mathbf{b}_{\mu}(\hat{\boldsymbol{x}}))\,.
\end{aligned} \tag{A-38}$$

Using eq. (A-33) and reminding that $\mu = \mu(\boldsymbol{u}, \hat{\boldsymbol{x}})$, we finallly obtain $\bar{\boldsymbol{\lambda}} = (\mu(\boldsymbol{u}, \hat{\boldsymbol{x}})\mathbf{I} - \mathbf{A}(\hat{\boldsymbol{x}}))^{-1}\mathbf{b}(\hat{\boldsymbol{x}}, 1, 1)$, as needed to be proved.

Appendix B.2.3 Result 3

We here prove result 3, which assumes that given any initial conditions $x(0) = x_0$ and $\boldsymbol{\lambda}(0) = \boldsymbol{\lambda}_0$ at $t = 0$ for the states and costates, the variables converge to a unique equilibrium, $\hat{\boldsymbol{x}} = \lim_{T\to\infty}\boldsymbol{x}(t)$, and $\hat{\boldsymbol{\lambda}} = \lim_{T\to\infty}\boldsymbol{\lambda}(t)$, which is hyperbollically stable (*sensu* Hirsch et al., 2004). Then, owing to the fact that $F_{\mathrm{a}}(T)$ is finite in eq. (A-25) and invoking the the ergodic theorem (Karlin and Taylor, 1975), which informs us that the time average is the spatial average, eq. (A-25) can be reduced to

$$s(u) = \left[\frac{\partial H\left(\boldsymbol{u}_{\bullet}, \hat{\boldsymbol{x}}, 1, \hat{\boldsymbol{\lambda}}, 0\right)}{\partial u_{\bullet}} + r(u)\frac{\partial H\left(\boldsymbol{u}_{\bullet}, \hat{\boldsymbol{x}}, 1, \hat{\boldsymbol{\lambda}}, 0\right)}{\partial u_{\circ}}\right]_{\boldsymbol{u}_{\bullet}=\mathbf{u}}, \tag{A-39}$$

which proves the required result, since from eq. (14) the equilibrium $\hat{\boldsymbol{x}} = \lim_{T\to\infty}\boldsymbol{x}(t)$ satisfies $g(\boldsymbol{u}, \hat{\boldsymbol{x}}) = 0$ and the equilibrium $\hat{\boldsymbol{\lambda}} = \lim_{T\to\infty}\boldsymbol{\lambda}(t)$ satisfies $\hat{\boldsymbol{\lambda}} = -\mathbf{A}(\hat{\boldsymbol{x}})^{-1}\mathbf{b}(\hat{\boldsymbol{x}}, 1, 0)$.

References

Akçay, E. 2020. Deconstructing evolutionary game theory: coevolution of social behaviors with their evolutionary setting. *The American Naturalist* 195:315–330.

Akçay, E. and J. Van Cleve. 2009. A theory for the evolution of other-regarding motivations integrating proximate and ultimate perspectives. *Proceedings of the National Academy of Sciences of the United States of America* 106:19061–19066.

Akçay, E. and J. Van Cleve. 2014. Pathways to social evolution: reciprocity, relatedness, and synergy. *Evolution* 68:2245–58.

Akçay, E. and J. Van Cleve. 2012. Behavioral responses in structured populations pave the way to group optimality. *American Naturalist* 179:257–269.

Alger, I., J. W. Weibull, and L. Lehmann. 2020. Evolution of preferences in structured populations: Genes, guns, and culture. *Journal of Economic Theory* 185:1–45.

Alon, U. 2020. An Introduction to System Biology. Taylor and Francis, London, 2nd edn.

André, J. B. and T. Day. 2007. Perfect reciprocity is the only evolutionarily stable strategy in the continuous iterated prisoner's dilemma. *Journal of Theoretical Biology* 247:11–22.

Aoki, K., J. Y. Wakano, and L. Lehmann. 2012. Evolutionarily stable learning schedules in discrete generation models. *Theoretical Population Biology* 81:300–309.

Arbib, M. A. 1987. Brains, Machines, and Mathematics. Springer Verlag, Berlin.

Aseev, S. M. and A. Kryazhimskiy. 2008. Shadow prices in infinite-horizon optimal control problems with dominating discounts. *Applied Mathematics and Computation* 204:519–531.

Astrom, K. J. and M. R. Murray. 2008. Feedback Systems: An Introduction for Scientists and Engineers. Princetion University Press, Princeton, NJ.

Athans, M. and P. L. Falb. 2007. Optimal Control: an Introduction to the Theory and its Applications. Dover Publications, New York (Original work published 1966).

Avila, P., L. Fromhage, and L. Lehmann. 2019. Sex-allocation conflict and sexual selection throughout the lifespan of eusocial colonies. *Evolultion* 73:1116–1132.

Avila, P., T. Priklopil, and L. Lehmann. 2021. Hamilton's rule, gradual evolution, and the optimal (feedback) control of reaction norms and other function-valued traits. *Journal of Theoretical Biology* 526:110602.

