Classifying the Role of Trade-offs in the Evolutionary Diversity of Pathogens

By Ivana Gudelj[1] †, Ciprian Coman[2] and Robert Beardmore[3]

[1]NERC Centre for Population Biology, Imperial College London, UK. [2]Dept. of Mathematics, University of Glasgow, Glasgow, UK. [3]Dept. of Mathematics, Imperial College London, UK.

In this paper we use a system of non-local reaction diffusion equations to study the effect of host heterogeneity on the phenotypic evolution of a pathogen population. The evolving phenotype is taken to be the transmission rate of the pathogen on the different hosts, and in our system there are two host populations present.

The central feature of our model is a trade-off relationship between the transmission rates on these hosts, which means that an increase in the pathogen transmission on one host will lead to a decrease in the pathogen transmission on the other. The purpose of the paper is to develop a classification of phenotypic diversity as a function of the shape of the trade-off relationship and this is achieved by determining the maximum number of phenotypes a pathogen population can support in the long-term, for a given form of the trade-off. Our findings are then compared with results obtained by applying classical theory from evolutionary ecology and the more recent *adaptive dynamics method* to the same host-pathogen system. We find our work to be in good agreement with these two approaches.

Keywords: host-pathogen systems, evolutionary diversity, trade-offs

1. Introduction

In Gudelj *et al.* (2004), motivated by the existence of sibling species among the fungal pathogens of agricultural crops, the authors studied the evolution of pathogen diversity in the following evolutionary host-pathogen system. Two sympatric crops act as hosts to a pathogen which is categorised into a number of possible types according to its transmission rate on the two hosts. Only one of these is used as an explicit measure of diversity, namely the transmission rate on host one, denoted x and taking on values in a discrete set contained within $[0, 1]$. This value is termed the phenotype of the pathogen.

The underlying biological feature of the system is the following trade-off relationship: pathogens with higher transmission rates on host one have lower transmission rates on host two. In order to represent this relationship, a monotonically decreasing function $f(x)$ is used to provide the transmission rate of a pathogen on host two directly from the value of x.

Gudelj *et al.* found that the number of pathogen types present and the values of their corresponding phenotypes depend on the shape of the trade-off function f. However, many of these results have been ascertained from numerical simulations

† corresponding author: tel: 020 75942351, fax: 01344 873173, e-mail: i.gudelj@imperial.ac.uk

conducted for particular trade-offs. In this paper we build on this work in order to obtain a qualitative description of what form of trade-off function will ensure the long-term presence of one, two or more phenotypes, where infinitely many are possible in principle. This ultimately provides a classification of pathogen diversity as a function of the trade-off relationship, something that would not be possible using the methods of Gudelj *et al.* (2004).

This is an important step because the experimental literature has for a long time acknowledged the importance of trade-offs in the evolution of host-parasite systems. Indeed, there are an increasing number of studies conducted to identify those components that are traded during the evolutionary process (for example see Kraaijeveld and Godfray 1997, Boots and Begon 1993 and Messenger *et al.* 1999). However, the shapes of these trade-offs are extremely difficult to obtain experimentally.

Let us now concentrate on the particular host-pathogen problem that we wish to study and consider the system of differential equations proposed in Gudelj *et al.* (2004):

$$\dot{P}_i = -\gamma P_i + H_1 x_i P_i + H_2 f(x_i) P_i, \tag{1.1a}$$

$$\dot{H}_1 = r_1 - \mu H_1 - H_1 \sum_i x_i P_i, \tag{1.1b}$$

$$\dot{H}_2 = r_2 - \mu H_2 - H_2 \sum_i f(x_i) P_i, \tag{1.1c}$$

where i is an index which counts the number of resident pathogen phenotypes present in the system. The state variables H_1 and H_2 represent densities of healthy tissue of the two host populations while P_i represents the density of host tissue infected with a pathogen of phenotype x_i. In the absence of pathogens, the host populations are governed by a simple *immigration and birth minus death* process of the form $r - \mu H$, which leads to the long-term persistence of each host. The parameter γ denotes the disease-induced mortality rate of the host tissue and it is assumed to be sufficiently small so that a non–trivial, positive steady state of (1.1) exists and is stable.

Equation (1.1) possesses a discrete number of phenotypes and the rules which describe the phenotypic mutation process and the conditions that define the fate of each mutant can be found in Geritz *et al.* (1998), this is known as the *adaptive dynamics* method. There are many other ways of representing mutations in phenotypic space and some examples can be found in Sasaki and Godfray (1999), Bowers and Hodgkinson (2001), Bonhoeffer and Nowak (1994) and more recently in Calsina and Cuadrado (2004). Calsina and Cuadrado (2004) consider the evolving trait to be a continuous variable and mutations are included by the use of an convolution operator with a mutation kernel giving the probability distribution of the offspring phenotype, for a given parental phenotype. This yields a system of integro-differential equations and the authors relate their findings to the discrete trait approach used in the adaptive dynamics method.

