universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

Evolutionary dynamics in structured populations: a life-cycle approach in continuous time

Verfasser

Mag. phil. Andreas Baumann

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer. nat.)

Danksagung

Dass ich ausgerechnet über die Sprachwissenschaft zur Biomathematik gestoßen bin — die mich dann wiederum zu meiner derzeitigen Beschäftigung am Institut für Anglistik gebracht hat —, mag zwar auf den ersten Blick verwundern, hat aber einen ganz einfachen Grund: Sowohl um die Evolution von Sprache als auch von biologischen Organismen formal zu beschreiben, werden zuweilen ganz ähnliche Methoden gebraucht. Gerade die Tatsache, alles Mögliche mit mathematischen Werkzeugen beschreiben zu können, hat für mich vom Beginn des Mathematikstudiums bis zur Fertigstellung dieser Arbeit nichts an Faszination eingebüßt.

So schön und inspirierend dieser Gedanke auch ist — es versteht sich von selbst, dass mir das Studium im Allgemeinen und das Schreiben dieser Arbeit im Besonderen nicht ohne Unterstützung möglich gewesen wäre. An allererster Stelle möchte ich mich deswegen bei Claus bedanken, der es mir zum einen ermöglicht hat, im Rahmen seines Forschungsprojektes diese Diplomarbeit zu schreiben, und zum anderen ein hervorragender Betreuer war. Ich weiß, dass es nicht selbstverständlich ist, so viel Zeit und Geduld in die Betreuung einer Arbeit zu stecken.

Meinen Kollegen am Institut für Anglistik danke ich für ihre Rücksicht während der letzten Monate und eine Vielzahl an offenen Ohren. Besonderer Dank gilt hier Niki für den vielen Freiraum, ohne den ich wohl diese Arbeit nie rechtzeitig zu einem Ende gebracht hätte.

Meiner Familie möchte ich für all die Unterstützung während der letzten siebeneinhalb Jahre danken: meinen Eltern für den finanziellen und moralischen Beistand und meinen Geschwistern für viele aufmunternde Worte — nicht nur während des wöchent-lichen Boulderkurses. An dieser Stelle ist auch ein großes Dankeschön an meinen Bruder Stefan für das Korrekturlesen angebracht.

Meine Freunde, allen voran Jakob und Jo, waren wohl ebenso nicht unwesentlich daran beteiligt, dass ich nie die Motivation verloren habe. Ein bisschen unfair wäre es auch, hier nicht Linda und Hubert zu erwähnen, die zuletzt beide während der Lernpausen notgedrungen immer für mich da waren. Ihnen widme ich Abbildung 1.

Theresa hat meine gesamte Studienzeit aus nächster Nähe miterlebt. Die Geduld und das notwendige Verständnis dafür aufzubringen, war sicher nicht immer leicht. Danke.

Contents

1 Introduction

Most species exhibit a certain population structure. Evidently, this holds for physiologically structured populations, where individuals could for instance differ in body size or available energy reserves (Diekmann et al., 2007). The ecological dynamics of structured populations have been studied extensively, in particular using mathematical tools (Metz and Diekmann, 1986; Cushing, 1998; Caswell, 2001). They are characterized by demographic parameters, namely death rates, maturation rates and rates of reproduction. We can go further and ask, *why* species differ in the demographic parameters and whether they can be understood from an evolutionary perspective. *Life history theory* (Stearns, 1989) tries to find answers to this question: due to natural limitations, at least some demographic parameters have to be traded off against each other. This means that reinforcing one demographic parameter automatically results in the mitigation of at least one other demographic parameter. Obviously, there is a wide range of different strategies of how to allocate given ressources, e.g. food or time, to the various demographic parameters, like maturation or reproduction. Species which show more efficient strategies, will be more successful.

In the past two decades, a considerable amount of research has been done on the evolutionary analysis of structured populations (De Roos, 2008; Hoyle et al., 2008; Bowers, 2011; Rueffler et al., 2013) in particular making use of the mathematical eco-evolutionary framework of *Adaptive Dynamics* (Dieckmann and Law, 1996; Metz et al., 1996; Geritz et al., 1998; Diekmann, 2004). Likewise, this holds for the present thesis. It aims to provide an eco-evolutionary analysis of structured-population models. Rueffler et al. (2013) analyze the evolution of life histories in the following way: evolutionary dynamics of structured-population models in discrete time are predicted by (i) structural properties of the underlying life histories, (ii) the choice of demographic parameters which are traded off against each other, as well as (iii) the choice of demographic parameters which depend on the density of the population. In this thesis, we go along the lines of Rueffler et al. (2013) and translate the results therein to population dynamical models in continuous time.

Mathematical models which are used to analyze physiologically structured populations, also find applications for the analysis of other types of population structure, like stages of infection in epidemiological models, spatial or social structure. Let us give some examples.

Examples of population structure: We consider populations that can be subdivided into several discrete developmental stages each of which is characterized by respective traits for survival and reproduction (Cushing, 1998). Juvenile-adult structures are likely to be the most basic examples of discrete population structures: the whole population consists of non-reproductive juveniles and reproductive adults. Juveniles develop into adults, which in turn give birth to juveniles. Clearly, population structure could be more fine-grained: the stage of juvenile individuals could be split up into multiple juvenile stages, each of which has its own survival rate, for instance in order to account for a decreasing mortality rate. Likewise, the adult stage could be split up into multiple classes characterized by differing rates of fecundity. Furthermore, post-reproductive stages could be considered.

Also more complex stage-structured populations can be found: The amoeba species *Dictyostelium discoideum* (Flowers et al., 2010; Li and Purugganan, 2011), for instance, shows three different modes of reproduction: at low densities, amoebae repro-
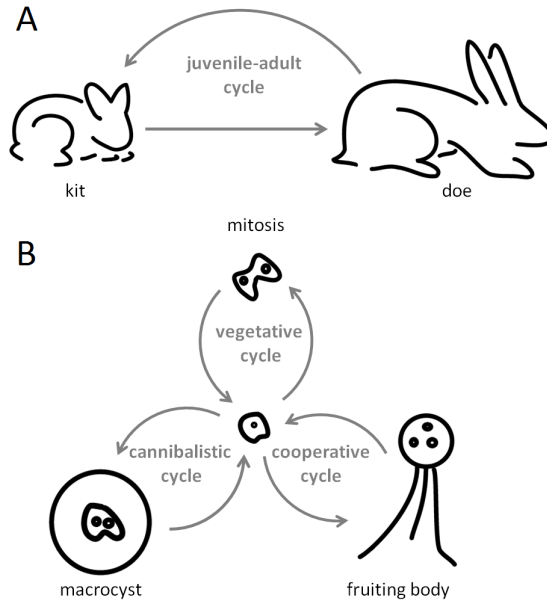
Figure 1: *Examples of stage-structured populations. Figure A shows a juvenile-adult cycle, as can be found in a huge number of species. Figure B shows the life-cycle structure of the amoeba* D. discoideum *with its three qualitatively different ways of reproduction.*

duce mitotically in a vegetative life cycle. If nearby bacteria, which function as food, become sparse, amoebae have the possibility to change to two rather different modes of reproduction: in the first, amoebae switch from asexual to sexual reproduction, thereby forming macrocysts. Macrocysts feed on amoebae and build a protecting shell, which allows them to produce a large number of (again asexual) amoebae. Alternatively, amoebae cooperatively aggregate and form a slug, which is capable of moving to areas of higher food supply. A fruiting body develops, which in turn produces amoebae. See Figure 1 for illustration.

Epidemiology provides another example of population structure: host populations can be divided into *susceptible* (compartment S) and *infected* (compartment I) individuals (Diekmann and Heesterbeek, 2000). Further compartments could be considered, like *exposed* (compartment E) individuals, which are infected but not yet infective, and *recovered* (compartment R) individuals. The choice of compartments depends on the considered disease. Structured epidemiological models are referred to as *SI-*, *SIR-*, *SEI-*, *SEIR*-models, depending on the respective compartments.

Furthermore, spatial structure (Doebeli and Dieckmann, 2004), with different habitat patches, or social structure (Cavalli-Sforza and Feldman, 1981; Doebeli, 2011) could be considered. Finally, we have to emphasize that all of the above mentioned types of population structures can be combined. That is, for example spatially structured *SI*-models with juvenile-adult cycles as well as spatially and socially structured *SEIR*-models could be considered.

Approach and outline: This thesis is concerned with the long-term evolutionary analysis of structured-population models in continuous time. While certain populations are better to be investigated under the assumption of discrete time, like models of annual plants, we are focusing on the case of continuous time. Rueffler et al. (2013) provide a thorough analysis of long-term evolution in discrete-time structured-population models. Although population dynamical modeling obviously differs in the continuous-time case, this thesis is pursuing the same approach as elaborated by Rueffler et al. (2013), where evolutionary predictions are based on an analysis of population structure itself, or more precisely, the configuration of life cycles therein. It has to be remarked, that Hoyle and Bowers (2008), Hoyle et al. (2008), and Bowers (2011) provide useful tools for the evolutionary analysis of structured populations, which are, however, based directly on the population dynamical model at hand, rather than on population structure. In this thesis, precisely like in Rueffler et al. (2013), it is the structural properties of a life cycle together with the choice of evolving and regulated demographic parameters, which serves as basis for evolutionary predictions, rather than properties of a mathematical model. In this respect, the results in this thesis will have slightly more direct biological interpretation.

Like in Rueffler et al. (2013), long-term evolution is investigated by employing the mathematical framework of *Adaptive Dynamics*. Using this framework, we can investigate the evolutionary trajectories of certain demographic parameters, which are perceived as phenotypic traits and are subject to mutations. The key feature of Adaptive Dynamics is the evolutionary *feed-back loop*. Populations influence the environment they are living in to a certain extent. If a rare mutant occurs, it has to cope with the environment set by the resident population. If the mutant population replaces the resident, it will shape the environment in a slightly different manner. It is now this new environment, a potential subsequent mutant population is going to face. In this way, evolution is modeled as a trait-substitution sequence. It is cruical, that in Adaptive Dynamics the fitness of a mutant is derived from an explicit population dynamical scenario and not assumed a priori.

We proceed as follows: in section 2 the class of structured-population models is defined. In this section, only ecological dynamics will be considered. Structured-population models will be described by demographic parameters and functions modifying demographic parameters which account for density regulation.

Section 3 introduces the Adaptive Dynamics approach. The set of evolutionary scenarios can be subdivided into two main types: *frequency independent* and *frequency dependent selection*. In models of the former type the evolutionary dynamics can be derived by analyzing an optimization criterion. That is, there are (more or less easy to interpret) measures of fitness, which are maximized by evolution (Mylius and Metz, in press). Under frequency dependence, selection can become disruptive, thereby giving rise to stable polymorphisms, i.e., stably coexisting differing species.

Finally, section 4 joins both parts together: the class of models introduced in section 2 is analyzed using the tools of Adaptive Dynamics given in section 3. This is done by providing an algebraically simple quantity to measure fitness, which in addition is tightly related to population structure. Using this so-called *fitness proxy*, we show which models support optimization, and derive conditions under which frequency dependent selection can be disruptive for a subset of our model family. The shape of *trade-offs* between evolving phenotypic traits will be cruical.

We will see, that with certain restrictions, the present results coincide with the discrete-time case presented in Rueffler et al. (2013). Throughout the whole thesis,

examples are provided, which shall help to illustrate both, introduced concepts as well as derived results.

2   Structured population models

This part is meant to provide an introduction of the basic mathematical tools which are needed to describe the ecological dynamics of structured populations. Population structure is assumed to be discrete. That is, we will assume that the whole population of individuals can be subdivided into a finite number of separate classes, stages or states (Keyfitz and Caswell, 2005). Such states can be interpreted in several ways, such as different stages of maturity or stages of infectivity. Likewise spatial structure could be considered. In contrast to Rueffler et al. (2013), we are focusing on population dynamics in continuous rather than discrete time. Consequently, population dynamics are descibed by systems of ordinary differential equations (ODEs) as compared to difference equations. Life histories will be described by three different types of demographic parameters: fertility rates, death rates and rates for transition from one state to another. The latter could represent maturation rates, rates of infection, or migration rates. Density dependence will be accounted for via regulatory functions, which modify demographic parameters.

Section 2.1 will introduce the model family. Section 2.2 gives a short outline of stability theory (Hofbauer and Sigmund, 1998). This will be useful, since for the evolutionary analysis we will by assumption demand the existence of stable point attractors of the ecological dynamics. Finally, 2.3 will introduce life-cycle graphs, i.e., graphical representations of the underlying system of ODEs. The structural properties of those life-cycle graphs will be exploited later on in order to facilitate an evolutionary analysis of structured-population models (Hoyle and Bowers, 2008; Rueffler et al., 2013). This, however, will be subject to a latter part, namely section 4.

2.1   Ecological dynamics of structured populations

Since we are interested in the evolutionary dynamics of life-histories, the relevant demographic parameters are fertility and survival rates of different life-history stages as well as rates at which individuals switch from one life-history stage to another one. For a given structured population model, these so-called transition rates implicitly define the set of possible life histories for each individual. It is assumed that certain demographic parameters can depend on the density of the whole population or on densities of subpopulations. For instance, the fertility of adults could decrease as the number of reproductive adults increases (as it is for instance the case for the European rabbit, see Rödel et al., 2004, cf. Example 4).

The mathematical description of the structured population is straightforwardly done in terms of a (generally nonlinear) dynamical system. This system then is converted into a more accessible matrix formulation. The core theory of dynamical systems can be found for instance in Hofbauer and Sigmund (1998).

Basic description via ODEs: The dynamics of populations with discrete population structure are described by k-dimensional dynamical systems in continuous time, where k is the number of different states (cf. Hoyle and Bowers, 2008). For each state $i \in \{1, ..., k\}$, the dynamics of the respective populations size $n_i(t)$ is defined by an

ordinary differential equation (ODE) of the form

$$\frac{\mathrm{d}n_i}{\mathrm{d}t} = \dot{n}_i = f_{ii} + \sum_{j=1, j \neq i}^{k} (t_{ij} + f_{ij})n_j - \sum_{j=1, j \neq i}^{k} t_{ji}n_i - d_i n_i, \qquad (1)$$

where f_{ji}, t_{ji}, and d_i are parameters which are relevant for the description of life histories (Keyfitz and Caswell, 2005):

□ **Fertility rates:** $f_{ji} \geq 0$, $i, j \in \{1, ..., k\}$, denotes the rate at which individuals in state i give birth to individuals in state j.

□ **Transition rates:** $t_{ji} \geq 0$, $i, j \in \{1, ..., k\}$ with $i \neq j$, denotes the rate at which individuals in state i change into state j.

□ **Death rates:** $d_i > 0$, $i \in \{1, ..., k\}$, denotes the rate at which individuals in state i die.

We will refer to f_{ji}, t_{ji}, and d_i as *demographic parameters.*

The first term in equation (1) is given by the intrastate fertility rate of state i. The second term collects all individuals that enter state i via reproduction or transition from a state j which is different from i. The third term collects all individuals that leave state i for other states j. Finally, the fourth term gives the number of i-individuals that die.

The use of deterministic models (like ODEs) is justified by the assumption that populations are sufficiently large to be insensitive to stochastic effects (Hofbauer and Sigmund, 1998; Dercole and Rinaldi, 2008). For employing differential equations (where time is measured continuously) rather than difference equations (where time is measured discretely), we need that each demographic event corresponding to a demographic parameter $c \in \{f_{ji}, t_{ji}, d_i\}$ occurs after a time-span following an exponential distribution around an average of $\frac{1}{c}$. This means, that the probability for a demographic event to occur in the time interval $[0, \tau]$ is $1 - e^{-c\tau}$ (Nowak, 2006, p. 11). This differs from discrete-time models, where all demographic events corresponding to a demographic parameter occur at exactly the same time. In practice, differential equations are used, if populations are relatively large, if demographic events are independent of each other, if the time-span between two demographic events is relatively short, or if generations are overlapping. For example, it is common practice to employ continuous-time models in epidemiological applications.

Example 1. Consider a population in which individuals are in either of two states, 1 and 2, where 1 denotes the juvenile non-reproductive state and 2 denotes the adult reproductive state, as depicted in Figure 1. Individuals can change from state 1 to state 2, i.e., they can develop from juvenile to adult, but not the reverse. Only adults can produce offspring of state 1. Both, juvenile and adult individuals die at certain rates. The resulting system of differential equations then reads:

$$\begin{aligned}
\dot{n}_1 &= f_{12}n_2 - t_{21}n_1 - d_1 n_1 \\
\dot{n}_2 &= t_{21}n_1 - d_2 n_2.
\end{aligned}$$

Regulatory functions: Demographic parameters can be either constant or depending on the density of all individuals or on the density in a subset of states. There are multiple reasons for why demographic parameters can be density dependent. For

instance, individuals could compete for limited resources, like food, space or nesting sites, hence resulting in increased rates of maturation or reproduction. Furthermore, mortality could increase due to increased probability of infection or an increased number of predators, both as a reflex of high population densities. Likewise, stress could cause increased death rates or decreased fertility rates.

If all demographic parameters f_{ij}, t_{ij}, and d_i (or equivalently r_i) remain constant, the system of ODEs given by equation (1) is linear. Otherwise it will become a nonlinear system of ODEs. For plausible modeling of population dynamics, population regulation is essential: otherwise we could not exclude unlimited exponential growth.

For the k-dimensional vector of state sizes let us write $\boldsymbol{n} = (n_1, ..., n_k)^T$.

Definition 1. A *regulatory function* corresponding to a demographic parameter x_{ij} is a continuously differentiable strictly monotone function $R_{x,ij} : \mathbb{R}_+^k \supseteq D \to (0, \infty)$, $\boldsymbol{n} \mapsto R_{x,ij}(\boldsymbol{n})$. We use the abbreviation $R_{x,i}$ for $R_{x,ii}$. A demographic parameter is called *multiplicatively regulated*, if it is of the form $\tilde{x}_{ij} := x_{ij} \cdot R_{x,ij}(\boldsymbol{n})$ and *additively regulated* if it is of the form $\tilde{x}_{ij} := x_{ij} + R_{x,ij}(\boldsymbol{n})$, where x_{ij} denotes the unregulated demographic parameter and $\tilde{x}_{ij}$ denotes the corresponding regulated demographic parameters. A regulatory function is called *multiplicative* and *additive* if it regulates a demographic parameter in the respective manner.

In the following, if not explicitly stated otherwise (as this is the case for density regulation by the law of mass action, see example below), we will assume that $\frac{\partial R_{x,ij}}{\partial n_l} \leq 0$ for all $l = 1, ..., k$, if x_{ij} is a fertility or transition rate and $\frac{\partial R_{x,i}}{\partial n_l} \geq 0$ for all $l = 1, ..., k$, if x_i is a death rate. The cruical difference between additively and multiplicatively regulated parameters $\tilde{x}_{ij}$ is, that while the former are bounded from below by $x_{ij} > 0$, the latter are in general bounded from below by 0. From now on, for general considerations we assume that any demographic parameter can be regulated. In specific examples, however, regulatory functions will be indicated as shown in the definition.

The way in which regulatory functions are defined above, allows for a wide range of applications. Due to the monotonicity requirement, however, Allee-effects are implicitly excluded.

Example 2. Let us give some examples of regulatory functions (for illustration see Figure 2):

1. For a demographic parameter x_{ij} and positive constants a and b, and $c_l \geq 0$, $l = 1, ..., k$, $R_{x,ij} : \boldsymbol{n} \mapsto \frac{a}{b + \sum_{1 \leq l \leq k} c_l n_l}$ defines a hyperbolic regulatory function. The constants c_l are state specific weighting parameters. The corresponding multiplicatively regulated demographic parameter reads $\tilde{x}_{ij} = \frac{a x_{ij}}{b + \sum_{1 \leq l \leq k} c_l n_l}$. This type of saturating regulatory functions is related to *Michaelis-Menten enzyme kinetics* (Murray, 2002) and *Holling type II functional response* (Otto and Day, 2007; Kot, 2001). If we consider the structured population given in the previous example, and if $x_{12} = f_{12}$ and $c_1 = c_2 = c$, then $\tilde{f}_{12} = \frac{a f_{12}}{b + c(n_1 + n_2)}$, which means that the fertility of adult individuals is (deceleratingly) decreasing in the overall size of the population (for instance due to competition for nesting sites).

2. Consider a demographic parameter x_{ij} and constants $a > 0$ and $b_l \geq 0$, $L = 1, ..., k$, $R_{x,ij} : \boldsymbol{n} \mapsto a \cdot \exp\left(-\sum_{1 \leq l \leq k} b_l n_l\right)$ defines a decreasing regulatory
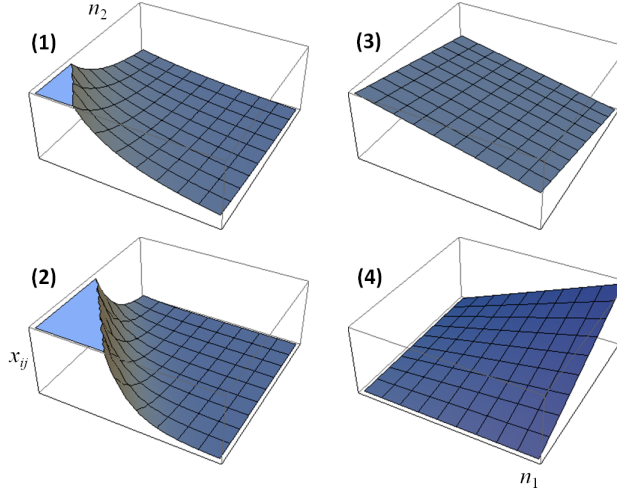
Figure 2: *Plots of regulatory functions $R_{x,ij}$ as in Example 2.1-4 for two dimensions n_1 and n_2, where weighting constants are chosen as 1 and 0.75, respectively, and $x_{ij} = 1$. Regulation is hyperbolic in (1), negatively-exponentially in (2), and logistic, i.e., linear, in (3), while (4) gives an example of parameter regulation via the law of mass action.*

function, where the constants b_l can be used to weight the effect of each class on the demographic parameter. It describes an exponential decrease in the demographic parameter as the sizes of classes are increasing. For instance, it has been incorporated into (discrete time) models of stage-structured insect populations, so-called *LPA-models* (as an abbreviation for *larvae-pupae-adult*, see Cushing, 1998; Henson, 1999; Otto and Day, 2007).

3. Clearly, $R_{x,ij} : \boldsymbol{n} \mapsto a \cdot \left(1 - \sum_{l \in \alpha} n_l / K\right)$, for some constants $a, K > 0$ and $\alpha \subseteq \{1, ..., k\}$, defines a decreasing logistic regulatory function. According to this function, a multiplicatively regulated parameter declines linearly, as the number of individuals in all states $l \in \alpha$ increases. Due to its simplicity, this type of regulatory functions is commonly used to implement density dependence into all sorts of population dynamical models (see Hofbauer and Sigmund, 1998; Otto and Day, 2007; Kot, 2001). Note, that logistic regulatory functions are in need of a restricted domain D in order to fulfill $R_{x,ij} > 0$.

4. Finally, as a somewhat more unconventional regulatory function, for a subset $\alpha \subseteq \{1, ..., k\}$ and $a_l > 0$ (for $l \in \alpha$) we have $R_{x,ij} : \boldsymbol{n} \mapsto \prod_{l \in \alpha} a_l n_l$, which, if applied multiplicatively, can be used to model parameter regulation via the *law of mass action*. That is, the considered parameter is proportional to the product of all sizes of classes in α, and hence increasing (Murray, 2002). It is important to emphasize, that this type of population regulation is rather different from the decreasing regulatory functions in the three examples above. This type of regulatory functions is often used in epidemiology, where transmission rates are supposed to be proportional to the number of contact events of susceptible and infected individuals (Diekmann and Heesterbeek, 2000; Dieckmann and Metz, 2006).

8

Matrix formulation: More conveniently, the system of ODEs given by (1) can equivalently be written as

$$\dot{\boldsymbol{n}} = \mathsf{A}(\boldsymbol{n}) \cdot \boldsymbol{n}, \tag{2}$$

where $\dot{\boldsymbol{n}} = (\dot{n}_1, ..., \dot{n}_k)^T$ and $\mathsf{A}(\boldsymbol{n})$ is a $k \times k$-matrix. The entry a_{ij} of A is given by

$$
\begin{aligned}
a_{ij}(\boldsymbol{n}) &= \tilde{f}_{ij} + \tilde{t}_{ij} & i \neq j \\
a_{ii}(\boldsymbol{n}) &= \tilde{f}_{ii} - \tilde{d}_i - \sum_{l \neq i} \tilde{t}_{li}
\end{aligned}
$$

We will refer to such a matrix A as a *life-cycle matrix* or simply *life cycle*. The entries of $\mathsf{A}(\boldsymbol{n})$ are (possibly constant) functions of $\boldsymbol{n}$, i.e., functions $a_{ij}(\boldsymbol{n})$, as they may contain regulatory functions.

Since all demographic parameters f_{ij} and t_{ij} are nonnegative, since all d_i are positive, and since all regulatory functions are strictlly positive, it follows that $a_{ij}(\boldsymbol{n}) \geq 0$ if $i \neq j$.

Example 3.

1. Consider the life-cycle of the amoeba *D. discoideum* (Flowers et al., 2010; Li and Purugganan, 2011; see also Figure 1): At high densities, amoebae either enter a cooperative life cycle or a cannibalistic life cycle, by moving to stage 2 or 3, respectively. Under the cooperative life cycle, amoebae form movable slug, so it is reasonable not to assume density dependence in stage 2. Amoebae in stage 3, on the other hand, form a zygote and become cannibalistic. Hence, amoebae in stage 3 affect the death rate d_1. The corresponding life-cycle matrix reads

$$
\mathsf{A} = \begin{pmatrix} f_{11} - \tilde{t}_{21} - \tilde{t}_{31} - \tilde{d}_1 & f_{12} & \tilde{f}_{13} \\ \tilde{t}_{21} & -d_2 & \\ \tilde{t}_{31} & & -d_3 \end{pmatrix}.
$$

2. The following epidemiological compartment model describes the dynamics of a disease with a latent but not infectious state, like for instance *tuberculosis* (see Ozcaglar et al., 2012 for an overview). The population is structured into the classes S (susceptible), E (exposed, i.e., infected but not infectious), and I (infected and infectious). We will use standard epidemiological notation. The variables S, E and I measure the densities of the respective subpopulations. Birth and death rates of the disease free susceptibles are denoted as b and d. Parameters related to the disease are denoted by Greek letters: rates of infection β_{ji}, recovery rates θ_{Si} from infected/exposed to susceptible (where $i \in \{S, E, I\}$) and ρ_{EI} from infected to exposed, are transition rates as defined above. Disease free death rates increase by α_i ($i \in \{E, I\}$) and depend on the state of infection. Assume that individuals always have to pass through the latent state E in order to reach I. Then, the dynamics are given by

$$
\begin{aligned}
\dot{S} &= [b - \beta_{ES}I - d]S + \theta_{SE}E + \theta_{SI}I \\
\dot{E} &= \beta_{ES}IS - [\beta_{IE}I + \theta_{SE} + (d + \alpha_E)]E + \rho_{EI}I \\
\dot{I} &= \beta_{IE}IE - [\rho_{EI} + \theta_{SI} + (d + \alpha_I)]I
\end{aligned}
$$

yielding the life-cycle matrix

$$
\mathsf{A} = \begin{pmatrix} b - \tilde{\beta}_{ES} - d & \theta_{SE} & \theta_{SI} \\ \tilde{\beta}_{ES} & -\tilde{\beta}_{IE} - \theta_{SE} - (d + \alpha_E) & \rho_{EI} \\ & \tilde{\beta}_{IE} & -\rho_{IE} - \theta_{SI} - (d + \alpha_I) \end{pmatrix},
$$

where the transition rates $\tilde{\beta}_{ES} = \beta_{ES}I$ and $\tilde{\beta}_{IE} = \beta_{IE}I$ are regulated via the law of mass action by the number of individuals in class I: it is assumed that the number of transitions from S to E only depends on the number of contacts of class S and class I individuals, since exposed individuals by definition are not infectious. Similar considerations hold for the number of transitions from E to I: here, $\tilde{\beta}_{IE}$ represents the rate of *exogenous reinfection*[1] depending on the number of contacts of exposed and infective individuals (Ozcaglar et al., 2012). Likewise it could be assumed that β_{IE} does not depend on contact events with infectious individuals, and is thereby unregulated. Then β_{IE} represents the rate of *endogenous reactivation.*

2.2 Ecological long-term behavior

For an evolutionary analysis of a structured population model finding an exact solution of the system of differential equations at hand is not of primary interest. We are only interested in the attractors of the system. As will be shown in section 3 we will assume the existence of nontrivial point attractors of the dynamical system in order to determine the fitness of a mutant in a resident structured population. Note, that for the existence of an internal equilibrium, regulatory functions are necessary. The structure of the life-cycle matrix A will play a key role in the evolutionary analysis of the structured population. Therefore, this section comprises a result about the structure of A when the system has reached its equilibrium on the ecological time scale.

Equilibria and stability: We are interested in the long-term dynamics of (2),

$$\dot{\boldsymbol{n}} = \mathsf{A}(\boldsymbol{n}) \cdot \boldsymbol{n}.$$

Although this system only rarely has an explicit solution we can make qualitative predictions about the behavior of the dynamical system by finding its equilibria and employing standard stability analysis (Hofbauer and Sigmund, 1998). As usual, for a k-dimensional dynamical system $\dot{n}_i(t) = f_i(\boldsymbol{n}(t))$, $i = 1, ..., k$, a point $\hat{\mathbf{n}}$ is a *population dynamic equilibrium*, if $f_i(\hat{\boldsymbol{n}}) = 0$ for all $i = 1, ..., k$. A population dynamical equilibrium $\hat{\boldsymbol{n}}$ is said to be *stable* if for all neighborhoods U of $\hat{\mathbf{n}}$ there is another neighborhood V so that all orbits $\boldsymbol{n}(t)$ starting in V remain in U for all t. It is *asymptotically stable* if in addition $\lim_{t\to\infty} \boldsymbol{n}(t) = \hat{\boldsymbol{n}}$ for all $\boldsymbol{n} \in V$.

The stability of a population dynamic equilibrium $\hat{\boldsymbol{n}}$ can be determined by linearizing the dynamical system around $\hat{\boldsymbol{n}}$ and calculating the eigenvalues of the corresponding Jacobian matrix evaluated at $\hat{\boldsymbol{n}}$:

$$\mathsf{J}|_{\hat{\boldsymbol{n}}} = \left(\frac{\partial f_i}{\partial n_j} \right)_{1 \le i,j \le k} \Bigg|_{\boldsymbol{n}=\hat{\boldsymbol{n}}}.$$

If for all eigenvalues of the Jacobian matrix, $\lambda(\mathsf{J}|_{\hat{\boldsymbol{n}}})$, the real part is negative, then $\hat{\boldsymbol{n}}$ is asymptotically stable, and $\hat{\boldsymbol{n}}$ is unstable if for at least one eigenvalue the real part is positive (in the case of $\max_{\lambda}\{\mathrm{Re}\lambda\} = 0$ higher order terms have to be considered).

---

[1]As opposed to *endogenous reactivation*, where the exposed individual becomes infective without external infection, "[e]xogenous reinfection is caused by a secondary external infection in which the new [Mycobacterium tuberculosis complex, i.e., the pathogen] makes the individual infectious, thereby causing the active [tuberculosis] infection" (Ozcaglar et al., 2012, 78).

Example 4. Consider the structured population given by

$$\dot{n}_1 = f_{12}n_2 - t_{21}n_1 - d_1 n_1$$
$$\dot{n}_2 = t_{21}n_1 - d_2 n_2.$$

This dynamical system has only one equilibrium, namely $\hat{\boldsymbol{n}} = (0,0)^T$, as the system is linear. In order to determine the stability of this equilibrium point we calculate the Jacobian matrix of the system given above:

$$\mathsf{J} = \begin{pmatrix} -t_{21} - d_1 & f_{12} \\ t_{21} & -d_2 \end{pmatrix}$$

Note, that in this case, since there are no entries depending on $\boldsymbol{n}$, J is just the life-cycle matrix of the structured population, A (an observation which will be useful later on in section 4). The Jacobian matrix does not depend on $\boldsymbol{n}$, hence $\mathsf{A} = \mathsf{J} = \mathsf{J}|_{\hat{\boldsymbol{n}}}$. We can directly calculate the eigenvalues of J obtaining

$$\lambda_{1,2} = \underbrace{-\frac{1}{2}(d_1 + d_2 + t_{21})}_{\alpha} \pm \underbrace{\frac{1}{2}\sqrt{(d_1 + d_2 + t_{21})^2 - 4(d_1 d_2 + d_2 t_{21} - f_{12} t_{21})}}_{\beta}.$$

The stability of $\hat{\boldsymbol{n}}$ then depends on whether $\alpha < \mathrm{Re}\beta$.

Now, let us incorporate a hyperbolic regulatory function as in Example 2.1. This means that the rate of adults giving birth to juveniles hyperbolically depends on the overall population size. The phenomenon of density dependent fertility regulation is widespread among mammals (Krebs, 1996) and particularly a characteristic feature of the European rabbit, *Oryctolagus cuniculus*, (Rödel et al., 2004): it has been shown that the average number of offspring is decreasing with increasing numbers of female rabbits. For simplicity, let us only keep track of the number of female rabbits. Accordingly, in the regulatory function $R_{f,12}$ as given in Example 2.1 we set $c_1 = 0$ and $c_2 = 1$. Furthermore, let us assume $a = b = 1$. Consequently, the system reads

$$\dot{n}_1 = \frac{f_{12}}{1 + n_2}n_2 - t_{21}n_1 - d_1 n_1$$
$$\dot{n}_2 = t_{21}n_1 - d_2 n_2.$$

This nonlinear system has two equilibria: $\hat{\boldsymbol{n}}_1 = (0,0)^T$ and a nonzero equilibrium. As described above, calculating the eigenvalues of the linearized system evaluated at $\hat{\boldsymbol{n}}_1$ and $\hat{\boldsymbol{n}}_2$ shows us, whether or not the equilibria are stable. Note, that $\hat{\boldsymbol{n}}_2 > 0$ only if $f_{12}t_{21} > d_2(d_1 + t_{21})$.

