When Does Evolution Optimise? On the relation between types of density dependence and evolutionarily stable life history parameters

J. (Hans) A.J. Metz [1,2],
Sido D. Mylius [1,*] &
Odo Diekmann [1,3,†]

[1] *Institute of Evolutionary and Ecological Sciences, Leiden University, P.O. Box 9516, NL – 2300 RA Leiden, The Netherlands*
[2] *International Institute for Applied Systems Analysis, A – 2361 Laxenburg, Austria*
[3] *Centre for Mathematics and Computer Science, P.O. Box 94079, NL – 1090 GB Amsterdam, The Netherlands*

International Institute for Applied Systems Analysis (IIASA)
Working Paper WP-96-004

March 1996

Abstract

In this paper we (i) put forward a simple notational device clarifying the, undeniable but generally ignored, role of density dependence in determining evolutionarily stable life histories, (ii) use this device to derive necessary and sufficient conditions for (a) the existence of an evolutionary extremisation principle, and (b) the reduction of such a principle to straight r- or R_0-maximisation, (iii) use the latter results to analyse a simple concrete example showing that the details of the population dynamical embedding may influence our evolutionary predictions to an unexpected extent.

* Author to whom correspondence should be addressed. *Present address*: Sido D. Mylius, Theoretical Biology, University of Groningen, P.O. Box 14, NL – 9750 AA Haren, The Netherlands. *E-mail*: mylius@biol.rug.nl

† *Present address*: O. Diekmann, Department of Mathematics, Utrecht University, P.O. Box 80010, NL – 3508 TA Utrecht, The Netherlands

1 Introduction

The literature is replete with statements that evolutionary predictions about behavioural, c.q. life history, parameters should be based on the maximisation of individual lifetime reproductive success, R_0 (Stearns, 1992; Roff, 1992; Charnov, 1993; Charlesworth, 1994, provide surveys), or else the intrinsic rate of natural increase, r (Stearns, 1992; Roff, 1992; Charlesworth, 1994; Caswell, 1989). In the former case it is often added, rather confusingly, that due to density dependence necessarily $R_0 = 1$ (see *e.g.* Charnov, 1993, and its review by Maynard Smith, 1993). No doubt most authors dealing with life history theory know how to interpret the last statement, and are aware of the implicit limitations of the traditional optimisation considerations. However, some asking around indicated that this awareness (i) has little diffusion among experimentalists, and (ii) appears rather dim even among most theorists. Our quick and dirty survey also revealed that probably the main cause of this small awareness is that advertising positive predictions gives more kudos than repeatedly spelling out their limitations. Yet we feel that precisely delimiting the applicability of particular evolutionary arguments is a worthy effort, not only for philosophical but also for practical reasons: By extending the limits as far as one can, one usually also extends the effective toolbox.

In this paper we put forward three closely related messages:
(i) We argue that adhering to a simple explicit notation fosters the awareness of some implicit limitations of life history arguments. Our notation only differs from the traditional one in that the roles of (a) the life history traits, and in particular (b) the environment, in determining the population dynamical behaviour of an individual, are made visible. This visibility also has the advantage of removing the minor confusion about R_0 simultaneously being maximised and kept equal to 1. We sincerely ask you to adopt this notation, or else to develop your own variant of it. The use of more simplified notations too often misleads!
(ii) We give necessary and sufficient conditions for the eventual outcome of the evolutionary process to be characterisable by some optimisation principle, and more in particular by straight r- or R_0-maximisation. These conditions are phrased in mathematical, structural, terms only. So far we haven't been able to delimit clear classes of corresponding physiological mechanisms. Dreaming up simple classes of mechanisms subsumed under our conditions is easy. We shall give some examples. But how wide exactly is the net?
(iii) We show how the details of the population dynamical embedding can influence the evolutionary predictions, by using the results from (ii) to analyse a particularly simple sample model, closely akin to traditional life history models. Our explicit notation also alerted us to the fact that for this example the life history parameters determined in the field show patterns which differ in a non-trivial way from the patterns in the parameters determined under laboratory conditions. This observation may act as an antidote to the, apparently common, belief that the message from (i) is for all practical purposes empty.

2 Setting the stage: fitness, density dependence, and ESS considerations

Our starting point is that there is one master fitness concept: the hypothetical average rate of exponential growth ρ which results from the thought experiment in which we let a clone of the type under consideration grow in a stationary environment (Charlesworth, 1980, 1994; Tuljapurkar, 1989, 1990; Caswell, 1989; Metz *et al.*, 1992; Rand *et al.*, 1994; Ferrière & Gatto, 1995).

Remark 2.1 The reasons for this particular choice of a definition are: (i) It is consistent with the use of the word fitness in the context of simple evolutionary scenarios on all points that count in a long term evolutionary context. (ii) For a large range of ecological scenarios it is sufficiently precise to yield a definite number. (iii) The number so defined is almost the minimal information necessary to deduce predictions about both evolutionary final states and non-equilibrium evolutionary patterns. See the arguments below, and Metz *et al.* (1995) for a further elaboration. □

Our verbal definition immediately brings out that ρ necessarily depends both on the type X of the clone and the environment E in which it supposedly lives. To keep our arguments, and our heads, clear we should explicitly account for this dependency in the notation, by writing

$$\rho(X, E) \tag{1}$$

(compare Diekmann & Metz, 1994, and Mylius & Diekmann, 1995).

A possible further potential source of confusion is that E necessarily refers to the environment as perceived by the individuals. This means that for instance density and types of conspecifics come as part and parcel of E (Michod, 1979; Metz & Diekmann, 1986; Pásztor, 1988; Metz & de Roos, 1992; Diekmann & Metz, 1994; Pásztor *et al.*, 1995). Yet in our thought experiment we considered those densities as given stationary random functions of time, not influenced by the growth of our clone.

The justification of this mental somersault is that we should think of fitness as the rate of invasion of a rare mutant multiplying amidst a large resident population. This presupposes that all evolutionarily relevant resident (sub)populations of the species are large, so that initially the influence of the mutant on the environment is properly diluted. The mutant heterozygote swarm reproduces faithfully by crossing with the residents. If dilution fails due to the interaction ranges of the individuals containing but a few more permanent sparring partners, we can sometimes take recourse to inclusive fitness considerations (Taylor, 1988a,b, 1989), but in ultimate generality the concept of fitness resists further extension. Luckily, the range of conditions covered is sufficiently large that we need not be overly bothered.

The corollary is that predictions about the trait values favoured by evolution should always derive from an ESS argument (*e.g.* Roughgarden, 1979; Charlesworth, 1994; Lessard, 1990):

1. Maximise $\rho(X, E)$ for each given E over all potential trait values, resulting in a function $X_{\mathrm{opt}}(E)$.
2. Determine for each trait value the environment which it generates as a resident, $E_{\mathrm{attr}}(X)$.
3. Vary X to find an evolutionarily unbeatable value X^*, *i.e.*, an X^* such that

$$X_{\mathrm{opt}}(E_{\mathrm{attr}}(X^*)) = X^* . \tag{2}$$

4. Ascertain that the set of trait values X_0 from which X^* is approximated with non-zero probability through a sequence $X_0, X_1, X_2, \ldots$, such that

$\rho(X_{i+1}, E_{\mathrm{attr}}(X_i)) > 0$, is sufficiently large to warrant consideration of X^* as a potential evolutionary trap.

The above description is only meant as a definition, not as a practical algorithm. The general procedure 1 to 4 has a habit of exceeding the available computer capacity, except in the simplest possible cases. Practical algorithms circumvent this by using special properties of particular cases.

One immediate general simplification is that even in the definition of an ESS we may restrict the attention to those E that can occur as $E_{\mathrm{attr}}(X)$ for some X. As this restriction becomes essential in the arguments below we introduce the

Convention Whenever we refer to E we shall mean only those E that can occur as $E_{\mathrm{attr}}(X)$ for some X.