In Calsina and Perelló (1994), Calsina *et al.* (1995), Tsmiring and Levine (1996) and Saldaña *et al.* (2003), the authors use an approach where a diffusion operator represents phenotypic mutations. At an informal level this corresponds to applying a continuum limit to (1.1) in the case where mutations are given by nearest-

phenotypic-neighbour connections. We also adopt this strategy and show that, in cases where a comparison is appropriate, our results agree with the findings of Gudelj *et al.* (2004) where the adaptive dynamics method is used, as well as many of the conclusions of classical evolutionary ecology (see Levins 1968 and Lawlor and Maynard-Smith 1976). Note that while classical theory assumes that populations evolve to maximise their fitness, neither the model (1.1) nor the one presented in this paper possess any global notion of fitness. Moreover, the methodology used in Gudelj *et al.* that is applied to (1.1) refers to a local notion of fitness of a phenotype, while there no concept of fitness is used in the remainder of this paper.

2. A non-local PDE

Let us begin this section with the formal definition of a trade-off function: if $f : [0,1] \rightarrow [0,1]$ satisfies

(F1) $f \in C^2[0,1], f(0) = 1, f(1) = 0, f'(x) < 0 \;\; \forall x \in (0,1)$.

then we call it a trade-off. From the mean-value theorem there is a $\theta \in (0,1)$ such that $f'(\theta) = -1$ and we also make the following restriction on the oscillatory properties of f:

(F2) the number $\Theta := \#\{\theta \in (0,1) : f'(\theta) + 1 = 0\}$ is finite and $\#\{x \in [0,1] : f''(x) = 0\} = \Theta - 1$, where a hash (#) symbol denotes set cardinality.

If $\Theta = 1$ then f is concave or convex. If $\Theta = 2$ then we shall say that f is *sigmoidal* and if $\Theta \geq 3$ then f is said to be a *staircase*. A sigmoidal trade-off is said to be *convex-concave* if there are intervals $I_1 = [0,a)$ and $I_2 = (a,1]$ such that f is convex on I_1 and concave on I_2. A concave-convex sigmoidal trade-off is similarly defined.

We now propose a non-local PDE system as our model of evolution by natural selection. Mutations in phenotypic space are represented by a diffusion process with a constant diffusion coefficient ϵ, which is motivated directly by the work of Calsina and Perelló (1995) and Calsina *et al.* (1995). The system represents the equations of motion for the probability density of the phenotype resident in the pathogen population as a function of evolutionary time.

Let us denote the state variables $H_i(t)$ for the density of host i, for $i \in \{1,2\}$, that is susceptible to infection at time t. Also $P(x,t)$ is the total density of hosts (both 1 and 2) infected with a pathogen with phenotypic trait x. Consequently, $H_1(t) + H_2(t) + \int_0^1 P(x,t)dx$ represents total biomass at time t. With this format we propose the following model:

$$(HP) \quad \begin{cases} P_t & = & \epsilon P_{xx} - \gamma P + H_1 x P + H_2 f(x) P, \\ \dot{H}_1 & = & r_1 - \mu H_1 - H_1 \int_0^1 x P(t,x)dx, \\ \dot{H}_2 & = & r_2 - \mu H_2 - H_2 \int_0^1 f(x) P(t,x)dx, \end{cases}$$

together with Neumann boundary conditions $P_x(0,t) = P_x(1,t) = 0$ such that $H_1(0) \geq 0, H_2(0) \geq 0$, where $P(x,0) = P_0(x) \in C^0[0,1]$ and $P_0(x) \geq 0$ for all $x \in [0,1]$. The remaining parameters in (HP) are defined as in (1.1).

Although the host dynamics of (HP) are very simple, namely when $P \equiv 0$, the results of this paper generalise to cover other types of interaction. We have used the particular form of (HP) in order to keep the exposition as straightforward as possible.

(a) Species

So as not to confuse terminology surrounding the term *species* with any currently in use we shall use the following, less ambitious terminology.

Definition 1. *By the term* modal phenotype at time t *for a solution of* (HP) *we understand a point $x \in [0,1]$ such that $P_x(x,t) = 0$ and $P_{xx}(x,t) < 0$. A modal phenotype $x \in (0,1)$ is said to be a* generalist *and a modal phenotype $x \in \{0,1\}$ is said to be a* specialist.

We consider the term *strategy* to be a synonym for phenotype.

(b) Critical Pathogenicity

Let us define some terminology. If X is any space of functions on $[0,1]$ containing all continuous functions, a subscript zero, as in X_0 denotes the fact that functions in X vanish at zero and one. Now define the Banach spaces $Y = \{u \in C^2[0,1] : u' \in C_0^1[0,1]\}$, $X := C^0[0,1]$ and $Z := X \times \mathbb{R}^2$ that will be needed, moreover Z_+ will denote the cone of non-negative functions in Z. The Sobolev spaces $W^{k,p}(0,1) = \{u \in C^{k-1}[0,1] : u^{(k)} \in L^p(0,1)\}$ for integer $k, p \geq 1$ will be used on occasion and we shall write $H^k(0,1)$ for $W^{k,2}(0,1)$. Here, $u^{(k)}$ denotes the k-th weak derivative and the operation of differentiating twice will sometimes be represented by the operator $\Delta : u \mapsto u''$, as in $\Delta : W^{k+2,p}(0,1) \to W^{k,p}(0,1)$ and $\Delta : Y \to X$.