Population dynamic equilibria and the sign structure of A: For simplicity, we will assume that for system (1) there is an equilibrium which is a single point attractor $\hat{\boldsymbol{n}}$.[2] We can make certain assertions about the sign structure of a life-cycle matrix $\mathsf{A}(\hat{\boldsymbol{n}})$. We can state the following proposition:

---

[2]This assumption will simplify the evolutionary analysis of the model family. Strictly speaking, under the assumption that mutational effects are small, the requirement that there is *exactly one* point attractor is not necessary. If there are multiple point attractors, one has to be specific about at which the population resides before the first mutant appears.

Proposition 1. *If the system $\dot{\boldsymbol{n}} = \mathsf{A}(\boldsymbol{n})\cdot\boldsymbol{n}$ is in a population dynamic equilibrium $\hat{\boldsymbol{n}}$, then the diagonal entries of $\mathsf{A}(\hat{\boldsymbol{n}})$ are nonpositive, i.e., $a_{ii}(\hat{\boldsymbol{n}}) \leq 0$, for all i. Moreover, if state i has (possibly regulated) incoming transition or fertility rates, i.e., there are $\tilde{t}_{ij}$ or $\tilde{f}_{ij} > 0$ for some $j \neq i$ and if each class has positive density at population dynamic equilibrium, i.e., $\hat{\boldsymbol{n}} > 0$, then $a_{ii}(\hat{\boldsymbol{n}}) < 0$.*

Proof. If $\hat{\boldsymbol{n}}$ is a population dynamic equilibrium of (1), then for all i we have

$$
\begin{aligned}
0 &= f_{ii}\hat{n}_i + \sum_{j=1}^{k}(\tilde{t}_{ij} + \tilde{f}_{ij})\hat{n}_j - \sum_{j=1}^{k}\tilde{t}_{ji}\hat{n}_i - d_i\hat{n}_i \\
&= \left[\left(\tilde{f}_{ii} - d_i - \sum_{j=1}^{k}\tilde{t}_{ji}\right)\hat{n}_i\right] + \underbrace{\left[\sum_{j\neq i}(\tilde{t}_{ij} + \tilde{f}_{ij})\hat{n}_j\right]}_{\geq 0}.
\end{aligned}
$$

Since all demographic parameters and all $\hat{n}_i$ are nonnegative, it immediately follows that $0 \geq \tilde{f}_{ii} - d_i - \sum_{j=1}^{k}\tilde{t}_{ij} = a_{ii}(\hat{\boldsymbol{n}})$. If there are incoming transition or fertility rates, then $\tilde{t}_{ij} + \tilde{f}_{ij} > 0$ for some $j \neq i$ and $\hat{\boldsymbol{n}} > 0$. Hence the first summand has to be strictly negative, whereby the second statement follows. $\qquad\square$

Note that this also applies if some demographic parameters are regulated, since regulatory functions are strictly positive. The proposition implies that

$$
a_{ij}(\hat{\boldsymbol{n}}) \begin{cases} \geq 0 & \text{if } i \neq j \\ \leq 0 & \text{if } i = j \end{cases}.
$$

2.3 Life-cycle graphs

Life-cycle graphs are a useful tool to investigate structured populations, as they provide a straight forward graphic representations of a given life cycle (see for instance Caswell, 2001, Otto and Day, 2007 and Kot, 2001). There is a natural one-to-one relation between life-cycle graphs and life-cycle matrices: roughly speaking, states are mapped to vertices and demographic parameters are mapped to edges. In this section we will define the loop transmission, which factorizes the rates that correspond to the states that are passed through in a loop in the life-cycle graph. As will be seen later on, the evolutionary dynamics of a structured population are closely linked to the structure of its corresponding life-cycle graph.

Graphs are formally defined as pairs of sets of vertices and edges (see Appendix A.1 for a precise definition). For a $n \times n$ life-cycle matrix $\mathsf{A} = (a_{ij})_{1 \leq i,j \leq n}$ the set of vertices S_A is defined by the set of states $\{1, ..., n\}$. The set of edges E_A collects all pairs of states (i, j) with $a_{ij} \neq 0$ $(i \neq j)$, and all pairs (i, i) for all i. The graph $G_\mathsf{A} = (S_\mathsf{A}, E_\mathsf{A})$ then is called a *life-cycle graph*. The weights on the edges are defined by the respective matrix entries, i.e., $w(i, j) = a_{ij}$.

Example 5. The directed graph correponding to the structured population given in Example 1 can be represented in the following way:
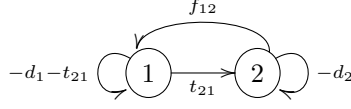
In a somewhat more intuitive way, it can be rearranged as a *flow diagram*, such that the death rates do not form self loops. Rather, death rates simply define outgoing arrows. Moreover outgoing transition rates do not contribute to the weight in the corresponding self loop (and do not establish self loops).
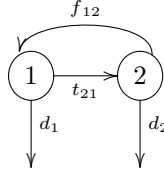
It is cruical that life-graphs as defined above directly represent its underlying life-cycle matrix, even if the type of a demographic parameter is not known. Without further information about the type of demographic parameters, this does not hold for flow diagrams. Hence, the latter way of representing a life-cycle matrix is ambiguous.

Sub-life-cycles and loops: A sub-life-cycle can be any subset of states of a structured population, whereas loops intuitively describe particular life-histories of individuals.

Definition 2. Let A be a $k \times k$ life-cycle matrix and let $\alpha \subset \{1, ..., k\}$ be a proper subset of the set of states. The *principal sub-matrix* of A that consists only of rows and columns with indices in α is denoted as A_α. A *sub-life-cycle* is then the life cycle corresponding to A_α. The principal sub-matrix of A, where all rows and columns with indices in α are deleted is denoted as $A_{\{1,...,k\}\setminus\alpha} =: A_{\setminus\alpha}$)

The sub-graph describing A_α is given by $G_{A_\alpha} := (\alpha, E_{A_\alpha})$. It contains exactly those edges $(s_i, s_j) \in E_A$ with $s_i, s_j \in \alpha$. Note that in the diagonal entries a_{ii} of A_α transition rates t_{ji} with a negative sign where $j \notin \alpha$ may appear. For biological interpretation, these should be added to the corresponding death rates d_i.

Loops are closed paths through a graph, where each vertex is traversed exactly once. For a formal definition of paths and loops see Appendix A.1. The length $\ell(G)$ of a loop G is the number of vertices in G. A self loop is a loop of length 1. By factorizing all weights along a loop, we get the corresponding *loop transmission*:

Definition 3. The *loop transmission* L of a loop $G = (S, E)$ is defined as the product of the weights of the edges of G, i.e., $L := \prod_{e \in E} w(e)$.

For loops $G = (\{i_1, i_2, ..., i_l\}, \{(i_1, i_2), (i_2, i_3), ..., (i_l i_1)\}) =: [i_1, i_2, ..., i_l, i_1]$ (see Appendix A.1) in a directed graph G_A induced by a life-cycle matrix A, the loop transmissions are then simply given by $L = a_{i_2 i_1} \cdot a_{i_3 i_2} \cdot ... \cdot a_{i_l i_i}$. Note, that loops and paths are graph-theoretic objects, while loop transmissions are products of matrix

entries. We can define $\mathcal{L}(\mathsf{A}) := \{L | L \text{ is the loop transmission of a loop } G \subseteq G_\mathsf{A}\}$ as the set of all loop transmissions in a directed graph generated by A. That is, $\mathcal{L}(\mathsf{A})$ collects all products along loops in a given life cycle. In a somewhat sloppy manner we will also use the term loop instead of loop transmission.

In a life cycle self loops are of the form $L = a_{ii} = f_{ii} - d_i - \sum_{l \neq i} t_{li}$ while all other loops are products of (sums of) fertility and transition rates, i.e., $L = x_{i_2 i_1} \cdot x_{i_3 i_2} \cdot ... \cdot x_{i_l i_i}$, where $x_{ij} = a_{ij} \in \{t_{ij}, f_{ij}, t_{ij} + f_{ij}\}$.

Alternatively, we could have defined loops as products of demographic parameters. Then each ii-entry would correspond to separated, pairwise connected loops of the form f_{ii}, $-d_i$, $-t_{li}$ ($l \neq i$). The definition of loops as products of *matrix entries* as opposed to products of *demographic parameters* is not unproblematic: First, as a consequence, factors in a loop can be sums of (possibly regulated) demographic parameters. This can make algebraic manipulations complicated. Second, at first sight it is biologically less intuitive.

We stick, however, to the above definition for the following reasons: First, if loops are defined as products of matrix entries, there is an injective correspondence between the set of pairwise unconnected loops and the set of disjoint permutation cycles in $\{1, ..., n\}$. This will simplify calculations.

Second, while reproductive self loops of the form f_{ii} are biologically intuitive, it is less plausible to perceive $-d_i$ as a self loop. Fertility rates and death rates are often simply combined to growth rates $r_i := f_{ii} - d_i$ (also in epidemiological applications). Furthermore, there is no clear biological interpretation of self loops which are generated by outgoing transition rates, i.e., $-t_{li}$. An intuitive option would be to perceive loops of the form $-t_{li}$ as quasi-death-rates. Then, as it is done in de Camino-Beck et al. (2009) for epidemiological models, we can combine death rates and outgoing transition rates to sums $m_i := d_i + \sum_{l \neq i} t_{li}$. Altogether, we can define the state-specific overall per-capita growth rate

$$\rho_i := f_{ii} - m_i = r_i - \sum_{l \neq i} t_{li} = a_{ii} = L,$$

which describes at which rate individuals in state i contribute to the growth of state i.

Third, additive regulatory functions are problematic, since under additive density regulation, self loops by definition show sums of demographic parameters and regulatory functions. This again complicates algebraic manipulations. Alternatively, we could perceive regulatory functions as self loops: for instance f_{ii} and $-R_{f,ii}$ could generate two connected self loops. Interpreting $-R_{f,ii}$ as a loop is, however, not entirely satisfying from a biological point of view.

Example 6.

1. The life-cycle graph for the stage-structured juvenile-adult model in Example 1 represented by
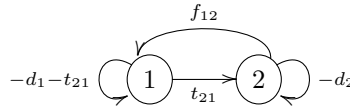
has two loops of length 1, $G_1 = [1,1]$, $G_2 = [2,2]$, and one loop of length 2, $G_3 = [1,2,1]$. The corresponding loop transmissions read $L_1 = -d_1 - t_{21}$, $L_2 = -d_2$, and $L_3 = t_{21}f_{12}$, respectively. Note that although the graph which goes from 1 to 2, from 2 back to 1, and finally from 1 to 1 (more precisely the graph $(\{1,2\}, \{(1,2),(2,1),(1,1)\})$) is a path, it is not a loop.

2. Consider the *SEI*-model from Example 3.2. If we are only interested in the dynamics of exposed and infected individuals, we have a similar life cycle as in the previous example. It corresponds to the submatrix $\mathsf{A}_{\{E,I\}}$.

We find two loops of length 1, $G_1 = [E,E]$, $G_2 = [I,I]$, and one loop of length 2, $G_3 = [I,E,I]$. The corresponding loop transmissions read $L_1 = -\theta_{SE} - (d+\alpha_E)$, $L_2 = -\rho_{EI} - \theta_{SI} - (d+\alpha_I)$, and $L_3 = \tilde{\beta}_{IS}\rho_{EI}$, respectively.

3. We have the following life-cycle graph for *D. discoideum* (see Example 3.1):

There are three pairwise unconnected loops of length 1, and two pairwise connected loops of length 2, one for each non-mitotic mode of reproduction. Loop $G_{\mathrm{vege}} = [1,1]$ represents the mitotic-vegetative life cycle, $G_{\mathrm{coop}} = [1,2,1]$ represents the cooperative life cycle and $G_{\mathrm{cann}} = [1,3,1]$ represents the cannibalistic life cycle. There are the 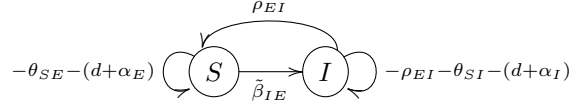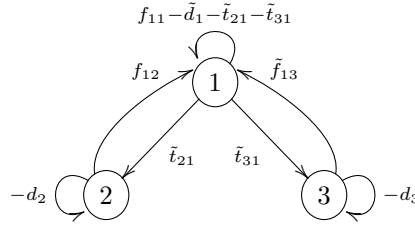following corresponding loop transmissions: $L_{\mathrm{vege}} = f_{11} - \tilde{d}_1 - \tilde{t}_{21} - \tilde{t}_{31}$, $L_{\mathrm{coop}} = \tilde{t}_{21}f_{12}$, and $L_{\mathrm{cann}} = \tilde{t}_{31}f_{13}$. Futhermore, there are two additional self loop transmissions: $-d_2$ and $-d_3$.

Irreducible and reducible life-cycles: Although the term *life cycle* suggests that an individual may pass through all stages in a cyclic manner, this is not necessarily the case: First, it is possible that an individual may choose among several different paths. Second, there can be stages from which an individual cannot reach certain other stages through transition or reproduction. The first possibility corresponds to different connected paths or loops. The second possibility refers to the reducibility of a life cycle.

Definition 4. In a graph, state i is *accessible* from state j, in short $j \to i$, if there exists a path from j to i. Two states i and j are *connected* if $j \to i$ and $j \to i$. A graph is *irreducible* if all of its states are pairwise connected, and *reducible* otherwise.

15

We say that a life cycle is reducible if this holds for its corresponding life-cycle graph. Likewise, the life-cycle matrix A is reducible, if its life-cycle graph is reducible, which is equivalent to the possibility to rearrange the rows and columns in A in such a way that

$$\mathsf{A} = \begin{pmatrix} \mathsf{B} & \varnothing \\ \mathsf{C} & \mathsf{D} \end{pmatrix},$$

as can be shown using standard Markov-chain theory (Feller, 1950).

Note that, if A is irreducible and if the population is at population dynamic equilibrium ($\hat{\boldsymbol{n}} > 0$), it follows from Proposition 1 and the fact that loop transmissions are only defined by non-zero factors, that

$$L \begin{cases} > 0 & \text{if } \ell(L) > 1 \\ < 0 & \text{if } \ell(L) = 1 \end{cases}. \tag{3}$$

Example 7. Consider the following life-cycle graphs G_1 and G_2:
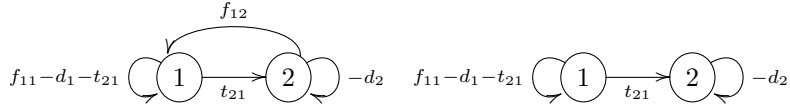
Clearly, the graph on the left hand side, G_1, is irreducible, since we can go from state 1 to state 2, and back to state 1 again. The graph on the right hand side is reducible: this life cycle has a post-reproductive class, namely stage 2. Its life-cycle matrix reads

$$\mathsf{A}_2 = \begin{pmatrix} f_{11} - t_{21} - d_1 & 0 \\ t_{21} & -d_2 \end{pmatrix}.$$

The purpose of this section was to collect all concepts which are necessary to describe the ecological dynamics of a structured population. The family of structured population models, as outlined above, resorts to two different ingredients: First, demographic parameters, and second, regulatory functions. The former fully describe the life history of a given population, the latter account for density dependence. Finally we have seen how to identify systems of ODEs, which model the dynamics of a population, with life-cycle graphs. We will come back to structured population models in section 4. Before that, the framework of Adaptive Dynamics shall be introduced, which will serve as foundation for the evolutionary analysis of life histories.

3   Adaptive Dynamics

Adaptive Dynamics is a mathematical framework, which has been developed to investigate long-term evolutionary dynamics of traits (Diekmann, 2004). It comprises elements of Evolutionary Game Theory and of population dynamics (Nowak and Sigmund, 1990; Nowak, 2006), thereby linking ecology and evolution in a unifying framework. More precisely, Adaptive Dynamics investigates the evolution of *phenotypic traits*, thereby ignoring all underlying genetic details. Sometimes, traits are also referred to as *strategies*. Furthermore, only phenotypic traits that can be identified with numerical characters are considered.

In Adaptive Dynamics, evolution is driven by two mechanisms, mutation and selection, while populations are assumed to be sufficiently large so that drift can be neglected. Mutations are random events, which modify phenotypic traits. Concerning mutations, Adaptive Dynamics relies on two basic assumptions (Dercole and Rinaldi, 2008, p. 65), namely that

1. mutations are rare, and

2. mutational effects are small.

If mutations are rare, on the ecological time scale populations have their time to settle down to population dynamical attractors, before a new mutant is going to appear. Hence, we can separate the evolutionary time scale from the ecological (demographic) time scale. The second assumption ensures that successful invasion of a rare mutant implies substitution of the resident by the mutant (Geritz et al., 2002). Furthermore, only under this assumption the evolution of a numerically measured phenotypic trait can be described as the trajectory of an ODE, namely, the so-called *canonical equation* of Adaptive Dynamics (Dieckmann and Law, 1996; Diekmann, 2004; Dercole and Rinaldi, 2008), which is a differential equation, modeling the change of a numerical trait by mutation and selection.

The concept of the *changing* (or *adaptive*) *fitness landscape* is cruical to Adaptive Dynamics: the fitness landscape a rare mutant faces is determined by the resident population, and − more importantly − generally differs for distinct resident populations. Each resident population shapes an environment (e.g. by consuming ressources therein). If now a mutant population occurs, it will be exposed to the environment set by the resident population. After successful invasion the mutant population, which by definition differs from the resident, becomes the new resident population and shapes a new environment, which need not be the same as before. This cyclic process is referred to as the *eco-evolutionary feedback loop* (Metz et al., 2008).

This section will present a short outline of Adaptive Dynamics. In section 3.1, the central concept of *invasion fitness* as well as a useful graphical method to study evolutionary dynamics for simple trait spaces will be introduced. Section 3.2 introduces *evolutionarily singular strategies*, the equilibria of the evolutionary dynamics. By identifying certain properties of evolutionarily singular strategies, one can immediately predict long-term evolutionary outcomes. Subsequently, section 3.3 shows under which conditions evolution can be perceived as process which optimizes traits. Finally, in section 3.4 we will see how evolutionary predictions can be made by solely inspecting the shape of trade-offs among various evolving traits. Later on, in section 4 the framework of Adaptive Dynamics will be applied to the structured population models which have been introduced previously.

Box 1. Eco-evolutionary assumptions

In the following list, all eco-evolutionary assumptions concerning the model family considered in this thesis shall be summarized:

General assumptions:

□ Populations are sufficiently large, so that ecological dynamics can be described by deterministic models.

□ The waiting time between two demographic events is exponentially distributed, so that ecological dynamics can be described by ODEs.

□ Reproduction is perceived as clonal.

Assumptions concerning population dynamics:

□ Populations can be decomposed into a finite number of pairwise disjoint states. That is, at each point in time, each individual belongs to exactly one state.

□ Realized demographic parameters can be regulated by population density. The resulting systems of ODEs is non-linear.

□ We restrict ourselves to modes of population regulation such that the population dynamics converge to a singe point attractor.

Assumptions concerning evolutionary dynamics:

□ The genes determining the demographic parameters are subject to mutations. Demographic parameters are numerically measured phenotypic traits. All underlying genetic details can be ignored.

□ Mutations are sufficiently rare, so that populations settle down on their single point attractor before the next mutant comes along.

□ Mutations are sufficiently small, so that invasion implies substitution.

3.1 The trait-substitution sequence

Let us start with an introduction of the basic concepts.

Invasion fitness: A population can be characterized by several traits, $x_1, ..., x_n$, which are collected in the so-called *trait vector* $\boldsymbol{x} = (x_1, ..., x_n)$, which lies in the *trait space* $\mathbb{X} \subseteq \mathbb{R}^n$. The resident population given by $\boldsymbol{x}$ fully determines the *environment* $\hat{\boldsymbol{E}}(\boldsymbol{x}) \in \mathbb{E}$, a new mutant population with trait vector $\boldsymbol{y}$ is exposed to, where $\mathbb{E} \subseteq \mathbb{R}^m$ is the set of all possible environments (Metz et al., 2008). The population density of the resident population determined by its population dynamic equilibrium may influence certain parameters which in turn influence the growth of the mutant population. No other factors are assumed to act upon the environment. We can now define the invasion fitness of a mutant (Metz et al., 1992, 2008; Metz, 2008):

Definition 5. The *invasion fitness* $s(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$ is the long-term exponential growth rate of a rare mutant population with strategy $\boldsymbol{y}$ in an environment $\hat{\boldsymbol{E}}(\boldsymbol{x})$ set by a resident population with strategy $\boldsymbol{x}$.

Invasion fitness is a function $s : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$, $(\boldsymbol{x}, \boldsymbol{y}) \mapsto s(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$. For notational simplicity, with some abuse of notation, we will also write $s(\boldsymbol{y}, \boldsymbol{x})$ instead of $s(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$, whenever the environment is not the primary focus of considerations.

The initial dynamics of the mutant population now are fully determined by its invasion fitness: if $s(\boldsymbol{y}, \boldsymbol{x}) < 0$ the mutant population will disappear, if $s(\boldsymbol{y}, \boldsymbol{x}) > 0$ it has a positive probability to to invade. In the case of equality, i.e., if $s(\boldsymbol{y}, \boldsymbol{x}) = 0$, the mutant population is selectively neutral with respect to the resident. Clearly, the resident is selectively neutral with respect to itself: by assumption, the resident population characterized by strategy $\boldsymbol{x}$ has reached its population dynamic equilibrium, hence $s(\boldsymbol{x}, \boldsymbol{x}) = s(\boldsymbol{x}, \hat{\boldsymbol{E}}(\boldsymbol{x})) = 0$.

Generally, we cannot make any further assertion about the future fate of an invading mutant population. There are two possible scenarios: First, the mutant could replace the resident population: under the assumption of small mutational effects, fixation of the mutant population is a consequence of its invasion. This is ensured by the so-called *Tube Theorem* (Geritz et al., 2002). Second, under certain circumstances the mutant and the resident may coexist. In the latter case, a subsequent mutant would face an environment determined by two resident populations. Its invasion fitness then is written as $s(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}_1, \boldsymbol{x}_2)) = s(\boldsymbol{y}, \boldsymbol{x}_1, \boldsymbol{x}_2)$.

The changing fitness landscape: If we assume that a mutant population is able to replace the resident population this leads to an evolutionary trajectory in the form of a trait-substitution sequence. Each given resident $\boldsymbol{x}$ defines a *fitness landscape* as a function of $\boldsymbol{y}$ by virtue of $s(\boldsymbol{y}, \boldsymbol{x})$. The graph of this function then is $\Phi_{\boldsymbol{x}} = \{((\boldsymbol{y}, s(\boldsymbol{y}, \boldsymbol{x})) : \boldsymbol{y} \in \mathbb{X}\}$. All strategies $\boldsymbol{y}$ that lie above the zero contour line of $\Phi_{\boldsymbol{x}}$ will have a positive probability to invade, and all strategies that lie below the zero contour line will go extinct. The fitness landscape changes with each new resident. We can think of the surface $\Phi = \{((\boldsymbol{x}, \boldsymbol{y}), s(\boldsymbol{y}, \boldsymbol{x})) : \boldsymbol{x}, \boldsymbol{y} \in \mathbb{X}\}$ as a *changing fitness landscape*.

The previously described surface can be nicely visualized if we restrict the trait vectors to be one-dimensional. Then invasion fitness is given by $s(y, x)$ and the changing fitness landscape reads $\Phi = \{((x, y), s(y, x)) : x, y \in \mathbb{X}\}$. Since $\mathbb{X} \subseteq \mathbb{R}$, Φ is a 2-dimensional surface over the set $\mathbb{X} \times \mathbb{X}$. For the investigation of the evolutionary
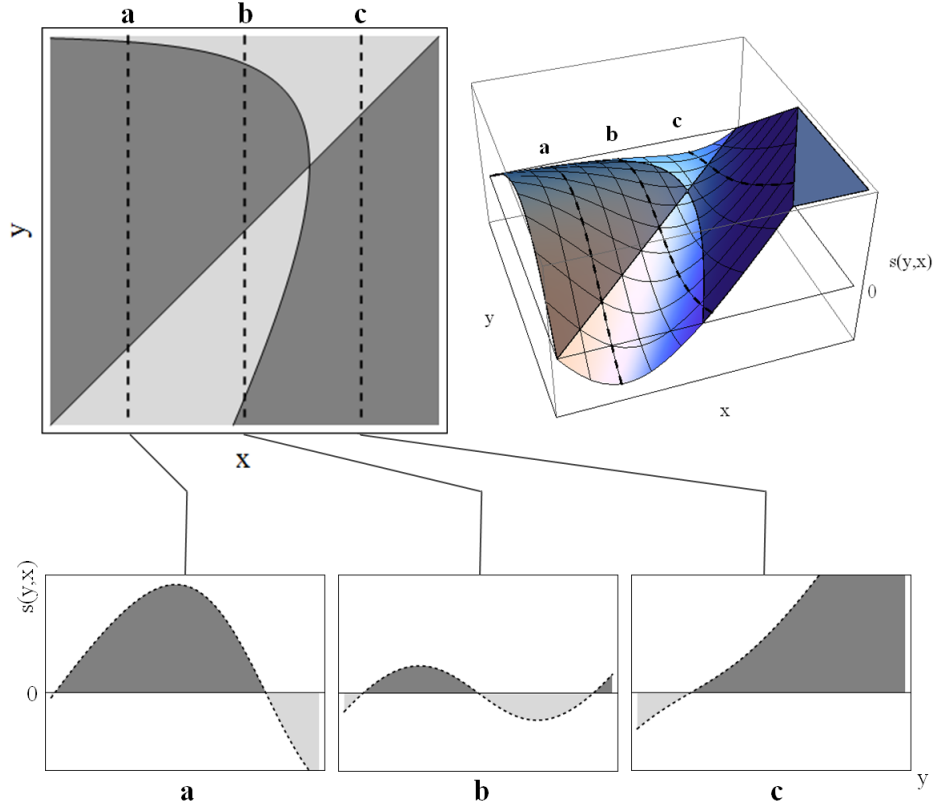
Figure 3: *The changing fitness landscape. The upper left figure shows a pairwise invasibility plot, which is the sign plot of $s(y, x)$, where x is the resident trait value and y is the mutant trait value. The plot in the upper right corner shows the plot of the invasion fitness function $s(y, x)$. Dark gray areas denote the positive region of the fitness landscape whereas light gray areas denote the negative region of the fitness landscape. For three different residents a, b, and c the respective fitness landscapes a rare mutant faces are shown in the bottom panel. The horizontal axes measure the trait value of the mutant y, the vertical axes measure invasion fitness. Each of the three fitness landscapes corresponds to a dashed straight line in the PIP on the left and a dashed curve in the changing fitness landscape on the right. As the trait value for the resident population changes from a to b to c, the shape of the fitness landscape, a rare mutant is exposed to, varies drastically. Note, that the horizontal axes measuring y in the lower panel are inverted.*

20

dynamics of the one-dimensional trait it is sufficient to know the sign of $s(y,x)$. Therefore, it is useful to look at the sign plot of $s(y,x)$, which is referred to as *pairwise invadibility plot* (PIP, Figure 4). In a PIP, the positive and the negative regions of Φ are separated by the zero contour lines, i.e., where $s(y,x) = 0$. Since by the definition of the invasion fitness we have $s(x,x) = 0$, the diagonal of the PIP is always such a zero contour line.

The entries x_i of the trait-substitution sequence $(x_i)_{i \in \{0,1,2,\dots\}}$ are derived by the following procedure:

1. Start with some strategy $x_i \in \mathbb{X}$ for $i = 0$ characterizing the initial resident population.

2. Choose a mutant strategy y_i in a small neighborhood of the resident strategy x_i.

3. If (x_i, y_i) does not lie in the positive region of the PIP, go back to step 2. Otherwise continue.

4. Define the new resident strategy as $y_i =: x_{i+1}$, and go to step 2.

3.2 Evolutionarily singular strategies

In the previous section we have seen that each resident with trait vector $\boldsymbol{x}$ defines a fitness landscape $\Phi_{\boldsymbol{x}}$. A mutant characterized by a trait vector $\boldsymbol{y}$ is able to invade, if the corresponding point on $\Phi_{\boldsymbol{x}}$ is positive (and is not able to invade in the negative case). The question to be asked now is: what if the fitness landscape $\Phi_{\boldsymbol{x}}$ lies entirely above or entirely below the zero contour line (apart from the point which corresponds to the resident itself)? Then, in the first case all mutants $\boldsymbol{y}$ can invade, and in the second case no mutant $\boldsymbol{y}$ can invade. Such points $\boldsymbol{x}$ are of special interest to our analysis. They are examples of so-called *evolutionarily singular strategies*. We will see in this section, that evolutionarily singular strategies can be classified with respect to two properties, viz. *invadibility* and *convergence stability*. In the remainder of this sub-section we will for simplicity restrict our considerations to one-dimensional trait spaces $\mathbb{X} \subseteq \mathbb{R}$. In this special case, both properties can be determined directly by inspecting the corresponding PIP locally around the evolutionarily singular strategy.

The direction of evolution: The direction of the evolution of a strategy x is given by the *local fitness gradient*

$$D(x) = \left. \frac{\partial s(y,x)}{\partial y} \right|_{y=x}.$$

It is easy to see that $D(x)$ determines, which mutants can invade: if $D(x) > 0$ nearby mutants with $y > x$ can invade. If $D(x) < 0$ nearby mutants with $y < x$ can invade (Geritz et al., 1998).

Evolution in the direction of the local fitness gradient continues as long as a so-called singular strategy is reached:

Definition 6. A strategy x^* is called an *evolutionarily singular strategy* if

$$D(x^*) = \left. \frac{\partial s(y,x^*)}{\partial y} \right|_{y=x^*} = 0.$$

21
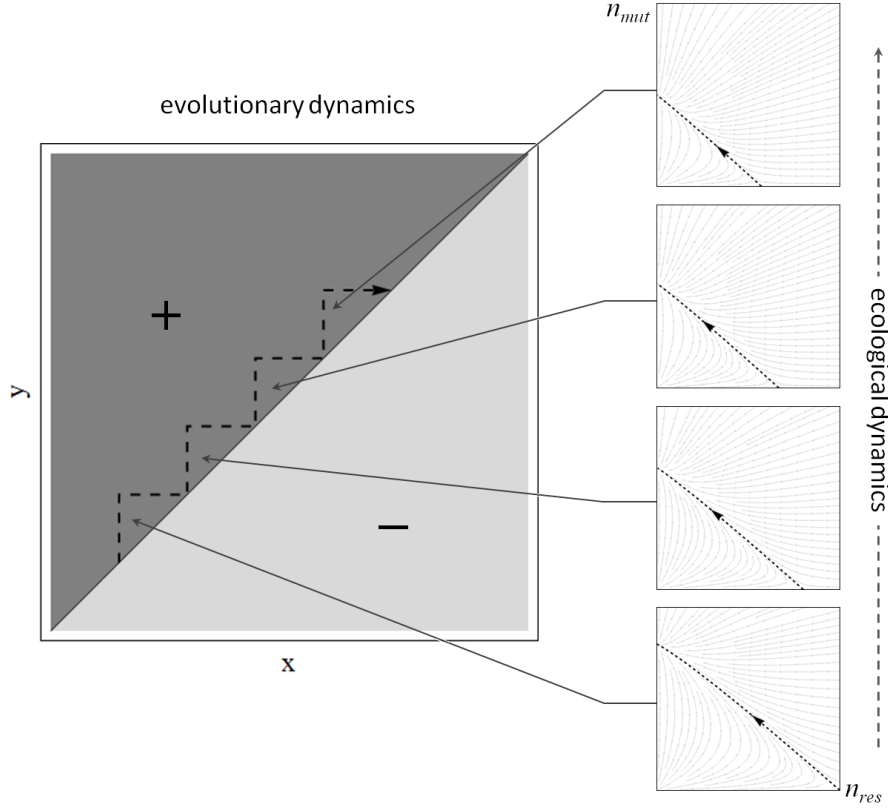
Figure 4: *Trait-substitution sequence. The figure on the left shows a stepwise trait-substitution sequence in a pairwise invadibility plot. On the horizontal axis, the trait x of the resident is measured while the vertical axis gives the trait value of the mutant y. The dark gray area denotes the positive region in the changing fitness landscape where $s(y,x) > 0$. The light gray area represents the negative region, i.e., where $s(y,x) < 0$. The pairs of strategies where $s(y,x) = 0$ are given by the solid black line. The trait-substitution sequence is illustrated by a step-shaped dashed arrow, which denotes the mutational steps on the evolutionary time scale. The figures on the right show the respective population dynamics on the ecological time scale as phase portraits. The vertical dashed arrow represents the evolutionary time scale. Note, that in the trait-substitution sequence shown in the figure, equilibrium population sizes are decreasing, as can be seen from the right panel.*

As long as the mutational step size is kept small and as long as the resident strategy x is not close to an evolutionarily stable x^*, invasion of a mutant characterized by strategy y will imply the substitution of y by x.

A singular strategies is a local extremum or saddle in the fitness landscape determined by itself. Here, evolutionary dynamics can show more interesting behavior than mere trait substitution events, as will be shown in the next paragraph.

Properties of singular strategies: In general, singular strategies are classified in terms of four different properties (Geritz et al., 1998; Diekmann, 2004). However, for our purpose the restriction to *invadibility* and *convergence stability* will suffice, since this allows for a division of evolutionarily singular points into four evolutionarily relevant and qualitatively well differing types (Rueffler et al., 2004). Both properties can be easily evaluated by inspecting PIPs.

If a singular strategy is uninvadible, then it cannot be invaded by nearby mutants: uninvadible singular strategies are maxima of the fitness landscape. Conversely, an invadible singular strategy lies at the bottom of a fitness landscape determined by the resident itself. Here, every mutant close to the resident will show a higher invasion fitness. The exact definition of invadibility reads as follows (see for instance Geritz et al., 1998; Otto and Day, 2007, ch. 12):

Definition 7. An evolutionarily singular strategy x^* is called *uninvadible* (or an *evolutionarily stable strategy*, short *ESS*) if it is a local maximum of the fitness landscape determined by itself, i.e., if

$$\left.\frac{\partial^2 s(y, x^*)}{\partial y^2}\right|_{y=x^*} < 0,$$

and *invadible* if it is a local minimum of the fitness landscape determined by itself, i.e., if

$$\left.\frac{\partial^2 s(y, x^*)}{\partial y^2}\right|_{y=x^*} > 0.$$

Note, that the locality condition is of great importance: it may well be that due to a very large mutational step, the evolutionary dynamics could leap over a valley from a peak (i.e., an ESS) to a higher position on the fitness landscape (not necessarily another peak). However, since we generally assume that the mutation step size is kept small in order to make fixation of the mutant possible, such cases are ignored for the definition of invadibility and noninvadibility.