Basar, T. and G. J. Olsder. 1999. Dynamic Noncooperative Game Theory. Siam, New York, 2nd edn.

Bryson, A. E. and Y.-C. Ho. 1975. Applied Optimal Control: Optimization, Estimation and Control. CRC Press, New York.

Day, T. and P. Taylor. 1998. The evolution of temporal patterns of selfishness, altruism, and group cohesion. *American Naturalist* 152:102–113.

Day, T. and P. D. Taylor. 1997. Hamilton's rule meets the Hamiltonian: Kin selection on dynamic characters. *Proceedings of the Royal Society of London Series B-Biological Sciences* 264:639–644.

Day, T. and P. D. Taylor. 2000. A generalization of Pontryagin's maximum principle for dynamic evolutionary games among relatives. *Theoretical Population Biology* 57:339–356.

Dockner, E., S. Jorgensen, N. V. Long, and G. Sorger. 2000. Differential Games in Economics and Management Science. Cambridge University Press, Cambridge.

Dorfman, R. 1969. An economic interpretation of optimal control theory. *American Economic Review* 59:817–831.

Dridi, S. and E. Akçay. 2018. Learning to cooperate: The evolution of social rewards in repeated interactions. *American Naturalist* 191:58–73.

Dridi, S. and L. Lehmann. 2014. On learning dynamics underlying the evolution of learning rules. *Theoretical Population Biology* 91:20–36.

Dridi, S. and L. Lehmann. 2015. A model for the evolution of reinforcement learning in fluctuating games. *Animal Behaviour* 104:1–28.

Enquist, M. and S. Ghirlanda. 2005. Neural Networks and Animal Behavior. Princeton University Press, Princeton.

Ewald, C.-O., J. McNamara, and A. Houston. 2007. Parental care as a differential game: A dynamic extension of the Houston–Davies game. *Applied Mathematics and Computation* 190:1450–1465.

Ezoe, I. and Y. Iwasa. 1997. Evolution of condition-dependent dispersal: a genetic-algorithm search for the ESS reaction norm. *Researches on Population Ecology* 39:127–137.

Fawcett, T., S. Hamblin, and L. A. Giraldeau. 2012. Exposing the behavioral gambit: the evolution of learning and decision rules. *Behavioral Ecology* 24:2–11.

Fisher, R. A. 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

Fleming, H. W. and H. M. Soner. 2006. Controlled Markov Processes and Viscosity Solutions. Springer, New York.

Frank, S. A. 1998. Foundations of Social Evolution. Princeton University Press, Princeton, NJ.

Fudenberg, D. and D. K. Levine. 1998. Theory of Learning in Games. MIT Press, Cambridge, MA.

Goodman, D. 1982. Optimal life histories, optimal notation, and the value of reproductive value. *The American Naturalist* 119:803–823.

Grafen, A. 2006. Optimization of inclusive fitness. *Journal of Theoretical Biology* 238:541–563.

Hamilton, W. D. 1964. The genetical evolution of social behaviour, 1. *Journal of Theoretical Biology* 7:1–16.

Haykin, S. 1999. Neural Networks : A Comprehensive Foundation. Prentice-Hall, 2nd edn.

Hirsch, M. W., S. Smale, and R. L. Devaney. 2004. Differential Equations, Dynamical Systems, and an Introduction to Chaos. Academic Press, San Diego, CA.

Houston, A. I. and J. M. McNamara. 1999. Models of Adaptive Behaviour: An Approach Based on State. Cambridge University Press, Cambridge, UK.

Irie, T. and Y. Iwasa. 2005. Optimal Growth Pattern of Defensive Organs: The Diversity of Shell Growth among Mollusks Optimal Growth Pattern of Defensive Organs: The Diversity of Shell Growth among Mollusks. *The American NaturalistAmerican Naturalist* 165:238–249.

Iwasa, Y. and J. Roughgarden. 1984. Shoot/root balance of plants: optimal growth of a system with many vegetative organs. *Theoretical Population Biology* 25:78–105.

Kamien, M. I. and N. L. Schwartz. 2012. Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management. Dover Publications, New York, 1th edn.

Karlin, S. and H. M. Taylor. 1975. A First Course in Stochastic Processes. Academic Press, San Diego.

Killingback, T. and M. Doebeli. 2002. The Continuous Prisoner's Dilemma and the Evolution of Cooperation through Reciprocal Altruism with Variable Investment. *American Naturalist* 160:421–438.

Kobayashi, Y., J. Y. Wakano, and H. Ohtsuki. 2019. Evolution of cumulative culture for niche construction. *J Theor Biol* 472:67–76.

Lehmann, L. and L. Keller. 2006. The evolution of cooperation and altruism - a general framework and a classification of models. *Journal of Evolutionary Biology* 19:1365–1376.

Lehmann, L. and F. Rousset. 2014. The genetical theory of social behaviour. *Philosophical Transactions of the Royal Society B* 369:1–18.