By γ_0, we denote the unique eigenvalue $\gamma \in \mathbb{R}$ of the problem of finding a $\phi \in H^2(0,1)$ such that $\phi' \in H_0^1(0,1)$ and

$$L\phi := \epsilon\phi'' + (r_1 x + r_2 f(x))\phi/\mu = \gamma\phi, \quad \phi > 0. \tag{2.1}$$

Clearly γ_0 is positive and by well-known properties of eigenvalues of elliptic problems, it is an algebraically simple eigenvalue. The value γ_0 will be called the *critical pathogenicity*. Let ϕ_0 denote the eigenfunction corresponding to γ_0 such that $\int_0^1 \phi_0^2 dx = 1$. Finally, let $\phi_1 \in H^2(0,1)$ be the unique positive function that satisfies $\phi_1' \in H_0^1(0,1)$, $\int_0^1 \phi_1^2 dx = 1$ and $\epsilon\phi'' + r_1 x\mu^{-1}\phi = \gamma_1\phi$, for some eigenvalue $\gamma_1 > 0$. Of course, the value of γ_0 depends on f and we may write $\gamma_0(f)$ to emphasize this. It is also true that $\gamma_0(f) > \gamma_1$ if $f > 0$ and that $\gamma_0(0) = \gamma_1$.

The existence and positivity of solutions of (HP) is a simple consequence of the maximum principle and the theory of analytic semigroups, and the following lemma gives some simple boundedness properties which show that the hosts are persistent.

Lemma 2.1. *Suppose that $(P_0, h_1, h_2) \in Z_+$ represents an initial datum for* (HP). *There is a unique non-negative solution of* (HP), $(P, H_1, H_2) \in C^1((0,\infty), Z_+)$ *with this initial datum. This solution satisfies*

$$H_i(t) \leq h_i e^{-\mu t} + r_i(1 - e^{-\mu t})/\mu, \quad \text{(for } i = 1, 2\text{)}. \tag{2.2}$$

If we define $m := \sup_{t \geq 0}(H_1(t) + H_2(t))$ (which is finite a-priori), then there is an upper bound

$$\int_0^1 P(x,t)dx \leq e^{-\gamma t} \int_0^1 P_0(x)dx + \frac{r_1 + r_2}{\gamma} + m + h_1 + h_2 \tag{2.3}$$

for all $t > 0$, so that total biomass is bounded above. If $M := \sup_{t \geq 0} \int_0^1 P(x, t)dx$ (which is finite a-priori) then there is a lower bound

$$h_i e^{-(M+\mu)t} + \frac{r_i}{M+\mu}(1 - e^{-(M+\mu)t}) \leq H_i(t), \quad i = 1, 2.$$

Consequently, the total biomass is eventually bounded below:

$$\liminf_{t \to \infty}(H_1(t) + H_2(t)) \geq \frac{r_1 + r_2}{M + \mu}.$$

Proof. The existence, uniqueness and positivity property is a simple excercise in well known theory of analytic semigroups (see Henry 1981, Theorem 3.3.3, p.54 and Exercise 8, p.61).

To prove (2.2), simply note that H_i satisfies the inequality $\dot{H}_i \leq r_i - \mu H_i$, which integrates to give the desired bound. To prove (2.3), write $n(t) = \int_0^1 P(x, t)dx$ and note that

$$\dot{n} + \gamma n = r_1 - \mu H_1 - \dot{H}_1 + r_2 - \mu H_2 - \dot{H}_2 \leq r_1 - \dot{H}_1 + r_2 - \dot{H}_2,$$

and this inequality can be integrated to show that

$$\begin{aligned} n(t) &\leq& e^{-\gamma t}n(0) + \frac{r_1 + r_2}{\gamma} + \gamma e^{-\gamma t} \int_0^t (H_1(s) + H_2(s))e^{\gamma s}ds + h_1 + h_2, \\ &\leq& e^{-\gamma t}n(0) + \frac{r_1 + r_2}{\gamma} + m + h_1 + h_2. \end{aligned}$$

To prove (2.4) for $i = 1$, note that $\langle x, P(x, t) \rangle \leq n(t) \leq M$ for some $M > 0$ depending on initial conditions and therefore

$$\dot{H}_1 = r_1 - \langle x, P \rangle H_1 - \mu H_1 \geq r_1 - M H_1 - \mu H_1.$$

This can be integrated to give

$$H_1(t) \geq h_1 e^{-(M+\mu)t} + \frac{r_1}{M+\mu}(1 - e^{-(M+\mu)t}).$$

The bound on H_2 is similarly obtained. □

The following lemma shows that if the pathogenicity has a greater value than the critical pathogenicity, the fate of all solutions of (HP) is to converge to the trivial steady-state whereby the pathogen distribution is the zero function.