In a PIP, the invadibility or noninvadibility can be determined by looking at the vertical line through the evolutionarily singular strategy x^*. If locally around the singular strategy, the vertical line completely lies in a positive region of the PIP, then x^* is invadible. On the contrary, if locally around x^*, the vertical line completely lies in a negative region of the PIP, then x^* is uninvadible, i.e., an ESS. In this case, the fitness landscape given by the resident x^* is decreasing around x^*.

Neutral stability captures the nongeneric case where the fitness landscape is flat around the singular point (but see Meszéna et al. 2001 for a discussion of the relation between neutrally stable strategies and ESSs in evolutionary matrix games). In a PIP

then the vertical line through x^* coincides with the nondiagonal boundary given by $s(y,x) = 0$ in some neighborhood around x^*.

The second property to be considered here is convergence stability. Roughly speaking, a singular strategy x^* is convergence stable if in some neighborhood, x is driven towards x^* (Otto and Day, 2007, p. 468). That is, the strategies in the trait-substitution sequence are approaching the convergence stable strategy: x^* is an evolutionary attractor.

In order to bring the concept of convergence stability in a more precise shape, we continue along the derivation in (Geritz et al., 1998, p. 39) by observing the local fitness gradient $D(x)$. Since evolution follows the direction of the local fitness gradient, $D(x)$ should be positive for strategies x which are less than x^* and negative for strategies x which are greater than x^* in some neighborhood of the singular strategy. We know that by definition $D(x^*) = 0$. By continuity it follows that locally around the singular strategy, the fitness gradient is a decreasing function of x, i.e.,

$$
\begin{aligned}
0 > \frac{\mathrm{d}D(x)}{\mathrm{d}x}\bigg|_{x=x^*} &= \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\partial s(y,x)}{\partial y}\bigg|_{y=x}\right]\bigg|_{x=x^*} \\
&= \left[\frac{\partial^2 s(y,x)}{\partial x \partial y}\bigg|_{y=x} + \frac{\partial^2 s(y,x)}{\partial y^2}\bigg|_{y=x}\right]\bigg|_{x=x^*} = \diamond
\end{aligned}
$$

Since constantly $s(y,x) = 0$ if $x = y$, the second-order directional derivative of $s(y,x)$ in the direction of $(x,x)^T$ (or more precisely: $\boldsymbol{v} = \frac{1}{\|(x,x)\|}(x,x)^T$) must vanish, i.e., $0 = \nabla_{\boldsymbol{v}}^2 s(y,x) = \frac{\partial^2 s(y,x)}{\partial x^2} + 2\frac{\partial^2 s(y,x)}{\partial x \partial y} + \frac{\partial^2 s(y,x)}{\partial y^2}$. Hence, we proceed with

$$
\begin{aligned}
\diamond &= \left[\frac{1}{2}\cdot\left(-\frac{\partial^2 s(y,x)}{\partial x^2}\bigg|_{y=x} - \frac{\partial^2 s(y,x)}{\partial y^2}\bigg|_{y=x}\right) + \frac{\partial^2 s(y,x)}{\partial y^2}\bigg|_{y=x}\right]\bigg|_{x=x^*} \\
&= \frac{1}{2}\cdot\left[-\frac{\partial^2 s(y,x)}{\partial x^2}\bigg|_{y=x} + \frac{\partial^2 s(y,x)}{\partial y^2}\bigg|_{y=x}\right]\bigg|_{x=x^*}
\end{aligned}
$$

from which we obtain the condition

$$
\frac{\partial^2 s(y,x)}{\partial x^2}\bigg|_{y=x=x^*} > \frac{\partial^2 s(y,x)}{\partial y^2}\bigg|_{y=x=x^*}. \tag{4}
$$

This motivates the following definition:

Definition 8. An evolutionarily singular strategy x^* is *convergence stable* if in a neighborhood the fitness gradient points towards x^*, which is the case if condition (4) holds.

The right hand side in (4) corresponds to the curvature of the changing fitness landscape Φ defined by $s(x,y)$ along the vertical line through x^*, whereas the left hand side corresponds to the curvature of Φ along the horizontal line though the singular strategy. With respect to the two curvatures, convergence stability is a relative matter.

More conveniently, one can simply look at the PIP and examine whether the region at the left of x^* over the 45° diagonal is positive and at the right of x^* is negative.

This is a direct translation of the condition that the fitness gradient has to point towards x^*: mutants with a trait value which is closer to the singular strategy are able to invade if the mutation step size is sufficiently small.

Note that the right part in condition (4) is the same as the left part in Definition 7. Thus, convergence stability and invadibility are related but nevertheless mutually independent properties, since the left part in condition (4) can vary independently of its right part.

Classification of singular strategies: Invadibility and convergence stability being two mutually independent binary-valued properties, one can distinguish between four different types of evolutionarily singular strategies (Geritz et al., 1998; Diekmann, 2004; Rueffler et al., 2004):

Definition 9. An evolutionarily singular strategy x^* is a *continuously stable strategy* (CSS) if it is uninvadible (an ESS) and convergence stable, a *branching point* if it is invadible (not an ESS) and convergence stable, a *Garden of Eden* point if it is uninvadible (an ESS) and not convergence stable, and an *evolutionary repellor* if it is invadible (not an ESS) and not convergence stable.

Continuously stable strategies are attractors and steady states of the evolutionary dynamics: the trait x evolves towards x^* and remains there.

Branching points are attractive as well in the sense that selection drives the trait towards x^*. However, once the dynamics are sufficiently close to the branching point, selection becomes disruptive and the once monomorphic population turns into a dimorphic population (Geritz et al., 1998; Rueffler et al., 2004). The rational behind this is the following: if x^* is a branching point we know that it is invadible as well as convergence stable. By Definition 7 and Definition 8 then evaluated at the singular strategy $\boldsymbol{x}^*$ it follows that

$$\frac{\partial^2 s(y,x)}{\partial y^2} > 0$$

and

$$\frac{\partial^2 s(y,x)}{\partial x^2} > \frac{\partial^2 s(y,x)}{\partial y^2}$$

hence again using $\nabla_{\boldsymbol{v}}^2 s(y,x) = 0$ as above we also have

$$\frac{\partial^2 s(y,x)}{\partial x^2} > -\frac{\partial^2 s(y,x)}{\partial y^2}$$

$$\Leftrightarrow \quad 2\cdot\left(\frac{\partial^2 s(y,x)}{\partial x^2} + \frac{\partial^2 s(y,x)}{\partial y^2}\right) > 0$$

$$\Leftrightarrow \quad 2\cdot\left(\frac{\partial^2 s(y,x)}{\partial x^2} + \frac{\partial^2 s(y,x)}{\partial y^2}\right)$$
$$\underbrace{-\left(\frac{\partial^2 s(y,x)}{\partial x^2} + 2\frac{\partial^2 s(y,x)}{\partial x \partial y} + \frac{\partial^2 s(y,x)}{\partial y^2}\right)}_{=\nabla_{\boldsymbol{v}} s(y,x)=0} > 0$$

$$\Leftrightarrow \quad \frac{\partial^2 s(y,x)}{\partial x^2} - 2\frac{\partial^2 s(y,x)}{\partial x \partial y} + \frac{\partial^2 s(y,x)}{\partial y^2} > 0$$

$$\Leftrightarrow \quad \nabla_{\mathbf{w}}^2 s(y,x) > 0,$$

where $\nabla_{\boldsymbol{w}}^2 s(y, x)$ is the second-order directional derivative of the fitness gradient in the direction of $\boldsymbol{w} = \frac{1}{\|(x, -x)\|}(x, -x)^T$, i.e., the -45° diagonal in the PIP. Along this line, the changing fitness landscape is convex.

Since by definition $s(x^*, x^*) = 0$, then the -45° diagonal lies in a positive region locally around x^*. But then there are pairs of nearby strategies x and y such that $s(y, x) > 0$ and $s(x, y) > 0$: if the axes in the PIP are swapped, the -45° diagonal will remain in the positive region of the changing fitness landscape. Hence, $\boldsymbol{x}$ and $\boldsymbol{y}$ can mutually invade each other whereby a dimorphic population will be established (Geritz et al., 1998, 40). This is the only configuration which allows to account for evolutionary branching.

A singular strategy which is referred to as a *Garden of Eden* point (Nowak and Sigmund, 1990; Nowak, 1990) is evolutionarily stable in that it cannot be invaded by rare mutants. However, it cannot be reached via small mutational steps. Rather it is only possible to reach a Garden of Eden point by precisely entering it. Any perturbation will drive the evolutionary dynamics away from the Garden of Eden point.

An *evolutionary repellor* simply is a strategy which is neither convergence stable nor immune to invasion. If the resident strategy happens to be an evolutionary repellor it can be successfully invaded by any nearby mutant and selection will drive the evolution of the trait directionally away from it.

Figure 5 shows examples of PIPs for each of the four scenarios.

3.3   Optimization and frequency dependence

The four evolutionary scenarios illustrated in Figure 5 are not possible for all models. Evolution could simply optimize a trait value. Under the assumptions of Adaptive Dynamics it is not a priori clear whether evolution optimizes a trait value or not. It seems intuitive that under optimization, evolutionary behavior is somewhat unspectacular. If a trait value is optimized it corresponds to a CSS (Metz et al., 2008). In particular, there can be no evolutionary branching. Models which support optimization are lacking what is referred to as *frequency dependent selection* (Heino et al., 1998), which roughly means that fitness depends on the frequencies of different types (a concept which is well known from Evolutionary Game Theory, see Nowak, 2006).

This section is meant to summarize the relevant notions and results about optimization. The first and the second paragraph give alternative necessary and sufficient conditions for optimization. We will see that both characterizations have their own raison d'être. The last paragraph then gives a definition of frequency dependence and links it to the previous results.

Optimization principles and monodimensional behavior: We start with the definition of optimization and pessimization principles:

Definition 10. An *optimization principle* or *optimization criterion* is a function $\xi : \mathbb{X} \to \mathbb{R}$ with the property that evolution maximizes ξ under any constraint on $\mathbb{X}$. A *pessimization principle, pessimization criterion* or *Verelendungs principle* is a function $\eta : \mathbb{E} \to \mathbb{R}$ with the property that evolution minimizes $\eta(\hat{\boldsymbol{E}}(\boldsymbol{x}))$, $\boldsymbol{x} \in \mathbb{X}$, under any constraint on $\mathbb{X}$.

This means that evolution drives the trait vector to a (local) maximum of ξ in $\mathbb{X}$ (or similarly to a local minimum of η in $\mathbb{E}$). In this way, an optimization principle
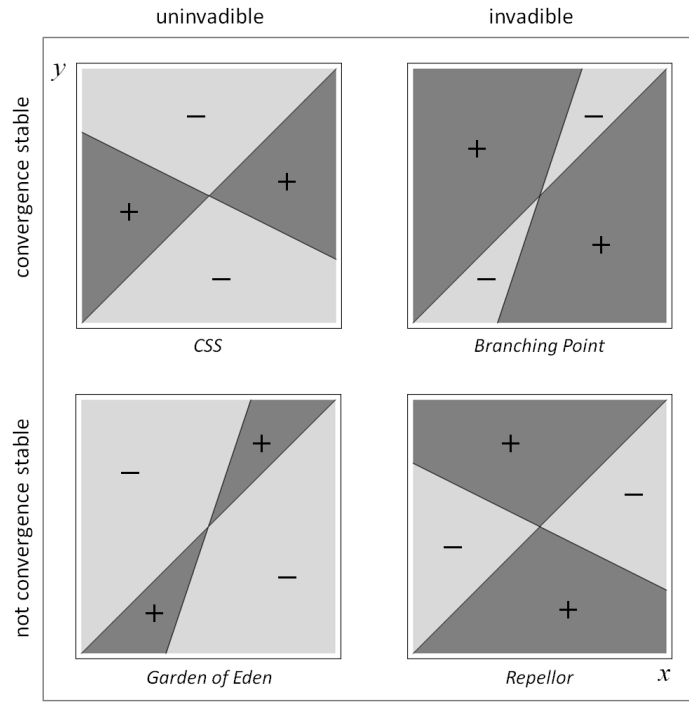
Figure 5: *Evolutionarily singular strategies. The PIPs corresponding to the four qualitatively different types of singular strategies are shown with respect to their convergence stability and invadibility properties. The dark gray area defines the positive region of the changing fitness landscape, while the light gray area corresponds to its negative region.*

(pessimization principle) qualifies as a Lyapunov function for the adaptive dynamics[3] in $\mathbb{X}$: by maximizing ξ (minimizing η) we can find the end point of the evolutionary dynamics in the trait space (Mylius and Metz, in press). Hence, having an optimization principle at hand for a given model, the evolutionary analysis turns out to be fairly straight forward.

Note that the phrase *under any constraint on* $\mathbb{X}$ in the definition above stresses that the optimization (pessimization) principle should not depend on how the trait space is restricted, i.e., constrained, for a given model (Mylius and Metz, in press, p. 635).

For notational simplicity we will write $\hat{\boldsymbol{E}} = \hat{\boldsymbol{E}}(\boldsymbol{x}) \in \mathbb{E}$ for the vector that denotes the environment (determined by some resident trait vector $\boldsymbol{x} \in \mathbb{X}$).

Definition 11. The *trait acts in a monotone and monodimensional manner* if there is a function $\xi : \mathbb{X} \to \mathbb{R}$ and a function $\alpha : \mathbb{R} \times \mathbb{E} \to \mathbb{R}$ increasing in its first argument such that
$$\operatorname{sign} s(\boldsymbol{y}, \hat{\boldsymbol{E}}) = \operatorname{sign} \alpha(\xi(\boldsymbol{y}), \hat{\boldsymbol{E}}).$$

The *environment acts in a monotone and monodimensional manner* if there is a function $\eta : \mathbb{E} \to \mathbb{R}$ and a function $\beta : \mathbb{X} \times \mathbb{R} \to \mathbb{R}$ increasing in its second argument such that
$$\operatorname{sign} s(\boldsymbol{y}, \hat{\boldsymbol{E}}) = \operatorname{sign} \beta(\boldsymbol{y}, \eta(\hat{\boldsymbol{E}})).$$

The diagram below (taken from Mylius and Metz, in press, with slight adaptations) visualizes the latter two concepts. Note that according to the previous definition this diagram commutes as long as the trait and the environment, respectively, act in a monotone and monodimensional manner:

The next proposition directly links optimization (pessimization) principles and monodimensional behavior of traits (environments), thereby providing a necessary and sufficient condition for the existence of an optimization (pessimization) principle:

Proposition 2. *A model has an optimization principle iff the trait acts in a monotone monodimensional manner. A model has a pessimization principle iff the environment acts in a monotone monodimensional manner.*

Proof. See Proposition 3.1 and Proposition 3.2 and the corresponding proofs in the appendix of Metz et al. (2008). □

The search for an optimization (pessimization) principle then boils down to finding functions ξ and α (or η and β, respectively) as specified in Definition 11. However, proving that there is *no* optimization principle for a given model turns out to be less convenient: one has to show that for all possible choices of candidate functions ξ and α

---

[3]Or more precisely: for the corresponding canonical equation of Adaptive Dynamics.

the above given sign equality does *not* hold. We will see in the subsequent paragraph that there is an alternative necessary and sufficient condition for the existence of an optimization principle, which is particularly useful for one-dimensional trait spaces.

The following proposition allows us to restrict ourselves to investigating optimization principles or pessimization principles only. We will see that this proposition is particularly useful, since for some models it is quite easy to find optimization principles but quite difficult to directly derive pessimization principles.

Proposition 3. *A model has an optimization principle iff it has a pessimization principle.*

Proof. For a given pessimization principle η set $\xi(\boldsymbol{x}) := -\eta(\hat{\boldsymbol{E}}(\boldsymbol{x}))$, and vice versa. See Proposition 3.3 in Metz et al. (2008). $\square$

Let us consider the case of one-dimensional trait spaces: for two trait values x and y, *being selectively neutral with respect to each other* defines an equivalence relationship on the trait space $\mathbb{X}$ (Mylius and Metz, in press). If a model supports an optimization principle, say ξ, by definition we have that a mutant with trait y can invade a resident with trait x, as long as $\xi(y) > \xi(x)$. This can be used to order all equivalence classes in $\mathbb{X}$ (with respect to selective neutrality). Then, in the corresponding PIP invasion boundaries must be symmetric with respect to the 45° diagonal. Furthermore, we have that $s(x,y) > 0 \Leftrightarrow s(y,x) < 0$. Thus, if a model supports an optimization principle, the corresponding PIP (or more sloppily put: the model) is skew-symmetric.

In general, we have that anti-symmetry, i.e., $s(\boldsymbol{x},\boldsymbol{y}) > 0 \Leftrightarrow s(\boldsymbol{y},\boldsymbol{x}) < 0$, is *necessary* for the existence of an optimization principle (but not sufficient, see Metz et al., 2008).

Optimization and rock-paper-scissors cycles: As we have seen, Proposition 2 is partially unpractical in the sense that it is difficult to show that optimization is *not* possible. Gyllenberg and Service (2011) provide another proposition which, assuming anti-symmetry, gives necessary and more importantly quite handy sufficient conditions for the nonexistence of an optimization criterion:

Proposition 4. *Let for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{X}$ hold that $s(\boldsymbol{x},\boldsymbol{y}) > 0 \Leftrightarrow s(\boldsymbol{y},\boldsymbol{x}) < 0$. Then there exists an optimization principle iff there are no weak rock-paper-scissors cycles (RPS), i.e., no ordered triple $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) \in \mathbb{X}^3$ with $s(\boldsymbol{x},\boldsymbol{y}) > 0$, $s(\boldsymbol{y},\boldsymbol{z}) > 0$, and $s(\boldsymbol{z},\boldsymbol{x}) > 0$ and $\geq$ for any two of the previous relations.*

Proof. See Theorem 4.3 in Gyllenberg and Service (2011). $\square$

In the previous paragraph we have seen that anti-symmetry is necessary for the existence of an optimization principle. Proposition 4 now shows, which models, although fulfilling anti-symmetry, actually support an optimization principle: namely those, which have no weak rock-paper-scissors cycle. Conversely, if we have a skew-symmetric model, we only need to find at least one such weak rock-paper-scissors cycle with trait values in $\mathbb{X}$ to prove that there exists no optimization principle.

The following proposition is a consequence of the previous one.

Proposition 5. *There exists an optimization principle iff the family $V = \{V_{\boldsymbol{x}} : \boldsymbol{x} \in \mathbb{X}\}$ where $V_{\boldsymbol{x}} = \{\boldsymbol{y} \in \mathbb{X} : s(\boldsymbol{x},\boldsymbol{y}) \geq 0\}$ is totally ordered by inclusion, i.e., for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{X}$ either $V_{\boldsymbol{x}_1} \subset V_{\boldsymbol{x}_2}$ or $V_{\boldsymbol{x}_2} \subset V_{\boldsymbol{x}_1}$.*

Proof. See Theorem 6.2 in Gyllenberg and Service (2011). □

For models with a one-dimensional trait space, this proposition asserts, that for models supporting an optimization principle, in the corresponding PIP we can draw a rectangle (with edges parallel to the horizontal and vertical axes) such that two diagonally opposed vertices lie in a positive region and the two remaining vertices lie in a negative region of the PIP.

Let us pause at this point and recall that we are interested in the quality of evolutionarily singular points. As illustrated in Figure 5, evolutionarily sigular points can be characterized as intersecting points of the zero contour lines of the changing fitness landscape, and the angle between the zero contour lines determines the evolutionary quality of the singular point. Anti-symmetry, as defined above, clearly means that in an evolutionarily singular point, this angle is precisely 90°. Hence, we have either a CSS or an evolutionary repellor. Since optimization requires skew-symmetry, these are the only two possible outcomes for models supporting an optimization principle.

The following proposition states that for the local study of evolutionarily singular points we do not have to bother about rock-paper-scissors cycles.

Proposition 6. *Let $\mathbb{X}$ be one-dimensional. For each evolutionarily singular strategy x^* there is a restricted trait space which is an interval $\mathbb{S} \subseteq \mathbb{X}$ with $x^* \in \mathbb{S}$, in such a way that a model supports an optimization principle in $\mathbb{S}$ iff it is skew-symmetric in $\mathbb{S}$.*

Proof. It is clear from the consideration above that skew symmetry follows from the existence of an optimization principle. Let us consider the opposite direction:

Choose an interval $\mathbb{S} \subseteq \mathbb{X}$, with $x^* \in \mathbb{S}$, so small, such that the boundary defined by the zero contour line which is not the 45° diagonal is a decreasing concave function $f : \mathbb{S} \to \mathbb{S}$. The convex case is completely analogue. We show that there is no optimization principle for s in $\mathbb{S}$ by demonstrating that there is no rectangle in such a way that two diagonally opposed vertices lie in a positive region and the two remaining vertices lie in a negative region of the PIP defined by s.

Without loss of generality, let us assume that in the PIP the region between the 45° diagonal and f is positive. Choose two points $p_1 = (x^* - a, x^* - b)$ and $p_2 = (x^* + c, x^* - b)$ in $\mathbb{S} \times \mathbb{S}$ with $s(x^* - b, x^* - a) > 0$ and $s(x^* - b, x^* + c) < 0$, $a, b, c > 0$. We have that $a > b$. Clearly p_1 lies in a positive and p_2 negative region. The line between p_1 and p_2 is the edge at the bottom of the rectangle. Let us assume for contradiction that we can complete the rectangle in the above described manner, i.e., that we can find appropriate vertices p_3 and p_4 in a negative and positive region of $\mathbb{S} \times \mathbb{S}$, respectively.

We can define $\delta_1 := f^{-1}(x^* - b) - x^*$. Then $s(x^* - b, x^* + \delta_1) = 0$. Because of the shape of f and since p_2 lies in a negative region, we know that the point $q_1 = (x^* + \delta_1, x^* - b)$ lies on the right of p_2, or more precisely, $\delta_1 > c$. Furthermore, define $\delta_2 := f^{-1}(x^* - a) - f^{-1}(x^* - b)$. Since f is decreasing, $\delta_2 > 0$. By definition, $s(x^* - a, x^* + \delta_1 + \delta_2) = 0$. From symmetry then follows that also $s(x^* + \delta_1 + \delta_2, x^* - a) = 0$. Due to skew-symmetry, above the evolutionarily singular point we have that $s(x^* + \delta_1 + \delta_2 + \delta_3, x^* - a) < 0$ iff $\delta_3 > 0$. Hence, the third vertex of the rectangle is of the form $p_3 = (x^* - a, x^* + \delta_1 + \delta_2 + \delta_3)$. Consequently, the remaining vertex must be $p_4 = (x^* + c, x^* + \delta_1 + \delta_2 + \delta_3)$. But then, since $\delta_1 + \delta_2 + \delta_3 > c$ and $s(x^* + c, x^* + c) = 0$, it follows that necessarily $s(x^* + \delta_1 + \delta_2 + \delta_3, x^* + c) < 0$, a contradiction.

The existence of an optimization principle then is a consequence of Proposition 5. $\qquad\qquad\square$

Frequency dependence: There are different ways to define frequency dependent selection (Heino et al., 1998). In its stricter sense, frequency dependent selection is defined as a process leading to ecologically stable polymorphisms, i.e., all evolutionary scenarios where we have $s(y,x) > 0$ and $s(x,y) > 0$ in the neighborhood of an evolutionarily singular strategy.

In a wider sense, frequency dependent selection occurs if "different individuals [...] have a different influence on the environment and, moreover, a different perception of the environment" (Heino et al., 1998, p. 370) by virtue of density dependence. That is, frequency dependence is related with multidimensionally behaving environments. Then it is natural to define frequency dependent selection via optimization:

Definition 12. A model is called *frequency dependent* in a connected subset $\mathbb{S} \subseteq \mathbb{X}$ if it does not support an optimization principle in $\mathbb{S}$.

For the one-dimensional case, Proposition 6 then entails that around an evolutionarily singular point, a model is frequency dependent, if and only if the corresponding PIP is not skew symmetric. Note in particular, that even in a small neighborhood of an evolutionarily singular point an angle of 90° between the two zero contour lines, or equivalently,

$$\frac{\partial^2 s(y,x)}{\partial x^2} = \frac{\partial^2 s(y,x)}{\partial y^2},$$

does not necessarily imply skew symmetry (see section 3.2), even if a restricted trait space $\mathbb{S} \subseteq \mathbb{X}$ is considered. In other words: stating whether or not a model shows frequency dependent selection in a restricted trait space, requires more than local analysis of the intersecting zero contour lines by means of derivatives.

With this (not very surprising) insight and the following diagram, which summarizes the relevant relations for one-dimensional trait spaces, we conclude this section.

$$
\begin{array}{ccccc}
\frac{\partial^2 s}{\partial x^2} = \frac{\partial^2 s}{\partial y^2} & \Leftarrow & \neg\,\text{frequency dep.} & & \\
\Uparrow & & \Updownarrow & & \\
\text{skew symm.} & \Leftrightarrow & \exists\,\text{opt.}\,\xi(x) & \Leftrightarrow & x\,\text{acts 1-dim.} \\
\Updownarrow & & \Updownarrow & & \\
\nexists\,\text{weak RPS} & & \exists\,\text{pess.}\,\eta(E) & \Leftrightarrow & E\,\text{acts 1-dim.}
\end{array}
$$
$$\underbrace{}_{\text{in a sufficiently small } \mathbb{S} \ni x^*}$$
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (5)$

3.4 Trade-offs and invasion boundaries

It is reasonable to assume that the set of possible combinations of trait values is limited. Otherwise traits which are linked to fitness would be driven by selection to ever larger values (or lower, in the case of death rates), and thereby would result in a so-called *Darwinian demon* (Law, 1978). For instance, it can be expected that increasing fecundity will result in decreased survival chances at later ages (see Law, 1978; Stearns, 1989). This brings us to the notion of the *trade-off*. We assume that, ultimately, the increase in one relevant trait automatically results in a decrease in some other traits (or reversed in the case of death rates).

31

The second concept, which is subject to this section, is the so-called *invasion boundary*, which is the curve defined by all combinations of trait values that are selectively neutral with respect to a given resident type at its populations dynamical attractor. Comparing the trade-off curve with the invasion boundary will allow us to make predictions about possible evolutionary outcomes (Levins, 1962; Rueffler et al., 2004; Hoyle et al., 2008).

Central concepts: In this section, we restrict our considerations to two-dimensional trait spaces, that is, trait vectors in $\mathbb{X} = \mathbb{X}_1 \times \mathbb{X}_2$ are of the form $\boldsymbol{x} = (x_1, x_2)$, where $x_i \in \mathbb{X}_i \subseteq \mathbb{R}^+$. We assume that the set of possible trait values is limited by a trade-off curve. In general, trade-offs on a n-dimensional trait space are $n-1$-dimensional (and for biological reasons compact) manifolds, i.e., hyperplanes. For the two-dimensional case we define trade-offs as follows:

Definition 13. A *trade-off* T is a strictly decreasing function $T : \mathbb{X}_1 \to \mathbb{X}_2$, $x_1 \mapsto T(x_1)$.

Strictly speaking, if either x_1 or x_2 is a death rate, *decreasing* has to be replaced with *increasing*. In the following section, we will define evolving traits in such a way, that tedious differentiations concerning death rates, and fertility or transition rates can be avoided.

Note, that T is defined as a *function*. The trade-off curve separates the trait space into biologically possible and impossible combinations of traits, i.e., trait vectors (x_1, x_2) which fulfill $x_2 \leq T(x_1)$ and $x_2 > T(x_1)$, respectively. The area in the first orthant of $\mathbb{R}^2$ below the trade-off curve is referred to as *feasibility set* (Rueffler et al., 2004). In the following, we assume that over the course of evolution all trait vectors have evolved in such a way that they have finally reached the trade-off curve. That is, without loss of generality, we assume $x_2 = T(x_1)$ for all pairs (x_1, x_2).

Definition 14. An *invasion boundary* is a function $I : \mathbb{X}_1 \to \mathbb{X}_2$, $x_1 \mapsto I(x_1)$ which is implicitly defined by $s((x_1, I(x_1)), \hat{\boldsymbol{E}}) = 0$.

Again, we are making the simplifying assumption, that I can be locally described as a function. Furthermore, we assume that I is strictly decreasing (or increasing, if again one of the two traits is a death rate). We observe, that in contrast to the trade-off curve, the invasion boundary depends on the environment $\hat{\boldsymbol{E}}$, and is thereby changing with each new resident.

Similar to the trade-off curve, the invasion boundary determined by a resident $\boldsymbol{x} = (x_1, x_2)$ separates the trait space into two areas: mutants $\boldsymbol{y} = (y_1, y_2)$ below the invasion boundary, i.e., $y_2 < I(y_1)$, are not able to invade: since fitness is assumed to increase in both traits, we have $0 = s((y_1, I(y_1)), \hat{\boldsymbol{E}}(\boldsymbol{x})) > s((y_1, y_2), \hat{\boldsymbol{E}}(\boldsymbol{x}))$. Likewise, mutants above the invasion boundary, i.e., $y_2 > I(y_1)$ can invade. By definition trait vectors with $y_2 = I(y_1)$ are selectively neutral with respect to the resident. In other words, the set of trait vectors below the invasion boundary corresponds to the subset of $\mathbb{X}$ where the fitness landscape $\Phi_{\boldsymbol{x}}$ is below the zero contour line, and the set of trait vectors above the invasion boundary corresponds to the subset of $\mathbb{X}$ where the fitness landscape $\Phi_{\boldsymbol{x}}$ is above the zero contour line.

Predicting evolutionary outcomes: At an evolutionarily singular point $\boldsymbol{x}^*$ the invasion boundary (determined by $\boldsymbol{x}^*$) and the trade-off curve are tangent (Rueffler

et al., 2004). If the invasion boundary lies locally in the inside of the feasibility set (except for the point of tangency), nearby pairs of traits (x_1, x_2) on the trade-off curve can invade, since then (x_1, x_2) is in the region where Φ_{x^*} is above the zero contour line. If, on the contrary, the invasion boundary lies in the outside of the feasibility set, then all nearby pairs of traits cannot invade (x_1, x_2), as they are located in the region where Φ_{x^*} is below the zero contour line. This shall be summarized in the following proposition:

Proposition 7. *A singular point x^* is uninvadible if* $\mathrm{d}^2 T / \mathrm{d} x_1^2|_{x_1 = x_1^*} < \mathrm{d}^2 I / \mathrm{d} x_1^2|_{x_1 = x_1^*}$ *and invadible if* $\mathrm{d}^2 T / \mathrm{d} x_1^2|_{x_1 = x_1^*} > \mathrm{d}^2 I / \mathrm{d} x_1^2|_{x_1 = x_1^*}$.

Proof. See Appendix in Rueffler et al. (2004). □

For linear trade-offs, it is clear that as a consequence from the previous proposition, x^* is uninvadible if I is convex and invadible if I is concave. Due to this observation, we will see in section 4.3 that Proposition 7 proves to be very useful to make predictions about the evolutionary behavior of structured population models.

Let us recapitulate: We have seen, that if mutations are assumed to be rare and small, evolution driven by selection and mutation can be modeled as a trait-substitution sequence. Adaptive Dynamics provides tools for the analysis of the long-term behavior of this sequence, by means of a classification of different evolutionarily singular strategies. We find, that the case of evolutionary branching points is of special interest. Evolutionary branching points are those singular points which are convergence stable and invadible. In some cases, models do not support the whole spectrum of evolutionary outcomes: rather they follow an optimization process. Finally, we have seen, that invadibility can also be determined by comparing invasion boundaries and trade-off curves. Let us proceed and apply the framework of Adaptive Dynamics to models of structured populations.

4 Evolutionary dynamics of structured populations

Up to this point preparations have been made. We were presenting concepts and results, which are necessary for an eco-evolutionary analysis of structured populations. Now, let us reap the fruits of our labor and apply the tools of Adaptive Dynamics given in section 3 to the model family outlined in section 2.

Models of structured populations in continuous time have been evolutionarily analyzed particularly by Hoyle et al. (2008) and Hoyle and Bowers (2008), who elaborated necessary criteria (see Hoyle and Bowers, 2008, 312) on the shape of the life-cycle matrix A for different evolutionary outcomes:

1. at least two rows of A are evolving (criterion A),

2. at least two columns of A are evolving (criterion B), and

3. at least two entries of A or two components of one entry of A are regulated differently (criterion C).

We will go one step further and derive necessary conditions on the life cycle itself, more precisely on its loop structure. Along the lines of Rueffler et al. (2013), where the discrete-time case has been worked out extensively, we will persue the following strategy: First, we shall establish an algebraically simple way to measure invasion fitness, a so-called *fitness proxy*. Hoyle and Bowers (2008) used the *next generation matrix* (Diekmann and Heesterbeek, 2000) to account for reduction of algebraic complexity. Rueffler et al. (2013), however, suggest an even simpler method for discrete-time models based on the determinant of A, which will be translated into the continuous-time case in section 4.1.

This fitness proxy will be applied with the following two objectives: First, in 4.2 it will be used to give sufficient conditions for optimization to occur. More precisely, we will calculate a list of explicit optimization principles for certain types of life cycles.[4] Second, the fitness proxy will be employed to predict evolutionary outcomes in the case of frequency dependence, based on the loop structure of a given life cycle. To this end, the methods introduced in 3.4 will be made use of. Throughout the whole section, models with multiplicative as well as additive regulatory functions will be considered, to extend the range of applications of the derived results.

4.1 Fitness proxies for continuous-time structured-population models

The goal of this section is to find an efficient way to determine whether or not a mutant of a given structured population can invade. We will see that, in a way, invasion fitness can be reduced to an expression which is tightly connected with the loop structure of a given life cycle. This will then enable us to make predictions about the behavior of the evolutionary dynamics just based on properties of the corresponding life cycle graph.

Consider a life-cycle matrix A. The components of the trait vector $\boldsymbol{x} = (x_1, ..., x_k)$, which is an element of a k-dimensional trait space $\mathbb{X}$, are identified with demographic parameters in A in the following way: if fertility rates f_{ij} or transition rates t_{ij} are *evolving*, then f_{ij} and t_{ij} are monotonically increasing positive functions of a component x_l of $\boldsymbol{x}$, $l \in \{1, ..., k\}$, i.e., $f_{ij} : \mathbb{X} \to \mathbb{R}^+$, $x_l \mapsto f_{ij}(x_l)$, and $t_{ij} : \mathbb{X} \to \mathbb{R}^+$,

---

[4]A word of caution: we will not show, that the list provided in 4.2 is exhaustive.

