

From the Price equation to the selection gradient in class-structured populations: a quasi-equilibrium route

Sébastien Lion

March 30, 2018

Centre d'Écologie Fonctionnelle et Évolutive (CEFE), CNRS, Université de Montpellier, Université Paul-Valéry Montpellier 3, EPHE, IRD. 1919, route de Mende 34293 Montpellier Cedex 5, France.

E-mail: sebastien.lion@cefe.cnrs.fr

Article published as

Lion, S. (2018) From the Price equation to the selection gradient in class-structured populations: a quasi-equilibrium route. *Journal of Theoretical Biology*. doi: [10.1016/j.jtbi.2018.03.033](https://doi.org/10.1016/j.jtbi.2018.03.033).

Abstract

Recent studies in theoretical evolutionary ecology have emphasised two approaches to modelling evolution. On the one hand, models based on a separation of time scales rely on the concept of invasion fitness. On the other hand, models based on the Price equation track the dynamics of a trait average, coupled with a description of ecological dynamics. The aim of this article is to show that, in class-structured populations, both approaches yield the same expression for the selection gradient under weak selection. Although the result is not new, I propose an alternative route to its derivation using the dynamics of scaled measures of between-class phenotypic differentiation. Under weak selection, these measures of phenotypic differentiation can be treated as fast variables compared to the trait mean, which allows for a quasi-equilibrium approximation. This suggests a different approach to calculating weak selection approximations of evolutionary dynamics, and clarifies the links between short- and long-term perspectives on evolution in structured populations.

Keywords: weak selection; adaptive dynamics; separation of time scales; reproductive value; population structure

1 Introduction

Most theoretical models in evolutionary ecology rely on a separation of time scales between ecology and evolution: ecological dynamics are assumed to take place on a fast time scale, while evolutionary dynamics unfold on a slow time scale (Metz et al., 1992, Lehmann and Rousset, 2014, Lion, 2018b). Such a separation of time scales arises when mutations are rare, so that the population can be assumed to reach an ecological attractor before the apparition of a new mutation. This result is used in invasion analyses, to compute the invasion fitness of a rare mutant in a resident population on its ecological attractor (Metz et al., 1992, Geritz et al., 1998, Metz, 2008). A separation of time scales is also observed under weak selection, i.e. when mutations have small phenotypic effects (Meszéna et al., 2005, Lehmann and Rousset, 2014, Van Cleve, 2015).

Whereas invasion analyses typically focus on long-term evolution, other methods have been proposed to study short-term evolution in the absence of a separation of time scales between ecology and evolution. For instance, Day and Proulx (2004), Day and Gandon (2006, 2007) have used an approach based on the Price equation and quantitative genetics to analyse the joint epidemiological and evolutionary dynamics of host-parasite interactions. This quantitative genetics formalism takes into account transient eco-evolutionary dynamics, but can also be used to recover the long-term predictions

of invasion analyses when weak selection is assumed (Abrams et al., 1993, Abrams, 2001, Day and Proulx, 2004, Day and Gandon, 2007, Lion, 2018b). In this case, the change in trait mean described by the Price equation is simply proportional to the selection gradient computed from invasion analyses.

In populations structured into discrete classes, a similar result holds (Taylor, 1990, Rousset, 1999, 2004, Otto and Day, 2007, Lehmann et al., 2016). However, the traditional route to this result is to calculate the dynamics of the weighted frequency of a mutant allele, using the reproductive values of each class as weights (Taylor, 1990, Rousset, 2004). This yields a modified Price equation where, instead of tracking the arithmetic mean of the trait of interest, one follows the dynamics of a weighted average of the trait (see also Lion (2018a) for a discussion in a more general setting).

My aim in this article is to describe an alternative route to this important result, based on a quasi-equilibrium approximation for the dynamics of measures of between-class phenotypic differentiation. The motivation is essentially methodological, but it should be of interest to researchers interested in bridging the gap between short-term theory based on the Price equation and long-term theory based on invasion fitness. In particular, I discuss an example where failing to take into account the effect of selection on between-class differentiation can produce wrong results. I also discuss potential applications of this approach to models where reproductive-value weighting is not straightforward, such as models of spatially structured populations analysed using spatial moment equations (Lion and Gandon, 2016).

The paper is organised as follows. In section 2, I present a general continuous-time model for the ecological dynamics of a polymorphic population. The model does not explicitly account for limited dispersal and demographic stochasticity, but potential extensions to spatially structured populations are examined in the discussion. In section 3, I give an overview of the key results of invasion analysis and Price equation approaches for class-structured populations. In section 4, I show that under weak selection, a separation of time scales allows for some simplifications of the coupled eco-evolutionary dynamics. In particular, a quasi-equilibrium approximation can be used to simplify the dynamics of the trait mean, which is then shown to be proportional to the classical form of the selection gradient in class-structured populations. This approach yields an expression of the individual reproductive values as a function of the transition matrix and vector of class frequencies computed at quasi-equilibrium. In section 5, an example is studied using an epidemiological model of a partially vaccinated host populations previously introduced by Gandon and Day (2007). Finally, I discuss the implications of this method, and its potential usefulness for other models, such as evolutionary models of spatially structured populations.

2 A model of polymorphic ecological dynamics

I consider a large population structured into K different classes and M clonally reproducing types. A type can represent an allele, or a phenotype, depending on the level of interest. I will follow the notation in Lion (2018a) and use the subscript i to refer to types and superscripts j and k to refer to classes. The total density of individuals in class k is thus n^k , and the density of type i individuals in class k is n_i^k . These densities are collected in the vectors $\mathbf{n}_i = (n_i^1 \dots n_i^K)^\top$ and $\mathbf{n} = (n^1 \dots n^K)^\top$.

Apart from clonal reproduction and large population densities, I will make only minimal ecological assumptions. The results are only expressed in terms of the transition rates r_i^{kj} of i individuals from class j to class k . These transitions may be due to reproduction, survival, maturation, or dispersal depending on the biological context. In general, the rates r_i^{kj} will depend on the vital rates of the focal type (fecundity, mortality, migration, infection...), but also on the vital rates of the other types.

In addition, the rates r_i^{kj} also depend on the environment \mathbf{E} , which collects the densities of the various types in the population, through the vectors \mathbf{n}_i , but also any ecological effects that are external to the focal population, which are collected in a vector \mathbf{e} . These external effects may represent predation, parasitism, interspecific competition, or changes in resources or abiotic factors. We thus have $\mathbf{E} = (\mathbf{n}_1 \dots \mathbf{n}_K \ \mathbf{e})^\top$.

In continuous time, the dynamics of the total densities in each class can be written in matrix form as

$$\frac{d\mathbf{n}}{dt} = \mathbf{R}(\mathbf{E})\mathbf{n}. \quad (1)$$

The matrix \mathbf{R} has element \bar{r}^{kj} on the k th line and j th column, where $\bar{r}^{kj} = \sum_i r_i^{kj} n_i^j / n^j$ is the average transition rate from class j to class k . The dynamics of the type-specific densities are further given by

$$\frac{d\mathbf{n}_i}{dt} = \mathbf{R}_i(\mathbf{E})\mathbf{n}_i. \quad (2)$$

Note that summing over i yields equation (1). Equation (2) can be coupled to an equation describing the dynamics of external variables \mathbf{e}

$$\frac{d\mathbf{e}}{dt} = \mathbf{T}(\mathbf{E}). \quad (3)$$

Equations (2)-(3) give the full ecological and evolutionary dynamics of the polymorphic population under study. Note that the matrices \mathbf{R} , \mathbf{R}_i and \mathbf{T} may also depend directly on time.

3 Two approaches to describing evolutionary change

In the literature, two main approaches have been used to analyse evolution in this general model. The first approach relies on the concept of invasion fitness, while the second tracks the dynamics of a trait mean using the Price equation. In this section, I summarise the key features of those two methods. This provides the basis for the derivation of section 4, which is the key result of this paper.

3.1 Invasion analysis and the selection gradient

A first (and nowadays standard) approach to studying evolution in the ecological model given by equations (2)-(3) is to assume a separation of time scales between ecology and evolution. This allows to analyse selection using the concept of invasion fitness (Tuljapurkar, 1989, Metz et al., 1992, Charlesworth, 1994, Rand et al., 1994, Otto and Day, 2007, Metz, 2008), defined as the long-term per-capita growth rate of a rare mutant type in the environment set by a community of residents.

Let us write $\mathbf{n}_m(t)$ the vector collecting the densities $n_m^k(t)$ of a focal mutant type in the different classes at type t . From equations (2), the dynamics of the vector $\mathbf{n}_m(t)$ satisfy the following matrix equation:

$$\frac{d\mathbf{n}_m}{dt} = \mathbf{R}_m(\mathbf{E})\mathbf{n}_m \quad (4)$$

If we assume the mutant type is initially rare, so that it does not affect the distribution of the types in the population, we can assume that the environment \mathbf{E} is solely determined by the other types, which we call the resident population. If we further assume that the environment is ergodic (i.e. if it shows no long-lasting trend), the multiplicative ergodic theorem tells us that the long-term growth rate can be computed as the dominant Lyapunov exponent ρ of $\mathbf{R}_m(\mathbf{E}(t))$. We then have

$$\lim_{t \rightarrow \infty} \frac{1}{t} \ln(n_m(t)) = \rho$$

where $n_m(t) = \sum_{k=1}^K n_m^k(t)$ is the total density of mutant individuals (Tuljapurkar, 1989, Metz et al., 1992, Rand et al., 1994, Ferrière and Gatto, 1995). When ρ is positive, the mutant population grows asymptotically at rate ρ , and the relative frequencies of the mutant in each class become independent of the initial condition. When ρ is negative, the mutant population goes extinct.

In particular, when the ecological attractor is an equilibrium, the environment is a constant vector $\hat{\mathbf{E}}$, and the dominant Lyapunov exponent is simply the dominant eigenvalue of the constant matrix $\hat{\mathbf{R}}_m = \mathbf{R}_m(\hat{\mathbf{E}})$. The dominant right eigenvector, \mathbf{u} , then gives the stable class frequencies and the dominant left eigenvector, \mathbf{v} , collects the individual reproductive values in each class (Taylor, 1990, Rousset, 2004).

Let us further assume that the trait distribution in the population is tightly clustered around the population mean, so that the trait of type i , z_i , can be written as $\bar{z} + \varepsilon \delta_i$, with ε a small parameter (in the two-type limit, we can consider a single resident type with trait z_w , and a mutant type with

trait $z_w + \varepsilon$). We can compute the selection gradient as the first-order perturbation of the neutral model where $\varepsilon = 0$ (or $z_m = z_w$ in the two-type model). We then have

$$s = \varepsilon \bar{\mathbf{v}}^\top \frac{d\hat{\mathbf{R}}_m}{d\varepsilon} \bar{\mathbf{u}} + O(\varepsilon^2) \quad (5)$$

where $d\hat{\mathbf{R}}_m/d\varepsilon$ collects all first-order effects of selection on the matrix $\hat{\mathbf{R}}_m$, evaluated at neutrality ($\varepsilon = 0$), while the eigenvectors $\bar{\mathbf{u}}$ and $\bar{\mathbf{v}}$ are evaluated in the neutral model. Equation (5) can be expanded to yield the classical expression for the selection gradient as a weighted sum of the effects of selection on class transitions (Taylor, 1990, Rousset, 1999, 2004)

$$s = \varepsilon \sum_j \sum_k \bar{u}^k \bar{v}^j \frac{dr_m^{jk}}{d\varepsilon} + O(\varepsilon^2). \quad (6)$$

Equation (6) is a very important result of evolutionary game theory and adaptive dynamics in class-structured populations.