The notional index *attr* alludes to the assumption that the population dynamics converges to an attractor. For later use we note that on this attractor necessarily

$$\rho(X, E_{\mathrm{attr}}(X)) \quad = \quad 0\,. \tag{3}$$

Remark 2.2 In general it cannot be excluded that the function E_{attr} is multi-valued. In theory this does not invalidate our arguments, except that our present phrasing is definitely lacking in the details. But it may considerably complicate attempts at applying them in practice. The wording of the special arguments in sections 3 and 4 happens to apply without change to the multi-valued case. □

Remark 2.3 In step 4, and only step 4, of the above algorithmic definition of an ESS, we implicitly invoked a genetical assumption. Whether or not convergence to the unbeatable strategy can occur will also depend on the (non-)presence of so-called genetic constraints. The simplest example is that the unbeatable phenotype can only be produced by a heterozygote, so that the population can never converge to a monomorphically X^* condition. We assume that the only constraints that are present are "physiological", *i.e.*, can be described in terms of a developmentally realisable subset of the trait space. The belief is that this assumption guarantees that convergence for the Mendelian case parallels that for the clonal case, as it allows us to dream up any needed mutations, including mutations that break up heterotic polymorphisms (compare Hammerstein & Selten, 1994, and Hammerstein, 1996). □

3 When does evolution optimise?

The outcome of the ESS calculation can only be reached by the straightforward application of some extremisation principle when the function $\rho(X, E)$ satisfies some rather stringent restrictions.

First we give some definitions. We shall say that *the trait vector acts one-dimensionally* whenever there exists a function ψ of X to the real numbers such that

$$\operatorname{sign}\rho(X, E) \quad = \quad \operatorname{sign}\alpha(\psi(X), E)\,, \tag{4}$$

for some function α which increases in its first argument. And we shall say that *the environment acts one-dimensionally* whenever there exists a function ϕ of E

to the real numbers such that

$$\operatorname{sign}\rho(X,E) \quad = \quad \operatorname{sign}\beta(X,\phi(E))\ , \tag{5}$$

for some function β which increases in its second argument.

Example 3.1 Assume that we only need to deal with constant environments. Whenever

$$R_0(X,E) \quad = \quad \phi(E)\, R_0(X,E_V)\ , \qquad V \text{ for } \textit{virgin}, \tag{6}$$

take

$$\alpha \ := \ \ln(R_0)\ , \qquad \beta \ := \ \ln(R_0)\ , \qquad \psi \ := \ R_0(X,E_V)\ . \tag{7}$$

(See section 4.) □

We shall call a function ψ of X to the real numbers with the property that evolution maximises ψ for any constraint on X an *optimisation principle.* And we shall call a function ϕ of E to the real numbers with the property that evolution minimises $\phi(E_{\text{attr}}(X))$ for any constraint on X, a *pessimisation* or *Verelendungs principle.*

Proposition 3.1 *Models in which the trait vector acts one-dimensionally have an optimisation principle, and vice versa.*

The forward implication is immediate. The somewhat unexpected reverse implication is spelled out in Appendix A.

Proposition 3.1 is of course a weakened form of the familiar justification for many of our commonly used optimisation principles: “Being more ‘efficient’ increases your fitness in any relevant environment.” However, the crucial phrase in that argument, “in any relevant environment”, rarely is mentioned explicitly.

Proposition 3.2 *Models in which the environment acts one-dimensionally have a pessimisation principle, and vice versa.*

The forward implication is immediate. The somewhat unexpected reverse implication is spelled out in Appendix A. In this proof we construct a β such that (5) holds true for the pessimisation principle ϕ. This construction also provided the heuristics for the term Verelendungs principle: Any ϕ satisfying (5), with β increasing in its second argument, allows a natural interpretation as a *measure of environmental quality*, as perceived through the physiology of our individuals.

Proposition 3.2 is of course nothing but the ultimate generalisation of two familiar evolutionary extremisation principles pertaining to the case of population dynamical equilibrium: (i) “Evolution minimises the availability of a limiting resource”, and (ii) “Evolution maximises total population density if the individual life history parameters are negatively affected by the total population density (and are unaffected by any other environmental variable influenced by the population)”.

Proposition 3.3 *Any pessimisation principle carries an optimisation principle in its wake and vice versa.*

This is easily proved by gauging the “ability to cope” to the “quality of the environment” through

$$\psi(X) \quad = \quad -\phi(E_{\text{attr}}(X))\ . \tag{8}$$

This recipe produces a ψ for any ϕ pried out of an expression for ρ, or vice versa. But beware, (8) usually doesn't hold true for a ψ and a ϕ arrived at separately. The strongest possible statement that can be made about two ψ's, or ϕ's, found by different means is that they are necessarily monotonically related.

The construction used to prove proposition 3.3 has as a corollary:

Proposition 3.4 *Whenever the trait vector acts one-dimensionally it is possible to find a function ϕ of E to the real numbers, or alternatively, whenever the environment acts one-dimensionally it is possible to find a function ψ of X to the real numbers, such that*

$$\operatorname{sign} \rho(X, E) \quad = \quad \operatorname{sign}\left(\psi(X) + \phi(E)\right) . \tag{9}$$

However, somewhat unexpectedly the aesthetically pleasing symmetry of (9) isn't very helpful, as usually at most one of the functions ϕ and ψ occurring in it can be expressed as an explicit formula. In contrast the more relaxed characterisations of one-dimensional action by means of either (4) or (5) often can be readily applied.

The arguments in Appendix A are only based on evolutionary unbeatability considerations. For completeness we summarise some immediately associated evolutionary attractivity properties as

Proposition 3.5 *When evolution operates in a context which allows an optimisation principle ψ which is at least piecewise continuous, and the supports of any mutation distribution contains at least the intersection of an ε-neighbourhood of the trait value of the progenitor X with the developmentally realisable subset $\mathbb{X}$ of the trait space, with ε independent of X:*

1. *A unique global optimum of ψ has a non-negligible basin of evolutionary attraction. Better still, it will often be a global evolutionary attractor. This happens for example when (a) the supports of the mutation distributions equal $\mathbb{X}$, or (b) ψ is continuous, and there are no local optima other than the global one.*
2. *When mutant trait values are restricted to a δ-neighbourhood of X, a particular non-isolated local optimum of ψ will have a non-negligible basin of evolutionary attraction whenever δ is sufficiently small.*

The application of (8) immediately yields the corresponding proposition for pessimisation principles.

As a final point we mention that for a one-dimensionally acting environment $\phi(E_{\text{attr}}(X))$ can be directly determined from

$$\beta(X, \phi(E_{\text{attr}}(X))) \quad = \quad 0 . \tag{10}$$

This allows the construction of a simple general algorithm for numerically analysing any model with a one-dimensionally acting environment: Numerically maximise ψ defined by (8), where $\phi(E_{\text{attr}}(X))$ is at each iteration step numerically determined from (10). This way the potentially unpleasant object $E_{\text{attr}}(X)$ is eliminated before the numerics.

We finish this section with three examples. The first example is essentially trivial. We put it in to demonstrate the various concepts in rigorous detail, unencumbered by technical distractions. Its second purpose is demonstrating how our formal definition of a one-dimensionally acting environment may somehow carry a wrong suggestion at the mechanistic level. The second example demonstrates why it may be difficult to find an explicit pessimisation principle from a given optimisation principle. The third example shows how it may be possible to find a pessimisation principle for non-equilibrium attractors, leading to an otherwise non-obvious optimisation principle.

Before starting on the examples we introduce one more piece of notation as this considerably simplifies their presentation: We shall denote the geometric mean operator as G,

$$G(z) \ := \ \lim_{T\to\infty} \sqrt[T]{\prod_{t=1}^{T} z(t)} \ , \tag{11}$$

and the logarithm of G as L, *i.e.*,

$$L(z) \ := \ \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} \ln(z(t)) \ . \tag{12}$$

In order not to unduly complicate the examples we shall moreover proceed as if reproduction were clonal.