34

$x_l \mapsto t_{ij}(x_l)$. Likewise, if a death rate d_i is *evolving*, then d_i is a monotonically decreasing positive function of a component x_l of $\boldsymbol{x}$, $l \in \{1, ..., k\}$, i.e., $d_i : \mathbb{X} \to \mathbb{R}^+$, $x_l \mapsto d_i(x_l)$. We assume that for each demographic parameter, there is exactly one component of $\boldsymbol{x}$ and vice versa. We will also say, that a component x_l of the trait vector is *acting on* a demographic parameter $c_{ij} \in \{f_{ij}, t_{ij}, d_i\}$, if c_{ij} is evolving, i.e., a function of x_l in the previously described manner. Evolving demographic parameters will be referred to as *traits*. For notational simplilcity, evolving demographic parameters c_{ij} will be directly identified with components x_l of $\boldsymbol{x}$, by setting $x_{ij} := c_{ij}(x_l)$. Dropping the argument, we then simply write $x_{ij} = c_{ij}$ for an evolving demographic parameter c_{ij}.[5]

Resorting to the Adaptive Dynamics approach as introduced in section 3, we have to calculate the invasion fitness for the mutant population characterized by $\boldsymbol{y} \in \mathbb{X}$ in the structured population model (2), i.e.,

$$\dot{\mathbf{n}}_{\boldsymbol{y}} = \mathsf{A}(\mathbf{n}_{\boldsymbol{y}}) \cdot \mathbf{n}_{\boldsymbol{y}}.$$

Invasion fitness is defined as the long-term exponential growth rate of a rare mutant in an environment determined by the resident. The non-zero entries of the life-cycle matrix may depend on the population-density vector $\boldsymbol{n}_{\boldsymbol{y}}$ via regulatory functions $R_{c,ij}$, where c can be any demographic parameter. Since by assumption the resident population is in a single point attractor $\hat{\boldsymbol{n}}_{\boldsymbol{x}}$ and since the mutant is assumed to be initially rare, the regulatory functions depend on the environment $\hat{\boldsymbol{E}}(\boldsymbol{x})$ as determined by the resident $\boldsymbol{x}$. Note, that the vector $\hat{\boldsymbol{E}}(\boldsymbol{x}) \in \mathbb{E} \subseteq \mathbb{R}^m$ is not the same as $\hat{\boldsymbol{n}}_{\boldsymbol{x}} \in \mathbb{R}^n$. The entries of $\hat{\boldsymbol{E}}(\boldsymbol{x})$ are (possibly constant) functions of the entries of $\hat{\boldsymbol{n}}_{\boldsymbol{x}}$ and $\boldsymbol{x}$. With notational generosity[6] we write $R_{c,ij}(\hat{\boldsymbol{E}}(\boldsymbol{x})) = R_{c,ij}(\hat{\boldsymbol{E}})$. Consequently, the life-cycle matrix of a rare mutant is characterized by evolving demographic parameters $\boldsymbol{y}$ and the environment $\hat{\boldsymbol{E}}(\boldsymbol{x}) = \hat{\boldsymbol{E}}$ by virtue of $R_{c,ij}(\hat{\boldsymbol{E}})$. This justifies the notation $\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$. Then, the system of ODEs describing the dynamics of a rare mutant subpopulation becomes

$$\dot{\boldsymbol{n}}_{\boldsymbol{y}} = \mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) \cdot \boldsymbol{n}_{\boldsymbol{y}}.$$

We know that in a constant environment the growth rate of a structured population model in continuous time is given by the stability modulus of the corresponding life-cycle matrix (Metz, 2008). By the Perron-Frobenius Theorem, Proposition B.3, we have that the stability modulus is real for life-cycle matrices. For the above given model, invasion fitness then is given as

$$s(\boldsymbol{y}, E_{attr}(\boldsymbol{x})) := \lambda_{\mathrm{d}}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) := \mu(\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))),$$

where $\lambda_{\mathrm{d}}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$ is the dominant eigenvalue, i.e., stability modulus, of $\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$. Subsequently, we will use the more economical notations $\lambda_{\mathrm{d}} = \lambda_{\mathrm{d}}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$ and $\mathbf{E} = \hat{\boldsymbol{E}}(\boldsymbol{x})$ whenever appropriate.

Unfortunately, explicitly calculating eigenvalues, and hence λ_{d}, is a complicated matter, if possible at all, and notably so for large life cycle matrices. In the remainder of this section we derive an algebraically more comfortable way to measure invasion fitness, by just demanding sign equivalence with respect to $s(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$. This is justified

---

[5]Note, that this detour is necessary, since death rates are assumed to be decreasing functions of a component of a trait vector. The notation $x_{ii} = c_{ii} = d_i$ should should be used with caution.

[6]In 2.1, regulatory functions were defined as functions of components of the population-density vector $\boldsymbol{n} \in \mathbb{R}^n$. Now, regulatory functions are seen as functions of the environment $\hat{\boldsymbol{E}} \in \mathbb{E} \subseteq \mathbb{R}^m$.
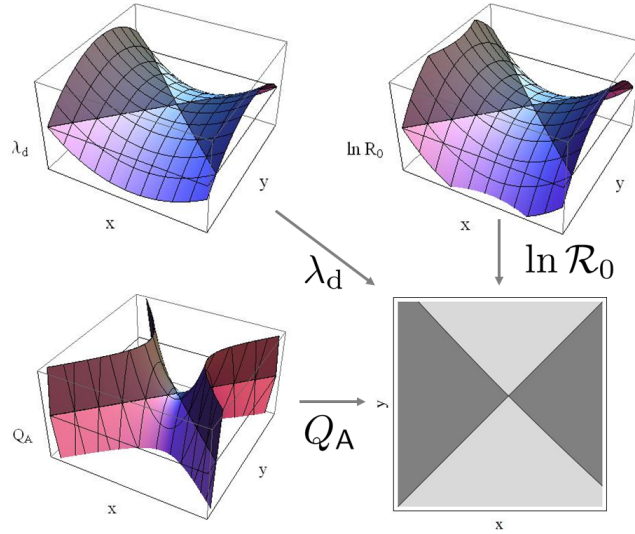
Figure 6: *Changing fitness landscapes defined by three sign equivalent measures of invasion fitness: the dominant real eigenvalue $\lambda_d(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$, the logarithm of the basic reproductive ratio $R_0(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$, and the fitness proxy $Q_A(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$. All three measures are leading to the same pairwise invadibility plot and thereby show to the same evolutionary dynamics.*

by the following argument: in Adaptive Dynamics primarily the dynamics on the evolutionary timescale are of particular interest. Due to the separation of time scales, we can ignore the actual population dynamic behavior of a mutant, as long as we know that it is successfully able to invade. This, however, solely depends on whether invasion fitness is positive or negative.

Definition 15. A *fitness proxy* is a function $p : \mathbb{X} \times \mathbb{E} \to \mathbb{R}$ for which

$$\operatorname{sign} p(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) = \operatorname{sign} s(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$$

holds for all $\boldsymbol{y} \in \mathbb{X}$ in a sufficiently small neighborhood of $\boldsymbol{x}$.

Example 8. The *basic reproductive ratio* R_0 denotes the expected number of offspring of an individual throughout its whole lifetime, as long as there is only one birth state (more broadly, the basic reproductive ratio can also be defined for several birth states, see Rueffler and Metz, 2013). Formally, R_0 is defined as the dominant eigenvalue of the next generation matrix corresponding to the given model (see Diekmann et al. (1990) and Diekmann and Heesterbeek (2000)).

In continuous-time structured-population models the next generation matrix reads $-\mathsf{F}\mathsf{T}^{-1}$, where for a given life-cycle matrix A, F comprises all fertility rates and T comprises all death and transition rates, such that $\mathsf{A} = \mathsf{F} + \mathsf{T}$. Hence, $R_0 = \rho(-\mathsf{F}\mathsf{T}^{-1})$. By the Perron-Frobenius Theorem (Proposition B.2 and Proposition B.3) one can show that $R_0 \gtreqqless 1$ iff $\lambda_d(\mathsf{A}) \gtreqqless 0$. Consequently, $\ln R_0(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$ and $R_0(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) - 1$ qualify as fitness proxies (see also Kon, 2007, for a comprehensive overview on R_0 in continuous-time structured-population models).

At first sight, calculating the dominant eigenvalue of the next generation matrix does not really seem much more beneficial than calculating the dominant eigenvalue

of the life-cycle matrix. If however F or T are sparse, this also holds for the next generation matrix, whereby the calculation of the dominant eigenvalue usually becomes simpler. More precisely, the number of birth states constitutes an upper bound of the degree of the characteristic polynomial of the next generation matrix. In particular, if there is only one birth state, the characteristic polynomial is linear, whereby R_0 can be easily calculated.

Analoguous to Rueffler et al. (2013) we derive a fitness proxy, which essentially reduces complexity to calculating a determinant rather than the dominant eigenvalue. From now on, we will assume that life-cycle matrices are *irreducible*.

Proposition 8. *For a given irreducible life-cycle matrix* $\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$ *let*

$$Q_\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) := -\det(-\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))).$$

Then $Q_\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$ *is a local fitness proxy.*

Proof. We know that the invasion fitness $s(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$ is given by the dominant eigenvalue $\lambda_\mathrm{d}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) = \lambda_\mathrm{d}$ of $\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$.

The characteristic polynomial of $\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$ is given by

$$\chi(\lambda, \boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) = \det(\lambda \mathsf{I} - \mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))),$$

where I is the identity matrix of appropriate dimension. We know that $\lambda_\mathrm{d} = 0$ for $\boldsymbol{y} = \boldsymbol{x}$. Clearly, $\chi(\lambda_\mathrm{d}, \boldsymbol{x}, \hat{\boldsymbol{E}}(\boldsymbol{x})) = 0$, as λ_d by definition is a root of χ. Furthermore, we know that the leading coefficient of χ equals $+1$. Since λ_d is the largest eigenvalue of $\mathsf{A}(\boldsymbol{x}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$, and since the life-cycle matrix is irreducible, we have that evaluated at the dominant eigenvalue $\frac{\partial}{\partial \lambda}\chi(\lambda, \boldsymbol{x}, \hat{\boldsymbol{E}}(\boldsymbol{x}))|_{\lambda=\lambda_\mathrm{d}=0} > 0$. Consequently, χ is increasing locally around λ_d.

Thus, if for some nearby $\boldsymbol{y}$ we have $\lambda_\mathrm{d}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) < 0$, then, since $\chi(0, \boldsymbol{x}, \hat{\boldsymbol{E}}(\boldsymbol{x})) = 0$, necessarily $\chi(0, \boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) > 0$. Conversely, if $\chi(0, \boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) > 0$ for some $\boldsymbol{y}$, from the local monotonicity around 0 follows that $\lambda_\mathrm{d}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) < 0$. Similarly, if $\lambda_\mathrm{d}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) > 0$ for some nearby $\boldsymbol{y}$, it follows that $\chi(0, \boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) < 0$ and vice versa. Hence, for all $\boldsymbol{y}$ in a sufficiently small neighborhood of $\boldsymbol{x}$, we have that $\chi(0, \boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) > 0$ if and only if $\lambda_\mathrm{d}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) < 0$.

Clearly, $\chi(0, \boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) = \det(0 \cdot \mathsf{I} - \mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))) = \det(-\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})))$. Since

$$Q_\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) = -\det(-\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))),$$

we find that $0 \gtrless Q_\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$ if and only if $0 \gtrless \lambda_\mathrm{d}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) = s(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$. Thus, $Q_\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$ and $s(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$ are sign equivalent, i.e., $Q_\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$ is a fitness proxy. $\qquad\square$

The formula provided by the preceding proposition proves to be very handy. Indeed, due to the immediate relation between the fitness proxy and the determinant of the life-cycle matrix, structural properties of the determinant directly translate into properties of the fitness proxy Q_A. In particular, A being an $n \times n$ matrix we find that

$$Q_\mathsf{A} = \begin{cases} -\det(\mathsf{A}) & \text{if } n \text{ even} \\ +\det(\mathsf{A}) & \text{if } n \text{ odd} \end{cases}$$

since the determinant is a multilinear function of the columns (or rows) of A.

Note, that the fitness proxy in the above introduced manner is not necessarily well defined for sub-life-cycles given by A_α, i.e., life cycles only containing states $i \in \alpha \subset \{1, ..., n\}$, since A_α is not necessarily irreducible. However, this can be easily repaired by defining

$$Q_{\mathsf{A}_\alpha} := \max\{Q_{\mathsf{A}_\beta} : \mathsf{A}_\beta \text{ is an irreducible principal submatrix of } \mathsf{A}_\alpha\}.$$

Then, the preceding proposition can be applied to each Q_{A_β} separately. Similarly, define $Q_{\mathsf{A}_{\setminus \alpha}}$ for a life cycle containing states $i \in \{1, ..., n\}\setminus\alpha$.

From Proposition B.4 we then immediately have the following very useful corollary:

Corollary 1. *Let $\alpha \subset \{1, ..., n\}$. Then $Q_{\mathsf{A}_\alpha}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x})) < Q_{\mathsf{A}}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$ and in particular $Q_{\mathsf{A}_\alpha}(\boldsymbol{x}, \hat{\boldsymbol{E}}(\boldsymbol{x})) < 0$.*

Remember, that the goal is to derive a fitness proxy which directly reflects the structure of the life cycle at hand. Therefore, we will further exploit the formula for the fitness proxy given in Proposition 8. The idea is to use the Leibniz formula for the determinant in order to restructure Q_{A} such that loops within the life cycle become visible.

Proposition 9. *For the fitness proxy Q_{A} of a given life-cycle matrix A, the following formula holds:*

$$Q_{\mathsf{A}} = \sum_{k=1}^{n}(-1)^{k+1} \sum_{\substack{\ell(L_1)+...+\ell(L_k)=n \\ L_i \cap L_j = \emptyset,\, i \neq j}} L_1 \cdots L_k \,. \tag{6}$$

Proof. From Proposition 8, multilinearity, and the Leibniz representation formula for the determinant follows that

$$
\begin{aligned}
Q_{\mathsf{A}} &= -\det(-\mathsf{A}) \\
&= (-1)^{n+1}\det(\mathsf{A}) \\
&= (-1)^{n+1}\sum_{\pi \in S_n}\operatorname{sgn}(\pi)\, a_{1\pi(1)}a_{2\pi(2)}\cdots a_{n\pi(n)}.
\end{aligned}
$$

From Proposition A.1 we know that the sign of a permutation can be written as $\operatorname{sgn}(\pi) = (-1)^{n-z(\pi)}$, where $z(\pi)$ is the number of disjoint permutation cycles in π.

The set of permutations S_n can be decomposed into pairwise disjunct subsets $C_k := \{\pi \in S_n | z(\pi) = k\}$, for $k = 1, ..., n$. So, for each k the set C_k just collects the permutations, which are composed by exactly k permutation cycles. Then, we can rewrite the Leibniz formula and continue the calculation:

$$
\begin{aligned}
&= (-1)^{n+1}\sum_{k=1}^{n}\sum_{\pi \in C_k}(-1)^{n-k}\, a_{1\pi(1)}\cdots a_{n\pi(n)} \\
&= (-1)^{2n}\sum_{\pi \in C_1} a_{1\pi(1)}\ldots a_{n\pi(n)} + (-1)^{2n-1}\sum_{\pi \in C_2} a_{1\pi(1)}\cdots a_{n\pi(n)} + \ldots \\
&\quad \ldots + (-1)^{n+1}\sum_{\pi \in C_n} a_{1\pi(1)}\ldots a_{n\pi(n)}
\end{aligned}
$$

38

A permutation $\pi \in C_k$ is composed by k disjoint permutation cycles. According to Appendix A.2, we can identify loop transmissions of (unconnected) loops with permuation cycles c of π via $L = a_{s_1 c(s_1)}...a_{s_l c(s_l)}$. The product in the sum in the formula above then is the product of k loop transmissions. This finally yields

$$
\begin{aligned}
= \quad & \sum_{\ell(L_1)=n} L_1 \quad - \sum_{\substack{\ell(L_1)+\ell(L_2)=n \\ L_1 \cap L_2 = \emptyset}} L_1 L_2 + \ldots \\
& \ldots + (-1)^{k+1} \sum_{\substack{\ell(L_1)+\ldots+\ell(L_k)=n \\ L_i \cap L_j = \emptyset,\, i \neq j}} L_1 \cdots L_k + \ldots \\
& \ldots + (-1)^{n+1} \sum_{\substack{\ell(L_i)=1,\, 1 \leq i \leq n \\ L_i \cap L_j = \emptyset,\, i \neq j}} L_1 \cdots L_n \\
= \quad & \sum_{k=1}^{n} (-1)^{k+1} \sum_{\substack{\ell(L_1)+\ldots+\ell(L_k)=n \\ L_i \cap L_j = \emptyset,\, i \neq j}} L_1 \cdots L_k .
\end{aligned}
$$

$\square$

The formula given above has the following interpretation: The fitness proxy can be decomposed into sums of products of unconnected loops. In each sum, lengths always add up to n, i.e., the number of states in the life cycle. The fitness proxy Q_A is linear in all loops and demographic parameters in A.

Sums of demographic parameters: Matrices for structured population models in continuous time often contain sums of demographic parameters. The loops (or more precisely: loop transmissions) in the formula for the fitness proxy derived in the previous proposition are products of matrix entries. Hence, loops can contain sums of demographic parameters as factors.

First, let us consider self loops, i.e., the diagonal entries of A, which are are of the form $a_{ii} = f_{ii} - d_i - \sum_{j \neq i} t_{ji}$. The following proposition shows how to expand Q_A in such a way, that diagonal entries show only single demographic parameters.

Proposition 10. *Let for $\alpha \subseteq \{1,...,n\}$, $|\alpha| = k$, $Q_{\mathsf{A}(v_{ii}+w_{ii}|i\in\alpha)}$ denote the fitness proxy of a life-cycle matrix where all ii-entries, $i \in \alpha$, are sums $a_{ii} = v_{ii} + w_{ii}$, i.e., a matrix which is, without loss of generality, of the form*

$$
\mathsf{A}(v_{ii}+w_{ii}|i\in\alpha) = \begin{pmatrix} v_{i_1 i_1} + w_{i_1 i_1} & \cdots & \star & \star & \cdots & \star \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ \star & \cdots & v_{i_k i_k} + w_{i_k i_k} & \star & \cdots & \star \\ \star & \cdots & \star & \star & \cdots & \star \\ \vdots & & \vdots & \vdots & & \vdots \\ \star & \cdots & \star & \star & \cdots & \star \end{pmatrix} .
$$

Then $Q_{\mathsf{A}(v_{ii}+w_{ii}|i\in\alpha)} = Q_{\mathsf{A}(v_{ii}|i\in\alpha)} - \sum_{i\in\alpha} w_{ii}Q_{\mathsf{A}_{\setminus i}}$, where

$$\mathsf{A}_{(v_{ii}|i\in\alpha)} = \mathsf{A}_{(v_{ii}+w_{ii}|i\in\alpha)} - \mathrm{diag}(w_{i_1 i_1},...,w_{i_k i_k},0,...,0).$$

Proof. We show the proof for $\alpha = \{i\}$. The assertion for arbitrary $\alpha \subseteq \{1,...,n\}$ then follows inductively.

The ith row of A is of the form $(\star,\ldots,\star,v_{ii}+w_{ii},\star,\ldots,\star)$. We use multilinearity to calculate

$$
\begin{aligned}
Q_{\mathsf{A}(v_{ii}+w_{ii})} &= -\det(-\mathsf{A}) = -\det\begin{pmatrix} -\star \\ -(\star,\ldots,\star,v_{ii}+w_{ii},\star,\ldots,\star) \\ -\star \end{pmatrix} \\
&= -\det\begin{pmatrix} -\star \\ -(\star,\ldots,\star,v_{ii},\star,\ldots,\star) \\ -\star \end{pmatrix} \\
&\quad -(-1)^n \det\begin{pmatrix} \star \\ (0,\ldots,0,w_{ii},0,\ldots,0) \\ \star \end{pmatrix},
\end{aligned}
$$

and applying Laplace's expansion rule we get

$$
\begin{aligned}
&= Q_{\mathsf{A}(v_{ii})} + (-1)^{n-1} w_{ii}\det(\mathsf{A}_{\setminus i}) \\
&= Q_{\mathsf{A}(v_{ii})} - w_{ii}Q_{\mathsf{A}_{\setminus i}}
\end{aligned}
$$

$\square$

Now, let us consider loop transmissions L of loops G with $\ell(G) > 1$ (note, that here G is a graph and L is a product of matrix entries). Assume, that all factors in L are of the form $a_{ij} = c_{ij}$, $c \in \{f,t\}$, except for one factor, which is of the form $a_{i_0 j_0} = f_{i_0 j_0} + t_{i_0 j_0}$. Then,

$$L = (f_{i_0 j_0} + t_{i_0 j_0}) \prod_{(j,i)\in E_G\setminus\{j_0,i_0\}} c_{ij} = f_{i_0 j_0} \prod_{(j,i)\in E_G\setminus\{j_0,i_0\}} c_{ij} + t_{i_0 j_0} \prod_{(j,i)\in E_G\setminus\{j_0,i_0\}} c_{ij}.$$

Hence, L is a sum of two products, each of which has $\ell(G)$ factors. If we now interpret the first product as loop transmission L_f of a loop going from j_0 to i_0 via $f_{i_0 j_0}$, and the second product as loop transmission L_t of a loop going from j_0 to i_0 via $f_{i_0 j_0}$, we find that L_f and L_t are two connected loops (more precisely: loop transmissions of two connected loops) of the same length. Hence, they do not appear together in a product of loops in the fitness proxy (6). Applying this argument repeatedly for all factors of the form $f_{ij} + t_{ij}$, we can assume that, without loss of generality, no loop in the fitness proxy contains a factor which is a sum of two demographic parameters. We will do so for the remainder of this thesis. Note, that according to the same logic we could have separated demographic parameters in the diagonal entries of the life-cycle matrix. We decided against this possibility, in order to keep the interpretation of $L = a_{ii}$ as the per capita rate at which indiduals in state i contribute to the growth of state i.

With Q_{A} we now have a useful tool to study the evolutionary dynamics of structured populations. There are two different strands: in section 4.2, we are going to

study the case of frequency independent selection. More precisely, optimization and pessimization criteria as presented in section 3.3 will be derived for certain types of life-cycle structures. Section 4.3 deals with the cases where frequency dependent selection is possible. This is done by employing the tools which have been introduced in 3.4, i.e., the analysis of invasion boundaries and trade-off curves (Rueffler et al., 2004).

4.2 Optimization

In this section we derive optimization criteria for continuous-time structured population models given by a life-cycle matrix A. Recall from section 3.3 that optimization criteria are certain functions which allow us to predict evolutionary outcomes from maximization arguments, thereby implicitly excluding frequency dependent behavior. To this end, we will make extensive use of the fitness proxy introduced in the previous section. By rearranging this fitness proxy, we give sufficient conditions for the existence of optimization criteria.

As we will see, finding optimization criteria is easier for models where regulatory functions are assumed to be multiplicative rather than additive. However, a short list of optimization criteria will also be derived for the latter class of models. Finally, we will derive optimization criteria which are based on the dominant eigenvalue of the life-cycle matrix and on the basic reproductive ratio R_0.

The following general proposition is directly taken from Rueffler et al. (2013):

Proposition 11. *Let p be a fitness proxy and let there be functions g_1, $g_2 : \mathbb{X} \to \mathbb{R}$ and e_1, $e_2 : \mathbb{E} \to \mathbb{R}$ with $g_2(\boldsymbol{x}) \neq 0$ for all $\hat{\boldsymbol{E}} \in \mathbb{E}$ such that*

$$p(\boldsymbol{y}, \hat{\boldsymbol{E}}) = \underbrace{g_1(\boldsymbol{y})e_1(\hat{\boldsymbol{E}})}_{a} + \underbrace{g_2(\boldsymbol{y})e_2(\hat{\boldsymbol{E}})}_{b}.$$

If $e_1(\hat{\boldsymbol{E}}) > 0$, then $\xi(\boldsymbol{x}) := g_1(\boldsymbol{x})/|g_2(\boldsymbol{x})|$ is an optimization criterion. If $e_1(\hat{\boldsymbol{E}}) < 0$, then $\xi(\boldsymbol{x}) := -g_1(\boldsymbol{x})/|g_2(\boldsymbol{x})|$ is an optimization criterion.

In particular the following holds:

1. *If $p(\boldsymbol{y}, \hat{\boldsymbol{E}}) = g_1(\boldsymbol{y})e_1(\hat{\boldsymbol{E}}) + e_2(\hat{\boldsymbol{E}})$ and if $e_1(\hat{\boldsymbol{E}}) \gtrless 0$, then $\xi(\boldsymbol{x}) := \pm g_1(\boldsymbol{x})$ is an optimization criterion.*

2. *If $p(\boldsymbol{y}, \hat{\boldsymbol{E}}) = g_1(\boldsymbol{y})e_1(\hat{\boldsymbol{E}}) + g_2(\boldsymbol{y})$ and $g_1(\boldsymbol{x}) \neq 0$, then $\xi(\boldsymbol{x}) := g_2(\boldsymbol{x})/|g_1(\boldsymbol{x})|$ is an optimization criterion. Similarly, if $g_2(\boldsymbol{x}) \neq 0$ and if $e_1(\hat{\boldsymbol{E}}) > 0$, then $\xi(\boldsymbol{x}) := g_1(\boldsymbol{x})/|g_2(\boldsymbol{x})|$ is an optimization criterion. If $e_1(\hat{\boldsymbol{E}}) < 0$, then $\xi(\boldsymbol{x}) := -g_1(\boldsymbol{x})/|g_2(\boldsymbol{x})|$ is an optimization criterion.*

3. *If $p(\boldsymbol{y}, \hat{\boldsymbol{E}}) = g_1(\boldsymbol{y}) + e_2(\hat{\boldsymbol{E}})$, then $\xi(\boldsymbol{x}) := g_1(\boldsymbol{x})$ is an optimization criterion.*

Proof. See Proposition 10 in Rueffler et al. (2013). $\square$

Proposition 12. *Let $\mathcal{M} \subset \mathcal{L}(\mathsf{A})$ be a set of loops in a given life cycle A. Then*

$$Q_{\mathsf{A}} = \underbrace{-\sum_{L \in \mathcal{M}} L \cdot Q_{\mathsf{A} \setminus \{L\}}}_{a} + \underbrace{R_{\setminus \mathcal{M}}}_{b}, \tag{7}$$

where $R_{\mathcal{L}(\mathsf{A}) \setminus \mathcal{M}}$ is an expression which does not depend on any of the loops in $\mathcal{M}$.

Proof. For a given permutation cycle c, let $L_c = a_{s_1 c(s_1)}...a_{s_l c(s_l)}$ denote the corresponding loop transmission, where $\{s_1,...,s_l\} \subseteq \{1,...,n\}$. Conversely, for a given loop transmission L, let c_L be the corresponding permutation cycle. The relation $c \trianglelefteq \sigma$ denotes that c is a cycle in the composition of disjoint permutation cycles of a permutation σ (cf. Appendix A.2). Similarly, $\sigma \trianglerighteq c$ shall denote a permutation containing a given permutation cycle c. Making use of the Leibniz formula for the determinant, we get from Proposition 8 that

$$
\begin{aligned}
Q_{\mathsf{A}} &= -\det(-\mathsf{A}) \\
&= -\sum_{\sigma \in S_{\{1,...,n\}}} \operatorname{sgn}(\sigma)(-a_{1\sigma(1)})...(-a_{n\sigma(n)}) \\
&= -\sum_{\sigma \in S_{\{1,...,n\}}} \operatorname{sgn}(\sigma) \prod_{c \trianglelefteq \sigma} (-1)^{\ell(c)} L_c \\
&= -\sum_{\sigma \in S_{\{1,...,n\}}, \sigma \trianglerighteq c_L \text{ for some } L \in \mathcal{M}} \operatorname{sgn}(\sigma) \prod_{c \trianglelefteq \sigma} (-1)^{\ell(c)} L_c \\
&\quad \underbrace{- \sum_{\tau \in S_{\{1,...,n\}}, \tau \ntrianglerighteq c_L \text{ for all } L \in \mathcal{M}} \operatorname{sgn}(\tau) \prod_{c \trianglelefteq \tau} (-1)^{\ell(c)} L_c.}_{=\triangle}
\end{aligned}
$$

The first sum ranges over all permutations σ which have a cycle c_L, $L \in \mathcal{M}$, in the composition of disjoint permutation cycles. In other words, σ can be decomposed into permutation cycles one of which corresponds to some loop L in $\mathcal{M}$.

The second sum $\triangle$ ranges over all permutations τ which have a composition into disjoint permutation cycles not including a cycle which corresponds to one of the loops L in $\mathcal{M}$. That is, no loop $L \in \mathcal{M}$ appears in one of the products in this sum. We then define $R_{\mathcal{L}(\mathsf{A}) \setminus \mathcal{M}} := -\triangle$, which is an expression independent of all $L \in \mathcal{M}$.

In the next step, we split the first sum into subsums, one for each of the loops in $\mathcal{M}$. Each of this subsums ranges over permutations that contain cycles corresponding to the respective loops. We use that if $\sigma = \tau_1 \tau_2$ is the product of two permutations, then $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau_1) \cdot \operatorname{sgn}(\tau_2)$. In particular this holds for compositions of disjoint permutation cycles $\sigma = c_1 \dots c_k$. Let $\sigma \setminus c_L$ denote the permutation which is composed of all permutation cycles of σ but c_L. Then $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma \setminus c_L) \cdot \operatorname{sgn}(c_L)$ and can we continue with

$$
\begin{aligned}
&= -\sum_{L \in \mathcal{M}} \sum_{\sigma \trianglerighteq c_L} \operatorname{sgn}(\sigma \setminus c_L) \cdot \operatorname{sgn}(c_L) \cdot (-1)^{\ell(c_L)} \prod_{z \trianglelefteq \sigma} L_z + R_{\setminus \mathcal{M}} \\
&= -\sum_{L \in \mathcal{M}} \sum_{\sigma \trianglerighteq c_L} \operatorname{sgn}(\sigma \setminus c_L) \cdot \operatorname{sgn}(c_L) \cdot L \cdot (-1)^{\ell(c_L)} \prod_{z \trianglelefteq \sigma, z \neq c_L} L_z + R_{\setminus \mathcal{M}} \\
&= \sum_{L \in \mathcal{M}} \operatorname{sgn}(c_L) \cdot (-1)^{\ell(c_L)} \cdot L \cdot \left(-\sum_{\sigma \trianglerighteq c_L} \operatorname{sgn}(\sigma \setminus c_L) \prod_{z \trianglelefteq \sigma, z \neq c_L} L_z \right) + R_{\setminus \mathcal{M}}.
\end{aligned}
$$

For the next step, we consider that summing over all possible permutations that include c_L and building the product of all loops generated by these permutations with the exception of L is precisely the same as summing over all possible permutations of the set of states reduced by all the states touched by L and then multiplying the relevant loops. Hence, we obtain

$$= \sum_{L \in \mathcal{M}} \mathrm{sgn}(c_L) \cdot (-1)^{\ell(c_L)} \cdot L \cdot \left(-\sum_{\pi \in S_{\{1,\dots,n\} \setminus \{L\}}} \mathrm{sgn}(\pi) \prod_{z \trianglelefteq \pi} L_z \right) + R_{\setminus \mathcal{M}}. \quad (8)$$

The expression within the brackets now is the fitness proxy for the sub-life-cycle $\mathsf{A}_{\setminus \{L\}}$, i.e., $Q_{\mathsf{A}_{\setminus \{L\}}}$. Finally, we make use of Proposition A.1 and conclude with

$$= \sum_{L \in \mathcal{M}} (-1)^{\ell(c_L)-1} (-1)^{\ell(c_L)} L \cdot Q_{\mathsf{A}_{\setminus \{L\}}} + R_{\setminus \mathcal{M}} \quad (9)$$

$$= -\sum_{L \in \mathcal{M}} L \cdot Q_{\mathsf{A}_{\setminus \{L\}}} + R_{\setminus \mathcal{M}},$$

which proves the proposition. $\qquad \square$

We will go back to this proof in some of the following results. The previous proposition can also be shown in a more direct way, resorting to Proposition 9. We present an alternative proof:

Proof. Using the formula given in Proposition 9, we find that

$$Q_{\mathsf{A}} = \sum_{k=1}^{n} (-1)^{k+1} \sum_{\substack{\ell(L_1) + \dots + \ell(L_k) = n \\ L_i \cap L_j = \emptyset, \, i \neq j \\ \exists i_0 \in \{1,\dots,k\} : L_{i_0} \in \mathcal{M}}} L_1 \cdots L_k$$

$$+ \underbrace{\sum_{k=1}^{n} (-1)^{k+1} \sum_{\substack{\ell(L_1) + \dots + \ell(L_k) = n \\ L_i \cap L_j = \emptyset, \, i \neq j \\ \forall i \in \{1,\dots,k\} : L_i \notin \mathcal{M}}} L_1 \cdots L_k}_{=: R_{\setminus \mathcal{M}}}.$$

Without loss of generality, let us assume that $L_{i_o} = L_k$ for each $k \in \{1,\dots,n\}$. We extract exactly one $L = L_{i_0} \in \mathcal{M}$ out of each sum. Then

$$Q_{\mathsf{A}} = \sum_{k=1}^{n} (-1)^{k+1} \sum_{L \in \mathcal{M}} L \underbrace{\sum_{\substack{\ell(L_1) + \dots + \ell(L_{k-1}) = n - \ell(L) \\ L_i \cap L_j = \emptyset, \, i \neq j, \, i,j \notin \{L\}}} L_1 \cdots L_{k-1}}_{(\star)} + R_{\setminus \mathcal{M}}$$

$$= -\sum_{L \in \mathcal{M}} L \underbrace{\sum_{k=1}^{n-|\{L\}|} (-1)^{k} \sum_{\substack{\ell(L_1) + \dots + \ell(L_{k-1}) = n - |\{L\}| \\ L_i \cap L_j = \emptyset, \, i \neq j, \, i,j \notin \{L\}}} L_1 \cdots L_{k-1}}_{Q_{A_{\setminus \{L\}}}} + R_{\setminus \mathcal{M}}.$$

43

The reduction from n to $n - |\{L\}|$ in the preceding step is possible, since $|\{L\}| = \ell(L)$ and since summands in $(\star)$ only exist for $k = 1, ..., n - |\{L\}|$, for each $L \in \mathcal{M}$. This gives formula (7) as desired. $\qquad\square$

Note, that in the previous proposition $R_{\backslash\mathcal{M}}$ may well contain loops that appear in $Q_{\mathsf{A}_{\backslash\{L\}}}$. That is, the loops in the first summand of the right-hand side of (7) are not generally exclusive to $\mathcal{M}$. If all $L \in \mathcal{M}$ are pairwise connected, then for each L, $Q_{\mathsf{A}_{\backslash\{L\}}}$ contains no other loops in $\mathcal{M}$.