3.2 Price equation for the dynamics of the trait mean

An alternative approach to modelling evolutionary dynamics is to track the dynamics of the mean of a phenotypic trait. Let z_i be the trait of type i . The trait mean in class k can then be computed as $\bar{z}^k = \sum_i z_i f_i^k$, where $f_i^k = n_i^k/n^k$ is the frequency of type i in class k . The average trait in the population, \bar{z} , can be calculated as a weighted sum of class means, $\bar{z} = \sum_k \bar{z}^k f^k$, where $f^k = n^k/n$ is the frequency of class k and n is the total density of individuals. With these definitions, the dynamics of the trait mean in the absence of mutation can be written as

$$\frac{d\bar{z}}{dt} = \sum_k f^k \text{cov}_k(z_i, r_i^{\bullet k}) + \sum_k (\bar{z}^k - \bar{z}) \bar{r}^{\bullet k} f^k \quad (7)$$

(Lion, 2018a). The first term of equation (7) is the weighted average, over all classes, of the within-class covariance between the trait and the total contribution of individuals of type i in class k , $r_i^{\bullet k} = \sum_j r_i^{jk}$. The within-class covariance for class k has the following expression

$$\text{cov}_k(z_i, r_i^{\bullet k}) = \sum_i z_i r_i^{\bullet k} f_i^k - \bar{z}^k \bar{r}^{\bullet k}. \quad (8)$$

and measures the effect of selection within class k .

The second term of equation (7) only depends on the mean transition rates $\bar{r}^{\bullet k}$ and on the phenotypic differentiation between the class means and the population mean, $\bar{z}^k - \bar{z}$. The dynamics of the phenotypic differentiation can be tracked using a similar equation,

$$\frac{d\bar{z}^k}{dt} = \sum_j \text{cov}_j(z_i, r_i^{kj}) \frac{f^j}{f^k} + \sum_j (\bar{z}^j - \bar{z}) \bar{r}^{kj} \frac{f^j}{f^k} - (\bar{z}^k - \bar{z}) \sum_j \bar{r}^{kj} \frac{f^j}{f^k}. \quad (9)$$

A crucial insight of equations (7) and (9) is that, even when the covariances are zero, the trait means \bar{z} and \bar{z}^k may still change due to between-class demographic transitions if there is some between-class phenotypic differentiation such that $\bar{z}^k - \bar{z} \neq 0$.

Equation (7) is a useful starting point for studying the short-term and long-term dynamics of the mean of a trait of interest. For instance, a variant of this equation has been used by Gandon and Day (2007) to investigate the evolutionary dynamics of pathogen virulence in a host population subdivided into a vaccinated class and a naive class (this example will be revisited in section 5). Equation (7) is valid at any time, and therefore can be used to make long-term predictions. However, it is not immediately clear how the selection gradient (6) can be recovered from a weak-selection approximation of equation (7). Simply approximating the covariances by derivatives $dr_m^{jk}/d\varepsilon$ does not work, because we then obtain equation (6) only in the limit where the reproductive values \bar{v}^j are all unity. Somehow, we need to take into account the second term which depends on the measures of phenotypic differentiation $\bar{z}^k - \bar{z}$. Showing how to do this is the aim of the next section.

4 The weak selection limit of the Price equation

In this section, I show that, under weak selection, the selection gradient (6) can be recovered from the Price equation (7) through a quasi-equilibrium approximation of the measures of phenotypic differentiation given by $\bar{z}^k - \bar{z}$. I first give a general description of the two time scales in the model under weak selection, before showing how a quasi-equilibrium approximation can be applied to calculate the selection gradient.

4.1 Fast-slow dynamics: general argument

At a given time t , the distribution of types is described by a vector $\mathbf{p}(t)$ collecting the frequencies of types in each class (the f_i^k 's, so \mathbf{p} has length $M \times K$), while the distribution of classes is described by the vector \mathbf{f} collecting the frequencies f^k . If the trait distribution is tightly clustered around the mean (e.g. we can write $z_i = \bar{z} + \varepsilon \delta_i$, where ε is a small parameter), the vector \mathbf{p} will change slowly compared to the demographic variables n , \mathbf{f} and \mathbf{e} , and so will the moments of the trait distribution such as \bar{z} , the variance $\omega = \text{var}(z_i)$ and so on. Crucially, even if \mathbf{p} changes on a slow time scale, the vector \mathbf{d} of differentiations between type frequencies, with elements $d_i^k = f_i^k - f_i$, changes on a fast time scale. A full justification of these different time scales will be given later on. For now, it suffices to note that this can be done by Taylor-expanding the rate of the change of the different variables. Introducing the following vector of fast variables,

$$\mathbf{q} = \begin{pmatrix} n \\ \mathbf{f} \\ \mathbf{e} \\ \mathbf{d} \end{pmatrix}, \quad (10)$$

the vectors \mathbf{q} and \mathbf{p} then satisfy a system of dynamical equations of the form

$$\frac{d\mathbf{q}}{dt} = \mathbf{g}(\mathbf{q}, \mathbf{p}, t), \quad (11a)$$

$$\frac{d\mathbf{p}}{dt} = \varepsilon \mathbf{h}(\mathbf{q}, \mathbf{p}, t), \quad (11b)$$

which is characteristic of slow-fast systems analysed in singular perturbation theory (Tikhonov, 1952, Fenichel, 1979, Verhulst, 2005). System (11) has the canonical form for the application of Tikhonov (1952)'s theorem, and can be used to investigate the behaviour of the model on both the fast and slow time scales.

Fast (ecological) dynamics In the limit $\varepsilon \rightarrow 0$, the system becomes

$$\frac{d\mathbf{q}}{dt} = \mathbf{g}(\mathbf{q}, \mathbf{p}, t) \quad (12a)$$

$$\frac{d\mathbf{p}}{dt} = 0 \quad (12b)$$

This means that, on a fast time scale, the distribution of types \mathbf{p} stays approximately constant (i.e. $\mathbf{p} = \mathbf{p}(0)$) while, if equation (12a) admits a stable equilibrium for fixed \mathbf{p} , the vector \mathbf{q} converges towards an equilibrium value \mathbf{q}^* given by the implicit equation $\mathbf{g}(\mathbf{q}^*, \mathbf{p}(0), t) = 0$.

Slow (evolutionary) dynamics We now introduce the rescaled time $\tau = t/\varepsilon$. This gives the slow system

$$\varepsilon \frac{d\mathbf{q}}{d\tau} = \mathbf{g}(\mathbf{q}, \mathbf{p}, t) \quad (13a)$$

$$\frac{d\mathbf{p}}{d\tau} = \mathbf{h}(\mathbf{q}, \mathbf{p}, t). \quad (13b)$$

In the limit $\varepsilon \rightarrow 0$, the vector \mathbf{p} moves slowly along the slow manifold characterised by $\mathbf{g}(\mathbf{q}, \mathbf{p}, t) = 0$. Assuming that the equation of the manifold can be written as $\mathbf{q} = \tilde{\mathbf{q}}(\mathbf{p}, t)$, Tikhonov (1952)'s theorem states that the solution of the full system will be close to the solution of

$$\frac{d\mathbf{p}}{dt} = \mathbf{h}(\tilde{\mathbf{q}}(\mathbf{p}, t), \mathbf{p}, t). \quad (14)$$

This justifies the following quasi-equilibrium approximation for an arbitrary distribution of types, provided the trait distribution is tightly clustered around the mean.

Importantly, on the slow manifold, the fast variables are not constant, but still vary with time: their time evolution is enslaved to the dynamics of the slow variable. Condition $\mathbf{g}(\mathbf{q}, \mathbf{p}, t) = 0$ must then be interpreted as a constraint on the evolutionary dynamics, rather than a true equilibrium condition.

4.2 Fast-slow dynamics for aggregate variables

Because the Price equation is written as a function of the mean phenotypes \bar{z}^k and \bar{z} , it makes more sense to work with aggregate variables rather than the full distribution \mathbf{p} . I will therefore compare the dynamics of the trait mean, \bar{z} , and that of the vector \mathbf{d} of scaled phenotypic differentiation, with elements $d_k = (\bar{z}^k - \bar{z})/\sigma$, where $\sigma = \sqrt{\omega}$ is the standard deviation in the trait. The reason for scaling the phenotypic differentiation by the standard deviation is to work with a quantity that does not vanish in the limit $\varepsilon \rightarrow 0$. As shown in Appendix A, the dynamics of \mathbf{d} are $O(1)$ and therefore take place on the fast time scale. In contrast, the dynamics of the trait mean are at least $O(\varepsilon)$, and thus \bar{z} is a slow variable (Appendix A).

To simplify the analysis, we rewrite the dynamics in matrix form. For the trait mean, equation (7) can be rewritten as

$$\frac{d\bar{z}}{dt} = \mathbf{1}^\top \mathbf{C}\mathbf{f} + \sigma \mathbf{1}^\top \mathbf{R}\mathbf{F}\mathbf{d}, \quad (15)$$

where $\mathbf{1}^\top$ is a K -dimensional row vector with ones, \mathbf{C} is the matrix of covariances with elements $C_{jk} = \text{cov}_k(z_i, r_i^{jk})$ and $\mathbf{F} = \text{diag}(\mathbf{f})$ is the diagonal matrix obtained from the vector of class frequencies.

We can combine equations (9) and (7) to obtain the dynamics of \mathbf{d} (Appendix A), which gives

$$\frac{d\mathbf{d}}{dt} = \frac{1}{\sigma} \mathbf{B}\mathbf{C}\mathbf{f} + \mathbf{A}\mathbf{d} - \frac{d \ln \sigma}{dt} \mathbf{d} \quad (16)$$

where $\mathbf{B} = \mathbf{F}^{-1} - \mathbf{1}\mathbf{1}^\top$, $\mathbf{A} = \mathbf{Q} - \mathbf{1}\mathbf{1}^\top \mathbf{R}\mathbf{F}$, and \mathbf{Q} is the matrix with elements $q_{jk} = \bar{r}^{jk} f^k / f^j$ for $j \neq k$ and $q_{kk} = -\sum_{j \neq k} q_{kj}$.

The crux of the quasi-equilibrium approximation of the Price equation is then to calculate the quasi-equilibrium expression of \mathbf{d} from equation (16) and then to plug it into equation (15).