Lemma 2.2. *The set $\mathcal{Z} := \{(P, H_1, H_2) \in Z_+ : P = 0\}$ is invariant for (HP) and any orbit in $\mathcal{Z}$ converges to $(0, r_1/\mu, r_2/\mu)$ as $t \to \infty$. If $\gamma > \gamma_0$ then any solution of (HP) converges eventually exponentially to $(0, r_1/\mu, r_2/\mu)$ as $t \to \infty$ in the norm on Z given by $\|(P, H_1, H_2)\| = \int_0^1 \phi_0|P|dx + |H_1| + |H_2|$.*

Proof. The first part of the lemma is obvious, so let us suppose that $\gamma > \gamma_0$ and define the functional $V(t) = \int_0^1 \phi_0(x)P(x, t)dx$, evaluated along a solution of (HP). Now V is differentiable for $t > 0$ and

$$\dot{V} = \int_0^1 \phi_0(\epsilon P_{xx} + P(-\gamma + xH_1 + f(x)H_2))dx,$$

$$= \int_0^1 P\phi_0(\gamma_0 - \gamma + x(H_1 - r_1/\mu) + f(x)(H_2 - r_2/\mu))dx,$$

and by Lemma 2.1, for every sufficiently small $\eta > 0$, there exists a $T > 0$ such that $H_1(t) < r_1/\mu + \eta$ and $H_2(t) < r_2/\mu + \eta$, for all $t > T$. Consequently, there exists an $\alpha < 0$ such that $\dot{V} < \alpha V$, for all $t \geq T$ and therefore $V(t) < e^{\alpha(t-T)}V(T)$ eventually. The result now follows. $\qquad\square$

3. Classifying Equilibrium States

Consider the equilibrium problem of (HP):

$$0 = \varepsilon P'' - \gamma P + \frac{r_1 x P}{\mu + \langle x, P \rangle} + \frac{r_2 f(x) P}{\mu + \langle f, P \rangle}, \qquad (3.1)$$

subject to the restrictions $P > 0, P'(0) = P'(1) = 0$. Throughout the paper we shall make use of the function

$$r := \varepsilon \frac{P''}{P} = \gamma - \frac{r_1 x}{\mu + \langle x, P \rangle} - \frac{r_2 f(x)}{\mu + \langle f, P \rangle}, \qquad (3.2)$$

defined whenever P is a positive solution of (3.1).

Lemma 3.1. *If (F1) holds and $f''(x) > 0$ for all $x \in [0,1]$ (f is strictly convex) then $\int_0^1 r(x)dx \geq 0$, r is concave and any zeros of r, of which there is at least one in $(0,1)$ and at most two in $[0,1]$, are necessarily transverse:*

$$x \in [0,1], \ r(x) = 0 \Rightarrow r'(x) \neq 0.$$

Moreover, r' has at most one zero in $[0,1]$ and it is transverse, so that r is either monotonic on $[0,1]$ or has exactly one local extremum in $[0,1]$ which is a global maximum.

Proof. Since $r''(x) = -\frac{r_2 f''(x)}{\mu + \langle f, P \rangle} < 0$, we see that r is strictly concave and so has at most two zeros. However, since $P'(0) = P'(1) = 0$, we find $P''(x_0) = 0$ for some $x_0 \in (0,1)$, whence $r(x_0) = 0$. Also $\int_0^1 r(x)dx = \int_0^1 \varepsilon P''/P dx = \varepsilon \int_0^1 (P'/P)^2 dx \geq 0$. The remainder of the proof is elementary and uses the mean-value theorem and the convexity of f. $\qquad\square$

Similarly, we have the following.

Lemma 3.2. *If (F1) holds and $f''(x) < 0$ for all $x \in [0,1]$ (f is strictly concave) then $\int_0^1 r(x)dx \geq 0$, r is convex and any zeros of r, of which there is at least one in $(0,1)$ and at most two in $[0,1]$, are necessarily transverse. Moreover, r' has at most one zero in $[0,1]$ and it is transverse, so that r is either monotonic on $[0,1]$ or has exactly one local extremum in $[0,1]$ which is a global minimum.*

Denote the right-hand side of (3.1) by the smooth operator $F(P, \gamma)$, where $F : Y \times \mathbb{R} \to X$, then we can find a $\delta > 0$ such that F is an analytic mapping on

the open set $K = \{(P, \gamma) \in Y \times \mathbb{R} : \gamma > -\delta, P \gg -\delta\}$, where $P \gg -\delta$ means that $P(x) > -\delta$ for all $x \in [0, 1]$. Let us also define a solution set,

$$S = \{(P, \gamma) \in Y \times \mathbb{R} : F(P, \gamma) = 0\},$$

a set of positive solutions, $S_+ = \{(P, \gamma) \in S : \gamma > 0, P > 0\}$, and a set of non-trivial solutions $N = \{(P, \gamma) \in S : P \neq 0\}$.

Note that if $f \in C^2$ as per assumption (F1) and P is a solution of (3.1), then $P \in C^4$ immediately follows. According to Lemma 2.2, there can be no positive solution of (3.1) for $\gamma > \gamma_0$; we now sketch a proof that in fact $\gamma = \gamma_0$ is a global bifurcation point to positive solutions of (3.1).

Let us first state a preliminary lemma.