If we now consider a structured-population model given by a life-cycle matrix A, we see that the environment $\hat{\boldsymbol{E}}$ acts on the fitness proxy $Q_{\mathsf{A}}(\boldsymbol{y}, \hat{\boldsymbol{E}})$ solely through regulatory functions $R_{c,ij}$ modifying (possibly evolving) demographic parameters $c \in \{f, t, d\}$ with indices ij. We collect all regulatory functions in a vector $\boldsymbol{R}$. Then, with some abuse of notation, we can write the fitness proxy as $Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R})$. If an expression h (in particular, for instance, a loop) depends on some traits in $\boldsymbol{y}$ or some regulatory functions in $\boldsymbol{R}$ we write $h(\boldsymbol{y})$ and $h(\boldsymbol{R})$ respectively, or $h(\boldsymbol{y}, \boldsymbol{R})$ if h depends on both, traits and regulatory functions. If there is no argument $(.)$, we assume that h does not depend on any traits or regulatory functions. Note, that strictly speaking $h(\boldsymbol{y})$, $h(\boldsymbol{R})$, and $h(\boldsymbol{y}, \boldsymbol{R})$ are not functions of their respective arguments. Rather $h(\boldsymbol{y})$ should be read as h *contains evolving traits*, and similarly for $h(\boldsymbol{R})$ and $h(\boldsymbol{y}, \boldsymbol{R})$.

Using this notation we have seven different ways to write the fitness proxy as a sum of two expressions depending on traits or regulatory functions, $Q_{\mathsf{A}} = a + b$:

$$Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = a(\boldsymbol{y}) + b(\boldsymbol{R}) \tag{10}$$
$$Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = a(\boldsymbol{R}) + b(\boldsymbol{y}) \tag{11}$$
$$Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = a(\boldsymbol{y}) + b(\boldsymbol{y}, \boldsymbol{R}) \tag{12}$$
$$Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = a(\boldsymbol{R}) + b(\boldsymbol{y}, \boldsymbol{R}) \tag{13}$$
$$Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = a(\boldsymbol{y}, \boldsymbol{R}) + b(\boldsymbol{y}) \tag{14}$$
$$Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = a(\boldsymbol{y}, \boldsymbol{R}) + b(\boldsymbol{R}) \tag{15}$$
$$Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = a(\boldsymbol{y}, \boldsymbol{R}) + b(\boldsymbol{y}, \boldsymbol{R}) \tag{16}$$

The subsequent proposition shows how the first summand a can be written as a product, i.e., $a = e \cdot g$.

Proposition 13. *Let all loops $L \in \mathcal{M}$ be pairwise connected and contain the same set of factors $\mathcal{F}_{\mathcal{M}}$. Let $\Pi := \prod_{f \in \mathcal{F}_{\mathcal{M}}} f$. Then*

$$a = -\sum_{L \in \mathcal{M}} L \cdot Q_{\mathsf{A}_{\backslash\{L\}}} = \underbrace{-\Pi}_{e} \cdot \underbrace{\sum_{L \in \mathcal{M}} \frac{L}{\Pi} Q_{\mathsf{A}_{\backslash\{L\}}}}_{g},$$

where all $f \in \mathcal{F}_{\mathcal{M}}$ only appear in the first factor e but not in g.

Proof. The proposition immediately follows from the fact that loops are products of matrix entries. $\qquad\square$

The following proposition will show when a factorization of $R_{\backslash\mathcal{M}} = b = e \cdot g$ is possible:

Proposition 14. *If there is a nontrivial partition α, β of $\{1,...,n\}$ such that each loop $L \in \mathcal{L}(\mathsf{A})\backslash\mathcal{M}$ either passes through states in α or through states in β, i.e., either $\{L\} \subseteq \alpha$ or $\{L\} \subseteq \beta$, then the rest term in (7) can be written as*

$$R_{\backslash\mathcal{M}} = \underbrace{-R_\alpha}_{e} \cdot \underbrace{R_\beta}_{g},$$

where R_α and R_β contain exclusively loops passing through α or β, respectively.

Proof. According to the proof of Proposition 12 we have

$$R_{\backslash\mathcal{M}} = - \sum_{\tau \in S_{\{1,...,n\}}, \tau \not\trianglelefteq c_L \text{ for all } L \in \mathcal{M}} \text{sgn}(\tau) \prod_{c \trianglelefteq \tau} (-1)^{\ell(c)} L_c.$$

By assumption all relevant loops either pass through α or through β. Hence, instead of summing over $\tau \in S_{\{1,...,n\}}$, we can equivalently consider compositions $\pi_\alpha \pi_\beta = \tau$, where without loss of generality $\pi_\alpha \in S_{\{1,...,|\alpha|\}}$, $\pi_\beta \in S_{\{n-|\beta|,...,n\}}$. In particular π_α and π_β are disjoint. Then we have

$$= - \sum_{\pi_\alpha \pi_\beta \in S_{\{1,...,n\}}, \pi_\alpha, \pi_\beta \not\trianglelefteq c_L \text{ for all } L \in \mathcal{M}} \text{sgn}(\pi_\alpha) \cdot \text{sgn}(\pi_\beta) \prod_{c \trianglelefteq \pi_\alpha \pi_\beta} (-1)^{\ell(c)} L_c$$

$$= - \left(\sum_{\pi_\alpha \in S_{\{1,...,|\alpha|\}}, \pi_\alpha \not\trianglelefteq c_L \text{ for all } L \in \mathcal{M}} \text{sgn}(\pi_\alpha) \prod_{c \trianglelefteq \pi_\alpha} (-1)^{\ell(c)} L_c \right)$$

$$\cdot \left(\sum_{\pi_\beta \in S_{\{n-|\beta|,...,n\}}, \pi_\beta \not\trianglelefteq c_L \text{ for all } L \in \mathcal{M}} \text{sgn}(\pi_\beta) \prod_{c \trianglelefteq \pi_\beta} (-1)^{\ell(c)} L_c \right)$$

$$=: -R_\alpha \cdot R_\beta.$$

$\square$

First, note that in contrast to Rueffler et al. (2013, 3.2), here R_α in general is not the same as the fitness proxy Q_{A_α}, since in particular R_α is lacking the product of loops of length 1, which is located in one of the $Q_{\mathsf{A}\backslash\{L\}}$, $L \in \mathcal{M}$. The same holds analogously for R_β.

Second, to be precise, the quantifying condition on the loops, i.e., *each loop $L \in \mathcal{L}(\mathsf{A})\backslash\mathcal{M}$*, is stricter than is necessary for some life cycles. In other words, for certain life cycles the proposition would not tell us that a factorization $R_{\backslash\mathcal{M}} = R_\alpha \cdot R_\beta$ is possible, although actually it is possible. This is, because there could exist loops $L \in \mathcal{L}(\mathsf{A})\backslash\mathcal{M}$ which do not appear in the expression $R_{\backslash\mathcal{M}}$ but only in one of the fitness proxies $Q_{\mathsf{A}\backslash\{M\}}$, for some $M \in \mathcal{M}$. Clearly, this holds for loops $N \in \mathcal{L}(\mathsf{A})\backslash\mathcal{M}$ connected to *all* other $L \in \mathcal{L}(\mathsf{A})\backslash\mathcal{M}$: since $\mathcal{M} \neq \emptyset$, it follows that $\ell(N) = n - |\bigcup_{M \in \mathcal{M}}\{M\}| < n$. Hence, N cannot be part of any product in $R_{\backslash\mathcal{M}}$. That is, it is sufficient to check whether *for each loop $L \in \mathcal{L}(\mathsf{A})\backslash\mathcal{M}$ not being connected to all other loops in $\mathcal{L}(\mathsf{A})\backslash\mathcal{M}$* holds that either $\{L\} \subseteq \alpha$ or $\{L\} \subseteq \beta$.

Optimization criteria based on Q_{A}: Recall that the goal of the previous paragraphs was to provide conditions for which the fitness proxy Q_{A} is of the form $Q_{\mathsf{A}} = a + b = g_1 e_1 + g_2 e_2$ such that a (or alternatively g_i, $i = 1, 2$) depends only

on the trait vector $\boldsymbol{x}$ and b (or alternatively e_i, $i = 1, 2$) depends only on the environment, acting through the vector of regulatory functions $\boldsymbol{R}$. *If* we find such a rearrangement of the fitness proxy, we can make use of Proposition 11 to prove the existence of an optimization criterion (and, which is no less important, thereby showing that selection is *not* frequency dependent in the model at hand). It is important to stress that the conditions on the life cycles which allow to rearrange the fitness proxy in the described manner are only *sufficient* and not necessary for the existence of an optimization criterion. That is, the list of optimization criteria presented below in this paragraph is by no means exhaustive. In this paragraph, only multiplicative regulatory functions will be considered.

Before we proceed, some additional notation shall be introduced: remember that via Proposition 9 we can express $Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R})$ in terms of loops, where $\mathcal{L}(\mathsf{A})$ denotes the set of loops in the life cycle defined by A. We will use the following notation: $\mathcal{E}$ shall denote the set of evolving loops, or more precisely $\mathcal{E} = \{L \in \mathcal{L}(\mathsf{A}) : L(\boldsymbol{y}) \text{ or } L(\boldsymbol{y}, \boldsymbol{R})\}$. Likewise, $\mathcal{R}$ shall denote the set of regulated loops, i.e., $\mathcal{R} = \{L \in \mathcal{L}(\mathsf{A}) : L(\boldsymbol{R}) \text{ or } L(\boldsymbol{y}, \boldsymbol{R})\}$. Furthermore, for a given loop L we will use the notation $\{L\}$ to denote the set of states traversed by L. For a set of loops $\mathcal{M}$ we define $\{\mathcal{M}\} = \bigcup_{L \in \mathcal{M}} \{L\}$.

So, this is the plan: Proposition 12 shows how to split the fitness proxy into $Q_{\mathsf{A}} = a + b$. Then, for each of the formulas (10) to (16) the conditions can be checked under which a or b can be written as products $g_i(\boldsymbol{y}) e_i(\boldsymbol{R})$. Here, we make extensive use of Propositions 13 and 14. Subsequently, by Proposition 11 we explicitly calculate optimization criteria.[7] This is done in the following list:

1. $Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = a(\boldsymbol{y}) + b(\boldsymbol{R}) = -\sum_{L \in \mathcal{M}} L(\boldsymbol{y}) \cdot Q_{\mathsf{A} \setminus \{L\}}(\boldsymbol{y}) + R_{\setminus \mathcal{M}}(\boldsymbol{R})$

   Conditions: Let $\mathcal{M} = \mathcal{E}$, such that $\mathcal{E} \cap \mathcal{R} = \emptyset$, $E \cap R \neq \emptyset$ for all $E \in \mathcal{E}$ and $R \in \mathcal{R}$. Note that it is not necessary for the $E \in \mathcal{E}$ to be pairwise connected. Apply Propositions 12 and 11.3.

   Optimization criterion: $\xi(\boldsymbol{x}) = a(\boldsymbol{x})$

2. $Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = a(\boldsymbol{R}) + b(\boldsymbol{y}) = -\sum_{L \in \mathcal{M}} L(\boldsymbol{R}) \cdot Q_{\mathsf{A} \setminus \{L\}}(\boldsymbol{R}) + R_{\setminus \mathcal{M}}(\boldsymbol{y})$

   Conditions: Let $\mathcal{M} = \mathcal{R}$, such that $\mathcal{E} \cap \mathcal{R} = \emptyset$, $E \cap R \neq \emptyset$ for all $E \in \mathcal{E}$ and $R \in \mathcal{R}$. Note that it is not necessary for the $R \in \mathcal{R}$ to be pairwise connected. Apply Propositions 12 and 11.3.

   Optimization criterion: $\xi(\boldsymbol{x}) = b(\boldsymbol{x})$

3. $Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = a(\boldsymbol{y}) + b(\boldsymbol{y}, \boldsymbol{R}) = -\sum_{L \in \mathcal{M}} L(\boldsymbol{y}) \cdot Q_{\mathsf{A} \setminus \{L\}}(\boldsymbol{y}) - R_{\alpha}(\boldsymbol{y}) \cdot R_{\beta}(\boldsymbol{R})$

   Conditions: Let $\mathcal{M} \subseteq \mathcal{E}$, such that $\mathcal{M} \cap \mathcal{R} = \emptyset$, $L \cap R \neq \emptyset$ for all $L \in \mathcal{M}$ and $R \in \mathcal{R}$. Furthermore, let there be a partition α, β of the set of states in such a way that $\{\mathcal{E} \setminus \mathcal{M}\} \subseteq \alpha$ and $\{\mathcal{R}\} \subseteq \beta$. Note that it is not necessary for the $E \in \mathcal{E}$ to be pairwise connected. Apply Propositions 12, 14 and 11.2.

---

[7]Note, that for some optimization criteria the sign is written as $\pm$ or $\mp$, which denotes that the actual sign ultimately depends on whether e_i or g_i is positive or negative as stated in Proposition 11.

Optimization criterion: $\xi(\boldsymbol{x}) = \mp a(\boldsymbol{x})/R_\alpha(\boldsymbol{x})$

4. $Q_\mathsf{A}(\boldsymbol{y},\boldsymbol{R}) = a(\boldsymbol{R}) + b(\boldsymbol{y},\boldsymbol{R}) = -\sum\limits_{L\in\mathcal{M}} L(\boldsymbol{R})\cdot Q_{\mathsf{A}\setminus\{L\}}(\boldsymbol{R}) - R_\alpha(\boldsymbol{y})\cdot R_\beta(\boldsymbol{R})$

Conditions: Let $\mathcal{M}\subseteq\mathcal{R}$, such that $\mathcal{E}\cap\mathcal{M}=\emptyset$, $E\cap L\neq\emptyset$ for all $E\in\mathcal{E}$ and $L\in\mathcal{M}$. Furthermore, let there be a partition α,β of the set of states in such a way that $\{\mathcal{E}\}\subseteq\alpha$ and $\{\mathcal{R}\setminus\mathcal{M}\}\subseteq\beta$. Note that it is not necessary for the $R\in\mathcal{R}$ to be pairwise connected. Apply Propositions 12, 14 and 11.1.

Optimization criterion: $\xi(\boldsymbol{x}) = \mp R_\alpha(\boldsymbol{x})$

5. $Q_\mathsf{A}(\boldsymbol{y},\boldsymbol{R}) = a(\boldsymbol{y},\boldsymbol{R}) + b(\boldsymbol{y}) = -\Pi(\boldsymbol{R})\cdot\sum\limits_{L\in\mathcal{M}}\frac{L(\boldsymbol{y})}{\Pi}Q_{\mathsf{A}\setminus\{L\}}(\boldsymbol{y}) + R_{\setminus\mathcal{M}}(\boldsymbol{y})$

Conditions: Let $\mathcal{M}=\mathcal{R}$, such that all $R\in\mathcal{R}$ are pairwise connected and contain the same product $\Pi(\boldsymbol{R})$ of all regulatory functions (which is only possible if there are no sums of traits and regulatory functions). Apply Propositions 12, 13 and 11.2.

Optimization criterion: $\xi(\boldsymbol{x}) = b(\boldsymbol{x})/|\sum_{L\in\mathcal{M}}\frac{L(\boldsymbol{x})}{\Pi}Q_{\mathsf{A}\setminus\{L\}}(\boldsymbol{x})|$

6. $Q_\mathsf{A}(\boldsymbol{y},\boldsymbol{R}) = a(\boldsymbol{y},\boldsymbol{R}) + b(\boldsymbol{R}) = -\Pi(\boldsymbol{y})\cdot\sum\limits_{L\in\mathcal{M}}\frac{L(\boldsymbol{R})}{\Pi}Q_{\mathsf{A}\setminus\{L\}}(\boldsymbol{R}) + R_{\setminus\mathcal{M}}(\boldsymbol{R})$

Conditions: Let $\mathcal{M}=\mathcal{E}$, such that all $E\in\mathcal{E}$ are pairwise connected and contain the same product $\Pi(\boldsymbol{y})$ of all traits (which is only possible if there are no sums of traits and regulatory functions). Apply Propositions 12, 13 and 11.1.

Optimization criterion: $\xi(\boldsymbol{x}) = \mp\Pi(\boldsymbol{x})$

7. $Q_\mathsf{A}(\boldsymbol{y},\boldsymbol{R}) = a(\boldsymbol{y},\boldsymbol{R}) + b(\boldsymbol{y},\boldsymbol{R}) = -\Pi(\boldsymbol{y})\cdot\sum\limits_{L\in\mathcal{M}}\frac{L(\boldsymbol{R})}{\Pi}Q_{\mathsf{A}\setminus\{L\}}(\boldsymbol{R}) - R_\alpha(\boldsymbol{y})\cdot R_\beta(\boldsymbol{R})$

Conditions: Let $\mathcal{M}\subseteq\mathcal{E}$, such that all $E\in\mathcal{M}$ are pairwise connected and contain the same product $\Pi(\boldsymbol{y})$ of all traits appearing in loops in $\mathcal{M}$ (which is only possible if there are no sums of traits and regulatory functions). Furthermore, let there be a partition α,β of the set of states in such a way that $\{\mathcal{E}\setminus\mathcal{M}\}\subseteq\alpha$ and $\{\mathcal{R}\}\subseteq\beta$. Apply Propositions 12, 13, 14 and 11.

Optimization criterion: $\xi(\boldsymbol{x}) = \pm\Pi(\boldsymbol{x})/R_\alpha(\boldsymbol{x})$

8. $Q_\mathsf{A}(\boldsymbol{y},\boldsymbol{R}) = a(\boldsymbol{y},\boldsymbol{R}) + b(\boldsymbol{y},\boldsymbol{R}) = -\Pi(\boldsymbol{R})\cdot\sum\limits_{L\in\mathcal{M}}\frac{L(\boldsymbol{x})}{\Pi}Q_{\mathsf{A}\setminus\{L\}}(\boldsymbol{x}) - R_\alpha(\boldsymbol{x})\cdot R_\beta(\boldsymbol{R})$

Conditions: Let $\mathcal{M}\subseteq\mathcal{R}$, such that all $R\in\mathcal{R}$ are pairwise connected and contain the same product $\Pi(\boldsymbol{R})$ of all regulatory functions appearing in loops in $\mathcal{M}$ (which is only possible if there are no sums of traits and regulatory functions). Furthermore, let there be a partition α,β of the set of states in such a way that $\{\mathcal{E}\}\subseteq\alpha$ and $\{\mathcal{R}\setminus\mathcal{M}\}\subseteq\beta$. Apply Propositions 12, 13, 14 and 11.

Optimization criterion: $\xi(\boldsymbol{x}) = \mp\sum\limits_{L\in\mathcal{M}}\frac{L(\boldsymbol{x})}{\Pi}Q_{\mathsf{A}\setminus\{L\}}(\boldsymbol{x})/R_\alpha(\boldsymbol{x})$

Remark on disjunct sets of loops: In contrast to formula (14) in Rueffler et al. (2013, p. 244), due to the structural nature of the fitness proxy, splitting Q_{A} up into two summands of which one contains exactly those loops in a set of loops $\mathcal{M}$ and the other one contains exactly those loops in a set of loops $\mathcal{N}$, turns out to be possible only for very simple partitions $\mathcal{M} \cup \mathcal{N} = \mathcal{L}(\mathsf{A})$: To see this, we first choose the sets of loops $\mathcal{M}$, $\mathcal{N}$ in such a way that for all $K \in \mathcal{M}$ and for all $L \in \mathcal{N}$ holds that K and L are connected. Then it is easy to see that we can write the fitness proxy as

$$Q_{\mathsf{A}} \;=\; \sum_{k=1}^{n}(-1)^{k+1}\!\!\!\sum_{\substack{\ell(K_1)+\ldots+\ell(K_k)=n \\ K_i\cap K_j=\emptyset,\, i\neq j \\ K_l\in\mathcal{M}}}\!\!\! K_1\ldots K_k$$

$$+\sum_{k=1}^{n}(-1)^{k+1}\!\!\!\sum_{\substack{\ell(L_1)+\ldots+\ell(L_k)=n \\ L_i\cap L_j=\emptyset,\, i\neq j \\ L_l\in\mathcal{N}}}\!\!\! L_1\ldots L_k\;.$$

Since $\mathcal{M}$, $\mathcal{N}$ qualify as a partition of $\mathcal{L}(\mathsf{A})$, $\mathcal{M} \cup \mathcal{N} = \mathcal{L}(\mathsf{A})$ and $\mathcal{M} \cap \mathcal{N} = \emptyset$ must hold. Now, let us assume without loss of generality that K_0 with $\ell(K_0) = 1$ is an element of $\mathcal{M}$. All loops of length 1 occur in (at least) one product of (6), since they are always pairwise unconnected. It follows that all length-1 loops K_i are in $\mathcal{M}$. All loops K_j with $1 < \ell(K_j) < n$ are necessarily unconnected with at least one of the length-1 loops K_i. Since $\ell(K_j) + \ell(K_i) \leq n$ it follows that all K_j are in $\mathcal{M}$. Hence, $\mathcal{N}$ contains only loops of length n (but not necessarily all of them). Thus,

$$Q_{\mathsf{A}} = \underbrace{\sum_{L_l\,\in\,\mathcal{N}} L_l}_{a} \;+\; \underbrace{\sum_{k=1}^{n}(-1)^{k+1}\!\!\!\sum_{\substack{\sum \ell(K_l)=n \\ K_i\cap K_j=\emptyset,\, K_l\in\mathcal{M}}}\!\!\! K_1\ldots K_k}_{b}\;.$$

Example 9. Consider the *D. discoideum* life cycle from Example 3. The corresponding fitness proxy without regulated parameters reads

$$Q_{\mathsf{A}} = -(f_{11} - t_{21} - t_{31} - d_1)(-d_2)(-d_3) + t_{21}f_{12}(-d_3) + t_{31}f_{13}(-d_2).$$

The set of loops is $\mathcal{L}(\mathsf{A}) = \{f_{11} - t_{21} - t_{31} - d_1, -d_2, -d_3, t_{21}f_{12}, t_{31}f_{13}\}$. As in Example 6 we define $L_{\text{vege}} = f_{11} - t_{21} - t_{31} - d_1$, $L_{\text{coop}} = t_{21}f_{12}$, $L_{\text{cann}} = t_{31}f_{13}$, $L_2 = -d_2$, $L_3 = d_3$. We now choose subsets $\mathcal{E}$ and $\mathcal{R}$ of $\mathcal{L}(\mathsf{A})$, such that there are optimization principles according to the cases 1 to 8 in the list above:

1. Set $\mathcal{E} = \{L_{\text{vege}}\}$ and $\mathcal{R} = \{L_{\text{coop}}, L_{\text{cann}}\}$. According to 1, $\xi(\boldsymbol{x}) = -L_{\text{vege}}(\boldsymbol{x})L_2 L_3$ is an optimization criterion.

2. Set $\mathcal{E} = \{L_{\text{vege}}, L_3\}$ and $\mathcal{R} = \{L_{\text{coop}}\}$. Choosing $\mathcal{M} = \{L_{\text{vege}}\} \subseteq \mathcal{E}$, and a partition $\alpha = \{3\}$, $\beta = \{1,2\}$, we find that according to 3, $\xi(\boldsymbol{x}) = -\frac{L_{\text{vege}}(\boldsymbol{x})L_2 L_3(\boldsymbol{x})}{L_3(\boldsymbol{x})} = -L_{\text{vege}}(\boldsymbol{x})L_2$ is an optimization criterion.

3. Set $\mathcal{E} = \{L_{\text{coop}}\}$, where f_{12} is evolving, and $\mathcal{R} = \mathcal{L}(\mathsf{A})$. According to 6, $\xi(\boldsymbol{x}) = f_{12}(\boldsymbol{x})$ is an optimization criterion.

4. Set $\mathcal{E} = \{L_{\text{vege}}, L_{\text{coop}}, L_3\}$, and $\mathcal{R} = \{L_3\}$. Choosing $\mathcal{M} = \{L_3\} \subseteq \mathcal{E}$, and a partition $\alpha = \{1, 2\}$, $\beta = \{3\}$, we find that according to 7, $\xi(\boldsymbol{x}) = L_3(\boldsymbol{x})/L_{\text{coop}}(\boldsymbol{x})$ is an optimization criterion.

Optimization of life cycles with additive components: In this paragraph, optimization principles for life cycles with additive components, in particular additive regulatory functions, will be derived. Up to this point, we have only considered regulatory functions which act upon demographic parameters in a multiplicative manner, i.e., entries of the life-cycle matrix of the form $c_{ij}R_{c,ij}$. Doing this, we have been ignoring models with additive regulatory functions, where the life-cycle matrix includes expressions like $c_{ij} + R_{ij}$, $c_{ii} + R_{y,ii}b_{ii}$, or $R_{c,ii}c_{ii} + R_{b,ii}b_{ii}$ where $c, b \in \{f, t, d\}$ and c_{ij}, b_{ij} are possibly evolving. Additive components are obviously relevant for two reasons: First, life cycle matrices contain diagonal elements of the form $a_{ii} = f_{ii} - d_i - \sum_{j \neq i} t_{ji}$. Here, demographic parameters could be regulated differently. Second, regulatory functions can be additive as well.

The strategy here is the same as for the mutliplicative case: The fitness proxy is split into two summands $a + b$ by Proposition 10 and 12. Proposition 10 captures the case when the diagonal entries of the life-cycle matrix show additive traits and regulatory functions. Proposition 12 provides the split by collecting all loops with additive traits and regulatory functions. Then under certain assumptions, a and b can be written as products of factors depending on either only evolving traits or only regulatory functions.

Unfortunately, additive components make the algebraic manipulation of the fitness proxy very complicated. Hence, the conditions on the existence of optimization criteria become very strict for additive-component models.

The following list shows a selection of optimization principles together with the conditions on their respective existence:

1. $Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = Q_{\mathsf{A}}(x_{ii} + R_{ii}|i \in \alpha) = Q_{\mathsf{A}}(x_{ii}|i \in \alpha) - \sum_{i \in \alpha} R_{ii}Q_{\mathsf{A}\setminus i}(\boldsymbol{y}) = Q_{\mathsf{A}}(x_{ii}|i \in \alpha) - R_{ii} \cdot \sum_{i \in \alpha} Q_{\mathsf{A}\setminus i}(\boldsymbol{y}) = a(\boldsymbol{y}) + e(\boldsymbol{R})g(\boldsymbol{y})$

   Conditions: $R_{ii} = R_{jj}$ for all $i, j \in \alpha$. All demographic parameters c_{ij}, $i \neq j$, may be evolving. Apply Propositions 10 and 11.2.

   Optimization criterion: $\xi(\boldsymbol{x}) = \sum_{i \in \alpha} Q_{\mathsf{A}\setminus i}(\boldsymbol{x})/Q_{\mathsf{A}}(x_{ii}|i \in \alpha)$

2. $Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = Q_{\mathsf{A}}(x_{ii} + R_{y,ii}y_{ii}|i \in \alpha) = Q_{\mathsf{A}}(x_{ii}|i \in \alpha) - \sum_{i \in \alpha} R_{y,ii}y_{ii}Q_{\mathsf{A}\setminus i}(\boldsymbol{y}) = Q_{\mathsf{A}}(x_{ii}|i \in \alpha) - R_{y,ii} \cdot \sum_{i \in \alpha} y_{ii}Q_{\mathsf{A}\setminus i}(\boldsymbol{y}) = a(\boldsymbol{y}) + e(\boldsymbol{R})g(\boldsymbol{y})$

   Conditions: $R_{y,ii} = R_{y,jj}$ for all $i, j \in \alpha$. All demographic parameters c_{ij}, $i \neq j$, may be evolving. Apply Propositions 10 and 11.2.

   Optimization criterion: $\xi(\boldsymbol{x}) = \sum_{i \in \alpha} y_{ii}Q_{\mathsf{A}\setminus i}(\boldsymbol{x})/Q_{\mathsf{A}}(x_{ii}|i \in \alpha)$

3. $Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = Q_{\mathsf{A}}(Rx_{ii} + R_{y,ii}y_{ii}|i \in \alpha) = Q_{\mathsf{A}}(Rx_{ii}|i \in \alpha) - \sum_{i \in \alpha} R_{y,ii}y_{ii}Q_{\mathsf{A}\setminus i}(\boldsymbol{y}) = R^n \cdot Q_{\mathsf{A}}(x_{ii}|i \in \alpha) - R_{y,ii} \cdot \sum_{i \in \alpha} y_{ii}Q_{\mathsf{A}\setminus i}(\boldsymbol{y}) = e_1(\boldsymbol{R})g_1(\boldsymbol{y}) + e_2(\boldsymbol{R})g_2(\boldsymbol{y})$

Conditions: $R_{y,ii} = R_{y,jj}$ for all $i, j \in \alpha$ and all demographic parameters but y_{ii} are regulated by the same function R. All off-diagonal demographic parameters may be evolving. Apply Propositions 10 and 11.

Optimization criterion: $\xi(\boldsymbol{x}) = \pm Q_{\mathsf{A}}(x_{ii}|i \in \alpha)/|\sum\limits_{i \in \alpha} y_{ii} Q_{\mathsf{A}_{\backslash i}}(\boldsymbol{x})|$

4. $Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = -\sum\limits_{L \in \mathcal{M}} L(x_L + R_L) \cdot Q_{\mathsf{A}_{\backslash \{L\}}} + R_{\backslash \mathcal{M}} = -\sum\limits_{L \in \mathcal{M}} L(x_L) \cdot Q_{\mathsf{A}_{\backslash \{L\}}} +$
$\left(-\sum\limits_{L \in \mathcal{M}} L(R_L) \cdot Q_{\mathsf{A}_{\backslash \{L\}}} + R_{\backslash \mathcal{M}}(\boldsymbol{R})\right) = a(\boldsymbol{y}) + b(\boldsymbol{R})$

Conditions: Let $\mathcal{M}$ be the set of all loops L with exactly one additive component of evolving traits and regulatory functions, i.e., $x_L + R_L$, such that all loops in $\mathcal{M}$ are pairwise connected. Note that not necessarily $x_L = x_M$ and $R_L = R_M$ for all $L, M \in \mathcal{M}$. Let there be no additional evolving traits. There may be a set of regulated loops $\mathcal{R} \subseteq \mathcal{L}(\mathsf{A}) \backslash \mathcal{M}$, but $L \cap R \neq \emptyset$ for all $L \in \mathcal{M}$ and $R \in \mathcal{R}$. Apply Propositions 12 and 11.3.

Optimization criterion: $\xi(\boldsymbol{x}) = -\sum\limits_{L \in \mathcal{M}} L(x_L) \cdot Q_{\mathsf{A}_{\backslash \{L\}}}$

5. $Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = -\sum\limits_{L \in \mathcal{M}} L(x_L + R_L) \cdot Q_{\mathsf{A}_{\backslash \{L\}}} + R_{\mathcal{L}(\mathsf{A}) \backslash \mathcal{M}} = -\sum\limits_{L \in \mathcal{M}} L(R_L) \cdot Q_{\mathsf{A}_{\backslash \{L\}}} +$
$\left(-\sum\limits_{L \in \mathcal{M}} L(x_L) \cdot Q_{\mathsf{A}_{\backslash \{L\}}} + R_{\backslash \mathcal{M}}(\boldsymbol{x})\right) = a(\boldsymbol{R}) + b(\boldsymbol{x})$

Conditions: Let $\mathcal{M}$ be the set of all loops L with exactly one additive component of evolving traits and regulatory functions, i.e., $x_L + R_L$, such that all loops in $\mathcal{M}$ are pairwise connected. Note that not necessarily $x_L = x_M$ and $R_L = R_M$ for all $L, M \in \mathcal{M}$. Let there be no additional regulated traits. There may be a set of evolving loops $\mathcal{E} \subseteq \mathcal{L}(\mathsf{A}) \backslash \mathcal{M}$, but $L \cap E \neq \emptyset$ for all $L \in \mathcal{M}$ and $E \in \mathcal{E}$. Apply Propositions 12 and 11.3.

Optimization criterion: $\xi(\boldsymbol{x}) = -\sum\limits_{L \in \mathcal{M}} L(x_L) \cdot Q_{\mathsf{A}_{\backslash \{L\}}} + R_{\backslash \mathcal{M}}(\boldsymbol{x})$

6. $Q_{\mathsf{A}}(\boldsymbol{y}, \boldsymbol{R}) = -\sum\limits_{L \in \mathcal{M}} L(x_L + R_L) \cdot Q_{\mathsf{A}_{\backslash \{L\}}} + R_{\mathcal{L}(\mathsf{A}) \backslash \mathcal{M}} = -R \cdot \sum\limits_{L \in \mathcal{M}} \frac{L}{R} \cdot Q_{\mathsf{A}_{\backslash \{L\}}}(\boldsymbol{y}) +$
$\left(-\sum\limits_{L \in \mathcal{M}} L(x_L) \cdot Q_{\mathsf{A}_{\backslash \{L\}}}(\boldsymbol{y}) + R_{\backslash \mathcal{M}}(\boldsymbol{y})\right) = e(\boldsymbol{R}) g(\boldsymbol{y}) + b(\boldsymbol{y})$

Conditions: Let $\mathcal{M}$ be the set of all loops L with exactly one additive component of evolving traits and regulatory functions, i.e., $x_L + R_L$, such that all loops in $\mathcal{M}$ are pairwise connected. Note that not necessarily $x_L = x_M$ but nevertheless $R_L = R_M = R$ for all $L, M \in \mathcal{M}$. That is, all loops in $\mathcal{M}$ are regulated by the same regulatory function R. Let there be no additional regulated traits. There may be a set of evolving loops $\mathcal{E} \subseteq \mathcal{L}(\mathsf{A}) \backslash \mathcal{M}$. Apply Propositions 12 and 11.2.