4.3 Quasi-equilibrium expressions of demographic variables

The dynamics of the fast demographic variables n and \mathbf{f} are given by the system

$$\begin{aligned} \frac{dn}{dt} &= \bar{r}(\mathbf{E})n \\ \frac{d\mathbf{f}}{dt} &= \mathbf{R}(\mathbf{E})\mathbf{f} - \bar{r}(\mathbf{E})\mathbf{f} \end{aligned}$$

along with equation (3) for external environmental variables (see also Lion (2018a)). Setting the right-hand sides to zero, this allows an implicit characterisation of the slow manifold through the equation

$$\mathbf{R}\mathbf{f} = \mathbf{0}. \quad (17)$$

4.4 Weak-selection approximation of phenotypic differentiations

The goal of this section is to characterise more precisely the slow manifold by giving an approximation of the vector of phenotypic differentiation \mathbf{d} . Because the first term of equation (15) is $O(\varepsilon^2)$ and σ is $O(\varepsilon)$, we seek an approximation of \mathbf{d} up to first order in ε . Using the expansion $\mathbf{d} = \mathbf{d}^{(0)} + \varepsilon \mathbf{d}^{(1)} + O(\varepsilon^2)$, we obtain the zeroth-order system

$$\frac{d\mathbf{d}^{(0)}}{dt} = \mathbf{A}^{(0)} \mathbf{d}^{(0)} \quad (18)$$

where $\mathbf{A}^{(0)}$ is the neutral value of \mathbf{A} (Appendix A). Note that the matrix $\mathbf{A}^{(0)}$ does not depend on the distribution of the mutant across classes, because $\mathbf{A}^{(0)}$ is a function of average transition rates evaluated at $\varepsilon = 0$, that is in a monomorphic population. Thus, $\mathbf{A}^{(0)}$ is only a function of the fast variables n , \mathbf{f} and \mathbf{e} . If these variables reach an equilibrium, the dynamics of \mathbf{d} are given in the long run by a linear ordinary differential equation, and, for biologically realistic models, \mathbf{d} will decrease exponentially to zero, yielding $\mathbf{d}^{(0)} = \mathbf{0}$ (Appendix B). In a two-type model, $\mathbf{d}^{(0)} = (f_m^{k,(0)} - f_m^{(0)}) / \sqrt{f_m^{(0)}(1 - f_m^{(0)})}$, where $f_m^{k,(0)}$ and $f_m^{(0)}$ are the neutral frequencies of a focal mutant allele. Hence, in the neutral model, ecological dynamics cause the distribution of types to become homogeneous, and at equilibrium the frequencies of the mutant in each class are equal to its global frequency. This occurs in the absence of selection and is simply due to the changes in the demographic variables.

Moving up to the first-order term, it is shown in Appendix A that the equation for $\mathbf{d}^{(1)}$ reads

$$\frac{d\mathbf{d}^{(1)}}{dt} = \frac{1}{\sigma^{(1)}} \mathbf{B}^{(0)} \mathbf{C}^{(2)} \mathbf{f}^{(0)} + \mathbf{A}^{(1)} \mathbf{d}^{(0)} + \mathbf{A}^{(0)} \mathbf{d}^{(1)} - \left(\frac{d \ln \sigma}{dt} \right)^{(1)} \mathbf{d}^{(0)} \quad (19)$$

where the superscripts refer to the order of the perturbation of the variable that is considered. Using $\mathbf{d}^{(0)} = \mathbf{0}$, this simplifies to

$$\frac{d\mathbf{d}^{(1)}}{dt} = \frac{1}{\sigma^{(1)}} \mathbf{B}^{(0)} \mathbf{C}^{(2)} \mathbf{f}^{(0)} + \mathbf{A}^{(0)} \mathbf{d}^{(1)} \quad (20)$$

Solving the latter equation requires some technicalities (we cannot simply calculate the inverse of $\mathbf{A}^{(0)}$, which is singular on the slow manifold, i.e. when equation (17) holds true), but in the end, it is shown in Appendix B that the quasi-equilibrium expression of \mathbf{d} can be written as

$$\tilde{\mathbf{d}} \approx -\frac{1}{\sigma} \mathbf{L} \mathbf{B} \mathbf{C} \mathbf{f} + O(\varepsilon^2),$$

where the matrix \mathbf{L} depends on the matrix \mathbf{R} and the vector of class frequencies \mathbf{f} (a full expression is given in Appendix B). Note that, for simplicity, I have dropped the superscripts pertaining to the perturbation order. Strictly speaking, to compute $\tilde{\mathbf{d}}$ up to first order, we only need to retain the neutral values of \mathbf{L} , \mathbf{B} and \mathbf{f} , while \mathbf{C} is evaluated using the standard second-order approximation of the covariance. For instance, for a focal mutant in a resident population, we have

$$\mathbf{C} = \boldsymbol{\Omega} \left. \frac{d\mathbf{R}_m}{d\varepsilon} \right|_{\varepsilon=0} + O(\varepsilon^3), \quad (21)$$

where $\boldsymbol{\Omega} = \text{diag}(\omega^1, \dots, \omega^K)$ is a diagonal matrix containing the class variances ω^j .

4.5 Slow dynamics on the evolutionary time scale

We then plug this quasi-equilibrium approximation into the equation for the change in trait mean, which yields

$$\frac{d\bar{z}}{dt} = \mathbf{v}^\top \mathbf{C} \mathbf{f} + O(\varepsilon^3) \quad (22)$$

where

$$\mathbf{v}^\top = \mathbf{1}^\top [\mathbf{I} - \mathbf{R} \mathbf{F} \mathbf{L} \mathbf{B}]. \quad (23)$$

The matrix \mathbf{C} and the vectors \mathbf{v} and \mathbf{f} are calculated using the quasi-equilibrium approximation (i.e., on the slow manifold). Note that, if selection is weak, the covariance matrix is $O(\varepsilon^2)$ and this justifies that we can approximate the vectors \mathbf{f} and \mathbf{v} by their values in the neutral model to obtain an approximation of $d\bar{z}/dt$ to second order.

4.6 Properties of \mathbf{v} and \mathbf{f} on the slow manifold

The vectors \mathbf{v}^\top and \mathbf{f} have important properties. First, they are co-normalised, that is $\mathbf{v}^\top \mathbf{f} = 1$. This is a direct consequence of equation (23) and does not require an assumption of neutrality or equilibrium (Appendix C).

Second, \mathbf{v}^\top and \mathbf{f} are linked on the slow manifold by the relationship:

$$\mathbf{v}^\top \mathbf{R} = \mathbf{R} \mathbf{f} = \mathbf{0}. \quad (24)$$

The second equality is simply the implicit characterisation of the slow manifold given by equation (17). The equality $\mathbf{v}^\top \mathbf{R} = \mathbf{0}$ follows directly from equations (23) and (17), although the proof is a bit lengthy (Appendix C). Thus, at any time on the slow manifold, \mathbf{f} is a right eigenvector of the matrix \mathbf{R} associated with eigenvalue 0, while \mathbf{v}^\top is a left eigenvector, i.e. it gives the individual reproductive values in each class (Taylor, 1990, Rousset, 2004). Importantly, equality (24) is valid for non-neutral populations, although the neutral values of \mathbf{v} and \mathbf{f} are sufficient to compute the desired approximation of $d\bar{z}/dt$. In a neutral population at equilibrium, the vectors \mathbf{v} and \mathbf{f} are equal to the eigenvectors $\bar{\mathbf{v}}$ and $\bar{\mathbf{u}}$ introduced in section 3.1.

As a simple illustration, it is helpful to look at the two-class model ($K = 2$), where we have

$$\mathbf{f} = \begin{pmatrix} f^A \\ f^B \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} r^{AA} & r^{AB} \\ r^{BA} & r^{BB} \end{pmatrix} \quad \text{and} \quad \mathbf{v}^\top = \left(\frac{f^A r^{BA}}{(f^A)^2 r^{BA} + (f^B)^2 r^{AB}} \quad \frac{f^B r^{AB}}{(f^A)^2 r^{BA} + (f^B)^2 r^{AB}} \right),$$

from which the conditions (24) are easily verified.

4.7 The selection gradient

To make a full connection with the selection gradient derived from invasion analyses, we need to assume that, under weak selection, the class variances will be approximately equal to the population variance ω , which gives

$$\frac{d\bar{z}}{dt} = \omega \sum_k \sum_j f^k v^j \left. \frac{dr_m^{jk}}{d\varepsilon} \right|_{\varepsilon=0} + O(\varepsilon^3). \quad (25)$$

If selection is weak, the dynamics of the trait mean will closely follow the trajectory indicated by the selection gradient, provided the class variances are approximately equal to the population variances. This latter approximation should be good under weak selection after initial transients have decayed, because, as we have seen, the dynamics of the neutral model tend to homogenise the distribution of types in each class. Turning to the two-type model for illustration, we have $\omega^j = \varepsilon^2 f_m^{j,(0)} (1 - f_m^{j,(0)}) + O(\varepsilon^3)$, and, because $\mathbf{d}^{(0)} = \mathbf{0}$ on the slow manifold, we have at quasi-equilibrium $f_m^{j,(0)} = f_m^{(0)}$, which implies $\omega^j = \omega + O(\varepsilon^3)$.

Of course, equation (25) is only the first of a hierarchy of moment equations, giving the change in the various moments of the trait distribution. Numerically integrating equation (25) will give meaningless results if we don't have a good approximation of the dynamics of the variance. Such approximations can be derived by writing a dynamical equations for the variance, under some suitable assumption on the distribution of the trait, such as normality (Slatkin, 1980, Taylor and Day, 1997, Day and Proulx, 2004, Wakano and Iwasa, 2013, Wakano and Lehmann, 2014, Débarre and Otto, 2016). However, the value of equation (25) is to show that the direction of selection is solely determined by the selection gradient. If we are interested in the *sign* of the change in trait mean, and not on the full trajectory of \bar{z} , the selection gradient is all we need.

4.8 Comparison with invasion analyses

In contrast to traditional invasion analyses, the derivation of equation (25) does not require that the system is poised at a demographic equilibrium. That is, we do not look at the perturbation of the equilibrium of a resident population resulting from the introduction of a mutant with small effect, as done in section 3.1. Equation (25) thus also gives the change in the trait mean during transient ecological dynamics, provided selection is sufficiently weak (see e.g. figure 1).

Nonetheless, the selection gradient takes the same form, and will give the same result as invasion analyses when the system reaches its equilibrium. In particular, if ε is vanishingly small, or if we start close to the equilibrium of the neutral population, we recover exactly equation (6), where the selective effects $dr_m^{jk}/d\varepsilon$ are weighted by the class frequencies and reproductive values in a neutral model at equilibrium. In this limit, the zeros of equation (25) also give the potential evolutionary endpoints.

There is one subtle apparent difference between equations (6) and (25): while the double sum in equation (6) is scaled by ε , the same quantity in (25) is scaled by ω , which is $O(\varepsilon^2)$. This is simply due to the fact that equation (6) derives the change in mutant frequency, while equation (25) derives the change in trait mean. For a unimodal trait distribution, the rate of change in these two quantities differ by a factor ε , as can be seen in a two-allele model with traits z_w and $z_m = z_w + \varepsilon$, and mutant frequency f_m . In this case, the trait average is $\bar{z} = z_w + \varepsilon f_m$, and therefore $d\bar{z}/dt = \varepsilon df_m/dt$.

5 Examples

5.1 A two-habitat model

To clarify the arguments given above, I now present a simple two-habitat example. Consider a model with two habitats, N and V , (for “neat” and “vile” respectively). The evolution of phenotypic traits was analysed in this simplified setting by Day and Gandon (2006), Gandon and Day (2007). From equation (7), we know that the dynamics of a trait mean \bar{z} is given by

$$\frac{d\bar{z}}{dt} = f^N \sigma_{zr}^N + f^V \sigma_{zr}^V + (\bar{z}^N - \bar{z})(\bar{r}^{NN} + \bar{r}^{VN})f^N + (\bar{z}^V - \bar{z})(\bar{r}^{NV} + \bar{r}^{VV})f^V, \quad (26)$$

where $\sigma_{zr}^N = \text{cov}_N(z_i, r_i^{NN} + r_i^{VN})$ and an analogous definition for σ_{zr}^V with superscripts N and V swapped. Thus, the change in trait mean can be partitioned into within-habitat and between-habitat components.