Example 3.2 Consider the following thought experiment. Birds are limited by the availability of nest sites. These sites have a density s. Only birds who have obtained a nest site in spring breed. The number of young M which they produce per capita is an increasing function of their ability to gather energy $\psi(X)$, where X is the trait which is assumed to be under evolutionary control. We measure this ability by the number of offspring it produces:

$$M \ = \ \psi(X) \ . \tag{13}$$

Old and young survive the winter with a probability p. Next spring, nest sites are allotted randomly among the survivors. Birds that fail to obtain a site are removed from the system.

An obvious choice for the condition of the environment in year t as perceived by a bird, is the total density n of winter survivors, of all trait types together,

$$E(t) \ = \ n(t) \ . \tag{14}$$

We shall present side by side a classical population dynamical calculation, and a calculation along the route laid out above. In neither calculation we take the obvious shortcuts as this would obstruct their comparison. Please bear with us, we only want to help you understand the full meaning of our previous considerations.

We shall distinguish the resident and mutant types by means of the indices 0 and 1. With this notation the population equations become, with $i \in \{0, 1\}$,

$$\begin{aligned} n_i(t+1) \ &= \ \left(p\,(1+M_i)\,\frac{s}{n(t)}\right) n_i(t) \\ &= \ \left(p\,s\,(1+\psi(X_i))\,\frac{1}{E(t)}\right) n_i(t) \ , \end{aligned} \tag{15}$$

with

$$n(t) \ = \ n_0(t) + n_1(t) \ . \tag{16}$$

(For notational simplicity we confine ourselves to initial conditions such that consistently $n(t) > s$.)

Applying the definition of ρ to (16) (without already confining the attention to the $E_{\mathrm{attr}}(X)$ which for this particular model necessarily are constant) results in

$$\begin{aligned} \rho(X,E) &= L\left(p\,s\,(1+\psi(X))\,\frac{1}{E}\right) \\ &= \ln(ps) + \ln(1+\psi(X)) - L(E)\,. \end{aligned} \tag{17}$$

Given the functional form of (17) and the verbal model description with which we started, one natural choice for ϕ is

$$\phi(E) := \frac{1}{G(E)}\,, \tag{18}$$

i.e., we measure the quality of the environment of a bird as the inverse of (the geometric mean of) the density of competitors which it encounters when it is searching for a nest site. With this definition we can write

$$\rho(X,E) = \ln(ps) + \ln(1+\psi(X)) + \ln(\phi(E))\,. \tag{19}$$

From this formula we see that both the trait and the environment act one-dimensionally, with

$$\alpha(\psi(X),E) := \rho(X,E) =: \beta(X,\phi(E))\,. \tag{20}$$

The conclusions that ψ is an optimisation, and ϕ a pessimisation principle, won't come as a surprise. Combining (8) with (10) and (19) leads to the, equivalent, optimisation principle

$$\psi'(X) := -\phi(E_{\mathrm{attr}}(X)) = -\frac{1}{p\,s\,(1+\psi(X))}\,. \tag{21}$$

Our first choice was to have our measure for the quality of the environment, ϕ, inversely proportional to the density of conspecifics. The matching "ability to cope", ψ', given by (21), is, of course, monotonically related to energy gathering ability ψ. A measure of environmental quality which for constant environments matches the optimisation principle ψ, is given by

$$\phi'(E) = 1 - \frac{1}{p\,s\,\phi(E)} = 1 - \frac{G(E)}{p\,s}\,. \tag{22}$$

For each of these pairs

$$\operatorname{sign}\rho(X,E) = \operatorname{sign}\left(\psi'(X)+\phi(E)\right) = \operatorname{sign}\left(\psi(X)+\phi'(E)\right)\,. \tag{23}$$

For the population dynamical invasion calculation we set $E(t) = n_0(t)$ to get

$$n(t) = n_0 = p\,s\,(1+\psi(X_0))\,, \tag{24}$$

and

$$\begin{aligned} n_1(t+1) &= p\,s\,(1+\psi(X_1))\,\frac{n_1(t)}{n_0} \\ &= \frac{1+\psi(X_1)}{1+\psi(X_0)}\,n_1(t)\,. \end{aligned} \tag{25}$$

(25) Tells that evolution leads to the optimisation of $\psi''(X) := 1 + \psi(X)$.

This example also shows how our unguided intuition may clash with our formal characterisations. Mechanistically the bird density is regulated by the nest sites, but structurally (*i.e.*, in terms of the mathematical relations connecting the various population dynamical variables) by the density of conspecifics competing for those sites. One should watch out for this type of discrepancy when applying proposition 3.2 to 3.4 in mechanistically formulated examples. □

Example 3.3 We make the following changes in the previous example. Losers of the lottery for nest sites aren't removed, and winter survival is variable. In that case

$$E(t) \quad = \quad (p(t), n(t)) \ , \tag{26}$$

and

$$\rho(X, E) \quad = \quad L(p) + L\left(1 + \frac{\psi(X)\, s}{n}\right) . \tag{27}$$

The fact that $(1 + \psi(X)\frac{s}{n(t)})$ increases in ψ, independent of $n(t)$, implies that $L(1 + \psi(X)\frac{s}{n})$ and therefore $\rho(X, E)$ increases whenever $\psi(X)$ increases. Since ψ is an optimisation principle, our model allows a pessimisation principle ϕ, by proposition 3.3. But it is clearly impossible to find any sort of explicit expression for $\phi(E)$. □

Example 3.4 Consider the population dynamical equations

$$n_i(t+1) \quad = \quad a_i\big(f(E(t))\big)^{b_i}\, n_i(t) \ , \qquad i = 0, \ldots, k \ , \tag{28}$$

with

$$E(t) \quad = \quad (c_0 n_0(t) + \cdots + c_k n_k(t)) \ , \tag{29}$$

all a_i, b_i and $c_i > 0$, and f decreasing from 1 to 0 for E increasing from 0 to ∞.

With the choice

$$f(E(t)) \quad = \quad (1 + E(t))^{-1} \ , \tag{30}$$

and $k = 0$, this model becomes the model launched into fashion by i.a. Hassell *et al.* (1976) as a touchstone for the appearance of chaotic fluctuations in single species population dynamics.

The trait vector appearing in (28) is

$$X \quad = \quad (a, b, c) \ . \tag{31}$$

The parameters a, $1/b$ and c can be interpreted in individual-based terms as respectively the per capita reproduction in a boom environment, the ability to cope with a bust environment, and the per capita impingement on the common environment.

From (28) we find

$$\rho(X, E) \quad = \quad L\big(a\,(f(E))^b\big) \quad = \quad \ln(a) + b\,\phi(E) \ , \tag{32}$$

with

$$\phi(E) \quad = \quad L(f(E)) \ . \tag{33}$$

From $\rho(X, E_{\text{attr}}(X)) = 0$ we deduce that

$$\phi(E_{\text{attr}}(X)) \quad = \quad -b^{-1}\ln(a) \ . \tag{34}$$

We conclude that evolution maximises

$$\psi(X) \;:=\; \frac{\ln(a)}{b} \;. \tag{35}$$

In accordance with propositions 3.3 and 3.4 we can define the functions a and b occurring in the definitions of one-dimensional action, as

$$\alpha(\psi(X), E) \;:=\; \psi(X) + \phi(E) \;=:\; \beta(X, \phi(E)) \;. \tag{36}$$

The point that we want to make is that the quantities α and β defined by (36) have the same sign as $\rho(X,E)$, but are not equal to $\rho(X,E)$, as was the case in the previous example. It can even be proved that for ρ given by (32) it is impossible to find pairs α and ψ, or β and ϕ, for which such an equality holds good. □

4 When does evolution maximise r or R_0?

In this section we shall consider the optimisation principles of classical life history theory, to wit r- and R_0-maximisation. Since r and R_0 are only defined for constant environments we shall from now on (i) assume that population dynamical equilibrium obtains, and (ii) have the symbol E refer alternatively to a potential condition of the environment at a particular time, or to constant functions of time having that condition of the environment as value. For constant environments

$$\rho(X,E) \;=\; r(X,E) \;. \tag{37}$$

Moreover,

$$r(X,E) \;\begin{matrix} > \\ = \\ < \end{matrix}\; 0 \qquad \text{if, and only if,} \qquad R_0(X,E) \;\begin{matrix} > \\ = \\ < \end{matrix}\; 1 \;, \tag{38}$$

allowing the replacement of $\rho(X,E)$ in the recipes of sections 2 and 3 by $\ln(R_0(X,E))$. (See *e.g.* Roughgarden, 1979; Charlesworth, 1994; Metz & Diekmann, 1986.)