Optimization criterion: $\xi(\boldsymbol{x}) = \pm \dfrac{\sum_{L \in \mathcal{M}} L(x_L) \cdot Q_{\mathsf{A}_{\backslash \{L\}}}(\boldsymbol{x}) - R_{\backslash \mathcal{M}}(\boldsymbol{x})}{|\sum_{L \in \mathcal{M}} \frac{L}{R} \cdot Q_{\mathsf{A}_{\backslash \{L\}}}(\boldsymbol{x})|}$

We see that the requirements on the structure of the life cycle are much stricter than in the case of multiplicative regulatory functions. In particular, the first three optimization criteria in the above list require that all relevant diagonal entries are

regulated in the same way, $R_{y,ii} = R_{y,jj}$ for all $i, j \in \alpha$, meaning that although the traits appearing in the diagonal entries differ, they have to be affected by the same regulatory functions. Such life cycles are not necessarily biologically meaningless: consider for instance a population where all regulated death rates and/or all (intra-state) fertility rates depend on the overall density of the population.

The next two optimization criteria in the list require that all evolving traits are additively regulated and appear in pairwise connected loops. In particular this is the case if for one state all outgoing transition rates are regulated additively. For instance, the transition rates could depend on the numbers of individuals in the state where the transitions are originating.

Finally, the last optimization criterion has similar conditions on the set of regulated evolving loops but requires that all regulating functions are identical. However, it also allows for additional evolving loops without strict requirements.

Optimization principles based on λ_{d}: We finally consider optimization and pessimization criteria which are directly based on the dominant eigenvalue λ_{d} of the life-cycle matrix A. The following proposition is the continuous-time version of Proposition 13 in Rueffler et al. (2013):

Proposition 15. *If there is a scalar function $\eta(\hat{\boldsymbol{E}})$ such that all regulatory functions $R_{f,ij}$ are monotonically increasing functions of $\eta(\hat{\boldsymbol{E}})$ and all multiplicative regulatory functions $R_{d,ii}$ are monotonically decreasing functions of $\eta(\hat{\boldsymbol{E}})$, then evolution will minimize $\eta(\hat{\boldsymbol{E}})$.*

Proof. By Proposition B.2 the dominant eigenvalue λ_{d} is monotonically increasing in all matrix entries $a_{ij} = \tilde{f}_{ij} + t_{ij} - \delta_{ij}(\tilde{d}_i + \sum t_{li})$. If only fertility and death rates are affected by regulation in the above described manner, clearly a_{ij} is increasing as $\eta(\hat{\boldsymbol{E}})$ increases. Let $\beta(\boldsymbol{y}, \eta(\hat{\boldsymbol{E}})) = \lambda_{\mathrm{d}}(\boldsymbol{y}, \eta(\hat{\boldsymbol{E}}))$ and apply Proposition 11. $\qquad\square$

Note that in contrast to the discrete-time case Rueffler et al. (2013, Proposition 13) transition rates may not be affected by regulation. The reason for this is that by increasing a transition rate the corresponding diagonal entry of the life-cycle matrix is decreasing. Hence, we cannot apply the Perron-Frobenius theorem (Proposition B.2).

Also note that Proposition 15 works for multiplicative as well as additive regulatory functions, i.e., demographic parameters of the form $\tilde{f}_{ij} = f_{ij} + R_{f,ij}(\eta(\hat{\boldsymbol{E}}))$ and $\tilde{d}_{ij} = d_{ij} + R_{d,ii}(\eta(\hat{\boldsymbol{E}}))$.

We continue with the next pessimization criterion as in Metz et al. (2008) and Rueffler et al. (2013):

Proposition 16. *If all demographic parameters c_{ij} are multiplicatively regulated by the same regulatory function $R = R_{c,ij}$, $c \in \{f, d, t\}$, then evolution minimizes $R(\hat{\boldsymbol{E}})$ and maximizes $\lambda_{\mathrm{d}}(\boldsymbol{x})$.*

Proof. The proof is similar to the one in Rueffler et al. (2013, Proposition 15). We have $\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}) = R(\hat{\boldsymbol{E}})\mathsf{A}(\boldsymbol{y})$, where $\mathsf{A}(\boldsymbol{y})$ denotes the unregulated analogue to the original life-cycle matrix. This is the same as the life-cycle matrix where all regulatory functions are set to 1. Hence, $\lambda_{\mathrm{d}}(\boldsymbol{y}, \hat{\boldsymbol{E}}) = R(\hat{\boldsymbol{E}})\lambda_{\mathrm{d}}(\boldsymbol{y})$. Set $\eta(\hat{\boldsymbol{E}}) = R(\hat{\boldsymbol{E}})$, $\xi(\boldsymbol{y}) = \lambda_{\mathrm{d}}(\boldsymbol{y})$ and $\beta(\boldsymbol{y}, \eta(\hat{\boldsymbol{E}})) = R(\hat{\boldsymbol{E}})\lambda_{\mathrm{d}}(\boldsymbol{y}) = \alpha(\xi(\boldsymbol{y}), \hat{\boldsymbol{E}})$ and apply Proposition 11 to see that the trait vector and the environment act in a monodimensional manner. $\qquad\square$

Optimization principles based on R_0: In this paragraph we use the basic reproduction ratio R_0 to derive optimization and pessimization principles. Recall from Example 8 that for a given life-cycle matrix A, the basic reproduction ratio is defined by $R_0 = \rho(-\mathsf{F}\mathsf{T}^{-1})$ where F is the matrix defined by all fertility rates and T is the matrix defined by all transition and death rates, such that $\mathsf{A} = \mathsf{F} + \mathsf{T}$. Then we can state the following proposition, which is related to Proposition 15 but in addition gives us an optimization criterion, namely $R_0(\boldsymbol{x})$. It has already been stated by Mylius and Diekmann (1995).

Proposition 17. *If all fertility rates f_{ij} are multiplicatively regulated by the same regulatory function $R_\mathsf{F} = R_{f,ij}$ and all transition rates t_{ij} and death rates d_i are unregulated then evolution minimizes $R_\mathsf{F}(\hat{\boldsymbol{E}})$ and maximizes $R_0(\boldsymbol{x})$.*

Proof. According to the assumptions we have that $\mathsf{F}(\boldsymbol{y}, \hat{\boldsymbol{E}}) = R_\mathsf{F}(\hat{\boldsymbol{E}})\mathsf{F}(\boldsymbol{y})$, where in $\mathsf{F}(\boldsymbol{y})$ all regulatory functions are set to 1. Then $R_0(\boldsymbol{y}, \hat{\boldsymbol{E}}) = \rho(-\mathsf{F}(\boldsymbol{y}, \hat{\boldsymbol{E}})\mathsf{T}(\boldsymbol{y})^{-1}) = R_\mathsf{F}(\hat{\boldsymbol{E}})\rho(-\mathsf{F}(\boldsymbol{y})\mathsf{T}(\boldsymbol{y})^{-1}) = R_\mathsf{F}(\hat{\boldsymbol{E}})R_0(\boldsymbol{y})$. Set $\eta(\hat{\boldsymbol{E}}) = R_\mathsf{F}(\hat{\boldsymbol{E}})$, $\xi(\boldsymbol{y}) = R_0(\boldsymbol{y})$ and $\beta(\boldsymbol{y}, \eta(\hat{\boldsymbol{E}})) = R(\hat{\boldsymbol{E}})R_0(\boldsymbol{y}) = \alpha(\xi(\boldsymbol{y}), \hat{\boldsymbol{E}})$. The assertion then follows from Proposition 11 and the fact that $\mathrm{sign}(\ln R_0) = \mathrm{sign}(\lambda_\mathrm{d})$. $\square$

Unfortunately, showing the existence of an optimization principle via R_0 if, in addition to $R_f = R_{f,ij}$, also transition and death rates are regulated, requires a regulating function $R_\mathsf{T} = R_{t,ij} = R_{d,ii}$ which is monotonically increasing. It is intuitive to assume that death rates increase as the number of individuals in one (or more) state(s) increases (decreasing death rates are biologically less plausible). However, the assumption of transition rates which are directly monotonically related to state densities is less intuitive.

We nevertheless find such kinds of density dependence in epidemiological models where parameters are regulated by the law of mass action. If then in a more general model all transition rates and death rates are equally regulated, we have the following refined version of Proposition 17, which, according to our knowledge, has not been stated previously:

Proposition 18. *If all fertility rates f_{ij} are multiplicatively regulated by the same regulatory function $R_\mathsf{F} = R_{f,ij}$ and all death and transition rates multiplicatively regulated by an increasing function $R_\mathsf{T} = R_{t,ij} = R_{d,ii}$, then evolution minimizes $R(\hat{\boldsymbol{E}}) = \frac{R_\mathsf{F}(\hat{\boldsymbol{E}})}{R_\mathsf{T}(\hat{\boldsymbol{E}})}$ and maximizes $R_0(\boldsymbol{x})$.*

Proof. Similar to the previously given proof we have that $\mathsf{F}(\boldsymbol{y}, \hat{\boldsymbol{E}}) = R_\mathsf{F}(\hat{\boldsymbol{E}})\mathsf{F}(\boldsymbol{y})$ and $\mathsf{T}(\boldsymbol{y}, \hat{\boldsymbol{E}}) = R_\mathsf{T}(\hat{\boldsymbol{E}})\mathsf{T}(\boldsymbol{y})$. Then $R_0(\boldsymbol{x}, \hat{\boldsymbol{E}}) = \rho(-\mathsf{F}(\boldsymbol{y}, \hat{\boldsymbol{E}})\mathsf{T}(\boldsymbol{x}, \hat{\boldsymbol{E}})^{-1}) = \frac{R_\mathsf{F}(\hat{\boldsymbol{E}})}{R_\mathsf{T}(\hat{\boldsymbol{E}})}\rho(-\mathsf{F}(\boldsymbol{y})\mathsf{T}(\boldsymbol{y})^{-1}) = R(\hat{\boldsymbol{E}})R_0(\boldsymbol{y})$. Set $\eta(\hat{\boldsymbol{E}}) = R(\hat{\boldsymbol{E}})$, $\xi(\boldsymbol{y}) = R_0(\boldsymbol{y})$ and $\beta(\boldsymbol{y}, \eta(\hat{\boldsymbol{E}})) = R(\hat{\boldsymbol{E}})R_0(\boldsymbol{y}) = \alpha(\xi(\boldsymbol{y}), \hat{\boldsymbol{E}})$. Again apply Proposition 11 and $\mathrm{sign}(\ln R_0) = \mathrm{sign}(\lambda_\mathrm{d})$. $\square$

Example 10. We consider a simple extension of the juvenile-adult model given in Example 4 and Figure 1, by adding social structure into the population of rabbits: it is known, that rabbit populations establish strict hierarchical structures (Rödel et al., 2004). The state of adults can be split into a lower social class and a higher social class. Rödel et al. (2004) show that the fertility rates of female rabbits are depending on (i) the density of the subpopulation of female rabbits and (ii) on the social status. Does, i.e., female rabbits, with a higher status on average produce more offspring that

does with a lower status. For simplicity, we assume that there are two different social classes of does and that rabbits can only increase their social status. We have the following extended version of the life cycle given in Example 4
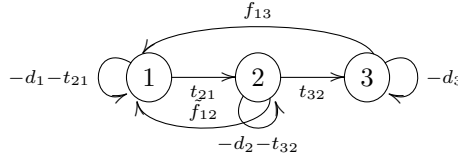
where state 2 corresponds to the lower and state 3 to the higher social class. Assume that t_{32} and f_{13} are evolving, hence we have a two dimensional trait vector $\boldsymbol{x} = (t_{32}, f_{13})$. This relates to the following biological question: is it better to restict access to the higher social class and benefit from increased fertility, or to facilitate entering the higher social class but accept the costs of reduced fertility. We parametrize the trait vecor along a trade-off curve by a scalar parameter ξ according to $\boldsymbol{x} = (2\xi^c, 3f_{12}(1 - \xi)^c)$. Thus, we can consider an effectively one-dimensional trait space. Let ξ and η denote the mutant and resident parameter, respectively.

Let us specify the regulatory functions acting on the fertility rates: First, let the fertility rates $\tilde{f}_{12}$ and $\tilde{f}_{31}$ be regulated logistically by $R_{f,12}(\boldsymbol{n}) = R_{f,31}(\boldsymbol{n}) = 1 - (n_2 + n_3)/K$, where K is a parameter which can be interpreted as breeding capacity, e.g. the number of breeding sites in a burrow. This means that reproductive rates for both social classes depend on the overall density of female adults. Let there be no other regulated parameters. From Proposition 17 it follows immediately, that there exists an optimization principle, namely $R_0(\boldsymbol{x})$. The existence of an optimization principle can also be seen from the skew-symmetric PIP in Figure 7.

Next, we consider the regulatory functions $R_{f,12}(\boldsymbol{n}) = 1 - n_2/K_1$ and $R_{f,31}(\boldsymbol{n}) = 1 - n_3/K_2$. In this case, fertility rates of lower class and higher class females depend on the respective subpopulation sizes. This can be interpreted as a distribution of breeding sites among the two classes. In each class, females compete for respective breeding sites. We see in Figure 7, that the PIP is not skew-symmetric around the singular point ξ^*. Selection is frequency dependent. Since locally around the singular point, above the 45° line through ξ^* the PIP is positive and since the vertical line through ξ^* lies in a negative region, we know that ξ^* is a CSS.

4.3 Trade-offs and invasion boundaries

We will now consider the case of frequency dependent selection. In particular, the goal is to show conditions under which evolutionary branching occurs. To this end, we will employ the methods presented in 3.4. This will be done in the following way: after some preparations, we will derive conditions for invasion boundaries to be linear, concave or convex. In the case of linear invasion boundaries, it is easy to determine whether an evolutionarily singular point is invadible or not. In our final result, we will use this to show that in two-dimensional trait spaces evolutionary branching can only occur, if locally around an evolutionarily singular point, the trade-off curve is slightly convex.

The derivatives of Q_{A}: From Proposition B.2 we know that the dominant eigenvalue (or more precisely, the stability modulus) of a life-cycle matrix is increasing in
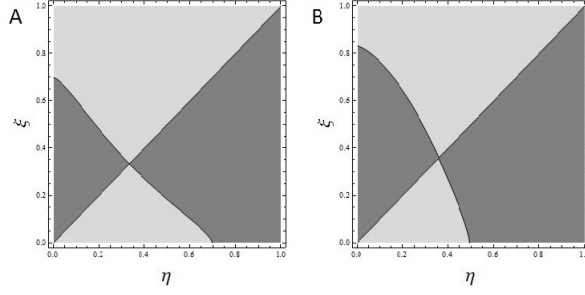
Figure 7: *PIP A on the left shows a case of optimization while PIP B on the right shows frequency dependent selection. In both cases, the singular strategy ξ^* is a CSS ($t_{12} = 2$, $d_1 = 2$, $d_2 = 1$, $d_3 = \frac{1}{2}$, $f_{12} = 4$, $c = \frac{3}{2}$, and $K = 20$).*

its entries. This however is not necessarily true for demographic parameters, which may appear in two different entries. This is shown by the following proposition:

Proposition 19. *Let Q_{A} be the fitness proxy corresponding to a life-cycle matrix $\mathsf{A}(\boldsymbol{y}, \hat{\boldsymbol{E}}(\boldsymbol{x}))$. Define $P_{\mathsf{A}_{\backslash ij}} := (-1)^{i+j+1}\det(-\mathsf{A}_{\backslash ij})$, where $\mathsf{A}_{\backslash ij}$ is the life-cycle matrix reduced by the ith row and jth column. Then*

1. *$\frac{\partial}{\partial f_{ij}} Q_{\mathsf{A}}(\boldsymbol{x}, \hat{\boldsymbol{E}}(\boldsymbol{x})) > 0$ for all $i, j \in \{1, ..., n\}$, and $Q_{A_{\backslash \alpha}}(\boldsymbol{x}, \hat{\boldsymbol{E}}(\boldsymbol{x})) < 0$, for all $\alpha \subset \{1, ..., n\}$,*

2. *$\frac{\partial}{\partial d_i} Q_{\mathsf{A}}(\boldsymbol{x}, \hat{\boldsymbol{E}}(\boldsymbol{x})) < 0$ for all $i \in \{1, ..., n\}$, and*

3. *$\frac{\partial}{\partial t_{ij}} Q_{\mathsf{A}}(\boldsymbol{x}, \hat{\boldsymbol{E}}(\boldsymbol{x})) \gtrless 0$ if $P_{\mathsf{A}_{\backslash ij}} \gtrless Q_{A_{\backslash j}}$, where $i, j \in \{1, ..., n\}$.*

Proof. From Theorem B.3 we know that λ_{d} is increasing in all entries of A. Since the matrix entry a_{ij} is increasing in f_{ij}, we have $\frac{\partial \lambda_{\mathrm{d}}}{\partial f_{ij}} > 0$. Due to the fact that Q_{A} is a local fitness proxy, by continuity we have that $\frac{\partial}{\partial f_{ij}} \lambda_{\mathrm{d}}(\boldsymbol{x}, \hat{\boldsymbol{E}}(\boldsymbol{x})) \gtrless 0$ iff $\frac{\partial}{\partial f_{ij}} Q_{\mathsf{A}}(\boldsymbol{x}, \hat{\boldsymbol{E}}(\boldsymbol{x})) \gtrless 0$. This shows the first part of 1.

Applying Laplace expansion to Q_{A}, it is easy to see that

$$0 < \frac{\partial}{\partial f_{ij}} Q_{\mathsf{A}} = (-1)^{n-1} \frac{\partial}{\partial f_{ij}} \det(\mathsf{A}) = (-1)^{n-1}(-1)^{i+j}\det(\mathsf{A}_{\backslash ij}) = -P_{\mathsf{A}_{\backslash ij}} \quad (17)$$

and setting $j = i$, from $P_{\mathsf{A}_{\backslash ii}} = Q_{\mathsf{A}_{\backslash i}}$ we obtain the second part of 1 (simply iterate the calculation for all $i \in \alpha$).

Since a_{ii} is decreasing in d_i, for all $i \in \{1, ..., n\}$, we analogously obtain the second assertion.

The transition rates t_{ij} always appear in two different entries of a life-cycle matrix, both of which lie in the jth column. Note particularly, that in the ij-entry, t_{ij} has a positive sign, and in the jj-entry t_{ij} has a negative sign. Let us expand the determinant of Q_{A} along the jth column:

$$\frac{\partial}{\partial t_{ij}} Q_{\mathsf{A}} = (-1)^{n-1} \frac{\partial}{\partial t_{ij}} \det(\mathsf{A}) = (-1)^{n-1}\left(-\det(\mathsf{A}_{\backslash jj}) + (-1)^{i+j}\det(\mathsf{A}_{\backslash ij})\right)$$
$$= Q_{A_{\backslash j}} - P_{\mathsf{A}_{\backslash ij}}$$

From inequality (17) we know that $Q_{A_{\backslash j}}, P_{A_{\backslash ij}} < 0$. Hence $\frac{\partial}{\partial t_{ij}} Q_A(\boldsymbol{x}, \hat{\boldsymbol{E}}(\boldsymbol{x})) \gtrless 0$ if $Q_{A_{\backslash j}} \lessgtr P_{A_{\backslash ij}}$, which shows the third assertion. $\qquad\square$

The condition for $\frac{\partial}{\partial t_{ij}} Q_A > 0$, i.e., $P_{A_{\backslash ij}} > Q_{A_{\backslash j}}$, can be interpreted in the following way: $Q_{A_{\backslash j}} = -\frac{\partial}{\partial a_{jj}} Q_A$ can be read as the change in fitness, as a_{jj} is changing. Fitness (i.e., Q_A) is increasing if t_{ij} (and therefore a_{ii}) is increasing, i.e., if more individuals stay in class j. Similarly, $P_{A_{\backslash ij}} = -\frac{\partial}{\partial a_{ij}} Q_A$ is the change in fitness, as $a_{ij} = t_{ij}$ is changing (in particular, a_{ij} is a linear function of t_{ij}). Consequently, fitness is increasing if less individuals stay in class j. Note, that since locally $\text{sign}\frac{\partial}{\partial a_{ij}} Q_A = \text{sign}\frac{\partial}{\partial a_{ij}} \lambda_d$, we can perceive $-P_{A_{\backslash ij}}$ as the *sensitivity* of the fitness proxy with respect to the ij-entry of the life-cycle matrix (Caswell, 1978; Keyfitz and Caswell, 2005). Proposition 19.3 then says, that if it is more beneficial to move from class j to class i than to stay there, then fitness is increasing in t_{ij}.

The curvature of the invasion boundary: In this paragraph we will present one of the main results of this thesis: as an analogue to Proposition 18 in Rueffler et al. (2013) we will give sufficient conditions for the invasion boundary to be convex, linear, or concave. We restrict the trait space to be 2-dimensional, i.e., $\boldsymbol{x} = (x_1, x_2)$. Recall, that the invasion boundary I is implicitly defined by $s((x_1, I(x_1)), \hat{\boldsymbol{E}}) = 0$.

Lemma 1. *Let $\mathcal{M}$ be a set of pairwise connected loops and let $(\mathcal{M}_1, ..., \mathcal{M}_k)$ be a partition of $\mathcal{M}$. Then*

$$Q_A = -\sum_{i=1}^{k} \sum_{L \in \mathcal{M}_i} L \cdot Q_{A_{\backslash \{L\}}} + R_{\backslash \mathcal{M}}.$$

Furthermore, let $\mathcal{N}$ be another set of pairwise connected loops. Let $(\mathcal{P}_1, ..., \mathcal{P}_k)$ be a partition of $\mathcal{P} := \mathcal{M} \otimes \mathcal{N} := \{(M, N) | M \in \mathcal{M}, N \in \mathcal{N}, M \text{ and } N \text{ unconnected}\}$. Then

$$Q_A = \sum_{i=1}^{k} \sum_{(M,N) \in \mathcal{P}_i} M \cdot N \cdot Q_{A_{\backslash \{M\} \cup \{N\}}} + R_{\backslash \mathcal{M} \otimes \mathcal{N}}.$$

Proof. The first part immediately follows from Proposition 12. For the second part, we calculate as in the proof of Proposition 12 (see in particular equation (8) and equation (9)):

$$
\begin{aligned}
Q_A &= \sum_{(M,N) \in \mathcal{M} \otimes \mathcal{N}} (-1)^{\ell(c_M)-1}(-1)^{\ell(c_M)} M \cdot \\
&\qquad \cdot (-1)^{\ell(c_N)-1}(-1)^{\ell(c_N)} N \cdot Q_{A_{\backslash \{M\} \cup \{N\}}} + R_{\backslash \mathcal{M} \otimes \mathcal{N}} \\
&= \sum_{(M,N) \in \mathcal{M} \otimes \mathcal{N} = \mathcal{P}} M \cdot N \cdot Q_{A_{\backslash \{M\} \cup \{N\}}} + R_{\backslash \mathcal{M} \otimes \mathcal{N}} \\
&= \sum_{i=1}^{k} \sum_{(M,N) \in \mathcal{P}_i} M \cdot N \cdot Q_{A_{\backslash \{M\} \cup \{N\}}} + R_{\backslash \mathcal{M} \otimes \mathcal{N}}.
\end{aligned}
$$

$\qquad\square$

Table 1: *Sufficient conditions for the invasion boundary $I(x_1) = x_2$ to be convex, linear, or concave.*

life-cycle type	trait x_1	trait x_2	curvature of I
1. pairwise connected single trait loops or multi trait loop with length 1	f_{ij}, d_j	f_{kl}, d_l	linear
	t_{ij}	f_{kl}, d_l	linear and all $L(x_2)$ pass through j
	t_{ij}	t_{kl}	for multi trait loops: linear and $j = l$
2. multi trait loops with length >1 and evolving loops are pairwise connected	f_{ij}	f_{kl}	convex
	t_{ij}	f_{kl}	$\begin{cases} \text{convex} & \text{if } \frac{\partial Q_A}{\partial x_1} > 0 \\ \text{concave} & \text{if } \frac{\partial Q_A}{\partial x_1} < 0 \end{cases}$ and all $L(x_2)$ pass through j
	t_{ij}	t_{kl}	impossible
3. only single trait loops and at least one pair of unconnected evolving loops	f_{ij}, d_j	f_{kl}, d_l	concave
	t_{ij}	f_{kl}, d_l	undetermined
	t_{ij}	t_{kl}	undetermined

Theorem 1. *Let $\mathbb{X}$ be two-dimensional, and let for $\boldsymbol{x} = (x_1, x_2)$, $I(x_1) = x_2$ denote the invasion boundary. Let both x_1, x_2 affect single demographic parameters f_{ij}, d_j, t_{ij} and f_{kl}, d_l, t_{kl} for some $i, j, k, l \in \{1, ..., n\}$. Furthermore, let $\frac{\partial}{\partial x_2} Q_A \neq 0$. Consider the following three classes of life-cycle structures as in (Rueffler et al., 2013, Proposition 9):*

1. *All evolving loops are either single trait loops or multi trait loops of length 1, and in addition pairwise connected.*

2. *There are evolving multi trait loops of length larger than 1 and all evolving loops are pairwise connected.*

3. *All evolving loops are single trait loops and there is at least one pair of unconnected evolving loops.*

Then the curvature of the invasion boundary I evaluated at a singular point, i.e., $\frac{\mathrm{d}^2 I}{\mathrm{d} x_1^2}(x_1^*)$ is given as in Table 1.

Proof. See Appendix C.1. □

As a direct consequence of Theorem 1 we obtain

Corollary 2. *Let $\mathbb{X}$ be two-dimensional and let $T(x_1) = x_2$ be a trade-off function. Consider the classes of life cycles as defined in Theorem 1.1-3. If in an evolutionarily singular point $\boldsymbol{x}^*$ the curvature of the trade-off function is known, invadibility properties of $\boldsymbol{x}^*$ are as shown in Table 2 and Table 3.*

Proof. This immediately follows from Proposition 7 on page 33 and Theorem 1. □

Since invadibility already gives us half of the information needed to classify an evolutionarily singular strategy (and in particular to determine whether a singular

Table 2: *Invadibility properties for evolutionarily singular points $\boldsymbol{x}^*$ implied by life-cycle structures and trade-off curvatures. Birth and death rates are evolving. A long dash denotes that uninvadible as well as invadible singular strategies are possible.*

$x_1, x_2 \in \{f_{ij}, d_k\}$		trade-off $\boldsymbol{T}$	
	linear	concave	convex
1. pairwise connected single trait loops or multi trait loop with length 1	*neutrally stable*	*uninvadible*	*invadible*
2. multi trait loops with length >1 and evolving loops are pairwise connected	*uninvadible*	*uninvadible*	—
3. only single trait loops and at least one pair of unconnected evolving loops	*invadible*	—	*invadible*

Table 3: *Invadibility properties for evolutionarily singular points $\boldsymbol{x}^*$ implied by life-cycle structures and trade-off curvatures. Transition and birth or death rates are evolving. A long dash denotes that uninvadible as well as invadible singular strategies are possible.*

$x_1 \in \{t_{ij}\}, x_2 \in \{f_{kl}, d_l\}$			trade-off $\boldsymbol{T}$	
		linear	concave	convex
1. pairwise connected single trait loops or multi trait loop with length 1		*neutrally stable*	*uninvadible*	*invadible*
2. multi trait loops with length >1 and evolving loops are pairwise connected	$\frac{\partial Q_A}{\partial x_1} > 0$	*uninvadible*	*uninvadible*	—
	$\frac{\partial Q_A}{\partial x_1} < 0$	*invadible*	—	*invadible*
3. only single trait loops and at least one pair of unconnected evolving loops		—	—	—

point is a branching point), we are left with examining convergence stability. At least for the case of pairwise connected single (or length 1 multi-) trait loops, as we will see in the subsequent paragraph, this turns out to be unproblematic.

Linear invasion boundaries: In this paragraph, we consider the special case of linear invasion boundaries. In this case, for the classification of singular points it suffices to investigate the (local) curvature of the trade-off curve. We use this in our final result, viz. Theorem 2, which is the continuous time analogue of Proposition 21 in Rueffler et al. (2013). We will see that under certain assumptions evolutionary scenarios can be predicted.

Again, we consider two-dimensional trait spaces, i.e., trait vectors of the form $(x_1, x_2) = \boldsymbol{x} \in \mathbb{X}$, where x_1 and x_2 are linked by a trade-off $T(x_1) = x_2$. As in Rueffler et al. (2013, Section 6), along the trade-off curve (x_1, x_2) are parametrized by a scalar on the unit interval. We will use ξ to denote the scalar value of the mutant, $\boldsymbol{y} := \boldsymbol{x}(\xi) = (x_1(\xi), x_2(\xi))$, and η to denote the scalar value of the resident, $\boldsymbol{x} := \boldsymbol{x}(\eta) = (x_1(\eta), x_2(\eta))$. We assume that $x_1(.)$ is increasing and that $x_2(.)$ is decreasing, at least locally around a singular point. Additionally, we assume that $x_1(.)$ and $x_2(.)$ are both convex or both concave. Under this assumption, we know from Rueffler et al. (2013, Appendix C) that $x_1(.)$ and $x_2(.)$ have the same curvature as T.

We are now ready to state the following result

Theorem 2. *Let the trait space be two-dimensional, i.e., there are two traits x_1 and x_2. Let all evolving loops be pairwise connected single trait loops or multi trait loops with length one. Assume that all evolving loops contain at most one regulatory function. Furthermore let for each trait x_i all evolving loops be regulated by the same regulatory function R_i. We allow for additive and multicative regulatory functions. Let the trait values of the mutant and the resident determining the environment be parametrized by ξ and $\eta \in [0, 1]$, respectively, along the trade-off curve T. We write $x_i = x_i(\xi)$ and $R_i = R_i(\eta)$. Finally, let ξ^* be the parameter corresponding to a singular point $\boldsymbol{x}^* = (x_1^*, x_2^*) = (x_1(\xi^*), x_2(\xi^*))$. We write $|_*$ for $|_{\xi=\eta=\xi^*}$.*

1. *Let x_1 affect a fertility rate and let x_2 affect a fertility or death rate. Let R_1 and R_2 be multiplicative regulatory functions with $\mathrm{d}R_1/\mathrm{d}\eta|_* < 0$, and $\mathrm{d}R_2/\mathrm{d}\eta|_* > 0$, or let R_1 be additive and R_2 be multiplicative with $\mathrm{d}R_2/\mathrm{d}\eta|_* > 0$. Then there is a constant $k > 0$ so that*

   (a) *$\boldsymbol{x}^*$ is an evolutionary repellor if $\mathrm{d}^2 T/\mathrm{d}x_1^2|_* > k$,*

   (b) *$\boldsymbol{x}^*$ is an evolutionary branching point if $k > \mathrm{d}^2 T/\mathrm{d}x_1^2|_* > 0$, and*

   (c) *$\boldsymbol{x}^*$ is a CSS if $0 > \mathrm{d}^2 T/\mathrm{d}x_1^2|_*$.*

2. *Let x_1 affect a transition rate and let x_2 affect a fertility or death rate. Let R_1 and R_2 be multiplicative regulatory functions with $\mathrm{d}R_1/\mathrm{d}\eta|_* < 0$, and $\mathrm{d}R_2/\mathrm{d}\eta|_* > 0$, or let R_1 be additive and R_2 be multiplicative with $\mathrm{d}R_2/\mathrm{d}\eta|_* > 0$, or let R_1 be multiplicative with $\mathrm{d}R_1/\mathrm{d}\eta|_* < 0$, and R_1 be additive. Without loss of generality, let us assume that the transition rate originates in state 1. Then there are constants $k, h > 0$ so that (a) to (c) apply as above, as long as $h > -Q_{\mathsf{A}\backslash\{1\}} > 0$.*

3. *Let x_1 affect a fertility or transition rate and let x_2 affect a fertility or death rate. If both regulatory functions are additive, evolution is not frequency dependent such that $\boldsymbol{x}^*$ is a repellor for $d^2 T/dx_1^2|_* > 0$ and a CSS for $d^2 T/dx_1^2|_* < 0$.*

Proof. See Appendix C.2. □

Note, that the constant h is defined as in (19) to (22) in Appendix C.2, respectively. Analogously to the preceding Theorem, one can show that if the signs of $\mathrm{d}R_1/\mathrm{d}\eta|_*$, and $\mathrm{d}R_2/\mathrm{d}\eta|_*$ are reversed, i.e., $\mathrm{d}R_1/\mathrm{d}\eta|_* > 0$ and $\mathrm{d}R_2/\mathrm{d}\eta|_* < 0$, then similar results apply, where (a) to (c) is replaced by

(a') $\boldsymbol{x}^*$ is an evolutionary repellor if $\mathrm{d}^2T/\mathrm{d}x_1^2|_* > 0$,

(b') $\boldsymbol{x}^*$ is a Garden of Eden point if $0 > \mathrm{d}^2T/\mathrm{d}x_1^2|_* > k$, and

(c') $\boldsymbol{x}^*$ is a CSS if $k > \mathrm{d}^2T/\mathrm{d}x_1^2|_*$.

A comparison of Theorem 2 and Proposition 21 in Rueffler et al. (2013) shows that for evolving fertility and death rates and multiplicative regulatory functions the results are precisely the same. Moreover, including additive rather than multiplicative regulatory functions simplifies the results, in that on the one hand conditions are relaxed and on the other hand predictions are stronger. The case of a mix of multiplicative and additive regulatory functions is of particular interest: here, we only have to account for the (local) slope of one regulatory function. Slightly changing the curvature of the trade-off then gives us the whole spectrum of evolutionary scenarios.

Fortunately, for a mix of evolving transition rates and fertility rates the results are not much worse. The only requirement is, that the fitness proxy of the reduced life cycle, i.e., $Q_{\mathsf{A}\setminus\{1\}}$, is sufficiently close to zero. Since $|Q_{\mathsf{A}\setminus\{1\}}| = |\partial Q_{\mathsf{A}}/\partial a_{11}|$, this is equivalent with requiring that fitness should only slightly depend on state 1. In other words, if fitness is sufficiently insensitive (cf. Caswell, 1978; Keyfitz and Caswell, 2005) with respect to the state where the transition is originating, we can treat the transition rate like a fertility rate.

5 Discussion

The aim of this thesis was to provide criteria which allow for predictions concerning the long-term evolution of structured populations. Those criteria were desired to be based on (i) the structure of the population in terms of life-cycle configurations as well as on (ii) the shape of trade-offs among evolving demographic parameters. Particularly, (i) differs from existing approaches to continuous-time structured-population models, in particular those of Hoyle and Bowers (2008), Hoyle et al. (2008), and Bowers (2011), where predictions are based on the model rather than on the population-structure itself.