For a two-habitat model, additional symmetries due to the fact that $f^N = 1 - f^V$ allow the equation to be rewritten in the following form

$$\frac{d\bar{z}}{dt} = f^N \sigma_{zr}^N + f^V \sigma_{zr}^V + (\bar{z}^N - \bar{z}^V) \left(\bar{r}^{VN} f^N - \bar{r}^{NV} f^V + \frac{df^N}{dt} \right) \quad (27)$$

which is equation (2.13) in Gandon and Day (2007) (note however that superscripts j and k in rates r_i^{jk} are swapped compared to the notations in the latter paper). If the frequency of individuals in habitat N is stable over time ($df^N/dt = 0$), the between-habitat component can be interpreted as a migration term: in the presence of between-habitat differentiation ($\bar{z}^N \neq \bar{z}^V$), migration between habitats (at rates $r^{VN} f^N$ and $r^{NV} f^V$) will cause a change in the trait mean \bar{z} measured at the population level. Changes in the distribution of individuals among habitats, through the term df^N/dt , introduce an additional cause of change in trait mean.

From equation (27), it is possible to recover the selection gradient under weak selection. However doing so requires some care. For instance, Gandon and Day (2007) consider that, provided mutation rates are low, the trait means in each habitat will be close, so that the term in $\bar{z}^N - \bar{z}^V$ can be neglected. Then, the equilibrium condition is given by $f^N \sigma_{zr}^N + f^V \sigma_{zr}^V = 0$. However, this approach is flawed because selection itself can produce differentiation between habitats, and we need to take this into account when measuring the long-term evolution of the trait.

The proper way of calculating the selection gradient is then to calculate the dynamics of the scaled differentiation $(\bar{z}^N - \bar{z}^V)/\sigma$, and to solve for the quasi-equilibrium value of this scaled differentiation under weak selection. We obtain

$$\frac{\bar{z}^N - \bar{z}^V}{\sigma} \approx \frac{1}{\sigma} \frac{\left[\sigma_{zr}^{NN} - \frac{f^N}{f^V} \sigma_{zr}^{VN} \right] - \left[\sigma_{zr}^{VV} - \frac{f^V}{f^N} \sigma_{zr}^{NV} \right]}{\frac{f^N}{f^V} \bar{r}^{VN} + \frac{f^V}{f^N} \bar{r}^{NV}}. \quad (28)$$

where $\sigma_{zr}^{XY} = \text{cov}_Y(z_i, r_i^{XY})$. We can then plug this result into equation (27), along with the condition $df^N/dt = 0$ (which comes from the quasi-equilibrium approximation (17)). We then obtain

$$\frac{d\bar{z}}{dt} \approx \omega \left[f^N \left(v^N \frac{\sigma_{zr}^{NN}}{\omega} + v^V \frac{\sigma_{zr}^{VN}}{\omega} \right) + f^V \left(v^N \frac{\sigma_{zr}^{NV}}{\omega} + v^V \frac{\sigma_{zr}^{VV}}{\omega} \right) \right] \quad (29)$$

which, when the right-hand side is set to zero, gives the correct equilibrium condition where newly produced individuals in each habitat are weighted by the reproductive value of the habitat:

$$v^N = \frac{\bar{r}^{VN} f^N}{(f^N)^2 \bar{r}^{VN} + (f^V)^2 \bar{r}^{NV}} \quad (30)$$

and a similar expression for v^V , with superscripts N and V swapped. Under weak selection, the ratios σ_{zr}^{XY}/ω can typically be approximated by derivatives, as in equation (6).

To understand the importance of taking into account the effects of selection on the dynamics of between-habitat differentiation, it is useful to note that simply setting $\bar{z}^N - \bar{z}^V$ to zero in equation (27) is equivalent to setting $v^N = v^V = 1$ in equation (29). Unless the two habitats have the same quality, the dynamics of $\bar{z}^N - \bar{z}^V$ cannot be neglected.

5.2 Evolution of virulence in class-structured host populations

Although the dynamics of the environmental variables have not been explored in this example, equations (26) and (29) can be applied to more complex ecological scenarios. For instance, with appropriate definitions for the rates r_i^{XY} , they can be used to study the evolution of virulence in partially vaccinated populations (Gandon and Day, 2007). From the parasite's point of view, naive (N) and vaccinated (V) hosts represent two different habitats, with clearly different qualities. However, because transmission events depend on the dynamics of susceptible hosts, we need to couple the Price equation (26) with a dynamical equation for the densities of naive and vaccinated susceptible hosts. These densities S^N and S^V are elements of the vector of external environmental variables \mathbf{e} . Following Gandon and Day (2007), we have the following equation for the densities of susceptible hosts

$$\begin{aligned} \frac{dS^N}{dt} &= (1 - \pi)b - \mu S^N - s^N h S^N \\ \frac{dS^V}{dt} &= \pi b - \mu S^V - s^V h S^V \end{aligned}$$

where π is the probability of vaccination at birth, b is the birth rate of susceptible hosts, μ is the background mortality rate, s^k is the susceptibility to the disease of host class k , and $h = \beta^N I^N + \beta^V I^V$ is the force of infection on susceptible hosts due to transmission from hosts in class k at rate β^k . As in Gandon and Day (2007), the following transition rates between host classes are defined for each parasite type i :

$$\begin{aligned} r_i^{NN} &= s^N \beta_i^N S^N - (\mu + \alpha_i^N) \\ r_i^{VV} &= s^V \beta_i^V S^V - (\mu + \alpha_i^V) \\ r_i^{NV} &= s^N \beta_i^V S^N \\ r_i^{VN} &= s^V \beta_i^N S^V \end{aligned}$$

where α_i^k is the virulence conferred by parasite strain i in a class- k host. Assuming that α_i^k and β_i^k are functions of a trait z (say, the within-host replication rate of the pathogen), we can calculate the quasi-equilibrium value of $(\bar{z}^N - \bar{z}^V)/\sigma$ using equation (28) and the dynamics of the trait mean using equation (29). Approximating the terms σ_{zr}^{kj}/ω using derivatives yields the following approximation, after some rearrangements:

$$\frac{d\bar{z}}{dt} = \omega (v^N s^N S^N + v^V s^V S^V) \left[\frac{I^N}{I} \left(\frac{d\beta^N}{dz}(\bar{z}) - \frac{\beta^N}{\mu + \alpha^N} \frac{d\alpha^N}{dz}(\bar{z}) \right) + \frac{I^V}{I} \left(\frac{d\beta^V}{dz}(\bar{z}) - \frac{\beta^V}{\mu + \alpha^V} \frac{d\alpha^V}{dz}(\bar{z}) \right) \right]. \quad (31)$$

where $f^N = I^N/I$ and $f^V = I^V/I$ are the frequency of naive and vaccinated infected hosts respectively. The direction of the change in the trait mean only depends on the term between square brackets, which is identical to expressions for the selection gradient obtained in other models of class-structured host-parasite interactions (Gandon et al., 2001, 2003, Lion, 2013). Note that v^k can be expressed as $(\beta^k/(\mu+\alpha^k))/[(\beta^N/(\mu+\alpha^N)f^N+(\beta^V/(\mu+\alpha^V)f^V)]$, and therefore $v^N s^N S^N + v^V s^V S^V = R_0/[(\beta^N/(\mu+\alpha^N)f^N + (\beta^V/(\mu+\alpha^V)f^V)]$, where R_0 is the basic reproduction ratio of an average infected host.

To check the validity of the quasi-equilibrium approximation, I conducted a numerical integration of the full multi-strain models, and compared the results with the prediction of equation (31). Figure 1a shows that the host densities $S^N(t)$, $S^V(t)$, $I^N(t)$ and $I^V(t)$ quickly reach the slow manifold (dashed lines). Figure 1c shows that this is also true for the scaled measure of phenotypic differentiation $(\bar{z}^N - \bar{z}^V)/\sigma$. Integrating numerically equation (31) yields a very good approximation of the trajectory of the trait mean (figure 1b). The discrepancy between the predicted and observed trajectories actually comes from errors in estimating the variance ω . For this figure, I simply used a rough non-linear fit of the dynamics of the variance when integrating (31), as shown in figure 1d (grey dashed line). A higher-order approximation could be obtained by deriving a dynamical equation for the class and global variances but this is beyond the scope of this paper. Figure 1d shows in particular that the class variances ω^N (dashed black line) and ω^V (dotted black line) initially do not coincide with the global variance ω (plain black line), but that this initial discrepancy disappears once the population starts following the slow manifold.

6 Discussion

The main result of this article is to provide a new route to the derivation of selection gradients in class-structured populations. Previous works have derived selection gradients in two main ways. First, the selection gradient can be derived from the invasion fitness of a rare mutant arising in a resident population at equilibrium: this is the invasion analysis route (Metz et al. (1992); see e.g. Otto and Day (2007), chapter 12). Second, the selection gradient has been computed under weak selection from the dynamics of the weighted frequency of a mutant allele, where the weights are the neutral class reproductive values (Taylor, 1990, Rousset, 2004, Lehmann and Rousset, 2014, Lehmann et al., 2016). In contrast to invasion analyses, the latter approach does not necessarily assume that the mutant is rare. The justification for using the reproductive values as weights is that this weighting guarantees that the observed change in trait mean is solely due to selection and not to any effects due to between-class demographic transitions (Taylor, 1990, Lehmann and Rousset, 2014, Engen et al., 2014, Grafen, 2015, Lion, 2018a). These demographic effects have been called “passive changes in trait mean” by Grafen (2015), and are often transient in populations with exponential growth. They contribute to the build-up and decay of phenotypic differentiation, captured by the second term of equation (7).

In this article, I take a different approach. Instead of weighting the frequencies of the focal allele in each class by reproductive values, I use an argument of separation of time scales to compute a quasi-equilibrium approximation of the second term of equation (7). The key point of the argument is that, in the limit where the population variance in the trait goes to zero, the measures of phenotypic differentiation $\bar{z}^k - \bar{z}$, scaled by the standard deviation \bar{z} in the trait, are a fast variable compared to the trait mean \bar{z} . More precisely, the scaled measures of phenotypic differentiation can be decoupled into a transient component and a fast component. The fast component ($d^{(1)}$) measures the part of the phenotypic differentiation that is due to selection, while the transient component ($d^{(0)}$) measures the neutral part of the phenotypic differentiation. Under weak selection, the transient component quickly decays, so that the only cause of variation among classes is due to selection. Hence, the quasi-equilibrium expressions of the d^k 's capture the effect of selection on the dynamics of phenotypic differentiation, after elimination of the passive changes.