Incidentally, although the usual definitions of r and R_0 are predicated upon all individuals being born equal, they can readily be extended to cater for variable birth states and spatial heterogeneity. The only proviso is that E should be constant in time. (See *e.g.* Diekmann *et al.*, 1990; Jagers, 1991, 1995; Kawecki & Stearns, 1993; Kozłowski, 1993; Diekmann & Metz, 1994.)

Below E_0 denotes some specially chosen fixed value of E.

The following proposition is an immediate corollary of proposition 3.1.

Proposition 4.1 *$r(X,E_0)$, or $R_0(X,E_0)$, is an optimisation principle for, and only for, combinations of life histories and ecological embedding, such that there exists a function α increasing in its first argument such that*

$$\operatorname{sign} r(X,E) \;=\; \operatorname{sign} \alpha(r(X,E_0), E) \;, \tag{39}$$

or

$$\operatorname{sign} \ln(R_0(X,E)) \;=\; \operatorname{sign} \alpha(\ln(R_0(X,E_0))) \tag{40}$$

respectively.

Remark 4.1 The result from proposition 3.4 allows us to replace the characterisations from proposition 4.1 by the characterisation that there should exist a function ϕ of E to the real numbers such that

$$\operatorname{sign} r(X,E) = \operatorname{sign}\left(r(X,E_0) + \phi(E)\right), \tag{41}$$

or

$$\operatorname{sign} \ln(R_0(X,E)) = \operatorname{sign}\left(\ln(R_0(X,E_0)) + \phi(E)\right) \tag{42}$$

respectively. However, this characterisation may in theory be equivalent to the characterisation from proposition 4.1, in practice it is less useful as ϕ rarely pops up as an explicit formula, whereas it is usually fairly easy to spot the α occurring in the characterisation from proposition 4.1. □

We shall say that *evolution just maximises* r, *or* R_0, whenever $r(X,E_0)$, respectively $R_0(X,E_0)$, is an optimisation principle for every choice of E_0.

Proposition 4.2 *Evolution just maximises* r, *or* R_0, *if and only if it deals with combinations of life histories and ecological embedding such that is possible to write*

$$r(X,E) = \alpha(r(X,E_0),E), \tag{43}$$

or

$$R_0(X,E) = \exp\left(\alpha(\ln(R_0(X,E_0)),E)\right) \tag{44}$$

respectively, with α *increasing in its first argument, and* E_0 *fixed, but otherwise arbitrary.*

A proof of this proposition can be found in Appendix A.

Example 4.1 Whenever the environment makes itself felt only through an additional death rate $\mu(E)$, acting equally on all individuals, $r(X,E)$ can be expressed as

$$r(X,E) = r(X,E_V) - \mu(E), \tag{45}$$

E_V the virgin environment. Therefore evolution within those confines just maximises r. □

Example 4.2 The confinement of the X- and E-dependence to non-overlapping life stages allows $R_0(X,E)$ to be expressed as

$$R_0(X,E) = \phi(E)\, R_0(X,E_V), \tag{46}$$

E_V the virgin environment. Therefore evolution within those confines just maximises R_0. □

5 The potential of the community dynamical feedback-loop for influencing life history predictions: an example

Consider the following simple family of life histories: Juveniles die at a rate μ_J and mature into adults at age T. Adults die at a rate μ_A and reproduce at a rate b. All these parameters may in principle be affected by E. Their values in the virgin environment E_{V} we shall indicate with an (additional) index V. The strategy parameter is the length of the juvenile period in the virgin environment, T_{V}. The adult reproduction rate b increases linearly with T_{V}; in the virgin environment

$$b(T_{\mathrm{V}},E_{\mathrm{V}}) \quad = \quad b_{\mathrm{V}}(T_{\mathrm{V}}) \quad = \quad \max(0, T_{\mathrm{V}} - 1)\ . \tag{47}$$

In addition we (i) brashly assume that population dynamical equilibrium obtains, and (ii) have the symbol E refer alternatively to a constant (with as value a condition the environment might be in at a particular time) or to a constant function of time.

We combine this basic scenario with six alternative environmental feedback rules:

1. E only equally additively affects the juvenile and adult mortality rates,

$$\mu_{\mathrm{J}}(E) \quad = \quad \mu_{\mathrm{JV}} + \gamma_1(E)\ , \qquad \mu_{\mathrm{A}}(E) \quad = \quad \mu_{\mathrm{AV}} + \gamma_1(E)\ , \tag{48}$$

(parameters for which nothing is specified are assumed always to take the value for the virgin environment, in this case $T(E) = T_{\mathrm{V}}$, $b(T_{\mathrm{V}},E) = b_{\mathrm{V}}(T_{\mathrm{V}})$),

2. E only additively affects the adult mortality rate,

$$\mu_{\mathrm{A}}(E) \quad = \quad \mu_{\mathrm{AV}} + \gamma_2(E)\ , \tag{49}$$

3. E only multiplicatively affects the reproduction rate,

$$b(T_{\mathrm{V}},E) \quad = \quad \frac{b_{\mathrm{V}}(T_{\mathrm{V}})}{\theta_3(E)}\ , \tag{50}$$

4. E only additively affects the age at maturation (without affecting the birth rate) in such a manner that for a constant environment

$$T(E) \quad = \quad T_{\mathrm{V}} + \gamma_4(E)\ , \tag{51}$$

5. E only multiplicatively affects the age at maturation (without affecting the birth rate), in such a manner that for a constant environment

$$T(E) \quad = \quad \theta_5(E)\, T_{\mathrm{V}}\ , \tag{52}$$

6. E only additively affects the juvenile mortality rate,

$$\mu_{\mathrm{J}}(E) \quad = \quad \mu_{\mathrm{JV}} + \gamma_6(E)\ , \tag{53}$$

with

$$\gamma_i(E) \quad \geq \quad \gamma_i(E_{\mathrm{V}}) \quad = \quad 0 \quad \text{for} \quad i \in \{1,2,4,6\}\ , \quad \text{and} \tag{54}$$

$$\theta_j(E) \quad \geq \quad \theta_j(E_{\mathrm{V}}) \quad = \quad 1 \quad \text{for} \quad i \in \{3,5\}\ . \tag{55}$$

For fixed values of T_{V} and E we can, directly from our initial model description, derive the characteristic equation,

$$\frac{b\, e^{-(r+\mu_{\mathrm{J}})\,T}}{r + \mu_{\mathrm{A}}} \quad = \quad 1\ , \tag{56}$$

as well as an explicit expression for R_0,

$$R_0 \quad = \quad \frac{b\, e^{-\mu_\mathrm{J} T}}{\mu_\mathrm{A}} \; . \tag{57}$$

Below we shall use a * to mark the ESS value of any quantity.

Feedback rule 1 makes our model a special case of the models considered in example 4.1. Therefore we can determine T_V^* by maximising $r(\cdot, E_\mathrm{V})$. In Appendix B we describe a simple way to calculate the, unique, maximum.