Lemma 3.3. *Suppose that (F1) holds, $(P, \gamma) \in N$ and $P \geq 0$ then $(P, \gamma) \in S_+$. More generally, if $(P, \gamma) \in N$ then the zeros of P are transverse and therefore finite in number. Consequently, if $C \subset N$ is a connected set that satisfies $C \cap S_+ \neq \emptyset$, then $C \subset S_+$.*

Proof. Suppose that $(P, \gamma) \in N$ satisfies $P(x_0) = P'(x_0) = 0$ for some $x_0 \in [0, 1]$, then define numbers $i_1 = r_1/(\mu + \langle x, P \rangle)$ and $i_2 = r_2/(\mu + \langle f, P \rangle)$. It follows that P is a solution of an ordinary differential equation $\epsilon P'' + (-\gamma + i_1 x + i_2 f(x))P = 0$ with initial condition $P(x_0) = P'(x_0) = 0$, but then $P(x) \equiv 0$. This is a contradiction since $(P, \gamma) \in N$ and therefore no such x_0 exists. This immediately implies that non-negative solutions of (3.1) are strictly positive.

Now suppose that $(P, \gamma) \in C \cap S_+$ and define a mapping $i : C \to \{0\} \cup \mathbb{N}$ by $i(P, \gamma) = \#\{x \in [0, 1] : P(x) = 0\}$. Then i is continuous on C by the transversality of zeros of P and integer valued. It is locally constant and therefore constant on the connected set C, so that $i(C) = 0$ as C contains an element also in S_+. Hence $C \subseteq S_+$. $\square$

Lemma 3.4. *For each $\gamma \in (0, \gamma_0)$, the Neumann problem (3.1) has at least one solution. If we denote a locus of such solutions by P_γ, then*

$$\lim_{\gamma \to \gamma_0} \|P_\gamma\|_{C^2} = 0 \text{ and } \lim_{\gamma \to 0} \|P_\gamma\|_{L^1} = \infty.$$

If $C(\gamma_0) \subset N$ is the maximal connected subset of N which contains the point $(P, \gamma) = (0, \gamma_0)$ in its closure, then $C(\gamma_0) \subset S_+$.

Proof. (sketch) First, note that $F(0, \gamma) \equiv 0$. Now, because $d_P F(0, \gamma)[h] = Lh - \gamma h$ where L is defined in (2.1) and $d^2_{P\gamma} F(0, \gamma)[h, \lambda] = -\lambda h$, by the defining property of γ_0 one may apply the theorem on bifurcation from a simple eigenvalue to obtain the existence of a local bifurcation. The Crandall and Rabinowitz global alternative (see Zeidler 1986) can then be used to demonstrate the existence of $C(\gamma_0)$ in S_+. We omit the details as they are straightforward.

Integrating (3.1) over $[0, 1]$ shows that $\gamma > 0$ must hold for any non-negative solution which is not identically zero. The second part then follows from a continuity argument because if there is a positive sequence (γ_n) such that $\gamma_n \to 0$ and (P_n) is a sequence of positive solutions of (3.1) which is L^1-bounded, the equation

$$-\varepsilon P_n'' + \alpha P_n = (\alpha - \gamma_n)P_n + \frac{r_1 x P_n}{\mu + \langle x, P_n \rangle} + \frac{r_2 f(x) P_n}{\mu + \langle f, P_n \rangle}, \quad (P_n'(0) = P_n'(1) = 0),$$

(where α is a constant chosen to ensure the invertibility of the linear operator $-\varepsilon\Delta + \alpha I$ independently of n in various spaces) yields a $W^{2,1}$-bound, hence a C^1 bound and using (F1), a C^3-bound follows for (P_n) that is independent of n. As a result, there is a subsequence of (P_n) that converges in C^2 to a non-negative solution of (3.1) when $\gamma = 0$, which is a contradiction as none exist.   □

(a) Steep Trade-Offs

The following result shows that if the trade-off function is sufficiently *steep* in the sense of having most of its mass concentrated around $x = 0$, then there is only one solution of (3.1) and it is close (in a C^1 sense) to a strictly monotonic increasing function on $(0,1)$ and so its mass is largely concentrated near $x = 1$. This corresponds to solutions of (HP) that are close to one that supports only one modal phenotype at the specialist point $x = 1$. Such behaviour was identified by Gudelj *et al.* (2004) (Fig. 3(b), p. 2190) using the adaptive dynamics approach and so we ought to expect to observe this result in (HP).

Theorem 3.5. *Suppose that $\gamma < \gamma_1$. There is a $\rho > 0$ such that if f is any trade-off function satisfying $\int_0^1 f(x)dx \leq \rho$, then there is exactly one positive solution $P(= P(f))$ of (3.1). Moreover, $P(f) \to p$ in $H^2(0,1)$ (and hence in $C^1[0,1]$) as $\int_0^1 f(x)dx \to 0$ where p is the strictly monotonic increasing function on $(0,1)$ that satisfies (3.1) with $f = 0$.*

Proof. Fix $\gamma < \gamma_1$ $(\leq \gamma_0(f))$ and first consider an auxiliary problem.

<u>Claim:</u> Under the assumptions of the theorem, the problem

$$0 = \varepsilon p'' - \gamma p + \frac{r_1 x p}{\mu + \langle x, p \rangle}, \qquad p \in Y, p > 0, \tag{3.3}$$

obtained by setting $f = 0$ in (3.1) has a unique positive solution and it is strictly monotonic increasing on $(0,1)$.