Reproductive-value weighting is an ubiquitous feature of models of selection in class-structured populations. However, this weighting has different meanings depending on the theoretical approach used to model evolutionary dynamics. If one is interested in computing the long-term growth rate of a rare mutant, one can weight each class by an arbitrary non-zero vector because the long-term growth rate of a weighted allele frequency does not depend on the choice of weights (Caswell, 2001,

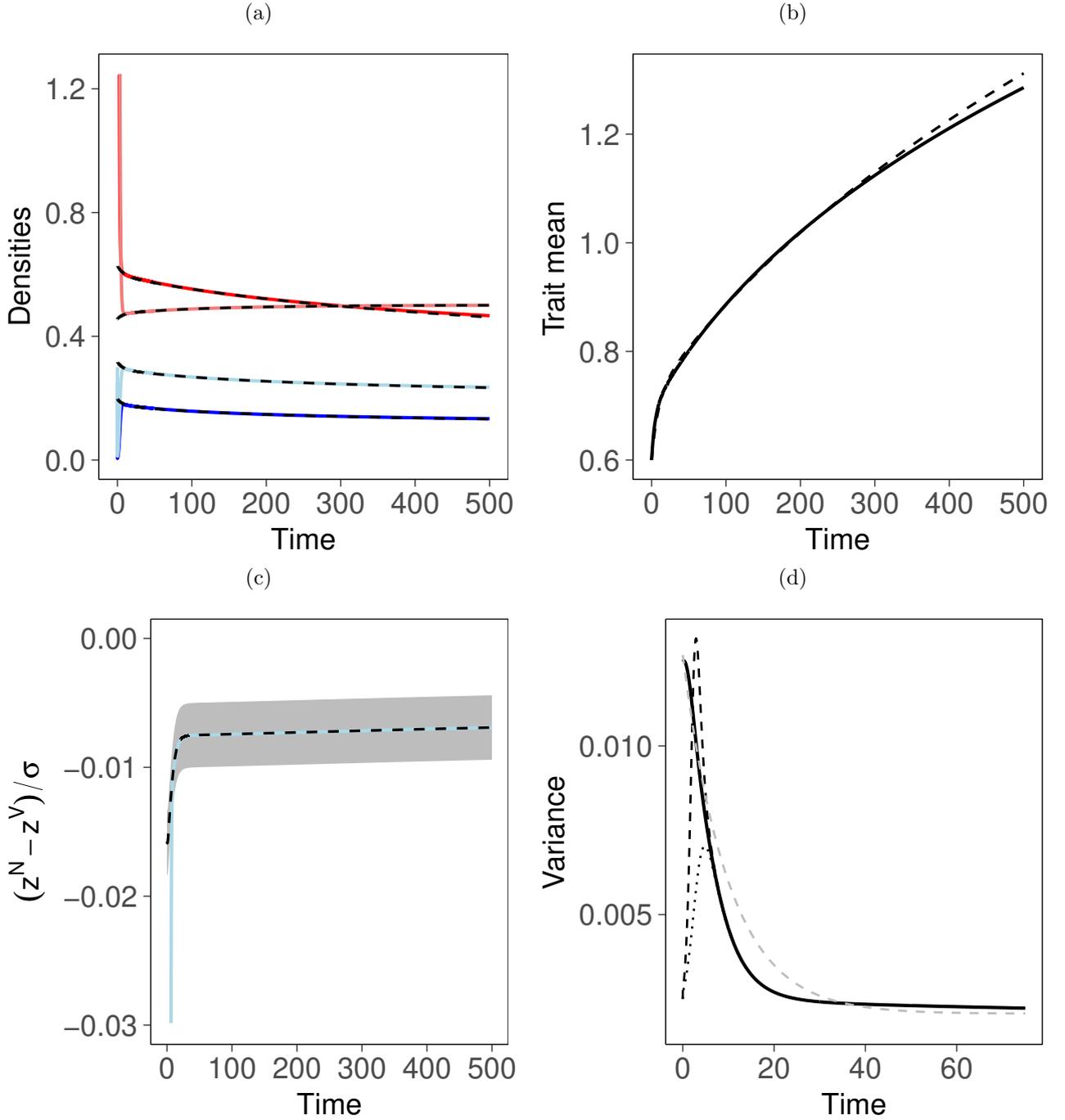


Figure 1: Epidemiological and evolutionary dynamics in a deterministic host-parasite interaction. (a) Dynamics of host densities: S^N (blue), S^V (light blue), I^N (red), and I^V (light red). (b) Dynamics of the trait mean in the full model (plain line) and using the quasi-equilibrium approximation (dashed line). (c) Dynamics of the scaled measure of phenotypic differentiation $(\bar{z}^N - \bar{z}^V)/\sigma$ in the full model (blue line) and as predicted by the quasi-equilibrium approximation (dashed line). The gray area represents a ε^2 neighbourhood of the quasi-equilibrium trajectory. (d) Dynamics of the variance: ω (black line), ω^N (dashed line), ω^V (dotted line). Note the different time scale in (d) compared with the other panels. The gray dashed line gives the approximation of the total variance used to compute the quasi-equilibrium approximation of the trait mean. – The model is given in the main text and follows the assumption of [Gandon and Day \(2007\)](#). Specifically, I assume $\beta^N(z) = \beta^V(z) = 10 \ln(1 + z)$, $\alpha^N(z) = z$, $\alpha^V(z) = (1 - c)z$. The parasite population consists of $M = 50$ types. I assume that the trait z is initially normally distributed in each class, with means $\bar{z}^N(0) = 0.5$ and $\bar{z}^V(0) = 0.7$ and standard deviation $\varepsilon = 0.05$. Initial densities are $S^N = 0.25$, $S^V = 0.3$, $I^N = I^V = 25$. Other parameters are as follows: $b = 2$, $\mu = 1$, $\pi = 0.4$, $s^N = 1$, $s^V = 0.3$, $c = 0.9$. To compute the change in trait mean, the variance is approximated as $\omega(t) \approx 0.002062 + 0.010633 \exp(-0.099892t)$.

Lehmann and Rousset, 2014, Lehmann et al., 2016). However, if we further assume weak selection, we can obtain a simple expression of the selection gradient by specifically choosing the reproductive values in the resident population as weights (see Appendix D for a discussion). In contrast, if we want to get rid of the passive changes in trait mean at any time for arbitrary strength of selection and distribution of types, we need to track the dynamics of a weighted allele frequency, where the weights are time-dependent reproductive values calculated from the average transition matrix (Lehmann and Rousset, 2014, Lion, 2018a).

The quasi-equilibrium approach I followed here both reconciles and extends these two usages of reproductive-value weighting. First, it shows that, under weak selection, the passive changes in trait mean decay on a fast time scale. The rest of the dynamics unfold on the slow manifold, where only the effect of selection drives the change in trait mean. Reproductive-value weighting appears here as an emergent property of the slow manifold. Second, in contrast to most previous analyses, the dynamics of the trait mean on the slow manifold are also valid away from equilibrium. The quasi-equilibrium approximation and resulting reproductive-value weighting can thus be used to investigate transient evolutionary dynamics, as observed for instance during ecological invasions or epidemics (Day and Proulx, 2004, Gandon and Day, 2007), provided the variance in the population is sufficiently small. Third, equation (23) provides an alternative definition of reproductive values as a function of the class frequencies, f^k , and of the average transition rates, \bar{r}^{kj} . At equilibrium, we recover the well-known result that the reproductive values are the elements of a left eigenvector of \mathbf{R} associated with eigenvalue 0. However, this property is shown to also hold true at quasi-equilibrium. This latter result can be seen as a special case of a more general dynamical definition of reproductive values (Lion, 2018a).

The result that the change in trait mean under weak selection is proportional to the selection gradient is a cornerstone of both quantitative genetics and adaptive dynamics approaches to modelling evolutionary change. In the canonical equation of adaptive dynamics, the selection gradient is scaled by the variance of the mutation kernel (Dieckmann and Law, 1996, Durinx et al., 2008, Metz and de Kovel, 2013), while in quantitative genetics a measure of standing variation plays this role (Lande, 1976, 1982, Lande and Arnold, 1983, Day and Proulx, 2004, Barfield et al., 2011). In both approaches, we therefore need to know how these measures of variance change over time. This has motivated some moment-based approaches to derive dynamical equations for higher-order moments of the trait distribution (Slatkin, 1980, Day and Proulx, 2004, Wakano and Iwasa, 2013, Wakano and Lehmann, 2014, Débarre and Otto, 2016). Since the prediction of equation (22) away from equilibrium will typically be poor if we lack an accurate description of the dynamics of the variance, this would be an interesting extension of this work.

There are interesting connections between this approach and other arguments routinely used in multi-locus genetics, when deriving quasi-linkage equilibrium approximations (Barton and Turelli, 1991, Kirkpatrick et al., 2002), and in spatially structured populations, where the dynamics of local measures of spatial differentiation typically reach a quasi-equilibrium on a fast time scale (Motro, 1982, Matsuda et al., 1992, van Baalen and Rand, 1998, Rousset, 2004, 2006, Roze and Rousset, 2008, Lion and Gandon, 2016). Both approaches look at scaled measures of differentiation that can be treated as a fast variable in a suitable limit (see also Van Cleve (2015)).

Beyond weak selection, the main assumptions used in this article are clonal reproduction and large population sizes, which are fairly standard assumptions in the literature based on invasion analyses. Other technical assumptions have been made for the sake of simplicity. For instance, as in most studies in evolutionary game theory and adaptive dynamics, I have assumed that the ecological attractor of the population dynamics is a fixed point, and not a more complicated object such as a limit cycle or chaotic attractor. However, the validity of the quasi-equilibrium approximation depends on some “nice” properties of the dynamics of the fast system. Although, for most typical models in the evolutionary literature, the fast system admits a stable equilibrium, many ecological models display more complex asymptotic behaviours, which can lead to bifurcations as the initial value of the slow variable changes (see e.g. Rinaldi and Scheffer (2000) for a non-technical discussion). In our model, this could happen for instance if the fast system converges to different equilibria depending on the initial distribution of traits. This topic certainly deserves deeper attention. Singular perturbation theory offers further mathematical techniques to analyse these more complex situations (Tikhonov,

1952, Fenichel, 1979, Verhulst, 2005).

In addition, although the present framework can be applied to a large class of deterministic non-spatial models described by systems of ordinary differential equations, it does not cover all ecological scenarios. Notably, stochastic spatial models of deme-structured populations with finite deme sizes, often analysed using inclusive fitness theory, also exhibit some separation of time scales properties that need to be analysed with other tools (Rousset, 2004, 2006, Van Cleve, 2015). In contrast, deterministic spatial models are implicitly covered by the framework. These include models with groups of infinite size (e.g. Svardal et al. (2015)) or stochastic models approximated using spatial moment equations, which lead to systems of ODEs similar to those used here.

The alternative approach used here could prove useful for analysing models where the reproductive values cannot be calculated easily. A specific case where a similar method has been applied is in spatial models analysed with deterministic spatial moment equations. For instance, Lion and Gandon (2016) have studied the dynamics of the mean virulence in a host population infected by a polymorphic population of parasites on a contact network. They derived an unclosed equation for the dynamics of the mean, taking the form of equation (7). However, the class structure in this model is not an assumption, but arises from the spatial moment expansion: the dynamics of mean virulence depends on the spatial differentiation between the mean virulence and the local mean virulence calculated in pairs of infected and susceptible hosts, $\bar{\alpha}^{IS} - \bar{\alpha}$. In turn, the dynamics of $\bar{\alpha}^{IS}$ depends on the mean virulence calculated in triplets of hosts, and so on. Despite this complexity, it is possible to calculate the selection gradient approximation from the Price equation by calculating the quasi-equilibrium value of the spatial differentiation scaled by the population variance (Lion and Gandon, 2016). Hence, even when the complexity of the ecological model makes a direct calculation of neutral reproductive values potentially difficult, the quasi-equilibrium approach can still be used to derive meaningful expressions for the dynamics of phenotypic traits.

Acknowledgements

Discussions with Sylvain Gandon over several years had a major influence on this work. I am also grateful to Mircea Sofonea for helpful discussions and comments, and for pointing out some mathematical inconsistencies in a previous version of the manuscript. Two referees provided excellent suggestions that greatly helped improve the clarity of the text. This work was supported by JCJC grant ANR-16-CE35-0012-01 from the Agence Nationale pour la Recherche (ANR).