Feedback rules 2 to 4 all lead to a formula for R_0 which, although the biological mechanism differs from that of the models considered in example 4.2, can be brought into the form (46), with

$$R_0(T_\mathrm{V}, E_\mathrm{V}) \quad = \quad \frac{b_\mathrm{V}(T_\mathrm{V})\, e^{-\mu_\mathrm{JV} T_\mathrm{V}}}{\mu_\mathrm{AV}} \; , \tag{58}$$

and

$$\text{rule } 2: \quad \phi(E) \quad = \quad \frac{\mu_\mathrm{AV}}{\mu_\mathrm{AV} + \gamma_2(E)} \; , \tag{59}$$

$$\text{rule } 3: \quad \phi(E) \quad = \quad \frac{1}{\theta_3(E)} \; , \tag{60}$$

$$\text{rule } 4: \quad \phi(E) \quad = \quad e^{-\mu_\mathrm{JV}\, \gamma_4(E)} \; . \tag{61}$$

(In Appendix C we show that it is possible to slightly reinterpret the model formulation such that cases 2 to 4 do become subsumed under example 4.2.)

Case 5 doesn't belong to any of the special cases considered in examples 4.1 or 4.2. However, it is easily seen from the interpretation that $\theta_5(E)$ monotonically affects R_0. Therefore we fall back on the general procedure for one-dimensionally acting environments, with $\ln(R_0)$ substituted for β, and $1/\theta_5$ for ϕ, *i.e.*, we set

$$R_0(T_\mathrm{V}, E_\mathrm{attr}) \quad = \quad \frac{b_\mathrm{V}(T_\mathrm{V})\, e^{-\mu_\mathrm{JV}\, \theta_5(E_\mathrm{attr})\, T_\mathrm{V}}}{\mu_\mathrm{AV}} \quad = \quad 1 \; , \tag{62}$$

in order to calculate the optimisation principle $\psi(T_\mathrm{V}) := \theta_5(E_\mathrm{attr}(T_\mathrm{V}))$. It turns out that we are lucky, and we end up with the explicit expression (after multiplying out the constant factor μ_JV)

$$\psi(T_\mathrm{V}) \quad = \quad \frac{\ln(b_\mathrm{V}(T_\mathrm{V})) - \ln(\mu_\mathrm{AV})}{T_\mathrm{V}} \; . \tag{63}$$

The story for case 6 is exactly the same as for case 5, with $-\gamma_6$ in the role of ϕ, even to the extent that we end up with the same optimisation principle.

Remark 5.1 In principle case 1 can be analysed by exactly the same procedure as cases 5 and 6, except that it isn't possible to find an explicit expression for $\gamma_1(E_\mathrm{attr}(T_\mathrm{V}))$. And our general results tell that anyway the resulting optimisation principle would be monotonically related to $r(\cdot, E_\mathrm{V})$. □

After the mathematics comes the interpretation problem. In the classic life history model this is less of a problem, as it is assumed that the life history parameters of an individual are constants, instead of being potentially under environmental control. In the case of the present model we shall distinguish

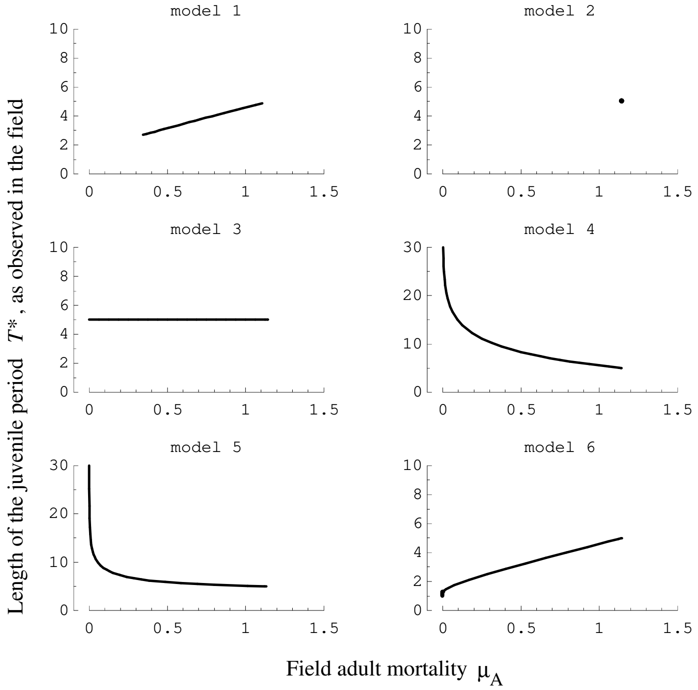

Figure 1: Correlations between the adult mortality rate μ_A and the duration of the evolutionarily stable juvenile period T^*, both "observed in the field", for the six models with alternative environmental feedback rules described in section 5. The value of the "physiological parameter" juvenile mortality in the virgin environment, μ_{JV}, was kept fixed at $\mu_{JV} = 0.25$.

two situations, called "laboratory" and "field". In the laboratory situation the environment is kept constant, whereas in the field situation the environment adjusts itself such that

$$R_0(T_V^*, E) = 1 . \tag{64}$$

For the feedback rules 1 to 6 the values of the life history parameters in the laboratory situation differ from those in the virgin environment by at most either an additive or a multiplicative constant. The field values are obtained by adjusting the virgin parameter values, where appropriate, by $\gamma_i(E)$ or $\gamma_j(E)$ determined from (64).

Figure 1 shows the correlations obtaining between the field observables T^* and μ_A, for a fixed value of μ_{JV}, for each of our six feedback rules.

The numbering of the panes refers to the feedback rules. The plotted field observables are determined by a combination of the "physiological parameters" μ_{AV} (the adult death rate in the virgin environment) and T_V^* (the ESS value of T_V, the juvenile period in the virgin environment), and the corresponding feedback rule. This amounts to plotting T_V^* versus $\mu_{AV} + \gamma_1(E_{attr}(T_V^*))$ (for model 1), T_V^* versus $\mu_{AV} + \gamma_2(E_{attr}(T_V^*))$ (for model 2), T_V^* versus μ_{AV} (for model 3), $T_V^* + \gamma_4(E_{attr}(T_V^*))$ versus μ_{AV} (for model 4), $\theta_5(E_{attr}(T_V^*))\, T_V^*$ versus μ_{AV} (for model 5), and T_V^* versus μ_{AV} (for model 6). For the computational details we refer to section 5 and Appendix B.

The, for all curves identical, upper limit of μ_A results from the fact that for higher values of μ_{AV} no strategy can invade into the virgin environment. Such values of μ_{AV} would lead in a, naive, calculation to $\gamma_i(E_{attr}(T_V^*)) < 0$ (in

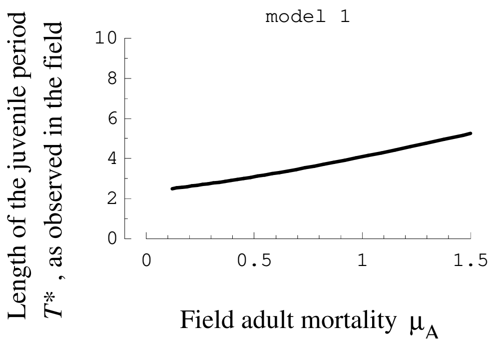

Figure 2: Correlation between the adult mortality rate μ_A and the evolutionarily stable duration of the juvenile period T^*, both "observed in the field", for feedback rule 1 from section 5. The difference with pane 1 of Figure 1 is that now the value of the observed juvenile mortality μ_J, instead of the physiological parameter μ_{JV}, was kept fixed, at $\mu_J = 0.5$.

models 1, 2, 4, or 6) or $\theta_j(E_{attr}(T_V^*)) < 1$ (in models 3 or 5), *i.e.*, values of γ_i or θ_j which were excluded *a priori* in our model specification. In pane 1 the lower limit of μ_A results from the additional mortality due to environmental feedback. In pane 2 we see that a feedback through the adult mortality by necessity exactly compensates for any difference in the adult mortality rate in the virgin environment.