Proof. (of claim) The *existence* part of this claim follows from Lemma 3.4. This shows that (3.3) has a global bifurcation to a branch of positive solutions emanating from the trivial solution at $\gamma = \gamma_1$ and that $0 < \gamma < \gamma_1$ holds along this branch.

The *uniqueness* part of the claim follows because if (3.3) has two positive solutions $p, q > 0$ and we consider their Wronskian $w := p'q - pq'$, then

$$\epsilon w' = \frac{r_1 x p q \langle x, p - q \rangle}{(\mu + \langle x, p \rangle)(\mu + \langle x, q \rangle)}.$$

Now $w(0) = w(1) = 0$ and so $w'(x_1) = 0$ follows for some $x_1 \in (0,1)$ and therefore $\langle x, p - q \rangle = 0$, so that $w(x) \equiv 0$. From this it follows easily that $p = q$.

The *monotonicity* part of the claim is proven as follows. Since $p'(0) = p'(1) = 0$ there is an $x_0 \in (0,1)$ such that $p''(x_0) = 0$ and therefore

$$\gamma p(x_0) = \frac{x_0 r_1 p(x_0)}{\mu + \langle x, p \rangle}, \qquad p(x_0) > 0.$$

This implies $x_0 = \gamma(\mu + \langle x, p \rangle)/r_1$ must be the unique point in $(0,1)$ where p'' can be zero. If $p'(x_0') = 0$ were true for some $x_0' \in (0,1)$, then there would be two points

in $(0,1)$ where p'' was zero, one in $(0, x_0')$ and one in $(x_0', 1)$. This contradiction ensures that no such x_0' exists. Since $\varepsilon p''(0) = \gamma p(0) > 0$ it follows that $x = 0$ is a local minimum for p, consequently p must be strictly monotonic increasing on $(0,1)$. □

Now for the proof of the theorem. Let us denote the solution of (3.3) by p and define the operator $N : L^2(0,1) \times H \to L^2(0,1)$ by

$$N(f, P) := \varepsilon P'' - \gamma P + \frac{r_1 x P}{\mu + \langle x, P \rangle} + \frac{r_2 f(x) P}{\mu + \langle f, P \rangle},$$

where $H := H^2(0,1) \cap \{u \in C^1[0,1] : u'(0) = u'(1) = 0\}$ is a closed subspace of $H^2(0,1)$. Now N is a C^1 mapping and from the above claim $N(0, p) = 0$, but $p \in H$ satisfies $p \in Y$ by a simple regularity argument and so provides the only solution of $N(0, P) = 0$. Moreover, the Fréchet derivative of N with respect to P at $(f, P) = (0, p)$ is

$$d_P N(f, P)[Q] = \varepsilon Q'' - \gamma Q + \frac{r_1 x Q}{\mu + \langle x, P \rangle} + \frac{r_2 f(x) Q}{\mu + \langle f, P \rangle}$$
$$- \frac{r_1 x P}{(\mu + \langle x, P \rangle)^2} \langle x, Q \rangle - \frac{r_2 f(x) P}{(\mu + \langle f, P \rangle)^2} \langle f(x), Q \rangle,$$

whence

$$d_P N(0, p)[Q] = \varepsilon Q'' - \gamma Q + \frac{r_1 x Q}{\mu + \langle x, p \rangle} - \frac{r_1 x p}{(\mu + \langle x, p \rangle)^2} \langle x, Q \rangle.$$

We may write $d_P N(0, p)[Q] = AQ - \langle x, Q \rangle g(x)$, where g is the smooth positive function $r_1 x p / (\mu + \langle x, p \rangle)^2$ and A is a self-adjoint, second-order linear differential operator such that $Ap = 0$. Since A is a Fredholm operator of index-0, so too is $d_P N(0, p)$ as it is a rank one perturbation of A and to show that this derivative is an isomorphism we only need prove injectivity. Now, $d_P N(0, p)[Q] = 0$ yields

$$0 = \int_0^1 p(AQ - \langle x, Q \rangle g) dx = \int_0^1 (Ap)Q - \langle x, Q \rangle pg) dx = -\langle x, Q \rangle \int_0^1 pg \, dx.$$

As a result, $\langle x, Q \rangle = 0$ and therefore $AQ = 0$ follows, from where $Q = \alpha p$ for some real constant α and the simplicity of the null-space of A. The same integral identity yields $\alpha = 0$ and we deduce that $d_P N(0, p)$ is an isomorphism from H to L^2, so that we can apply the implicit function theorem to solve $N(f, P) = 0$ locally to $(0, p)$ for $P = P(f)$.

Local uniqueness is thus established, but for global uniqueness suppose that there are two distinct solutions of (3.1) P_n^1 and P_n^2 say, that are found for $f = f_n$ where $\int_0^1 f_n(x) dx \to 0$ as $n \to \infty$. Note that $\|f_n\|_{L^p} \le \left(\int_0^1 f_n(x) dx \right)^{1/p}$ holds by (F1) and so (f_n) tends to zero in each L^p-space for $p \ge 1$.