References

- Abrams, P., Harada, Y., Matsuda, H., 1993. On the relationship between quantitative genetic and ESS models. *Evolution* 47, 982–985.
- Abrams, P.A., 2001. Modelling the adaptive dynamics of traits involved in inter- and intraspecific interactions: An assessment of three methods. *Ecol. Lett.* 4, 166–175. URL: <http://dx.doi.org/10.1046/j.1461-0248.2001.00199.x>, doi:10.1046/j.1461-0248.2001.00199.x.
- van Baalen, M., Rand, D.A., 1998. The unit of selection in viscous populations and the evolution of altruism. *J. theor. Biol.* 193, 631–648. URL: <http://dx.doi.org/10.1006/jtbi.1998.0730>, doi:10.1006/jtbi.1998.0730.
- Barfield, M., Holt, R.D., Gomulkiewicz, R., 2011. Evolution in stage-structured populations. *Am. Nat.* 177, 397–409. URL: <http://dx.doi.org/10.1086/658903>, doi:10.1086/658903.
- Barton, N.H., Turelli, M., 1991. Natural and sexual selection on many loci. *Genetics* 127, 229–255.
- Caswell, H., 2001. *Matrix population models: Construction, analysis, and interpretation*. Sinauer Associates.
- Charlesworth, B., 1994. *Evolution in age-structured populations*. Cambridge University Press, Cambridge, UK. doi:10.1017/CB09780511525711.

- Day, T., Gandon, S., 2006. Insights from Price's equation into evolutionary epidemiology, in: Feng, Z., Dieckmann, U., Levin, S. (Eds.), *Disease evolution: models, concepts and data analyses*. American Mathematical Society. volume 71 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pp. 23–43.
- Day, T., Gandon, S., 2007. Applying population-genetic models in theoretical evolutionary epidemiology. *Ecol. Lett.* 10, 876–888. URL: <http://dx.doi.org/10.1111/j.1461-0248.2007.01091.x>, doi:10.1111/j.1461-0248.2007.01091.x.
- Day, T., Proulx, S.R., 2004. A general theory for the evolutionary dynamics of virulence. *Am. Nat.* 163, E40–E63. URL: <http://dx.doi.org/10.1086/382548>, doi:10.1086/382548.
- Débarre, F., Otto, S.P., 2016. Evolutionary dynamics of a quantitative trait in a finite asexual population. *Theor. Pop. Biol.* 108, 75–88. URL: <http://dx.doi.org/10.1016/j.tpb.2015.12.002>, doi:10.1016/j.tpb.2015.12.002.
- Dieckmann, U., Law, R., 1996. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of mathematical biology* 34, 579–612. URL: <http://dx.doi.org/10.1007/BF02409751>, doi:10.1007/BF02409751.
- Durinx, M., Metz, J.A.J., Meszéna, G., 2008. Adaptive dynamics for physiologically structured populations models. *J. math. Biol.* 56, 673–742. URL: <http://dx.doi.org/10.1007/s00285-007-0134-2>, doi:10.1007/s00285-007-0134-2.
- Engen, S., Kvalnes, T., Sæther, B.E., 2014. Estimating phenotypic selection in age-structured populations by removing transient fluctuations. *Evolution* 68, 2509–23. URL: <http://dx.doi.org/10.1111/evo.12456>, doi:10.1111/evo.12456.
- Fenichel, N., 1979. Geometric singular perturbation theory for ordinary differential equations. *Journal of Differential Equations* 31, 53–98.
- Ferrière, R., Gatto, M., 1995. Lyapunov exponents and the mathematics of invasion in oscillatory or chaotic populations. *Theor. Pop. Biol.* 48, 126–171. URL: <http://dx.doi.org/10.1006/tpbi.1995.1024>, doi:10.1006/tpbi.1995.1024.
- Gandon, S., Day, T., 2007. The evolutionary epidemiology of vaccination. *J. R. Soc. Interface* 4, 803–817. URL: <http://dx.doi.org/10.1098/rsif.2006.0207>, doi:10.1098/rsif.2006.0207.
- Gandon, S., Mackinnon, M.J., Nee, S., Read, A.F., 2001. Imperfect vaccines and the evolution of pathogen virulence. *Nature* 414, 751–756. URL: <http://dx.doi.org/10.1038/414751a>, doi:10.1038/414751a.
- Gandon, S., Mackinnon, M.J., Nee, S., Read, A.F., 2003. Imperfect vaccination: some epidemiological and evolutionary consequences. *Proc. R. Soc. B* 270, 1129–1136. URL: <http://dx.doi.org/10.1098/rspb.2003.2370>, doi:10.1098/rspb.2003.2370.
- Geritz, S.A.H., Kisdi, E., Meszéna, G., Metz, J.A.J., 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.* 12, 35–57. URL: <http://dx.doi.org/10.1023/A:1006554906681>, doi:10.1023/A:1006554906681.
- Grafen, A., 2015. Biological fitness and the Price equation in class-structured populations. *J. theor. Biol.* 373, 62–72. URL: <http://dx.doi.org/10.1016/j.jtbi.2015.02.014>, doi:10.1016/j.jtbi.2015.02.014.
- Kirkpatrick, M., Johnson, T., Barton, N., 2002. General models of multilocus evolution. *Genetics* 161, 1727–1750.
- Lande, R., 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30, 314–334.

- Lande, R., 1982. Elements of a quantitative genetic model of life history evolution, in: Dingle, H., Hegmann, J.P. (Eds.), *Evolution and genetics of life histories*. Springer, Berlin, pp. 21–29.
- Lande, R., Arnold, S.J., 1983. The measurement of selection on correlated characters. *Evolution* 37, 1210–1226.
- Lehmann, L., Mullan, C., Akçay, E., Van Cleve, J., 2016. Invasion fitness, inclusive fitness, and reproductive numbers in heterogeneous populations. *Evolution* 70, 1689–1702. URL: <http://dx.doi.org/10.1111/evo.12980>, doi:10.1111/evo.12980.
- Lehmann, L., Rousset, F., 2014. The genetical theory of social behaviour. *Phil. Trans. Roy. Soc. B* 369. URL: <http://dx.doi.org/10.1098/rstb.2013.0357>, doi:10.1098/rstb.2013.0357.
- Lion, S., 2013. Multiple infections, kin selection and the evolutionary epidemiology of parasite traits. *J. evol. Biol.* 26, 2107–2122. URL: <http://dx.doi.org/10.1111/jeb.12207>, doi:10.1111/jeb.12207.
- Lion, S., 2018a. Class structure, demography and selection: reproductive-value weighting in non-equilibrium, polymorphic populations. *Am. Nat.* doi:<https://doi.org/10.1101/155879>. (In press).
- Lion, S., 2018b. Theoretical approaches in evolutionary ecology: environmental feedback as a unifying perspective. *Am. Nat.* 191, 21–44. URL: <https://dx.doi.org/10.1086/694865>, doi:10.1086/694865.
- Lion, S., Gandon, S., 2016. Spatial evolutionary epidemiology of spreading epidemics. *Proc. R. Soc. B* 283, 20161170. URL: <http://dx.doi.org/10.1098/rspb.2016.1170>, doi:10.1098/rspb.2016.1170.
- Matsuda, H., Ogita, N., Sasaki, A., Sato, K., 1992. Statistical mechanics of population. *Prog. Theor. Phys.* 88, 1035–1049. URL: <http://dx.doi.org/10.1143/PTP.88.1035>, doi:10.1143/PTP.88.1035.
- Meszéna, G., Gyllenberg, M., Jacobs, F.J., Metz, J.A.J., 2005. Link between population dynamics and dynamics of Darwinian evolution. *Phys. Rev. Lett.* 95. URL: <http://dx.doi.org/10.1103/PhysRevLett.95.078105>, doi:10.1103/PhysRevLett.95.078105.
- Metz, J.A.J., 2008. Fitness, in: Jorgensen, S.E., Fath, B.D. (Eds.), *Encyclopedia of Ecology*. Academic Press, Oxford, UK, pp. 1599 – 1612. doi:10.1016/B978-008045405-4.00792-8.
- Metz, J.A.J., de Kovel, C.G., 2013. The canonical equation of adaptive dynamics for mendelian diploids and haplo-diploids. *Interface focus* 3, 20130025. URL: <http://dx.doi.org/10.1098/rsfs.2013.0025>, doi:10.1098/rsfs.2013.0025.
- Metz, J.A.J., Nisbet, R.M., Geritz, S.A.H., 1992. How should we define 'fitness' for general ecological scenarios? *Trends Ecol. Evol.* 7, 198–202. URL: [http://dx.doi.org/10.1016/0169-5347\(92\)90073-K](http://dx.doi.org/10.1016/0169-5347(92)90073-K), doi:10.1016/0169-5347(92)90073-K.
- Motro, U., 1982. Optimal rates of dispersal I. Haploid populations. *Theor. Pop. Biol.* 21, 394–411.
- Otto, S.P., Day, T., 2007. *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. Princeton University Press, Princeton, NJ.
- Rand, D.A., Wilson, H.B., McGlade, J.M., 1994. Dynamics and evolution : evolutionarily stable attractors, invasion exponents and phenotype dynamics. *Phil. Trans. Roy. Soc. B* 343, 261–283. URL: <http://dx.doi.org/10.1098/rstb.1994.0025>, doi:10.1098/rstb.1994.0025.
- Rinaldi, S., Scheffer, M., 2000. Geometric analysis of ecological models with slow and fast processes. *Ecosystems* 3, 508–521. URL: <https://dx.doi.org/10.1007/s100210000045>, doi:10.1007/s100210000045.

- Rousset, F., 1999. Reproductive value vs sources and sinks. *Oikos* 86, 591–596.
- Rousset, F., 2004. Genetic structure and selection in subdivided populations. Princeton University Press, Princeton, NJ, USA.
- Rousset, F., 2006. Separation of time scales, fixation probabilities and convergence to evolutionarily stable states under isolation by distance. *Theor. Pop. Biol.* 69, 165–179. URL: <http://dx.doi.org/10.1016/j.tpb.2005.08.008>, doi:10.1016/j.tpb.2005.08.008.
- Roze, D., Rousset, F., 2008. Multilocus models in the infinite island model of population structure. *Theor. Pop. Biol.* 73, 529–542. URL: <http://dx.doi.org/10.1016/j.tpb.2008.03.002>, doi:10.1016/j.tpb.2008.03.002.
- Slatkin, M., 1980. Ecological character displacement. *Genetics* 61, 163–177.
- Svardal, H., Rueffler, C., Hermisson, J., 2015. A general condition for adaptive genetic polymorphism in temporally and spatially heterogeneous environments. *Theor. Pop. Biol.* 99, 76–97. URL: <http://dx.doi.org/10.1016/j.tpb.2014.11.002>, doi:10.1016/j.tpb.2014.11.002.
- Taylor, P.D., 1990. Allele-frequency change in a class-structured population. *Am. Nat.* 135, 95–106. URL: <http://dx.doi.org/10.1086/285034>, doi:10.1086/285034.
- Taylor, P.D., Day, T., 1997. Evolutionary stability under the replicator and the gradient dynamics. *Evol. Ecol.* 11, 579–590. URL: <http://dx.doi.org/10.1007/s10682-997-1513-2>, doi:10.1007/s10682-997-1513-2.
- Tikhonov, A.N., 1952. Systems of differential equations containing a small parameter multiplying the derivative. *Mat. Sb.* 31, 575–586. (In Russian).
- Tuljapurkar, S., 1989. An uncertain life: Demography in random environments. *Theor. Pop. Biol.* 35, 227–294.
- Van Cleve, J., 2015. Social evolution and genetic interactions in the short and long term. *Theor. Pop. Biol.* , 2–26 URL: <http://dx.doi.org/10.1016/j.tpb.2015.05.002>, doi:10.1016/j.tpb.2015.05.002.
- Verhulst, F., 2005. Methods and applications of singular perturbations, boundary layers and multiple timescale dynamics. Springer-Verlag, Berlin.
- Wakano, J.Y., Iwasa, Y., 2013. Evolutionary branching in a finite population: deterministic branching vs. stochastic branching. *Genetics* 193, 229–41. URL: <http://dx.doi.org/10.1534/genetics.112.144980>, doi:10.1534/genetics.112.144980.
- Wakano, J.Y., Lehmann, L., 2014. Evolutionary branching in deme-structured populations. *J. theor. Biol.* 351, 83–95. URL: <http://dx.doi.org/10.1016/j.jtbi.2014.02.036>, doi:10.1016/j.jtbi.2014.02.036.