Proceeding along the lines of Rueffler et al. (2013), we were persuing the following strategy: In section 2 we defined our model family, namely the family of continuous-time matrix-population models, which are described by demographic parameters as well as regulatory functions. In order to widen the scope of the present approach, mutliplicative as well as additive regulatory functions were permitted. We allowed for various types of regulatory functions, including population regulation via the law of mass action, as this is of particular relevance to the study of evolutionary epidemiology.

The main analytical tools for the evolutionary analysis of the models are derived within the framework of Adaptive Dynamics. We presented all concepts relevant to our study in section 3, two of them shall be emphasized: First, we presented a review of results related to optimization. If a model supports an optimization principle, the evolutionary dynamics is expected to approach a set of trait values that maximizes the optimization principle. Second, we prepared all results which are necessary in order to determine the invadibility of evolutionarily singular points right from the curvature of trade-off curves.

The tools presented in section 3 were then used to investigate the family of structured population models. First, following Rueffler et al. (2013) we presented a fitness proxy Q_A for our class of models, which reduces computational effort to the calculation of a determinant. This has three advantages: First, it is algebraically simple. Second, we can directly use results from the theory of determinants to calculate fitness: via the Leibniz rule for the determinant, we were able to identify loops in a life-cycle graph with products in the determinant of the corresponding life-cycle matrix. Most of the following results in this thesis are based on this fitness proxy. Third, in contrast to λ_d, it can always be expressed analytically. We have to stress that although the fitness proxy Q_A is algebraically simple, models of structured populations in continuous time are to a certain extent problemantic in comparison to their discrete-time analogues. This is due to the transition rates, which always show up in two separate entries of the life-cycle matrix and hence also in two separate loops, which are pairwise connected.

In section 4.2, we made use of the fitness proxy in order to calculate explicit optimization principles for a class of models. This is done by splitting the set of loops in a life cycle into approproate subsets of regulated loops, evolving loops or loops which are evolving and regulated. We also found optimization principles using the basic reproductive number R_0, which is widely used in mathematical epidemiology.

Finally, we considered the case of frequency dependent selection in section 4.3. As in Rueffler et al. (2013), for two-dimensional trait spaces, sufficient conditions for the invasion boundary to be concave, linear, or concave could be given (Theorem 1 as a counterpart of Proposition 18 in Rueffler et al., 2013). The linear case is of special interest: it allows to predict invadibility of an evolutionarily singular point solely by

inspecting the trade-off curve. We used this fact to show under which conditions evolutionary branching becomes possible (Theorem 2 as a counterpart of Proposition 21 in Rueffler et al., 2013).

Comparing the results to the discrete-time case (Rueffler et al., 2013): Due to the structural properties of continuous-time matrix models, namely transition rates appearing in two matrix entries, the conditions in both theorems become stricter as the number of evolving transition rates increases. Hence, the results are less elegant than in the discrete-time case. Note however, that according to Theorem 2, we obtain similar predictions as in the discrete-time case, as long as the fitness proxy is sufficiently insensitive with respect to originating states of evolving transition rates, i.e., as long as changes in the relevant diagonal entries of $\mathbf{A}$ do not contribute much to changes in the fitness proxy. Keeping this in mind, we can make similar predictions as in Rueffler et al. (2013) concerning the evolution of traits in structured populations:

First, if only death rates or fertility rates are evolving and if these rates appear in a pair of unconnected evolving loops, then the invasion boundary is concave. This entails that evolutionarily singular strategies are likely to be repellors or branching points (Rueffler et al., 2004, 2013). In the first case, by selection the loop transmission of one of the two loops is increasing, while the loop transmission of the other one is decreasing. In the second case, two coexisting species arise, both of which increase the respective loop transmission and decrease the other one. If fertility rates (not located in a self loop) are decreased to zero, the corresponding loop vanishes. Thus, species with unconnected loops containing evolving fertility rates are predicted to be unlikely to occur. Furthermore, since in the case of continuous-time models, death rates (which are assumed to be non-zero) establish unconnected self loops, there are two possible predictions: either, one of the two death rates increases while the other decreases, thereby resulting in species which specialize in the survival of one state, or death rates of separate states are unlikely to be traded off against each other (but see Stearns, 1989, who argues for intergenerational mortality trade-offs).

Second, if there are only mutli-trait loops which in addition share exactly the same evolving demographic parameters as factors, then by virtue of the optimization principles derived in section 4.2, evolutionary branching is not possible. Hence, in accordance with Rueffler et al. (2013), closely related coexisting species are predicted to differ in traits which are not exclusively located in the same set of loops.

Third, in the case of pairwise connected single-trait loops it can be predicted by virtue of Theorem 2, that evolutionary branching is possible only if evolving loops are regulated differently. Note, that the conditions on the respective regulatory functions are different for multiplicative and additive density regulation. For evolutionary branching to be possible, it is required that not all regulatory functions are additive.

Comparing the results to Hoyle et al. (2008), Hoyle and Bowers (2008), and Bowers (2011): In the approach of Hoyle et al. (2008), Hoyle and Bowers (2008), and Bowers (2011), conditions are directly based on properties of the life-cycle matrix rather than on the corresponding loop structures (Hoyle and Bowers, 2008, p. 312; see also section 2). Recall, that the criteria established therein assert the following:

□ Criterion A: at least two rows of $\mathbf{A}$ are evolving.

□ Criterion B: at least two columns of $\mathbf{A}$ are evolving.

Table 4: *Classification of evolutionary outcomes depending on* $\mathrm{d}^2 T/\mathrm{d}x_1^2$ *according to Hoyle and Bowers (2008). In this table,* $+$ *denotes that a criterion is fulfilled,* $-$ *denotes that a criterion is not fulfilled.* I_1 *is the invasion boundary defined by* $s(\boldsymbol{y}, \boldsymbol{x}) = 0$, *while* I_2 *is the invasion boundary defined by* $s(\boldsymbol{x}, \boldsymbol{y}) = 0$.

Case	I_1	I_2	Possible evolutionary outcomes	A	B	C
I	linear	linear	CSS/repellor	$-$	$-$	$-$
II	curved	curved	CSS/repellor	$+$	$+$	$-$
III	linear	curved	CSS/BP/repellor *or* repellor/GoE/CSS	$+$	$-$	$+$
IV	curved	curved	CSS/BP/repellor *or* repellor/GoE/CSS	$+$	$+$	$+$

□ Criterion C: at least two entries of A or two components of one entry of A are regulated differently.

Evolutionary scenarios are classified with respect to the trade-off curve T, the invasion boundary I_1, implicitly defined by $s(\boldsymbol{y}, \boldsymbol{x}) = 0$ (i.e., $I_1 = I$), and an additional curve I_2, which is the invasion boundary implicitly defined by $s(\boldsymbol{x}, \boldsymbol{y}) = 0$. Hoyle and Bowers (2008) show the following:

□ If A is not fulfilled, then I_1 and I_2 are linear and superimposed.

□ If B is not fulfilled, then I_1 is linear.

□ If C is not fulfilled, then selection is frequency-independent.

Let us compare this to our results: The first part of the first result follows from Theorem 1: if A is not fulfilled, then only demographic parameters in one, say the jth, row of A are evolving. Then all evolving demographic parameters correspond to edges in the life-cycle graph which go to state j. Consequently, all evolving loops are pairwise connected single-trait loops, or there is a multi-trait self loop. Hence, I_1 is linear. A similar argument applies to the second result: if B is not fulfilled, then all evolving demographic parameters are located in the ith column, hence corresponding to edges which go away from state i. Again, all evolving loops are pairwise connected single trait loops, or there is a multi-trait self loop, whereby I_1 is linear. We see that Theorem 1 provides a more general result by widening the class of life-cycles for which I_1 is linear. Note, however, that the conditions on the life cycle in Theorem 1 for I_1 to be linear are still only sufficient and not necessary, since for life-cycles with unconnected evolving loops, the invasion boundary can be linear (although only in non-generic cases). Furthermore, Theorem 1 gives conditions for I_1 to be convex or concave. Obviously, Theorem 1 makes no assertion about the curvature of I_2.

The third result asserts, that there exists an optimization principle, if criterion C is not fulfilled. Then all demographic parameters, if regulated, are regulated in the same way. It is still to be shown whether or not this special case is captured by one of the optimization principles which have been derived in section 4.2.

Table 4 shows the classification of possible evolutionary outcomes together with necessary conditions on the criteria A, B and C. Let us consider the four different cases: in case I and case II only CSSs or repellors are possible evolutionarily singular strategies. According to Hoyle and Bowers (2008), both require that criterion C is not fulfilled, and hence frequency-independent selection. Whether the singular strategy

is a CSS or a repellor depends on the curvature of T. Given that there exists an optimization principle, case I and case II directly follow from Proposition 7.

Case III corresponds to Theorem 2: particularly, an evolutionarily singular point is a branching point if the trade-off curve is locally slightly convex. Necessary assumptions are however different: A and not B is a special case of the condition that all evolving loops must be pairwise connected: if entries in at least two different rows are evolving (criterion A) and if those entries must not be located in different columns (not criterion B), then all entries correspond to outgoing loops of one state i and thereby to pairwise connected loops through i. Then, by Theorem 1, the invasion boundary is linear and Theorem 2 applies. That is, Theorems 1 and 2 taken together give more general results, which are ultimately based on the loop structure of the life cycle. Note, that case IV has not been considered in this thesis.

Altogether, we conclude, that the classification via loops rather than via properties of the life-cycle matrix in terms of the criteria A, B and C, provides a more general picture of the evolutionary dynamics in structured populations.

The analysis presented in this thesis can be extended in multiple ways. First, higher-dimensional trait spaces could be considered. Second, it could be checked whether redefining loops as products of demographic parameters and regulatory functions as opposed to products of matrix entries simplifies the results. Third, it would be desirable to find *necessary* conditions for optimization principles to exist, which are based on the loop configuration of the life cycle at hand (but see Theorem 2). Fourth, one could investigate conditions for evolutionary branching in the case of multi-trait loops or unconnected single-trait loops, thereby accounting for case IV in Hoyle and Bowers (2008, see also Table 4 on the preceding page), in order to shed more light on the diversification of structured populations.

A   Elementary combinatorics

A.1   Graphs

Definition A.1. The *directed graph* is defined as an ordered pair $G := (S, E)$ such that $E \subseteq S \times S$. If $T \subseteq S$ and $F \subseteq E$, then $H := (S, E)$ is a *subgraph* of G. Similarly, for two directed graphs $G = (S, E)$ and $H = (T, F)$ we set $G \cap H := (S \cap T, E \cap F)$ and $G \cup H := (S \cup T, E \cup F)$. If $S \cap T$ is empty, then G and H are *disjoint*. A *weighted* directed graph is a triple $G_w := (S, E, w)$, where w is a function from the set of edges to $\mathbb{R}$.

Definition A.2. *A* path *is a directed graph* $P = (S, E)$ *with*

$$S = \{s_1, ..., s_l\},\ E = \{(s_1, s_2), ..., (s_{l-2}, s_{l-1}), (s_{l-1}, s_l)\}.$$

A loop *is a path* $C = (S, E)$ *with*

$$S = \{s_1, ..., s_l\},\ E = \{(s_1, s_2), ..., (s_{l-1}, s_l), (s_l, s_1)\}$$

and $|S| = |E|$, *i.e., no state is entered more than once. The number of states in S is called the* length *of the path. A* self loop *is a loop of length 1.*

Since in a loop no state is entered more than once, we can use the simpler notation

$$C = (S, E) = (\{s_1, ..., s_l\}, \{(s_1, s_2), ..., (s_{l-1}, s_l), (s_l, s_1)\}) =: [s_1, s_2, ..., s_l, s_1].$$

A.2   Permutation cycles

This appedix introduces permutation cycles, which function as analogues to unconnected loops in life-cycle graphs.

Definition A.3. A *permutation* of a set $S = \{1, ..., n\}$ is a bijective function $\pi : S \to S$. We use the notation

$$\pi = \left(\begin{array}{ccccc} 1 & 2 & 3 & \ldots & n \\ \pi(1) & \pi(2) & \pi(3) & \ldots & \pi(n) \end{array} \right).$$

The set of possible permutations of a set $\{1, ..., n\}$ is labelled S_n. An *inversion* in a permutation π is a pair (i, j) with $i < j$ and $\pi(i) > \pi(j)$. The *sign* of a permutation π, $\mathrm{sgn}\pi$, is the number of its inversions. A *permutation cycle* is a permutation of the form

$$c = \left(\begin{array}{ccccc} s_1 & s_2 & s_3 & \ldots & s_l \\ s_2 & s_3 & s_4 & \ldots & s_1 \end{array} \right),$$

where l is called the *length* of the permutation cycle. Each permutation can be written as a *composition of disjoint permutation cycles* $\pi = c_1 \ldots c_k$. If $\pi = c_1 \ldots c_k$ is a composition of disjoint permutation cycles and $c = c_i$ for some $i \in \{1, ..., k\}$, we write $c \trianglelefteq \pi$. We define $z(\pi)$ as the number of cycles in the composition of disjoint permutation cycles of π.

If S is the set of states in a given life cycle, it is clear from the previous definition that there is a one-to-one relation between the set of *pairwise disconnected* loops in the

corresponding life-cycle graph and the set of permutation cycles of a permutation on S: to a loop $G = (\{s_1, ..., s_l\}, \{(s_1, s_2), ..., (s_l, s_1)\})$ there corresponds a permutation cycle

$$c := \left(\begin{array}{ccccc} s_1 & s_2 & s_3 & \dots & s_l \\ s_2 & s_3 & s_4 & \dots & s_1 \end{array} \right).$$

Conversely, given c we define $G = (\{s_1, ..., s_l\}, \{(s_1, c(s_1)), ..., (s_l, c(s_l))\})$. Similarly, by setting $L = a_{s_1 c(s_1)}...a_{s_l c(s_l)}$, there is a one-to-one relation between permutation cycles and loop transmissions.

This identification allows us to use results about permutations for the investigation of the loop structure of given life cycles. In particular we will use the following proposition (for a proof, see for instance Cameron, 1994, ch. 5):

Proposition A.1. *Let $\pi \in S_n$ and c be a cycle with length $\ell(c)$. Then*

1. $\operatorname{sgn}(\pi) = (-1)^{n-z(\pi)}$ *and*

2. $\operatorname{sgn}(c) = (-1)^{\ell(c)-1}$.

B   Matrix analysis

Matrices that are used to describe structured populations usually have a special structure. In discrete time models of structured populations, such matrices are nonnegative, since all demographic parameters (if appearing in the entries of the matrix) are positive. Nonnegative matrices have been well studied: for a survey on nonnegative matrices see the reference works Berman and Plemmons (1979) and Horn and Johnson (1985; 1991). As it is defined in section 2.1, a life-cycle matrix that models the dynamics of structured populations in continuous time shows a slightly different structure. In the so-called essentially nonnegative matrices negative entries may appear, albeit only in diagonal entries. This is clear, as in a life-cycle matrix the diagonal entries comprise explicit death rates and outgoing transition rates in contrast to its analogue in the discrete time case.

However, much of the theory of nonnegative matrices can be translated into results about essentially nonnegative matrices. This section collects some of the basic results for each of the two types.

Nonnegative matrices: We first state basic definitions and standard results for general nonnegative square matrices. For a matrix A we write $\mathsf{A} \geq 0$ (or $\leq, >, <$) if for all of its entries we have $a_{ij} \geq 0$ (or $\leq, >, <$, respectively).

Definition B.1. The *spectral radius* of a $k \times k$ square matrix A with eigenvalues $\lambda_1, ..., \lambda_k$ is defined as $\rho(\mathsf{A}) := \max_{1 \leq i \leq k}(|\lambda_i|)$.

Proposition B.1. *Let A be a nonnegative square matrix and for a proper subset of indices α, let A_α be the principal submatrix of A. Then $\rho(\mathsf{A}_\alpha) < \rho(\mathsf{A})$.*

Proof. See Varga (2000, p. 35, Lemma 2.6). $\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proposition B.2. Perron-Frobenius Theorem for nonnegative matrices. *Let A be a nonnegative irreducible $k \times k$ matrix. Then*

1. A *has a positive real eigenvalue $\lambda_0 > 0$ with $\lambda_0 = \rho(\mathsf{A})$.*

2. $\rho(\mathsf{A})$ is a simple eigenvalue of A.

3. There is a positive eigenvector $\boldsymbol{u}_0 > 0$ corresponding to λ_0.

4. $\frac{\partial \rho(\mathsf{A})}{\partial a_{ij}} > 0$ for all $1 \leq i, j \leq k$.

Proof. See Varga (2000, p. 35, Theorem 2.7). □

Essentially nonnegative matrices: As the diagonal entries of a life-cycle matrix can be negative, we have to reconsider the previously stated propositions. In general, the results about nonnegative matrices can be translated into quite similar propositions about matrices with negative diagonal entries.

Definition B.2. A $k \times k$ matrix A is called *essentially nonnegative* if $a_{ij} \geq 0$ for all $i \neq j$, $i, j \in \{1, ..., k\}$.

It is clear that a life-cycle matrix given by $a_{ij} = f_{ij} + t_{ij} - \delta_{ij}(d_i + \sum_{l=1}^{k} t_{li})$ fulfills the criteria for being essentially nonnegative (this holds even if the population is not in its population dynamic equilibrium).

Definition B.3. The *stability modulus* of a $k \times k$ square matrix A with eigenvalues $\lambda_1, ..., \lambda_k$ is defined as $\mu(\mathsf{A}) := \max_{1 \leq i \leq k} (\mathrm{Re}\lambda_i)$.

Proposition B.3. Perron-Frobenius Theorem for essentially nonnegative matrices. *Let A be an essentially nonnegative irreducible $k \times k$ matrix. Then*

1. A has a real eigenvalue $\lambda_0 = \mu(\mathsf{A})$.

2. $\mu(\mathsf{A})$ is a simple eigenvalue of A.

3. There is a positive eigenvector $\boldsymbol{u}_0 > 0$ corresponding to λ_0.

4. $\frac{\partial \mu(\mathsf{A})}{\partial a_{ij}} > 0$ for all $i, i \in \{1, ..., k\}$.

Proof. See Varga (2000, 282, Theorem 8.3). □

Proposition B.4. *Let A be an essentially nonnegative, irreducible square matrix and for a proper subset of indices α, let A_α be the principal submatrix of A. Then $\mu(\mathsf{A}_\alpha) < \mu(\mathsf{A})$.*

Proof. Since A is essentially nonnegative, there exists a real $s > 0$ such that $\mathsf{B} = \mathsf{A} + s\mathsf{I}_k$, where I_k denotes the k-dimensional identity matrix, is nonnegative. The Perron-Frobenius theorem for nonnegative matrices implies that there exists a vector $\boldsymbol{u}$ such that

$$\mathsf{B}\boldsymbol{u} = \rho(\mathsf{B})\boldsymbol{u}$$
$$\Leftrightarrow \quad (\mathsf{A} + s\mathsf{I}_k)\boldsymbol{u} = \rho(\mathsf{A} + s\mathsf{I}_k)\boldsymbol{u}$$
$$\Leftrightarrow \quad \mathsf{A}\boldsymbol{u} = \underbrace{(\rho(\mathsf{A} + s\mathsf{I}_k) - s)}_{=:\mu(\mathsf{A})}\boldsymbol{u} = \mu(\mathsf{A})\boldsymbol{u}$$

which defines the stability modulus of the Perron-Frobenius Theorem for essentially nonnegative matrices.

Set $l = |\alpha|$ and $\mathsf{B}_\alpha := \mathsf{A}_\alpha + s\mathsf{I}_l$. analogously we can define $\rho(\mathsf{A}_\alpha + s\mathsf{I}_l) - s =: \mu(\mathsf{A}_\alpha)$. Clearly, B_α is nonnegative and a principal submatrix of B. From Proposition B.1 then follows that $\rho(\mathsf{B}_\alpha) < \rho(\mathsf{B})$. Hence, $\mu(\mathsf{A}_\alpha) = \rho(\mathsf{A}_\alpha + s\mathsf{I}_l) - s = \rho(\mathsf{B}_\alpha) - s < \rho(\mathsf{B}) - s = \rho(\mathsf{A} + s\mathsf{I}_k) - s = \mu(\mathsf{A})$, which proves the proposition. □

C   Proofs

In this appendix, the proofs of the two main results, Theorem 1 and Theorem 2 shall be given.

C.1   Proof of Theorem 1

The first steps in this proof are completely analogous to the proof of Proposition 18 in Rueffler et al. (2013). Since Q_A is a fitness proxy, we have that the invasion boundary fulfills $Q_A((x_1, I(x_1)), \hat{E}) = s((x_1, I(x_1)), \hat{E}) = 0$. We differentiate twice for x_1 yielding

$$\frac{\partial^2}{\partial x_1^2} Q_A + 2\frac{\partial}{\partial x_1} I \cdot \frac{\partial^2}{\partial x_1 \partial x_2} Q_A + \frac{\partial^2}{\partial x_2^2} Q_A \cdot \left(\frac{\partial}{\partial x_1} I\right)^2 + \frac{\partial}{\partial x_2} Q_A \cdot \frac{\partial^2}{\partial x_1^2} I = 0.$$

Recall, that $\frac{\partial^2}{\partial x_1^2} Q_A = \frac{\partial^2}{\partial x_2^2} Q_A = 0$. After some rearrangement we then obtain

$$\frac{\partial^2}{\partial x_1^2} I = -\frac{2\frac{\partial}{\partial x_1} I \cdot \frac{\partial^2}{\partial x_1 \partial x_2} Q_A}{\frac{\partial}{\partial x_2} Q_A}. \tag{18}$$

Let us consider all parts of the expression on the right-hand side separately. First, we have $\frac{\partial}{\partial x_1} I < 0$ for all x_1 which are increasing birth rates f_{ij} and decreasing death rates d_i. Second, Proposition 19 gives us conditions for the sign of the denominator $\frac{\partial}{\partial x_2} Q_A$.

Third, we calculate the sign of $\frac{\partial^2}{\partial x_1 \partial x_2} Q_A$. This requires some additional considerations:

Case 1. Let x_1 affect either f_{ij} or d_j, and let x_2 affect either f_{kl} or d_l for some $i, j, k, l \in \{1, ..., n\}$. From Proposition 19.1 and 2 then follows, that $\frac{\partial}{\partial x_2} Q_A > 0$, as x_2 is assumed to increases f_{ij} and decreases d_i.

If all evolving loops are single trait loops and pairwise connected, x_1 and x_2 do not appear together in one summand of Q_A. Hence $\frac{\partial^2}{\partial x_1 \partial x_2} Q_A = 0$ and I is linear. If there is only a multi trait loop of length 1, it is necessarily of the form $L = a_{ii} = f_{ii} - d_i - \sum t_{hi}$ where f_{ii} and d_i are evolving. Then clearly also $\frac{\partial^2}{\partial x_1 \partial x_2} Q_A = 0$ (by multilinearity split Q_A into two summands).

If the set $\mathcal{M}(x_1, x_2)$ of evolving multi trait loops of length $\ell > 1$ is non-empty, i.e., $\mathcal{M}(x_1, x_2) := \mathcal{L}(x_1, x_2) \backslash \{L | \ell(L) = 1\} \neq \emptyset$ and all evolving loops are pairwise connected, we apply Lemma 1 and get

$$\begin{aligned}
\frac{\partial^2}{\partial x_1 \partial x_2} Q_A &= -\frac{\partial^2}{\partial x_1 \partial x_2} \sum_{L \in \mathcal{M}(x_1, x_2)} L(x_1, x_2) \cdot Q_{A \backslash \{L\}} + 0 \\
&= -\sum_{L \in \mathcal{M}(x_1, x_2)} \frac{\partial^2 x_1 x_2}{\partial x_1 \partial x_2} \cdot K_1 \cdot Q_{A \backslash \{L\}}
\end{aligned}$$

for some $K_1 > 0$. Since here we are only interested in Q_A near the singular point $\boldsymbol{x}^*$ we get from Proposition 19.1 that $Q_{A \backslash \{L\}} < 0$. Then $\frac{\partial^2}{\partial x_1 \partial x_2} Q_A > 0$ and the invasion boundary is locally convex.

If all loops are single trait loops and the set of pairs of unconnected evolving loops $\mathcal{L}(x_1) \otimes \mathcal{L}(x_2)$ is non-empty, we apply Lemma 1 and find that

$$\frac{\partial^2}{\partial x_1 \partial x_2} Q_{\mathsf{A}} = \frac{\partial^2}{\partial x_1 \partial x_2} \sum_{(L,M) \in \mathcal{L}(x_1) \otimes \mathcal{L}(x_2)} L(x_1) \cdot M(x_2) \cdot Q_{\mathsf{A} \setminus \{L\} \cup \{M\}} + 0$$

$$= \sum_{(L,M) \in \mathcal{L}(x_1) \otimes \mathcal{L}(x_2)} \frac{\partial^2 x_1 x_2}{\partial x_1 \partial x_2} \cdot K_2 \cdot Q_{\mathsf{A} \setminus \{L\}} < 0$$

since $K_2 > 0$ is a constant depending solely on positive factors. Thus, the invasion boundary is concave.

Case 2. Let x_1 affect t_{ij} and x_2 affect either f_{kl} or d_l. Again, from 19.1 and 2 we get $\frac{\partial}{\partial x_2} Q_{\mathsf{A}} > 0$.

Let all evolving loops be single trait loops and pairwise connected. Since t_{ij} generates at least two connected evolving loops, one of which is the length-1 loop $L = a_{jj} = f_{jj} - d_j - t_{ij} - \sum_{h \neq i,j} t_{hj}$, this entails that all loops in $\mathcal{L}(x_1)$ pass through j. Then x_1 and x_2 do not appear together in one summand of Q_{A} and $\frac{\partial^2}{\partial x_1 \partial x_2} Q_{\mathsf{A}} = 0$, i.e., the invasion boundary is linear. If there is a multi trait loop of length 1, it is of the form $L = a_{jj} = f_{jj} - d_j - t_{ij} - \sum_{h \neq i,j} t_{hj}$ where t_{ij} and f_{jj}, or t_{ij} and d_i are evolving. Hence, again $\frac{\partial^2}{\partial x_1 \partial x_2} Q_{\mathsf{A}} = 0$.

Now, let the set of evolving multi trait loops with length $\ell > 1$, $\mathcal{M}(x_1, x_2)$, be non-empty and let all evolving loops be pairwise connected. Again, this entails that all $\mathcal{L}(x_2)$ pass through j. As in *Case 1* we obtain $\frac{\partial^2}{\partial x_1 \partial x_2} Q_{\mathsf{A}} > 0$. Since the sign of $\frac{\partial Q_{\mathsf{A}}}{\partial x_1}$ is a priori not known, we have that I is locally convex if $\frac{\partial Q_{\mathsf{A}}}{\partial x_1} > 0$ and concave if $\frac{\partial Q_{\mathsf{A}}}{\partial x_1} < 0$.

If all evolving loops are single trait loops and there are unconnected evolving loops, i.e., $\mathcal{L}(x_1) \otimes \mathcal{L}(x_2) \neq \emptyset$ and $\mathcal{L}(-x_1) \otimes \mathcal{L}(x_2) \neq \emptyset$, we can expand the fitness proxy in the following way:

$$\frac{\partial^2}{\partial x_1 \partial x_2} Q_{\mathsf{A}} = \frac{\partial^2}{\partial x_1 \partial x_2} \sum_{(L,M) \in \mathcal{L}(x_1) \otimes \mathcal{L}(x_2)} L(x_1) \cdot M(x_2) \cdot Q_{\mathsf{A} \setminus \{L\} \cup \{M\}}$$

$$+ \frac{\partial^2}{\partial x_1 \partial x_2} \sum_{(L,M) \in \mathcal{L}(-x_1) \otimes \mathcal{L}(x_2)} L(-x_1) \cdot M(x_2) \cdot Q_{\mathsf{A} \setminus \{L\} \cup \{M\}}$$

Here, the first part contains all pairs of loops $(L(x_1), M(x_2)) \in \mathcal{L}(x_1) \otimes \mathcal{L}(x_2)$, where $\ell(L(x_1)) > 1$. The second part contains all pairs of loops where the first loop has length 1. In these loops t_{ij} has a negative sign, which is denoted by $L(-x_1)$. Note, that since there are no multi-trait loops, there are in particular no multi-trait loops of length 1. Hence, for $L \in \mathcal{L}(-x_1)$, we have $\frac{\partial}{\partial x_1} L(-x_1) = -\frac{\partial}{\partial x_1} x_1 \cdot \text{const} + 0$, since all other summands in $L = a_{jj}$ vanish. Thus, we can find positive constants K_3 and K_4 such that

$$\frac{\partial^2}{\partial x_1 \partial x_2} Q_{\mathsf{A}} = \sum_{(L,M) \in \mathcal{L}(x_1) \otimes \mathcal{L}(x_2)} \frac{\partial^2 x_1 x_2}{\partial x_1 \partial x_2} \cdot K_3 \cdot Q_{\mathsf{A} \setminus \{L\} \cup \{M\}}$$

$$- \sum_{(L,M) \in \mathcal{L}(x_1) \otimes \mathcal{L}(x_2)} \frac{\partial^2 x_1 x_2}{\partial x_1 \partial x_2} \cdot K_4 \cdot Q_{\mathsf{A} \setminus \{L\} \cup \{M\}}$$

where both parts are positive. From this follows, that we cannot predict the curvature of I without explicit calculation of the difference

$$\triangle := \sum_{(L,M)\in\mathcal{L}(x_1)\otimes\mathcal{L}(x_2)} K_3 \cdot Q_{\mathsf{A}\setminus\{L\}\cup\{M\}} - \sum_{(L,M)\in\mathcal{L}(x_1)\otimes\mathcal{L}(x_2)} K_4 \cdot Q_{\mathsf{A}\setminus\{L\}\cup\{M\}}.$$

Case 3. Let x_1 affect t_{ij} and x_2 affect either t_{kl}. The case of pairwise connected single trait loops is not possible, since both t_{ij} and t_{kl} generate loops of length 1: x_1 and x_2 necessarily appear in the multi trait loop $L = a_{jj} = f_{jj} - d_j - t_{ij} - t_{kl} - \sum t_{hj}$ where $l = j$. Clearly, L has length 1. In this case, as in *Case 1* and 2 we find that $\frac{\partial^2}{\partial x_1 \partial x_2} Q_{\mathsf{A}} = 0$, hence the invasion boundary is linear, as long as $\frac{\partial}{\partial x_2} Q_{\mathsf{A}} \neq 0$, and undefined otherwise.

Also the case of multi trait loops with length $\ell > 1$ and pairwise connected single trait loops is not possible: otherwise, we would have two single trait loops $L_1(-t_{ij}) = a_{jj}$ and $L_2(-t_{kl}) = a_{ll}$ which are only connected if $j = l$. But then a_{jj} constitues a multi trait loop of length 1.

Finally, let $\mathcal{L}(\pm x_1)\otimes\mathcal{L}(\pm x_2) \neq \emptyset$ for all possible combinations of signs. Expanding the fitness proxy, we find

$$\begin{aligned}
\frac{\partial^2}{\partial x_1 \partial x_2} Q_{\mathsf{A}} &= \frac{\partial^2}{\partial x_1 \partial x_2} \sum_{(L,M)\in\mathcal{L}(x_1)\otimes\mathcal{L}(x_2)} L(x_1) \cdot M(x_2) \cdot Q_{\mathsf{A}\setminus\{L\}\cup\{M\}} \\
&+ \frac{\partial^2}{\partial x_1 \partial x_2} \sum_{(L,M)\in\mathcal{L}(-x_1)\otimes\mathcal{L}(x_2)} L(-x_1) \cdot M(x_2) \cdot Q_{\mathsf{A}\setminus\{L\}\cup\{M\}} \\
&+ \frac{\partial^2}{\partial x_1 \partial x_2} \sum_{(L,M)\in\mathcal{L}(x_1)\otimes\mathcal{L}(-x_2)} L(x_1) \cdot M(-x_2) \cdot Q_{\mathsf{A}\setminus\{L\}\cup\{M\}} \\
&+ \frac{\partial^2}{\partial x_1 \partial x_2} \sum_{(L,M)\in\mathcal{L}(-x_1)\otimes\mathcal{L}(-x_2)} L(x_1) \cdot M(x_2) \cdot Q_{\mathsf{A}\setminus\{L\}\cup\{M\}}
\end{aligned}$$

and further

$$\begin{aligned}
&= \sum_{(L,M)\in\mathcal{L}(x_1)\otimes\mathcal{L}(x_2)} K_5 \cdot Q_{\mathsf{A}\setminus\{L\}\cup\{M\}} \\
&- \sum_{(L,M)\in\mathcal{L}(-x_1)\otimes\mathcal{L}(x_2)} K_6 \cdot Q_{\mathsf{A}\setminus\{L\}\cup\{M\}} \\
&- \sum_{(L,M)\in\mathcal{L}(x_1)\otimes\mathcal{L}(-x_2)} K_7 \cdot Q_{\mathsf{A}\setminus\{L\}\cup\{M\}} \\
&+ \sum_{(L,M)\in\mathcal{L}(-x_1)\otimes\mathcal{L}(-x_2)} K_8 \cdot Q_{\mathsf{A}\setminus\{L\}\cup\{M\}}
\end{aligned}$$

for some constants $K_5, ..., K_8 > 0$, whereby the curvature of I is a priori undetermined.

C.2   Proof of Theorem 2

All evolving loops are pairwise connected single trait loops or multi-trait loops of length one. Hence, from Corollary 2 we know that $\boldsymbol{x}^*$ is uninvadible if the trade-off is locally concave, i.e., $\mathrm{d}^2T/\mathrm{d}x_1^2|_* < 0$, and invadible if the trade-off is locally convex,

i.e., $\mathrm{d}^2 T/\mathrm{d}x_1^2|_* > 0$. So, in order to fully classify the singular point, we have to check for convergence stability (see 3.2). Since locally around the singular point invasion fitness can be replaced by the fitness proxy Q_{A}, we have to evaluate

$$\frac{\partial^2}{\partial \xi^2} Q_{\mathsf{A}}(\xi, \eta)|_* + \frac{\partial^2}{\partial \xi \partial \eta} Q_{\mathsf{A}}(\xi, \eta)|_* =: \triangle.$$

If $\triangle < 0$, then $\boldsymbol{x}^*$ is convergence stable, and if $\triangle > 0$, then $\boldsymbol{x}^*$ is repelling.