Apparently different feedback rules can lead to radically different patterns. Pane 1 of Figure 1 differs from Figure 2 by whether we plot cases with matching values of μ_{JV} (Figure 1) or matching values of μ_J (Figure 2). The second picture applies to a protocol in which we select species on the basis of their equality of the observed value of μ_J, the first picture to the more usual protocol where we select them for their *a priori* expected similarity with respect to μ_{JV}. Although conceptually different, the two protocols induce similar model predictions. In cases 2 to 6 the predictions for the two protocols are even exactly the same. In cases 2 to 5 this is due to the assumption that $\mu_J = \mu_{JV}$, in case 6 to what appears to be just an algebraic quirk (see Appendix B).

Remark 5.2 We refrained from including plots for all different possible parameter combinations: The plots of T_V^* against μ_{AV}, with μ_{JV} fixed, are less spectacular. The plots for cases 1, 3 and 6 look like the corresponding panes in Figure 1, those for cases 2 and 4 like pane 3 of Figure 1, and the plot for case 5 is equal to that for case 6. The plots of T^* against μ_J, with μ_{AV} fixed, all show a roughly hyperbolically decreasing relation, like in pane 4 of Figure 1. The plots of T_V^* against μ_{JV}, with μ_{AV} fixed, show either a decreasing relation, in cases 1 to 4, or a horizontal line in cases 5 and 6. □

6 Concluding remarks

The main relevance of our propositions is that they rigorously show that on an abstract level the suite of simple examples 3.2 to 3.4 are representative of all population dynamical scenarios allowing an evolutionary extremisation principle. These scenarios can only differ in the, unfortunately often quite horrible, technical details of the calculations.

Our propositions also show that having an extremisation principle really is a rather special property.

In the intuitively obvious case we can point to an intermediate scalar quantity which when increased, increases fitness *in all relevant environments.* As it turns out the environments that matter are those stationary environments that can potentially be generated by the family of communities under consideration as reactions to particular values of the trait vector. A technical elaboration moreover shows that the initial requirement can be weakened by replacing the word "fitness" by the phrase "some quantity that is sign-equivalent with fitness". This technical variant we have dubbed "one-dimensional action" of the trait vector (or strategy parameters, if your leaning is ecological instead of taxonomical).

The other, slightly less obvious, scenario, is that the environment acts one-dimensionally (in the aforementioned technical sense). We have proved that these two cases are effectively only one case, and, what is more, the *only* case allowing an evolutionary extremisation principle. Proposition 3.4 tells moreover that in that case the trait vector and the environment by necessity act not only one-dimensionally but also, in a certain technical sense, independently.

It is our conviction that it is only our own, unwitting or deliberate, moulding of evolutionary scenarios that leads to the frequent occurrence of extremisation principles in the life history models studied in the literature. For more complicated feedback rules shortcuts in the form of an optimisation principle don't exist!

The next step should be to analyse scenarios where there is not one but two essential scalar components of environmental action. The initial stages of such an analysis can be found in Meszéna (1995).

The cases where evolution just maximises r or R_0 are considerably rarer still. First of all the community should generate only constant environments. Secondly the dependencies of r or R_0 on the trait vector in these different environments should be monotonically related.

The example from section 5 furthermore shows how the details of the environmental feedback loop can have a non-trivial influence on the predicted relationships between life-history parameters, even when we restrict ourselves to scenarios where evolution just maximises R_0 (cases 2 to 4).

In conclusion, the choice of a single optimisation criterion, be it R_0 or r or still something else, always entails, often fairly special, assumptions about the nature of the environmental feedback loop. The current literature consistently underemphasises this aspect.

Acknowledgements

The work of Sido Mylius was supported by the Life Sciences Foundation (SLW), which is subsidized by the Netherlands Organization for Scientific Research (NWO). Peter Klinkhamer, James McAllister and Jan Sevenster gave useful comments on a previous version of the manuscript.

References

H. Caswell, 1989. *Matrix Population Models: Construction, Analysis, and Interpretation*. Sinauer Associates, Sunderland, Massachusetts, U.S.A.

B. Charlesworth, 1994. *Evolution in Age-Structured Populations*, vol. 13 of *Cambridge Studies in Mathematical Biology*, second edition. Cambridge University Press, Cambridge, U.K.

B. Charlesworth, 1980. *Evolution in Age-Structured Populations*, vol. 1 of *Cambridge Studies in Mathematical Biology*. Cambridge University Press, Cambridge, U.K.

E.L. Charnov, 1993. *Life History Invariants: Some Explorations of Symmetry in Evolutionary Ecology*. Oxford Series in Ecology and Evolution. Oxford University Press, Oxford, U.K.

O. Diekmann, J.A.P. Heesterbeek & J.A.J. Metz, 1990. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382.

O. Diekmann & J.A.J. Metz, 1994. On the reciprocal relationship between life histories and population dynamics. In: S.A. Levin (ed.), *Frontiers of Mathematical Biology*, pp. 263–279. Springer-Verlag, Berlin, Germany.

R. Ferrière & M. Gatto, 1995. Lyapunov exponents and the mathematics of invasion in oscillatory or chaotic populations. *Theor. Popul. Biol.* **48**, 126–171.

P. Hammerstein, 1996. Darwinian adaptation, population genetics and the streetcar theory of evolution. *J. Math. Biol.* **34**, 511–532.

P. Hammerstein & R. Selten, 1994. Game theory and evolutionary biology. In: R.J. Aumann & S. Hart (eds.), *Handbook of Game Theory with Economic Applications*, vol. 2, pp. 929–993. North-Holland, Amsterdam, The Netherlands.

M.P. Hassell, J.H. Lawton & R.M. May, 1976. Patterns of dynamical behaviour in single species systems. *J. Anim. Ecol.* **45**, 471–486.

P. Jagers, 1991. The growth and stabilization of populations. *Stat. Sci.* **6**, 269–283.

P. Jagers, 1995. The deterministic evolution of general branching populations. In: O. Arino, D. Axelrod & M. Kimmel (eds.), *Mathematical Population Dynamics Vol. III*. Würz, Winnipeg, Canada.

T.J. Kawecki & S.C. Stearns, 1993. The evolution of life histories in spatially heterogeneous environments: optimal reaction norms revisited. *Evol. Ecol.* **7**, 155–174.

J. Kozłowski, 1993. Measuring fitness in life-history studies. *Trends Ecol. Evol.* **8**, 84–85.

S. Lessard, 1990. Evolutionary stability: One concept, several meanings. *Theor. Popul. Biol.* **37**, 159–170.

J. Maynard Smith, 1993. Life history, symmetry and evolution. *Q. Rev. Biol.* **68**, 557–559.

J.A.J. Metz & A.M. de Roos, 1992. The role of physiologically structured population models within a general individual-based modelling perspective. In: D.L. DeAngelis & L.J. Gross (eds.), *Individual-based Models and Approaches in Ecology*, pp. 88–111. Chapman & Hall, New York, U.S.A.

J.A.J. Metz & O. Diekmann (eds.), 1986. *The Dynamics of Physiologically Struc-*

tured Populations, vol. 68 of *Lecture Notes in Biomathematics*. Springer-Verlag, Berlin, Germany.

J.A.J. Metz, S.A.H. Geritz, G. Meszéna, F.J.A. Jacobs & J.S. van Heerwaarden, 1995. Adaptive dynamics: A geometrical study of the consequences of nearly faithful reproduction. Working paper WP-95-099, IIASA, Laxenburg, Austria.

J.A.J. Metz, R.M. Nisbet & S.A.H. Geritz, 1992. How should we define "fitness" for general ecological scenarios? *Trends Ecol. Evol.* **7**, 198–202.

R.E. Michod, 1979. Evolution of life-histories in response to age-specific mortality factors. *Am. Nat.* **113**, 531–550.