Appendix A: Derivation of the perturbed system

Equation (1) shows that the rate of change of \mathbf{n} depends on the matrix of average rates \bar{r}^{kj} . These rates are necessarily $O(1)$ because population densities must still be able to change when there is no selection. In practice, a Taylor expansion of the vital rates around the trait mean can be used to calculate the effect of selection on the dynamics of densities. A similar argument holds for the dynamics of the external variables, \mathbf{e} .

We now consider the dynamics of \bar{z} . The first term of equation (15) depends on the class frequencies \mathbf{f} , which are $O(1)$, and on the covariance matrix which is $O(\varepsilon^2)$ for a unimodal distribution. The second term of equation (15) depends on the standard deviation σ , which is $O(\varepsilon)$, on the matrix product $\mathbf{R}\mathbf{F}$,

which is $O(1)$, and on \mathbf{d} . By construction, \mathbf{d} is $O(1)$, thus the second term (and therefore $d\bar{z}/dt$) is at least $O(\varepsilon)$. Hence, we do have a separation of time scales between \bar{z} and the ecological variables \mathbf{n} and \mathbf{f} . However, we will see in Appendix B that the approximation of \mathbf{d} on the slow manifold is $O(\varepsilon)$, so that the dynamics of \mathbf{z} on the slow manifold is $O(\varepsilon^2)$. Again, an approximation to the right-hand side of equation (15) can be obtained through a Taylor expansion.

We now turn to the dynamics of \mathbf{d} . From equations (7) and (9), we obtain the following equation

$$\frac{d}{dt}(\bar{z}^k - \bar{z}) = \sum_j f^j \text{cov}_j \left(z_i, \frac{r_i^{kj}}{f^k} - r_i^{\bullet j} \right) + \sum_j (\bar{z}^j - \bar{z}) \left(\bar{r}^{kj} \frac{f^j}{f^k} - \bar{r}^{\bullet j} f^j \right) - (\bar{z}^k - \bar{z}) \sum_j \bar{r}^{kj} \frac{f^j}{f^k}, \quad (\text{A.1})$$

which can be written more compactly as

$$\frac{d(\sigma \mathbf{d})}{dt} = \mathbf{B} \mathbf{C} \mathbf{f} + \mathbf{A}(\sigma \mathbf{d}) \quad (\text{A.2})$$

where the matrices \mathbf{A} and \mathbf{B} are defined in the main text. This finally yields equation (16). The leading-order term of the right-hand side of equation (16) is $\mathbf{A} \mathbf{d}$, which is $O(1)$ because the matrix \mathbf{A} depends on the average transition rates \bar{r}^{kj} , which are $O(1)$. Hence, \mathbf{d} changes on a fast time scale compared to \bar{z} .

We can Taylor-expand the right-hand side of equation (16) to obtain an approximation up to first order in ε . This yields

$$\frac{d\mathbf{d}}{dt} = \mathbf{A}^{(0)} \mathbf{d}^{(0)} + \varepsilon \left[\frac{1}{\sigma^{(1)}} \mathbf{B}^{(0)} \mathbf{C}^{(2)} \mathbf{f}^{(0)} + \mathbf{A}^{(0)} \mathbf{d}^{(1)} + \mathbf{A}^{(1)} \mathbf{d}^{(0)} - \mathbf{d}^{(0)} \left(\frac{d \ln \sigma}{dt} \right)^{(1)} \right] + O(\varepsilon^2) \quad (\text{A.3})$$

where the notation $x^{(m)}$ represents the m th term of the Taylor series of the x variable, $x = x^{(0)} + \varepsilon x^{(1)} + \varepsilon^2 x^{(2)} + \dots$. A similar analysis can be carried out for the change in the distribution of types, \mathbf{p} , and for the change in the between-class differences in frequencies, $f_i^k - f_i$.

Appendix B: Quasi-equilibrium value of between-class differentiation

To first order in ε , the quasi-equilibrium value of \mathbf{d} can be calculated by setting the right-hand side of equation (20) to zero, along with the condition (17). Doing so would require to invert matrix $\mathbf{A}^{(0)}$. However, when condition (17) holds, the matrix \mathbf{A} is singular. To see this, note that \mathbf{A} has elements $a_{kj} = \bar{r}^{kj} f^j / f^k - f^j \sum_\ell \bar{r}^{\ell j}$ and $a_{kk} = -\sum_{j \neq k} \bar{r}^{kj} f^j / f^k - f^k \sum_\ell \bar{r}^{\ell k}$. As a result, the sum of the elements of each row is $-\sum_j \sum_\ell \bar{r}^{\ell j} f^j = -(1/n) dn/dt = \bar{r}$, and we have

$$\mathbf{A} \mathbf{1} = -\bar{r} \mathbf{1},$$

so the total growth rate of the population, $-\bar{r}$, is an eigenvalue of \mathbf{A} with associated eigenvector $\mathbf{1}$. This means in particular that $\mathbf{A}^{(0)}$ is not invertible at quasi-equilibrium.

To solve for the quasi-equilibrium, we therefore need to eliminate at least one equation. This can be done using the fact that $\sum_k f^k (\bar{z}^k - \bar{z}) = 0$ by definition. Assuming for convenience that the matrix \mathbf{A} has rank $K - 1$, we can restrict the analysis to a vector $\boldsymbol{\delta}$ of length $K - 1$. For instance, we eliminate the equation for z^K and write

$$\mathbf{d} = \begin{pmatrix} \boldsymbol{\delta} \\ d_K \end{pmatrix}. \quad (\text{B.1})$$

We have the relationship

$$0 = \mathbf{f}^\top \mathbf{d} = (\mathbf{P} \mathbf{f})^\top \mathbf{d} + f^K d_K \quad (\text{B.2})$$

where \mathbf{P} is a $(K - 1) \times K$ matrix that removes the K element of a vector, i.e.

$$\mathbf{P} = (\mathbf{I}_{K-1} \quad \mathbf{0}) \quad (\text{B.3})$$

where \mathbf{I}_{K-1} is the identity matrix with size $K - 1$. We thus have

$$d_K = -\frac{(\mathbf{P} \mathbf{f})^\top}{f^K} \boldsymbol{\delta} \quad (\text{B.4})$$

which leads to

$$\frac{d(\bar{z}^k - \bar{z})}{dt} + = \sum_{j \neq K} (\bar{z}^j - \bar{z}) [\bar{r}^{kj} - \bar{r}^{kK}] \frac{f^j}{f^k} - (\bar{z}^k - \bar{z}) \sum_j \bar{r}^{kj} \frac{f^j}{f^k} - \sum_{j \neq K} (\bar{z}^j - \bar{z}) [\bar{r}^{\bullet j} - \bar{r}^{\bullet K}] f^j$$

so the matrix $\mathbf{\Gamma}$ has elements

$$\gamma_{kj} = (\bar{r}^{kj} - \bar{r}^{kK}) \frac{f^j}{f^k} - (\bar{r}^{\bullet j} - \bar{r}^{\bullet K}) f^j \quad \text{for } j \neq k \quad (\text{B.12})$$

$$\gamma_{kk} = \bar{r}^{kk} - \bar{r}^{kK} - (\bar{r}^{\bullet k} - \bar{r}^{\bullet K}) f^k - \sum_j \bar{r}^{kj} \frac{f^j}{f^k} \quad (\text{B.13})$$

In the quasi-equilibrium approximation, equation (17) holds true, so that $\sum_j \bar{r}^{kj} f^j = 0$. This gives:

$$\forall j, k, \quad \gamma_{kj} = (\bar{r}^{kj} - \bar{r}^{kK}) \frac{f^j}{f^k} - (\bar{r}^{\bullet j} - \bar{r}^{\bullet K}) f^j. \quad (\text{B.14})$$

Appendix C: Reproductive values

Here, I show that the vector $\mathbf{v}^\top = \mathbf{1}^\top [\mathbf{I} - \mathbf{RFLB}]$ defined in Appendix B is co-normalised with the vector of class frequencies and is a left eigenvector of the matrix \mathbf{R} associated with eigenvalue 0 on the slow manifold. The first property can be written as $\mathbf{v}^\top \mathbf{f} = 1$. This is easily checked by noting that

$$\mathbf{v}^\top \mathbf{f} = \mathbf{1}^\top \mathbf{f} - \mathbf{1}^\top \mathbf{RFLBf} = 1$$

because $\mathbf{1}^\top \mathbf{f} = 1$ and $\mathbf{Bf} = \mathbf{F}^{-1} \mathbf{f} - \mathbf{1}(\mathbf{1}^\top \mathbf{f}) = \mathbf{0}$

To show that $\mathbf{v}^\top \mathbf{R} = \mathbf{0}$ on the slow manifold, let us first note that this is equivalent to showing that

$$(\mathbf{1}^\top \mathbf{R}) \mathbf{Y} = \mathbf{1}^\top \mathbf{R} \quad (\text{C.1})$$

where $\mathbf{Y} = \mathbf{FGP}^\top \mathbf{\Gamma}^{-1} \mathbf{PBR}$. To calculate the inverse of $\mathbf{\Gamma}$, we remark that we can rewrite $\mathbf{\Gamma}$ as follows:

$$\mathbf{\Gamma} = (\mathbf{H} - \mathbf{x} \mathbf{1}^\top) \mathbf{F}_{K-1} \quad (\text{C.2})$$

where \mathbf{H} is the $(K-1) \times (K-1)$ matrix with elements $h_{kj} = \bar{r}^{kj}/f^k - \bar{r}^{\bullet j}$, \mathbf{x} is a vector of length $K-1$ with elements $x_k = \bar{r}^{kK}/f^k - \bar{r}^{\bullet K}$ and \mathbf{F}_{K-1} is the restriction of \mathbf{F} to the $K-1$ first rows and columns. Note that, in equation C.2, the vector $\mathbf{1}$ has length $K-1$. Throughout this appendix, I use the convention that $\mathbf{1}^\top$ is a row vector of ones with length conformable for right multiplication by a matrix or vector.