Let us collect all evolving (and by assumption pairwise connected) loops in a set $\mathcal{E}$. In the following we will make use of Lemma 1 in order to simplify $\triangle$. We will need some additional

Notation. Let $\mathcal{M}(x_i \cdot R_i)$ denote the set of loops containing x_i and the corresponding multiplicative regulatory function, let $\mathcal{M}(x_i + R_i)$ denote the set of loops containing x_i and the corresponding additive regulatory function, and let $\mathcal{M}(x_i \backslash R_i)$ denote the set of loops containing x_i and no regulatory function. We will assume that x_i is positive, i.e., we have $\mathcal{M}(-x_i \cdot R_i)$, $\mathcal{M}(-x_i - R_i)$, and $\mathcal{M}(-x_i \backslash R_i)$ for denoting the sets of loops affected by x_i in a negative way (in particular, this is the case for self loops affected by outgoing transition rates). For $\star \in \{\cdot, +, \backslash\}$ we use the notation

$$\sum_{\mathcal{M}(x_i \star R_i)} = -\sum_{L \in \mathcal{M}(x_i \star R_i)} L \cdot Q_{\mathsf{A} \backslash \{L\}}$$

$$\sum_{\mathcal{M}(x_i \star R_i)} \times = -\sum_{L \in \mathcal{M}(x_i \star R_i)} L/(x_i \star R_i + \mathrm{C}) \cdot Q_{\mathsf{A} \backslash \{L\}}$$

where C denotes some (not necessarily positive) constant not depending on any trait or regulatory function. Furthermore, for the sake of clarity we use $\bigcirc$ instead of $\sum$, if there is only one summand which in addition is a self loop, i.e., a loop of length 1. Finally, let $R(\boldsymbol{R})$ denote some expression depending on regulatory functions but not on evolving traits.

We have $\sum_{\mathcal{M}(x_i \star R_i)} \times > 0$ and $\bigcirc_{\mathcal{M}(x_i \star R_i)} \times > 0$ since the fitness proxy is evaluated sufficiently close to the singular point and since $L/(x_i \star R_i + \mathrm{C}) > 0$. By Lemma 1 we can write

$$Q_{\mathsf{A}}(\xi, \eta) = \sum_{\mathcal{E}} + R(\mathbf{R}) = \sum_{\mathcal{M}_1} + ... + \sum_{\mathcal{M}_k} + R(\mathbf{R})$$

for any partition $\mathcal{M}_1, ..., \mathcal{M}_k$ of $\mathcal{E}$. In the following case-by-case analysis $\mathcal{E}$ will be decomposed into suitable subsets.

Case 1. Assume that only death and fertility rates are evolving and that R_1, R_2 are multiplicative regulatory functions. Let $\mathcal{E} = \mathcal{M}(x_1 \cdot R_1) \cup \mathcal{M}(x_2 \cdot R_2) \cup \mathcal{M}(x_1 \backslash R_1) \cup \mathcal{M}(x_2 \backslash R_2)$. Then

$$
\begin{aligned}
Q_{\mathsf{A}}(\xi, \eta) &= \sum_{\mathcal{M}(x_1 \cdot R_1)} + \sum_{\mathcal{M}(x_2 \cdot R_2)} + \sum_{\mathcal{M}(x_1 \backslash R_1)} + \sum_{\mathcal{M}(x_2 \backslash R_2)} + R(\mathbf{R}) \\
&= (x_1 R_1 + \mathrm{C}) \sum_{\mathcal{M}(x_1 \cdot R_1)} \times + (x_2 R_2 + \mathrm{C}) \sum_{\mathcal{M}(x_2 \cdot R_2)} \times \\
&\quad + (x_1 + \mathrm{C}) \sum_{\mathcal{M}(x_1 \backslash R_1)} \times + (x_2 + \mathrm{C}) \sum_{\mathcal{M}(x_2 \backslash R_2)} \times + R(\mathbf{R})
\end{aligned}
$$

70

and

$$\triangle \;=\; \frac{\mathrm{d}^2 x_1}{\mathrm{d}\xi^2} R_1 \bigg|_* \sum_{\mathcal{M}(x_1 \cdot R_1)} \times \; + \frac{\mathrm{d}^2 x_2}{\mathrm{d}\xi^2} R_2 \bigg|_* \sum_{\mathcal{M}(x_2 \cdot R_2)} \times \; + \frac{\mathrm{d}^2 x_1}{\mathrm{d}\xi^2} \bigg|_* \sum_{\mathcal{M}(x_1 \backslash R_1)} \times$$

$$+ \frac{\mathrm{d}^2 x_2}{\mathrm{d}\xi^2} \bigg|_* \sum_{\mathcal{M}(x_2 \backslash R_2)} \times \; + \frac{\mathrm{d}x_1}{\mathrm{d}\xi} \frac{\mathrm{d}R_1}{\mathrm{d}\eta} \bigg|_* \sum_{\mathcal{M}(x_1 \cdot R_1)} \times \; + \frac{\mathrm{d}x_2}{\mathrm{d}\xi} \frac{\mathrm{d}R_2}{\mathrm{d}\eta} \bigg|_* \sum_{\mathcal{M}(x_2 \cdot R_2)} \times \; .$$

If T is concave, i.e., $\mathrm{d}^2 T/\mathrm{d}x_1^2|_* < 0$, then $x_i(\xi)$ is concave for $i = 1, 2$, whereby all summands containing double derivatives are negative. Under the conditions in Theorem 2.1. all mixed derivatives are negative, since $\mathrm{d}x_1/\mathrm{d}\xi > 0$ and $\mathrm{d}x_2/\mathrm{d}\xi < 0$. Thus, $\triangle < 0$ and $\boldsymbol{x}^*$ is convergence stable, and hence a CSS. If T is convex, then all summands containing double derivatives are positive. The summands containing mixed derivatives remain negative. That is, there exists a threshold $k > 0$ such that $\triangle < 0$, i.e., convergence stable, for $k > \mathrm{d}^2 T/\mathrm{d}x_1^2|_* > 0$, and $\triangle > 0$, i.e., repelling, for $\mathrm{d}^2 T/\mathrm{d}x_1^2|_* > k$. Thus, $\boldsymbol{x}^*$ is an evolutionary branching point in the former case and an evolutionary repellor in the latter case.

Note, that with some algebra, one can easily show that the case of linear combinations of the form $\mathcal{M}(x_1 \cdot R_1 + x_2 \cdot R_2)$ leads to precisely the same result.

Case 2. Assume that only death and fertility rates are evolving, and that R_1, R_2 are multiplicative regulatory functions. Let $\mathcal{E} = \mathcal{M}(x_1 + R_1) \cup \mathcal{M}(x_2 + R_2) \cup \mathcal{M}(x_1 \backslash R_1) \cup \mathcal{M}(x_2 \backslash R_2)$. Then

$$\begin{aligned}
Q_\mathsf{A}(\xi, \eta) &= \sum_{\mathcal{M}(x_1 + R_1)} + \sum_{\mathcal{M}(x_2 + R_2)} + \sum_{\mathcal{M}(x_1 \backslash R_1)} + \sum_{\mathcal{M}(x_2 \backslash R_2)} + R(\mathbf{R}) \\
&= (x_1 + R_1 + \mathrm{C}) \sum_{\mathcal{M}(x_1 + R_1)} \times \; + (x_2 + R_2 + \mathrm{C}) \sum_{\mathcal{M}(x_2 + R_2)} \times \\
&\quad + (x_1 + \mathrm{C}) \sum_{\mathcal{M}(x_1 \backslash R_1)} \times \; + (x_2 + \mathrm{C}) \sum_{\mathcal{M}(x_2 \backslash R_2)} \times \; + R(\mathbf{R})
\end{aligned}$$

The assertion then follows from an optimization argument: it is easy to see that the fitness proxy can be rearranged as $Q_\mathsf{A}(\xi, \eta) = a(\boldsymbol{x}) + b(\boldsymbol{R})$. By Proposition 11, there exists an optimization principle. Hence, the evolutionarily singular strategy is convergence stable iff it is uninvadible, i.e., it is a CSS if $\mathrm{d}^2 T/\mathrm{d}x_1^2|_* < 0$ and a repellor if $\mathrm{d}^2 T/\mathrm{d}x_1^2|_* > 0$.

Case 3. Assume that only death and fertility rates are evolving, and that R_1 is a multiplicative and R_2 an additive regulatory function. Let $\mathcal{E} = \mathcal{M}(x_1 \cdot R_1) \cup \mathcal{M}(x_2 + R_2) \cup \mathcal{M}(x_1 \backslash R_1) \cup \mathcal{M}(x_2 \backslash R_2)$. Then

$$\begin{aligned}
Q_\mathsf{A}(\xi, \eta) &= \sum_{\mathcal{M}(x_1 \cdot R_1)} + \sum_{\mathcal{M}(x_2 + R_2)} + \sum_{\mathcal{M}(x_1 \backslash R_1)} + \sum_{\mathcal{M}(x_2 \backslash R_2)} + R(\mathbf{R}) \\
&= (x_1 R_1 + \mathrm{C}) \sum_{\mathcal{M}(x_1 \cdot R_1)} \times \; + (x_2 + R_2 + \mathrm{C}) \sum_{\mathcal{M}(x_2 + R_2)} \times \\
&\quad + (x_1 + \mathrm{C}) \sum_{\mathcal{M}(x_1 \backslash R_1)} \times \; + (x_2 + \mathrm{C}) \sum_{\mathcal{M}(x_2 \backslash R_2)} \times \; + R(\mathbf{R})
\end{aligned}$$

and

$$\triangle = \left.\frac{\mathrm{d}^2 x_1}{\mathrm{d}\xi^2} R_1\right|_* \sum_{\mathcal{M}(x_1 \cdot R_1)}\times + \left.\frac{\mathrm{d}^2 x_2}{\mathrm{d}\xi^2}\right|_* \sum_{\mathcal{M}(x_2 + R_2)}\times + \left.\frac{\mathrm{d}^2 x_1}{\mathrm{d}\xi^2}\right|_* \sum_{\mathcal{M}(x_1 \setminus R_1)}\times$$

$$+ \left.\frac{\mathrm{d}^2 x_2}{\mathrm{d}\xi^2}\right|_* \sum_{\mathcal{M}(x_2 \setminus R_2)}\times + \left.\frac{\mathrm{d} x_1}{\mathrm{d}\xi}\frac{\mathrm{d} R_1}{\mathrm{d}\eta}\right|_* \sum_{\mathcal{M}(x_1 \cdot R_1)}\times \ .$$

Hence, the same results as in *Case 1* apply, but we only need $\mathrm{d}R_1/\mathrm{d}\eta < 0$.

Case 4. We now turn over to combinations of transition rates, and fertility or death rates as evolving traits. Assume that x_1 is affecting a transition rate, that x_2 is affecting a fertility or death rate and that both regulatory functions are multiplicative. Then

$$Q_\mathsf{A}(\xi,\eta) = \sum_{\mathcal{M}(x_1 \cdot R_1)} + \bigcirc_{\mathcal{M}(-x_1 \cdot R_1)} + \sum_{\mathcal{M}(x_2 \cdot R_2)} + \sum_{\mathcal{M}(x_1 \setminus R_1)}$$

$$+ \bigcirc_{\mathcal{M}(-x_1 \setminus R_1)} + \sum_{\mathcal{M}(x_2 \setminus R_2)} + R(\mathbf{R})$$

$$= (x_1 R_1 + \mathrm{C})\left[\sum_{\mathcal{M}(x_1 \cdot R_1)}\times - \bigcirc_{\mathcal{M}(-x_1 \cdot R_1)}\times\right] + (x_2 R_2 + \mathrm{C})\sum_{\mathcal{M}(x_2 \cdot R_2)}\times$$

$$+ (x_1 + \mathrm{C})\left[\sum_{\mathcal{M}(x_1 \setminus R_1)}\times - \bigcirc_{\mathcal{M}(-x_1 \setminus R_1)}\times\right] + (x_2 + \mathrm{C})\sum_{\mathcal{M}(x_2 \setminus R_2)}\times + R(\mathbf{R})$$

and

$$\triangle = \left.\frac{\mathrm{d}^2 x_1}{\mathrm{d}\xi^2} R_1\right|_* \left[\sum_{\mathcal{M}(x_1 \cdot R_1)}\times - \bigcirc_{\mathcal{M}(-x_1 \cdot R_1)}\times\right] + \left.\frac{\mathrm{d}^2 x_2}{\mathrm{d}\xi^2} R_2\right|_* \sum_{\mathcal{M}(x_2 \cdot R_2)}\times$$

$$+ \frac{\mathrm{d}^2 x_1}{\mathrm{d}\xi^2}\left[\sum_{\mathcal{M}(x_1 \setminus R_1)}\times - \bigcirc_{\mathcal{M}(-x_1 \setminus R_1)}\times\right] + \left.\frac{\mathrm{d}^2 x_2}{\mathrm{d}\xi^2}\right|_* \sum_{\mathcal{M}(x_2 \setminus R_2)}\times$$

$$+ \left.\frac{\mathrm{d} x_1}{\mathrm{d}\xi}\frac{\mathrm{d} R_1}{\mathrm{d}\eta}\right|_* \left[\sum_{\mathcal{M}(x_1 \cdot R_1)}\times - \bigcirc_{\mathcal{M}(-x_1 \cdot R_1)}\times\right] + \left.\frac{\mathrm{d} x_2}{\mathrm{d}\xi}\frac{\mathrm{d} R_2}{\mathrm{d}\eta}\right|_* \sum_{\mathcal{M}(x_2 \cdot R_2)}\times$$

where $\bigcirc_{\mathcal{M}(-x_1 \cdot R_1)}\times \neq 0$ iff $\bigcirc_{\mathcal{M}(-x_1 \setminus R_1)}\times = 0$, since the self loop containing $-x_1$ is either regulated or unregulated, but not both. If $\bigcirc_{\mathcal{M}(-x_1 \cdot R_1)}\times > 0$ is sufficiently small so that

$$h := \sum_{\mathcal{M}(x_1 \cdot R_1)}\times > \bigcirc_{\mathcal{M}(-x_1 \cdot R_1)}\times = -Q_{\mathsf{A}\setminus\{L(-x_1 \cdot R_1)\}} > 0, \tag{19}$$

the same as in *Case 1* applies, and similarly if $\bigcirc_{\mathcal{M}(-x_1 \setminus R_1)} > 0$ in such a way that

$$h := \sum_{\mathcal{M}(x_1 \setminus R_1)}\times > \bigcirc_{\mathcal{M}(-x_1 \setminus R_1)}\times = -Q_{\mathsf{A}\setminus\{L(-x_1 \setminus R_1)\}} > 0. \tag{20}$$

72

Case 5. Assume that x_1 is affecting a transition rate, that x_2 is affecting a fertility or death rate and that both regulatory functions are additive. Then

$$
\begin{aligned}
Q_{\mathsf{A}}(\xi,\eta) \;=\;& \sum_{\mathcal{M}(x_1+R_1)} + \bigcirc_{\mathcal{M}(-x_1-R_1)} + \sum_{\mathcal{M}(x_2+R_2)} + \sum_{\mathcal{M}(x_1\backslash R_1)} \\
& + \bigcirc_{\mathcal{M}(-x_1\backslash R_1)} + \sum_{\mathcal{M}(x_2\backslash R_2)} + R(\mathbf{R}) \\
\;=\;& (x_1+R_1+\mathrm{C})\left[\sum_{\mathcal{M}(x_1+R_1)}\times \;-\; \bigcirc_{\mathcal{M}(-x_1-R_1)}\times\right] + (x_2+R_2+\mathrm{C})\sum_{\mathcal{M}(x_2+R_2)}\times \\
& + (x_1+\mathrm{C})\left[\sum_{\mathcal{M}(x_1\backslash R_1)}\times \;-\; \bigcirc_{\mathcal{M}(-x_1\backslash R_1)}\times\right] + (x_2+\mathrm{C})\sum_{\mathcal{M}(x_2\backslash R_2)}\times \;+\; R(\mathbf{R})
\end{aligned}
$$

and after some rearrangements we can apply an optimization argument as in *Case 2.* Hence, $\mathrm{d}^2T/\mathrm{d}x_1^2|_* < 0$ iff $\triangle < 0$.

Case 6. Assume that x_1 is affecting a transition rate, that x_2 is affecting a fertility or death rate and that R_1 is additive and R_2 is multiplicative. Then

$$
\begin{aligned}
Q_{\mathsf{A}}(\xi,\eta) \;=\;& \sum_{\mathcal{M}(x_1+R_1)} + \bigcirc_{\mathcal{M}(-x_1-R_1)} + \sum_{\mathcal{M}(x_2\cdot R_2)} + \sum_{\mathcal{M}(x_1\backslash R_1)} \\
& + \bigcirc_{\mathcal{M}(-x_1\backslash R_1)} + \sum_{\mathcal{M}(x_2\backslash R_2)} + R(\mathbf{R}) \\
\;=\;& (x_1+R_1+\mathrm{C})\left[\sum_{\mathcal{M}(x_1+R_1)}\times \;-\; \bigcirc_{\mathcal{M}(-x_1-R_1)}\times\right] + (x_2 R_2+\mathrm{C})\sum_{\mathcal{M}(x_2\cdot R_2)}\times \\
& + (x_1+\mathrm{C})\left[\sum_{\mathcal{M}(x_1\backslash R_1)}\times \;-\; \bigcirc_{\mathcal{M}(-x_1\backslash R_1)}\times\right] + (x_2+\mathrm{C})\sum_{\mathcal{M}(x_2\backslash R_2)}\times \;+\; R(\mathbf{R})
\end{aligned}
$$

and

$$
\begin{aligned}
\triangle \;=\;& \frac{\mathrm{d}^2x_1}{\mathrm{d}\xi^2}\Bigg|_*\left[\sum_{\mathcal{M}(x_1+R_1)}\times + \sum_{\mathcal{M}(x_1\backslash R_1)}\times - \bigcirc_{\mathcal{M}(-x_1+R_1)}\times - \bigcirc_{\mathcal{M}(-x_1\backslash R_1)}\times\right] \\
& + \frac{\partial^2 x_2}{\partial\xi^2}\Bigg|_*\left[R_2\sum_{\mathcal{M}(x_2+R_2)}\times + \sum_{\mathcal{M}(x_2\backslash R_2)}\times\right] + \frac{\mathrm{d}x_2}{\mathrm{d}\xi}\frac{\mathrm{d}R_2}{\mathrm{d}\eta}\Bigg|_*\sum_{\mathcal{M}(x_2\cdot R_2)}\times \;.
\end{aligned}
$$

Then the same results as in *Case 3,* with the exception that only $\mathrm{d}R_2/\mathrm{d}\eta > 0$ is required, apply as long as

$$
h := \sum_{\mathcal{M}(x_1+R_1)}\times \;+\; \sum_{\mathcal{M}(x_1\backslash R_1)}\times \;>\; \bigcirc_{\mathcal{M}(-x_1-R_1)}\times \;=\; -Q_{\mathsf{A}\backslash\{L(-x_1\cdot R_1)\}} \tag{21}
$$

or

$$
h := \sum_{\mathcal{M}(x_1+R_1)}\times \;+\; \sum_{\mathcal{M}(x_1\backslash R_1)}\times \;>\; \bigcirc_{\mathcal{M}(-x_1\backslash R_1)}\times \;=\; -Q_{\mathsf{A}\backslash\{L(-x_1\backslash R_1)\}}. \tag{22}
$$

$Case\ 7.$ Finally, consider the case of x_1 affecting a transition rate, and x_2 affecting a fertility or death rate, R_1 being multiplicative and R_2 being additive. Then

$$Q_{\mathsf{A}}(\xi,\eta) \;=\; \sum_{\mathcal{M}(x_1\cdot R_1)} + \bigcirc_{\mathcal{M}(-x_1\cdot R_1)} + \sum_{\mathcal{M}(x_2+R_2)} + \sum_{\mathcal{M}(x_1\backslash R_1)}$$
$$+ \bigcirc_{\mathcal{M}(-x_1\backslash R_1)} + \sum_{\mathcal{M}(x_2\backslash R_2)} + R(\mathbf{R})$$
$$=\; (x_1 R_1 + \mathrm{C})\left[\sum_{\mathcal{M}(x_1\cdot R_1)}\times\;-\;\bigcirc_{\mathcal{M}(-x_1\cdot R_1)}\times\right] + (x_2 + R_2 + \mathrm{C})\sum_{\mathcal{M}(x_2+R_2)}\times$$
$$+(x_1 + \mathrm{C})\left[\sum_{\mathcal{M}(x_1\backslash R_1)}\times\;-\;\bigcirc_{\mathcal{M}(-x_1\backslash R_1)}\times\right] + (x_2 + \mathrm{C})\sum_{\mathcal{M}(x_2\backslash R_2)}\times\;+\;R(\mathbf{R})$$

and

$$\triangle \;=\; \left[\frac{\mathrm{d}^2 x_1}{\mathrm{d}\xi^2}R_1 + \frac{\mathrm{d}x_1}{\mathrm{d}\xi}\frac{\mathrm{d}R_1}{\mathrm{d}\eta}\right]_* \left[\sum_{\mathcal{M}(x_1\cdot R_1)}\times\;-\;\bigcirc_{\mathcal{M}(-x_1\cdot R_1)}\times\right]$$
$$+\frac{\mathrm{d}^2 x_1}{\mathrm{d}\xi^2}\left[\sum_{\mathcal{M}(x_1\backslash R_1)}\times\;-\;\bigcirc_{\mathcal{M}(-x_1\backslash R_1)}\times\right] + \frac{\mathrm{d}^2 x_2}{\mathrm{d}\xi^2}\bigg|_*\left[\sum_{\mathcal{M}(x_2+R_2)}\times\;+\;\sum_{\mathcal{M}(x_2\backslash R_2)}\times\right]$$

so that again for $0 < -Q_{\mathsf{A}\backslash\{L(-x_1\cdot R_1)\}} < h$, as defined in equation (19) and equation (20), the same results as in $Case\ 3$ apply.

74

References

Berman, A., Plemmons, R. J., 1979. Nonnegative matrices in the mathematical sciences. Academic Press, New York.

Bowers, R. G., 2011. On the determination of evolutionary outcomes directly from population dynamics of the resident. Journal of Mathematical Biology 62, 901–924.

Cameron, P. J., 1994. Combinatorics. Cambridge University Press, Cambridge.

Caswell, H., 1978. A general formula for the sensitivity of population growth rate to changes in life history parameters. Journal of Theoretical Biology 14, 215–230.

Caswell, H., 2001. Matrix population models: construction, analysis, and interpretation. Sinauer Associates, Sunderland.

Cavalli-Sforza, L., Feldman, M., 1981. Cultural transmission and evolution: a quantitative approach. Princeton University Press, Princeton.

Cushing, J. M., 1998. An introduction to structured population dynamics. SIAM, Philadelphia.

de Camino-Beck, T., Lewis, M. A., van den Driesche, P., 2009. A graph-theoretic method for the basic reproduction number in continuous time epidemiological models. Journal of Mathematical Biology 59, 503–516.

De Roos, A. M., 2008. Demographic analysis of continuous-time life-history models. Ecology Letters 11, 1–15.

Dercole, F., Rinaldi, S., 2008. Analysis of evolutionary processes. Princeton University Press, Princeton, New York.

Dieckmann, U., Law, R., 1996. The dynamical theory of coevolution: a derivation from stochastic ecological processes. Journal of Mathematical Biology 34, 579–612.

Dieckmann, U., Metz, J. A. J., 2006. Surprising evolutionary predictions from enhanced ecological realism. Theoretical Population Biology 69, 263–281.

Diekmann, O., 2004. A beginner's guide to adaptive dynamics. Banach Center Publications 63, 47–68.

Diekmann, O., Gyllenberg, M., Metz, J. A. J., 2007. Physiologically structured population models: Towards a general mathematical theory. In: Takeuchi, Y., Iwasa, Y., Sato, K. (Eds.), Mathematics for ecology and environmental sciences. Springer, Heidelberg, pp. 5–20.

Diekmann, O., Heesterbeek, J. A. P., 2000. Mathematical epidemiology of infectiuous diseases. Whiley, New York.

Diekmann, O., Heesterbeek, J. A. P., Metz, J. A. J., 1990. On the definition and the computation of the basic reproductive ratio R_0 in models for infectious diseases in heterogeneous populations. Journal of Mathematical Biology 28, 365–382.

Doebeli, M., 2011. Adaptive diversification. Princeton University Press, Princeton, Oxford.

Doebeli, M., Dieckmann, U., 2004. Adaptive dynamics of speciation: spacial structure. In: Dieckmann, U., Doebeli, M., Metz, J. A. J., Tautz, D. (Eds.), Adaptive speciation. Cambridge University Press, Cambridge, pp. 140–167.

Feller, W., 1950. An introduction to probability theory and its applications. Whiley, New York.

Flowers, J., Malik, H., Li, S. I., Stathos, A., Saxer, G., Ostriwski, E., Queller, D., Strassmann, J., Purugganan, M. D., 2010. Variation, sex, and social cooperation: Molecular population genetics of the social amoeba dictyostelium discoideum. PLoS Genetics 6, e1001013.

Geritz, S. A., Gyllenberg, M., Jacobs, F. J., Parvinen, K., 2002. Invasion dynamics and attractor inheritance. Journal of Mathematical Biology 44, 548–560.

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

Gyllenberg, M., Service, R., 2011. Necessary and sufficient conditions for the existence of an optimisation principle in evolution. Journal of Mathematical Biology 62, 359–369.

Heino, M., Metz, J. A. J., Kataitala, V., 1998. The enigma of frequency-dependent selection. Trends in Ecology and Evolution 13, 367–370.

Henson, S., 1999. A continuous, age-structured insect population model. Journal of Mathematical Biology 39, 217–243.

Hofbauer, J., Sigmund, K., 1998. Evolutionary games and population dynamics. Cambridge University Press, Cambridge.

Horn, R., Johnson, C. A., 1985. Matrix analysis. Cambridge University Press, Cambridge.

Horn, R., Johnson, C. A., 1991. Topics in matrix analysis. Cambridge University Press, Cambridge.

Hoyle, A., Bowers, R. G., 2008. Can possible evolutionary outcomes be determined directly from the population dynamics? Theoretical Population Biology 74, 311–323.

Hoyle, A., Bowers, R. G., White, A., Boots, M., 2008. The influence of trade-off shape on evolutionary behaviour in classical evolutionary scenarios. Journal of Theoretical Biology 250, 498–511.

Keyfitz, N., Caswell, H., 2005. Applied mathematical demography. Springer, New York.

Kon, R., 2007. Permanence of structured population models governed by ODEs and the basic reproduction number. Japan Journal of Industrial and Applied Mathematics 24, 17–37.

Kot, M., 2001. Elements of mathematical ecology. Cambridge University Press, Cambridge.

Krebs, C., 1996. Population cycles revisited. Journal of Mammalogy 77, 8–24.

Law, R., 1978. Optimal life histories under age-specific predation. The American Naturalist 114, 399–417.

Levins, R., 1962. Theory of fitness in a heterogeneous environment. i. the fitness set and adaptive function. The American Naturalist 96, 361–373.

Li, S. I., Purugganan, M. D., 2011. The cooperative amoeba: Dictyostelium as a model for social evolution. Trends in Genetics 27, 48–54.

Meszéna, G., Kisdi, E., Dieckmann, U., Geritz, S. A. H., Metz, J. A. J., 2001. Evolutionary optimisation models and matrix games in the unified perspective of adaptive dynamics. Selection 2, 193–210.

Metz, J. A. J., 2008. Fitness. In: Jørgensen, S., Fath, B. (Eds.), Evolutionary ecology. Vol. 2 of Encyclopedia of Ecology. Elsevier, Oxford, pp. 1599–1612.

Metz, J. A. J., Diekmann, O., 1986. The Dynamics of physiologically structured populations. Lecture Notes in Biomathematics. Springer, Berlin.

Metz, J. A. J., Geritz, S. A. H., Meszéna, G., Jacobs, F. J. A., Van Heerwaarden, J. S., 1996. Stochastic and spatial structures of dynamical systems. In: van Strien, S. J., Verduyn Lunel, S. M. (Eds.), Proceedings of the Royal Dutch Academy of Science. North Holland, Dordrecht.

Metz, J. A. J., Mylius, S. D., Diekmann, O., 2008. When does evolution optimize? Evolutionary Ecology Research 10, 629–654.

Metz, J. A. J., Nisbet, R. M., Geritz, S. A. H., 1992. How should we define 'fitness' for general ecological scenarios? Trends in Ecology and Evolution 7, 198–202.

Murray, J. D., 2002. Mathematical biology 1: an introduction. Springer, Berlin, Heidelberg.

Mylius, S. D., Diekmann, O., 1995. On evolutionarily stable life histories, optimization and the need to be specific about density dependence. Oikos 74, 218–224.

Mylius, S. D., Metz, J. A. J., in press. When does evolution optimize? On the relationship between evolutionary stability, optimization and density dependence. In: Dieckmann, U., Metz, J. A. J. (Eds.), Elements of adaptive dynamics. Cambridge University Press, Cambridge.

Nowak, M., 1990. An evolutionarily stable strategy may be inaccessible. Journal of Theoretical Biology 142, 237–241.

Nowak, M., 2006. Evolutionary dynamics: exploring the equations of life. Belknap, Harvard.

Nowak, M., Sigmund, K., 1990. The evolution of stochastic strategies in the prisoner's dilemma. Acta Applicandae Mathematicae 20, 247–265.

Otto, S. P., Day, T., 2007. A biologist's guide to mathematical modeling in ecology and evolution. Princeton University Press, Princeton.

Ozcaglar, C., Shabbeer, A., Vandenberg, S. L., Yener, B., Bennett, K. P., 2012. Epidemiological models of Mycobacterium tuberculosis complex infections. Mathematical Biosciences 236, 77–96.

Rödel, H., Bora, A., Kaiser, J., Kaetzke, P., Khaschei, M., von Horst, D., 2004. Density-dependent reproduction in the European rabbit: a consequence of individual response and age-dependent reproductive performance. OIKOS 529, 529–539.

Rueffler, C., Metz, J. A. J., 2013. Necessary and sufficient conditions for R_0 to be a sum of contributions of fertility loops. Journal of Mathematical Biology 66, 1099–1122.

Rueffler, C., Metz, J. A. J., Van Dooren, T. J. M., 2013. What life cycle graphs can tell about the evolution of life histories. Journal of Mathematical Biology 66, 225–279.

Rueffler, C., Van Dooren, T. J. M., Metz, J. A. J., 2004. Adaptive walks on changing landscapes: Levins' approach extended. Theoretical Population Biology 65, 165–178.

Stearns, S. C., 1989. Trade-offs in life-history evolution. Functional Ecology 3, 259–268.

Varga, R. S., 2000. Matrix iterative analysis. Springer, Berlin.

Abstract

We investigate the evolutionary dynamics of life-history traits in structured populations, under the assumption of discrete population structure. The model family analyzed in this thesis consists of deterministic models in continuous time, that is, we consider systems of ordinary differential equations. Life cycles are characterized by demographic parameters and regulatory functions, which account for density dependence. We employ the framework of Adaptive Dynamics in order to conduct the evolutionary analysis of the model family, thereby investigating phenotypic evolution. An algebraically simple fitness proxy is derived, which allows to predict long-term evolutionary outcomes based on the configuration of loops, evolving and regulated demographic parameters in a life cycle. We derive a list of sufficient conditions on the structure of life cycles for frequency-independent selection. More precisely, we provide a range of explicit optimization criteria. If for a given life cycle there is an optimization criterion, then phenotypic variation is selected against. Furthermore, we consider frequency-dependent selection. In the case of two evolving demographic parameters which are traded off against each other, we derive sufficient conditions for different evolutionary outcomes. For this, the curvature of the trade-off between the two evolving parameters plays a cruical role. In particular, we show under which conditions evolutionary branching occurs.

Zusammenfassung

Es wird die evolutionäre Dynamik von Lebenszyklusparametern (*life-history traits*) in strukturierten Populationen unter Annahme diskreter Populationsstruktur untersucht. Die in dieser Arbeit analysierte Modellfamilie besteht aus deterministischen Modellen in kontinuierlicher Zeit, es werden also Systeme gewöhnlicher Differentialgleichungen betrachtet. Lebenszyklen werden durch demografische Parameter und regulierende Funktionen, welche Dichteabhängigkeit in die Modelle einbringen, charakterisiert. Für die evolutionäre Analyse der Modellfamilie wird das Adaptive-Dynamics-Framework angewandt. Dabei ist die Evolution phänotypischer Merkmale Gegenstand der Untersuchungen. Es wird ein algebraisch einfacher Fitnessproxy hergeleitet, welcher es ermöglicht, über das evolutionäre Langzeitverhalten Vorhersagen zu treffen, die auf der Anordnung von Schleifen (*loops*), evolvierenden und regulierten demografischen Parametern in einem Lebenszyklus basiert. Es wird eine Liste hinreichender Bedingungen für frequenzunabhängige Selektion erarbeitet, oder genauer, eine Liste von expliziten Optimierungskriterien (*optimization principles*). Gibt es ein solches für einen Lebenszyklus, so ist phänotypische Variation unwahrscheinlich. Weiters wird frequenzabhängige Selektion betrachtet. Im Falle von zwei evolvierenden demografischen Parametern, welche einem Trade-off unterliegen, werden hinreichende Bedingungen für verschiedene evolutionäre Szenarien hergeleitet. Dabei spielt die Krümmung der Trade-off-Kurve eine bedeutende Rolle. Insbesondere wird gezeigt, unter welchen Umständen evolutionäre Verzweigungen (*evolutionary branching*) vorkommen.

Curriculum vitae

09/1992—06/1996	Primary school, Volksschule, Scheibenbergstraße 63, 1180 Vienna
09/1996—06/2004	Grammar school specializing in humanities, Bundesgymnasium XVIII, Klostergasse 25, 1180 Vienna
10/2004—09/2005	Community service, St John Ambulance, 1180 Vienna
Since 10/2005	Studies of Mathematics, Vienna University
10/2005—01/2010	Studies of Applied and Theoretical Linguistics, Vienna University
Since 03/2010	Doctoral studies of Linguistics, Vienna University
Since 05/2011	University assistant (prae doc), Department of English and American Studies, Vienna University