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

E. Pásztor, 1988. Unexploited dimensions of optimization life history theory. In: G. de Jong (ed.), *Population Genetics and Evolution*, pp. 19–32. Springer-Verlag, Berlin, Germany.

L. Pásztor, G. Meszéna & É. Kisdi, 1995. Identical optimization models of life-history based on different fitness measures: a simple example with an important message. Manuscript.

D.A. Rand, H.B. Wilson & J.M. McGlade, 1994. Dynamics and evolution: evolutionarily stable attractors, invasion exponents and phenotype dynamics. *Phil. Trans. R. Soc. Lond.* B **343**, 261–283.

D.A. Roff, 1992. *The Evolution of Life Histories: Theory and Analysis*. Chapman & Hall, New York, U.S.A.

J. Roughgarden, 1979. *Theory of Population Genetics and Evolutionary Ecology: An Introduction*. Macmillan, New York, U.S.A.

S.C. Stearns, 1992. *The Evolution of Life Histories*. Oxford University Press, Oxford, U.K.

P.D. Taylor, 1988a. An inclusive fitness model for dispersal of offspring. *J. theor. Biol.* **130**, 363–378.

P.D. Taylor, 1988b. Inclusive fitness models with two sexes. *Theor. Popul. Biol.* **34**, 145–168.

P.D. Taylor, 1989. Evolutionary stability in one-parameter models under weak selection. *Theor. Popul. Biol.* **36**, 125–143.

S. Tuljapurkar, 1989. An uncertain life: Demography in random environments. *Theor. Popul. Biol.* **35**, 227–294.

S. Tuljapurkar, 1990. *Population Dynamics in Variable Environments*, vol. 85 of *Lecture Notes in Biomathematics*. Springer-Verlag, Berlin, Germany.

Appendices

A Theorems underlying the statements in sections 3 and 4

In the main text we presented our propositions in an order which seemed natural in view of their interpretation and/or application. The order in which these results are naturally deducible is rather different. Therefore we make a fresh start. The propositions of the main text should be seen primarily as but a convenient summary of the results from the arguments below.

Convention Whenever we refer to r or R_0 we implicitly restrict ourselves to community dynamical scenarios for which $E_{\text{attr}}(X)$ is time-constant for all relevant X. Otherwise we only require E to be ergodic (and realisable as $E_{\text{attr}}(X)$ for some X). The virgin environment will be denoted as E_{V}.

The following four theorems and corollaries are trivial. The crux are the questions that follow them.

Theorem 1 *If there exist functions ψ of X, and α of ψ and E, to the real numbers, with α increasing in ψ, such that*

$$\operatorname{sign}\alpha(\psi(X),E) \quad = \quad \operatorname{sign}\rho(X,E)$$

then evolution maximises $\psi(X)$ (or equivalently $\alpha(\psi(X),E)$ for any fixed E).

Theorem 2 (universal Verelendungs principle) *If there exist functions ϕ of E, and β of X and ϕ, to the real numbers, with β increasing in ϕ, such that*

$$\operatorname{sign}\beta(X,\phi(E)) \quad = \quad \operatorname{sign}\rho(X,E)$$

then evolution minimises $\phi(E_{\text{attr}}(X))$.

Corollary 3 *If we can write $r(X,E)$ in the form*

$$r(X,E) \quad = \quad \alpha(\psi(X),E)\ ,$$

with α increasing in ψ, then evolution maximises $r(X,E_{\text{V}})$ (and, more generally, $r(X,E_0)$ for any fixed E_0).

Corollary 4 *If we can write $R_0(X,E)$ in the form*

$$R_0(X,E) \quad = \quad \exp\left(\alpha(\psi(X),E)\right)\ ,$$

with α increasing in ψ, then evolution maximises $R_0(X,E_{\text{V}})$ (and, more generally, $R_0(X,E_0)$ for any fixed E_0).

Questions

1. Is there any relation between theorems 1 and 2?
2. Can theorems 1 and 2 be made into “if and only if” statements, e.g. by requiring that the extremisation principle should hold independent of the particular choice we may still make for a constraint on X?

3. Is this also possible for the corollaries?

Theorem 5 (answer to question 1) *The assumptions of both theorems 1 and 2 are equivalent to: There exist functions ϕ of E, and ψ of X to the real numbers, such that*

$$\operatorname{sign}\left(\psi(X) + \phi(E)\right) = \operatorname{sign}\rho(X, E)\ . \tag{65}$$

Proof: Theorem 1: Define the function ϕ of E to the real numbers by $\alpha(-\phi(E), E) = 0$. Then

$$\operatorname{sign}\left(\psi(X) + \phi(E)\right) = \operatorname{sign}\alpha(\psi(X), E) = \operatorname{sign}\rho(X, E).$$

Therefore the assumption of theorem 1 implies the assumption made above. The converse implication is obvious.

Theorem 2: Let $\psi(X) := -\phi(E_{\text{attr}}(X))$. As $\beta(X, \phi(E_{\text{attr}}(X))) = 0$

$$\begin{aligned} \operatorname{sign}\left(\phi(E) + \psi(X)\right) &= \operatorname{sign}\left(\phi(E) - \phi(E_{\text{attr}}(X))\right) = \\ \operatorname{sign}\beta(X, \phi(E)) &= \operatorname{sign}\rho(X, E). \end{aligned}$$

Therefore the assumption of theorem 2 implies the assumption made above. The converse implication is obvious. □

Apparently we may without loss of essential information replace $\alpha(\psi, E)$ by $\psi + \phi(E)$ respectively $\beta(X, \phi)$ by $\psi(X) + \phi$, with ϕ respectively ψ defined above.

Remark 1.1 The reasoning underlying theorem 5 does not extend to corollaries 3 and 4: From $r(X, E) = \alpha(\psi(X), E)$ we cannot even conclude that there exist functions ϕ' of E and ψ' of X such that $r(X, E) = \psi'(X) + \phi'(E)$. Neither can we conclude from $R_0(X, E) = \exp(\alpha(\psi(X), E))$ that there exist functions ϕ' of E and ψ' of X such that $R_0(X, E) = \exp(\psi'(X) + \phi'(E))$. □

The next theorem is again trivial. However, it forms a natural introduction to the somewhat unexpected, though on second thought equally trivial, theorem 7.

Theorem 6 (first part of the answer to question 2)

(1) If we require that we can determine the ESS under any possible constraint by maximising a function ψ of X then this function is uniquely determined up to an increasing transformation.

(2) If we require that that we can determine the ESS under any possible constraint by minimising a function ϕ of $E \in E_{\text{attr}}(X)$ then this function is uniquely determined up to an increasing transformation.

Theorem 7 (second part of the answer to question 2)

(1) If there exists a function ψ of X to the real numbers such that we can determine the ESS value of X by maximising ψ, independent of any choice that we may still make for a constraint on X, then there exists a function ϕ of E such that (65) applies.

(2) If there exists a function ϕ of E to the real numbers such that we can determine the ESS value of X by minimising $\phi(E_{\text{attr}}(X))$, independent of any choice that we may still make for a constraint on X, then there exists a function ψ of X such that (65) applies.

(3) The functions ϕ respectively ψ are uniquely determined by their counterparts.

Proof: In case (1) we define ϕ by $\phi(E_{\text{attr}}(X)) := -\psi(X)$. In case (2) we define $\psi(X) := -\phi(E_{\text{attr}}(X))$. (65) is derived by considering all possible constraints of the type $X \in \{X_1, X_2\}$. Maximising $\psi(X)$ or minimising $\phi(E_{\text{attr}}(X))$ will only predict the right ESS for this constraint if $\operatorname{sign}(\psi(Xi) + \phi(E_{\text{attr}}(X_j))) = \operatorname{sign}\rho(X_i, E_{\text{attr}}(X_j))$ for all values of i and j. Uniqueness of ϕ respectively ψ follows from the fact that $\operatorname{sign}(\psi(X) + \phi(E_{\text{attr}}(X)))$ should be 0. □

Apparently any optimisation principle ψ automatically carries a pessimisation principle ϕ in its wake, and vice versa.