Let P be a solution of (3.1). By integrating over $[0,1]$ the f-independent L^1-bound $\int_0^1 P(x) dx \le (r_1 + r_2)\gamma^{-1}$ holds. Integrating from 0 to $z \in (0,1)$ we obtain

$$\left| -\epsilon P'(z) + \gamma \int_0^z P(x) dx \right| = \left| \int_0^z -\epsilon P'' + \gamma P \, dx \right| \le \int_0^1 \frac{r_1 x P}{\mu + \langle x, P \rangle} + \frac{r_2 f(x) P}{\mu + \langle f, P \rangle} dx$$

which is bounded above by $r_1 + r_2$. The elementary inequality $||x| - |y|| \le |x - y|$ now gives

$$\left| |\epsilon P'(z)| - \left| \gamma \int_0^z P(x)dx \right| \right| \le r_1 + r_2,$$

as a result $|P'(z)| \le 2(r_1 + r_2)\epsilon^{-1}$ and since $P(x) - P(y) = \int_x^y P'(z)dz$, we find $|P(x)| \le |P(y)| + 2(r_1 + r_2)\epsilon^{-1}|x - y|$. Integrating this gives

$$|P(x)| \le \int_0^1 |P(y)|dy + 2(r_1 + r_2)\epsilon^{-1} \int_0^1 |x - y|dy,$$

so that $\|P\|_{C^0} \le (r_1 + r_2)(\gamma^{-1} + 4\epsilon^{-1})$. Choosing an $\alpha \in \mathbb{R}$ such that $-\epsilon\Delta + \gamma + \alpha : H \to L^2(0,1)$ is an isomorphism, a solution P of (3.1) satisfies

$$\|P\|_{H^2} \le \|(-\epsilon\Delta + \gamma + \alpha)^{-1}\|_{BL(L^2,H)} \left\| \frac{r_1 x P}{\mu + \langle x, P \rangle} + \frac{r_2 f(x) P}{\mu + \langle f, P \rangle} + \alpha P \right\|_{L^2}$$

$$\le \|(-\epsilon\Delta + \gamma + \alpha)^{-1}\|_{BL(L^2,H)}(\alpha + (r_1 + r_2)\mu^{-1})(r_1 + r_2)(\gamma^{-1} + 4\epsilon^{-1})$$

where we have used $0 \le f(x) \le 1$ that comes from (F1).

As H is a Hilbert space we can extract weakly convergent subsequences so that $P_n^1 \rightharpoonup P^1$ and $P_n^2 \rightharpoonup P^2$ in H as $n \to \infty$ for some $P^1, P^2 \in H$ and by compact embedding results, we can choose a new subsequence if necessary to ensure that $P_n^i \to P^i$ strongly in $H^1(0,1)$ as $n \to \infty$. Since $f_n \to 0$ in L^2, it follows from (3.1) that

$$\varepsilon(P_n^i)'' = \gamma P_n^i - \frac{r_1 x P_n^i}{\mu + \langle x, P_n^i \rangle} - \frac{r_2 f_n(x) P_n^i}{\mu + \langle f, P_n^i \rangle} \to \gamma P^i - \frac{r_1 x P^i}{\mu + \langle x, P^i \rangle}$$

in L^2 for each $i = 1, 2$. As a result $P_n^i \to P^i$ strongly in $H^2(0,1)$ and because of the continuity of $N : L^2(0,1) \times H \to L^2(0,1)$, P^1 and P^2 must both be solutions of $N(f, P) = 0$ when $f = 0$. But then $P^1 = P^2$ and it immediately follows from the previously proven local uniqueness result that $P_n^1 = P(f_n) = P_n^2$ for all sufficiently large n as both sequences must eventually lie in H-neighbourhoods of p. □

One could restrict the trade-off further in the statement of Theorem 3.5 if we wished in such a way that the resulting operator $P(f)$ would yield monotonic solutions of (3.1), rather than solutions that are C^1-close to being monotonic. However, we have not done this for brevity.

4. Convex and Concave Trade-Offs

Let $f \in C^2[0,1]$ and $P \in Y \cap C^4[0,1]$ be a corresponding positive solution of (3.1). In Gudelj *et al.* (2004), the authors deduced that the convexity or concavity properties of particular trade-offs resulted in the presence of one or two pathogen *strains*. The purpose of this and the following section is to use the properties of the function r defined in (3.2) to deduce how many modal phenotypes a solution of (3.1) can support. Indeed, we deduce a similar result.

Lemma 4.1. *Assuming (F1) and $f''(x) \ne 0$ for all $x \in [0,1]$ (so that (F2) holds) then any solution of (3.1) has at most three distinct local extrema in $[0,1]$. Consequently, there is at most one local extremum of such a solution in $(0,1)$.*

Proof. Suppose that a solution P of (3.1) has four, or more, local extrema $0 = x_1 < x_2 < x_3 < x_4 = 1$, so that $P'(x)$ vanishes at all of these points. This yields three zeros of P'' in $(0, 1)$ and therefore r too has three zeros. Now $r'(x)$ has at most one zero since

$$r'(x) = -\frac{r_1}{\mu + \langle x, P \rangle} - \frac{r_2 f'(x)}{\mu + \langle f, P \rangle} \qquad (4.1)$$

and f' is monotonic, a contradiction. Since the boundary conditions on P are Neumann, the boundary points provide two local extrema. $\qquad \square$

It immediately follows from Lemma 4.1 that any steady-state solution of (HP) has either one modal phenotype, which we call monomorphism, or two modal phenotypes, termed dimorphism. The manner in which monomorphism and dimorphism occurs when $f''(x) \neq 0$ is quantified in the following two propositions.