Leimar, O. 2021. The Evolution of Social Dominance through Reinforcement Learning. *The American Naturalist* 197:560–575.

León, J. A. 1976. Life histories as adaptive strategies. *Journal of Theoretical Biology* 60:301–335.

Liberzon, D. 2011. Calculus of Variations and Optimal Control Theory: a Concise Introduction. Princeton University Press, Princeton.

Lotem, A. 2012. Learning to avoid the behavioral gambit. *Behavioral Ecology* 24:13.

Mangel, M., C. W. Clark, et al. 1988. Dynamic modeling in behavioral ecology. Princeton University Press.

McFarland, D. and A. Houston. 1981. Quantitative Ethology: The State Space Approach. Pitman, Boston.

McFarland, D. J. and R. M. Sibly. 1975. The behavioural final common path. *Philosophical Transactions of the Royal Society B* 270:265–93.

McNamara, J. M., Z. Barta, and A. I. Houston. 2004. Variation in behaviour promotes cooperation in the prisoner's dilemma game. *Nature* pp. 745–748.

McNamara, J. M., C. E. Gasson, and A. I. Houston. 1999. Incorporating rules for responding into evolutionary games. *Nature* 401:368–71.

McNamara, J. M. and A. I. Houston. 1999. Integrating function and mechanism. *Trends in Ecology and Evolution* 24:670–675.

Michod, R. E. 1982. The theory of kin selection. *Annual Review of Ecology and Systematics* 13:23–55.

Michod, R. E. and W. D. Hamilton. 1980. Coefficients of relatedness in sociobiology. *Nature* 288:694–697.

Mullon, C., L. Keller, and L. Lehmann. 2016. Evolutionary stability of jointly evolving traits in subdivided populations. *American Naturalist* 188:175–195.

Mullon, C. and L. Lehmann. 2019. An evolutionary quantitative genetics model for phenotypic (co)variances under limited dispersal, with an application to socially synergistic traits. *Evolution* 73:1695–1728.

Nakamura, M. and H. Ohtsuki. 2016. Optimal Decision Rules in Repeated Games Where Players infer an opponent's mind via simplified belief calculation. *Games* 7:1–23.

Perrin, N. 1992. Optimal resource allocation and the marginal value of organs. *American Naturalist* 139:1344–1369.

Rousset, F. 2004. Genetic Structure and Selection in Subdivided Populations. Princeton University Press, Princeton, NJ.

Rousset, F. and S. Billiard. 2000. A theoretical basis for measures of kin selection in subdivided populations: finite populations and localized dispersal. *Journal of Evolutionary Biology* 13:814–825.

Roze, D. and F. Rousset. 2003. Selection and drift in subdivided populations: A straightforward method for deriving diffusion approximations and applications involving dominance, selfing and local extinctions. *Genetics* 165:2153–2166.

Roze, D. and F. Rousset. 2004. The robustness of Hamilton's rule with inbreeding and dominance: Kin selection and fixation probabilities under partial sib mating. *American Naturalist* 164:214–231.

Siegelman, A. T. and E. D. Sontag. 1995. On the computational power of neural nets. *Journal of computer and system sciences* 50:132–150.

Sontag, E. D. 1998. Mathematical Control Theory. Springer Science & Business Media, New York.

Stearns, S. 1992. The Evolution of Life Histories. Oxford University Press, Oxford.

Sydsaeter, K., P. Hammond, A. Seierstad, and A. Strøm. 2008. Further Mathematics for Economic Analysis. Prentice Hall, Essex, 2th edn.

Taylor, P. D. and T. Day. 2004. Stability in negotiation games and the emergence of cooperation. *Proceedings of the Royal Society of London Series B-Biological Sciences* 271:669–674.

Taylor, P. D., T. Day, and G. Wild. 2007. From inclusive fitness to fixation probability in homogeneous structured populations. *Journal of Theoretical Biology* 249:101–110.

Taylor, P. D. and S. A. Frank. 1996. How to make a kin selection model. *Journal of Theoretical Biology* 180:27–37.

Van Cleve, J. 2015. Social evolution and genetic interactions in the short and long term. *Theoretical Population Biology* 103:2–26.

Van Cleve, J. 2020. Building a synthetic basis for kin selection and evolutionary game theory using population genetics. *Theoretical Population Biology* 133:65–70.

Wakano, J. Y. and C. Miura. 2014. Trade-off between learning and exploitation: The Pareto-optimal versus evolutionarily stable learning schedule in cumulative cultural evolution. *Theoretical Population Biology* 91:37–43.

Wakano, J. Y., H. Ohtsuki, and Y. Kobayashi. 2012. A mathematical description of the inclusive fitness theory. *Theoretical Population Biology* 84:46–55.

Weber, T. A. 2011. Optimal Control Theory with Applications in Economics. MIT press Cambridge, Cambridge, MA.

Wild, G. 2011. Direct fitness for dynamic kin selection. *Journal of Evolutionary Biology* 24:1598–610.