Corollary 8 (last part of the answer to question 2) *We may replace the opening "if"s of theorems 1 and 2 by "iff"s.*

Corollary 9 (first part of the answer to question 3)

(1) If we can determine the ESS value of X by maximising $r(X, E_0)$ for some special value E_0 of E, independent of any choice that we may still make for a constraint on X, then there exists a function ϕ of E such that

$$\operatorname{sign}(r(X, E_0) + \phi(E)) = \operatorname{sign} r(X, E) .$$

(2) If we can determine the ESS value of X by maximising $R_0(X, E_0)$ for some special value E_0 of E, independent of any choice that we may still make for a constraint on X, then there exists a function ϕ of E such that

$$\operatorname{sign}(\ln(R_0(X, E_0)) + \phi(E)) = \operatorname{sign}\ln(R_0(X, E)) .$$

It is not possible to get any representation of $r(X, E)$ or $R_0(X, E)$ under the, weak, condition that there is at least one E_0 such that evolution maximises $r(X, E_0)$ respectively $R_0(X, E_0)$. We need to make a stronger assumption about the sense in which evolution maximises r respectively R_0:

Theorem 10 (last part of the answer to question 3)

(1) If the maximisation principle from corollary 9 (1) holds good for all possible choices of E_0, then it is possible to write

$$r(X, E) = \alpha(\psi(X), E) ,$$

with α increasing in its first argument and $\psi(X) = r(X, E_0)$ for some, arbitrary but fixed, E_0.

(2) If the maximisation principle from corollary 9 (2) holds good for all possible choices of E_0, then it is possible to write

$$R_0(X, E) = \exp(\beta(\psi(X), E)) ,$$

with β increasing in its first argument and $\psi(X) = \ln(R_0(X, E_0))$ for some, arbitrary but fixed, E_0.

Proof: The maximisation of, say, $\gamma(X, E)$, E fixed, can only lead to the same value of the maximum as the maximisation of $\gamma(X, E_0)$ for all possible constraints if $\gamma(X, E_0)$ and $\gamma(X, E)$, considered as functions of X, are related by an increasing function: $\gamma(X, E) = f(\gamma(X, E_0), E, \gamma)$, where the last argument is at this stage only notional. For any given E (and γ) this function is necessarily unique. In cases (1) and (2) we define $\alpha(\psi, E) := f(\psi, E, r)$ respectively $\beta(\psi, E) := \ln(f(\psi, E, R_0))$. □

B Analysis of the example from section 5

Case 1 We consider the maximisation of r defined by

$$g(r, T_V) = 1\,, \tag{66}$$

with

$$g(r, T_V) = \frac{b_V(T_V)\, e^{(r+\mu_{JV})\, T_V}}{r + \mu_{AV}}\,. \tag{67}$$

Implicit differentiation of (66) gives

$$\frac{\partial r}{\partial T_V}\frac{\partial g}{\partial r} = -\frac{\partial g}{\partial T_V}\,. \tag{68}$$

From (68) we see immediately that g decreases in r. Therefore $\frac{\partial g}{\partial r} < 0$. It is also easy to see (i) that $\frac{\partial g}{\partial T_V} < 0$ for T_V sufficiently large, and (ii) that the fact that $b_V(1) = 0$, and that b_V increases in T_V, imply that $\frac{\partial g}{\partial T_V} > 0$ for $T_V = 1$. Therefore r has at least one maximum in $(1, \infty)$.

To calculate that maximum we set $\frac{\partial r}{\partial T_V} = 0$ in (68). This tells us that at $T_V = T_V^*$

$$\frac{\partial g}{\partial T_V} = 0\,.$$

By differentiating (67) for T_V we find that

$$\frac{\partial g}{\partial T_V} = \frac{\partial b_V}{\partial T_V}\frac{g}{b_V} - (r + \mu_{JV})\, g\,.$$

Substitution of the resulting relation

$$(r + \mu_{JV}) = \frac{d\ln(b_V)}{dT_V}$$

in (66) with (67) gives

$$b_V(T_V)\, \exp\left(-\frac{d\ln(b_V)}{dT_V}T_V\right) = \frac{d\ln(b_V)}{dT_V} + (\mu_{AV} - \mu_{JV}) \tag{69}$$

together with

$$\frac{d\ln(b_V)}{dT_V} > (\mu_{JV} - \mu_{AV})\,.$$

The next step is to substitute (47). This reduces (69) to

$$(T_V - 1)\ \exp\left(-\frac{T_V}{T_V - 1}\right) \quad = \quad \frac{1}{T_V - 1} + (\mu_{AV} - \mu_{JV})\ . \tag{70}$$

The introduction of

$$y \quad := \quad \frac{1}{T_V^* - 1} \tag{71}$$

lets us replace (70) by

$$\frac{e^{-(1+y)}}{y} - y \quad = \quad \mu_{AV} - \mu_{JV}\ . \tag{72}$$

The left hand side of (72) decreases from ∞ at $y = 0$ to $-\infty$ at $y = \infty$. We conclude that r has a unique optimum T_V^* which can easily be determined from (72) with (71).

Formulas (71) and (72) moreover allow us immediately to plot the relation between T_V^* and μ_A at fixed μ_{JV}. The relation of T_V^* with μ_A for fixed μ_J, can be plotted as a parametric curve, with y as a parameter. □

Cases 2 to 4 From $\frac{\partial R_0}{\partial T_V} = 0$ we find that

$$T_V^* \quad = \quad 1 + \frac{1}{\mu_{JV}}\ .$$

Apparently T_V^* is independent of μ_{AV}. This is clearly brought out in pane 3 of Figure 1, where the environmental feedback loop acts through the birth rate b. The decreasing relation in pane 4 derives entirely from the effect of the environmental feedback loop on $T^* = T_V^* + \gamma_4(E)$. In pane 2 we see the effect of the environmental feedback loop keeping μ_A constant, independent of μ_{AV}. □

Cases 5 and 6 Setting $\frac{\partial \psi}{\partial T_V} = 0$ leads to

$$(T_V - 1)\ \exp\left(-\frac{T_V}{T_V - 1}\right) \quad = \quad \mu_{AV}\ . \tag{73}$$

When T_V increases from 1 to ∞ the left hand side of (73) increases from 0 to ∞. Therefore (73) has a unique solution.

In case 5 we plot the relation between $T^* = \theta_5(E)T_V^*$ and μ_A as a parametric curve with T_V^* as parameter. Although in case 6 the feedback loop influences μ_J, it makes no difference whether we keep μ_{JV} or μ_J constant, as by (73) T^* is independent of μ_{JV}. □

C Bringing cases 3 to 4 from the example in section 5 in line with example 4.2

We can, by slightly reinterpreting the model formulation, make each of the cases 2 to 4 from the example from section 5 into a special case of example 4.2. This is done by introducing a third stage which is either the only stage affected by E, and isn't affected by T_V, or the only stage affected by T_V, and isn't affected by E. We shall consider the cases in opposite order.

Case 4: We split the juvenile period into a basic juvenile period of length T_V, and a subadult period of length $\gamma_4(E)$.

Case 3: We introduce an infinitesimally short nursery stage before the juvenile stage. Adults reproduce according to $b_\mathrm{V}(T_\mathrm{V})$. Nursery survival is $1/\theta_3(E)$.

Case 2: We again apply the nursery stage trick, except that we now assume that the adult reproduction rate and nursery survival are

$$b_\mathrm{M} \;:=\; \max_{T_\mathrm{V}} \left(b_\mathrm{V}(T_\mathrm{V})\right) , \qquad \text{and} \qquad \frac{b_\mathrm{V}(T_\mathrm{V})}{b_\mathrm{M}} \qquad \text{respectively} . \tag{74}$$

(In case 3 we first consider models with a maximum adult lifespan, to let (74) make sense, and then use a limit argument.)