Assuming that \mathbf{H} is invertible, the matrix $\mathbf{H} - \mathbf{x} \mathbf{1}^\top$ is invertible iff $1 - \mathbf{1}^\top \mathbf{H}^{-1} \mathbf{x} \neq 0$. The latter inequality will be proved later, and will be assumed for the time being. If it holds true, the inverse of $\mathbf{\Gamma}$ is given by

$$\mathbf{\Gamma}^{-1} = \mathbf{F}_{K-1}^{-1} (\mathbf{H} - \mathbf{x} \mathbf{1}^\top)^{-1}. \quad (\text{C.3})$$

The second inverse matrix can be calculated using the Sherman-Morrison formula, which gives:

$$\mathbf{\Gamma}^{-1} = \mathbf{F}_{K-1}^{-1} \left(\mathbf{H}^{-1} + \frac{\mathbf{H}^{-1} \mathbf{x} \mathbf{1}^\top \mathbf{H}^{-1}}{1 - \mathbf{1}^\top \mathbf{H}^{-1} \mathbf{x}} \right) \quad (\text{C.4})$$

Now using the fact that \mathbf{PBR} can be written as the block matrix

$$\mathbf{PBR} = (\mathbf{H} \quad \mathbf{x}), \quad (\text{C.5})$$

we have

$$\mathbf{\Gamma}^{-1} \mathbf{PBR} = \mathbf{F}_{K-1}^{-1} \left(\mathbf{I}_{K-1} + \frac{\mathbf{H}^{-1} \mathbf{x} \mathbf{1}^\top}{1 - \mathbf{1}^\top \mathbf{H}^{-1} \mathbf{x}} \right) \mathbf{H}^{-1} (\mathbf{H} \quad \mathbf{x}) \quad (\text{C.6})$$

$$= \mathbf{F}_{K-1}^{-1} \left(\mathbf{I}_{K-1} + \frac{\mathbf{H}^{-1} \mathbf{x} \mathbf{1}^\top}{1 - \mathbf{1}^\top \mathbf{H}^{-1} \mathbf{x}} \right) (\mathbf{I}_{K-1} \quad \mathbf{H}^{-1} \mathbf{x}) \quad (\text{C.7})$$

$$= \mathbf{F}_{K-1}^{-1} \left(\mathbf{I}_{K-1} + \frac{\mathbf{H}^{-1} \mathbf{x} \mathbf{1}^\top}{1 - \mathbf{1}^\top \mathbf{H}^{-1} \mathbf{x}} \quad \mathbf{H}^{-1} \mathbf{x} + \frac{\mathbf{H}^{-1} \mathbf{x} \mathbf{1}^\top \mathbf{H}^{-1} \mathbf{x}}{1 - \mathbf{1}^\top \mathbf{H}^{-1} \mathbf{x}} \right) \quad (\text{C.8})$$

$$= \mathbf{F}_{K-1}^{-1} \left(\mathbf{I}_{K-1} + \frac{\mathbf{H}^{-1} \mathbf{x} \mathbf{1}^\top}{1 - \mathbf{1}^\top \mathbf{H}^{-1} \mathbf{x}} \quad \frac{\mathbf{H}^{-1} \mathbf{x}}{1 - \mathbf{1}^\top \mathbf{H}^{-1} \mathbf{x}} \right). \quad (\text{C.9})$$

Pre-multiplying by \mathbf{FGP}^\top leads to

$$\mathbf{Y} = \begin{pmatrix} \mathbf{I}_{K-1} \\ -\mathbf{1}^\top \end{pmatrix} \left(\mathbf{I}_{K-1} + \frac{\mathbf{H}^{-1}\mathbf{x}\mathbf{1}^\top}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} \quad \frac{\mathbf{H}^{-1}\mathbf{x}}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} \right) \quad (\text{C.10})$$

$$= \begin{pmatrix} \mathbf{I}_{K-1} + \frac{\mathbf{H}^{-1}\mathbf{x}\mathbf{1}^\top}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} & \frac{\mathbf{H}^{-1}\mathbf{x}}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} \\ -\mathbf{1}^\top - \frac{\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}\mathbf{1}^\top}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} & \frac{-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} \end{pmatrix} \quad (\text{C.11})$$

$$= \begin{pmatrix} \mathbf{I}_{K-1} + \frac{\mathbf{H}^{-1}\mathbf{x}\mathbf{1}^\top}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} & \frac{\mathbf{H}^{-1}\mathbf{x}}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} \\ \frac{-\mathbf{1}^\top}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} & \frac{-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} \end{pmatrix} \quad (\text{C.12})$$

and we finally have

$$\begin{aligned} \mathbf{I} - \mathbf{Y} &= \begin{pmatrix} -\frac{\mathbf{H}^{-1}\mathbf{x}\mathbf{1}^\top}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} & -\frac{\mathbf{H}^{-1}\mathbf{x}}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} \\ \frac{\mathbf{1}^\top}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} & 1 + \frac{\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} \end{pmatrix} \\ &= \begin{pmatrix} -\frac{\mathbf{H}^{-1}\mathbf{x}\mathbf{1}^\top}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} & -\frac{\mathbf{H}^{-1}\mathbf{x}}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} \\ \frac{\mathbf{1}^\top}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} & \frac{1}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} \end{pmatrix} \\ &= \frac{1}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} \begin{pmatrix} -\mathbf{H}^{-1}\mathbf{x}\mathbf{1}^\top & -\mathbf{H}^{-1}\mathbf{x} \\ \mathbf{1}^\top & 1 \end{pmatrix} \end{aligned}$$

Noting the row vector $\mathbf{1}^\top\mathbf{R}$ as $(\mathbf{u}_1 \quad u_2)$, where u_2 is the scalar $\bar{r}^{\bullet K}$ and \mathbf{u}_1 the vector with length $K-1$ and elements $\bar{r}^{\bullet k}$, we find that

$$\mathbf{1}^\top\mathbf{R}(\mathbf{I} - \mathbf{Y}) = \frac{-\mathbf{u}_1^\top\mathbf{H}^{-1}\mathbf{x} + u_2}{1-\mathbf{1}^\top\mathbf{H}^{-1}\mathbf{x}} \mathbf{1}^\top \quad (\text{C.13})$$

To prove that \mathbf{v}^\top is a left eigenvector of \mathbf{R} associated with eigenvalue zero, we thus need to show that $\mathbf{u}_1^\top\mathbf{H}^{-1}\mathbf{x} = u_2$. Writing η_{kj} the elements of the matrix \mathbf{H}^{-1} , we have

$$\mathbf{u}_1^\top\mathbf{H}^{-1}\mathbf{x} = \sum_{k=1}^{K-1} \bar{r}^{\bullet k} \sum_{\ell=1}^{K-1} \eta_{k\ell} \left(\frac{\bar{r}^{\ell K}}{f^\ell} - \bar{r}^{\bullet K} \right). \quad (\text{C.14})$$

On the slow manifold, we have $\sum_{j=1}^K \bar{r}^{\ell j} f^j = 0$ (equation (17)), and this gives the two relationships

$$\begin{aligned} \bar{r}^{\ell K} &= -\frac{1}{f^K} \sum_{j=1}^{K-1} \bar{r}^{\ell j} f^j \\ \bar{r}^{\bullet K} &= -\frac{1}{f^K} \sum_{j=1}^{K-1} \bar{r}^{\bullet j} f^j. \end{aligned}$$

Equation (C.14) can then be rewritten as

$$\begin{aligned} \mathbf{u}_1^\top\mathbf{H}^{-1}\mathbf{x} &= \sum_{k=1}^{K-1} \bar{r}^{\bullet k} \sum_{\ell=1}^{K-1} \eta_{k\ell} \frac{1}{f^K} \sum_{j=1}^{K-1} \left(-\frac{\bar{r}^{\ell j} f^j}{f^\ell} + \bar{r}^{\bullet j} f^j \right) \\ &= -\sum_{j=1}^{K-1} \frac{f^j}{f^K} \sum_{k=1}^{K-1} \bar{r}^{\bullet k} \sum_{\ell=1}^{K-1} \eta_{k\ell} \left(\frac{\bar{r}^{\ell j}}{f^\ell} - \bar{r}^{\bullet j} \right) \end{aligned}$$

The last sum gives the elements of $\mathbf{H}^{-1}\mathbf{H} = \mathbf{I}_{K-1}$, and therefore

$$\sum_{\ell=1}^{K-1} \eta_{k\ell} \left(\frac{\bar{r}^{\ell j}}{f^\ell} - \bar{r}^{\bullet j} \right) = \delta_{kj} \quad (\text{C.15})$$

where δ_{kj} is Kronecker's symbol. This yields

$$\mathbf{u}_1^\top \mathbf{H}^{-1} \mathbf{x} = - \sum_{j=1}^{K-1} \frac{f^j}{f^K} \sum_{k=1}^{K-1} \bar{r}^{\bullet k} \delta_{kj} = - \sum_{j=1}^{K-1} \frac{f^j}{f^K} \bar{r}^{\bullet j} = \bar{r}^{\bullet K} = u_2$$

which proves the result. The above suite of equalities can also be used to check our assumption that $\mathbf{1}^\top \mathbf{H}^{-1} \mathbf{x} \neq 1$. Replacing the vector \mathbf{u}_1 by the vector $\mathbf{1}$, we obtain

$$1 - \mathbf{1}^\top \mathbf{H}^{-1} \mathbf{x} = 1 + \sum_{j=1}^{K-1} \frac{f^j}{f^K} \sum_{k=1}^{K-1} \delta_{kj} = 1 + \sum_{j=1}^{K-1} \frac{f^j}{f^K} = 1 + \frac{1 - f^K}{f^K} = \frac{1}{f^K} > 0. \quad (\text{C.16})$$

Appendix D: Reproductive-value weighting in invasion analyses

If the resident population is at equilibrium, the invasion fitness ρ of a rare mutant is the dominant eigenvalue of the matrix $\mathbf{M} = \hat{\mathbf{R}}_{\mathbf{m}}$. Then, noting \mathbf{u} the associated right eigenvector, we have $\mathbf{M}\mathbf{u} = \rho\mathbf{u}$. Pre-multiplying by an arbitrary non-zero row vector \mathbf{v}^\top co-normalised with \mathbf{u} (i.e. $\mathbf{v}^\top \mathbf{u} = 1$) leads to

$$\rho = \mathbf{v}^\top \mathbf{M}\mathbf{u},$$

which shows that we don't need to choose a specific vector of class weights to calculate the long-term growth rate of the mutant population. If selection is weak, a standard perturbation analysis of equation $\mathbf{M}\mathbf{u} = \rho\mathbf{u}$ shows that

$$\frac{d\mathbf{M}}{d\varepsilon} \bar{\mathbf{u}} + \bar{\mathbf{M}} \frac{d\mathbf{u}}{d\varepsilon} = \rho \bar{\mathbf{u}} + O(\varepsilon^2).$$

where $\bar{\mathbf{M}}$ and $\bar{\mathbf{u}}$ are evaluated at $\varepsilon = 0$. Pre-multiplying by an arbitrary non-zero vector \mathbf{v}^\top leads to

$$\mathbf{v}^\top \frac{d\mathbf{M}}{d\varepsilon} \bar{\mathbf{u}} + \mathbf{v}^\top \bar{\mathbf{M}} \frac{d\mathbf{u}}{d\varepsilon} = \rho \mathbf{v}^\top \bar{\mathbf{u}} + O(\varepsilon^2).$$

Now, if we specifically choose \mathbf{v}^\top as the left eigenvector $\bar{\mathbf{v}}^\top$ of $\bar{\mathbf{M}}$ associated with 0, and such that $\bar{\mathbf{v}}^\top \bar{\mathbf{u}} = 1$, we have $\bar{\mathbf{v}}^\top \bar{\mathbf{M}} = \mathbf{0}$ and therefore

$$\rho = \bar{\mathbf{v}}^\top \frac{d\mathbf{M}}{d\varepsilon} \bar{\mathbf{u}} + O(\varepsilon^2).$$

Thus, although under weak selection we could still use any arbitrary non-zero vector of weights to calculate ρ , only the vector of reproductive values in the resident monomorphic population at equilibrium leads to a simple expression where only the effect of the mutation on the matrix \mathbf{M} , and not on \mathbf{u} and \mathbf{v} , needs to be calculated.