Proposition 4.2. *If (F1) holds and $f''(x) > 0$ for all $x \in [0, 1]$ then any solution of (3.1) is either monotonic or has a unique global minimum and two local maxima at $x = 0$ and $x = 1$.*

Proof. By Lemma 4.1, either there exists an $x_0 \in (0, 1)$ such that $P'(x_0) = 0$ or there does not. In the latter case P is monotonic, so let us suppose that such an x_0 exists and note that it must be unique by Lemma 4.1. From Lemma 3.1 there are four cases to consider:

1. r has two zeros in $(0, 1)$,
2. r has one zero in $(0, 1)$ and $r(0) = 0, r(1) \neq 0$,
3. r has one zero in $(0, 1)$ and $r(0) \neq 0, r(1) = 0$ and
4. r has one zero in $(0, 1)$ and $r(0) \neq 0, r(1) \neq 0$.

In case (1), since f is convex r is concave and therefore has *sign pattern* $- + -$ on $[0, 1]$, as shown in the left-hand, topmost diagram of Figure 1. Suppose that $r(x) > 0$ on an interval $I \subset (0, 1)$ and r is zero on ∂I, if $x_0 \notin I$ then P'' must vanish outside $\overline{I}$, and therefore so too must r. This contradiction ensures that $x_0 \in I$. Since $\text{sgn}(P''(x)) = \text{sgn}(r(x))$, we see that $P'(x_0) = 0$ and $P''(x_0) > 0$, so that $x = x_0$ is a local minimum for P. Since P is monotonic on $(0, x_0)$ and $(x_0, 1)$, P must have a global minimum at $x = x_0$. The boundary conditions ensure that P has two local maxima at the boundary of $[0, 1]$.

In cases (2-4), the mean-value theorem ensures that P' cannot have a zero in $(0, 1)$ and P is therefore monotonic on $(0, 1)$. $\qquad \square$

Proposition 4.3. *If (F1) holds and $f''(x) < 0$ for all $x \in [0, 1]$ then any solution of (3.1) is either monotonic or has a unique global maximum in $(0, 1)$ and two local minima at $x = 0$ and $x = 1$.*

Proof. Using Lemma 3.2, the proof is entirely analogous to that of Proposition 4.2. The only possible non–monotonic solution is illustrated in Figure 2. $\qquad \square$

(a) Phenotypic Structure as $\epsilon \to 0$

In this section we perform an asymptotic analysis to locate the modal phenotype of (3.1) in the case where the trade-off function f is concave. We concentrate on

Figure 1. The horizontal axis is x in both left and right-hand figures. The only possible non-monotonic shape for a solution of (3.1) when f is convex is shown in the topmost diagram (two specialist phenotypes). Also, r is shown in the left-hand figures (the thick black line denotes zero) and the corresponding solution P in the right-hand figures.
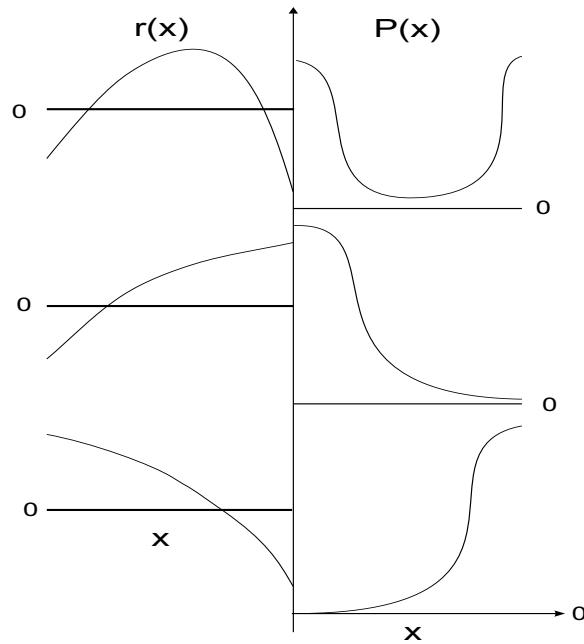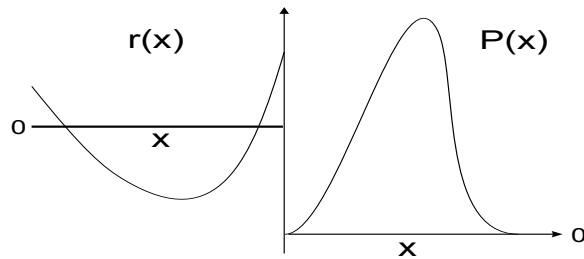

Figure 2. The only possible non-monotonic shape for a solution of (3.1) when f is concave, illustrating a unique generalist evolutionary stable strategy. (see also Figure 1)

a specific form of trade-off but, provided f is smooth, our analysis can be carried out for more general trade-off functions. We therefore have in mind equation (3.1) where

$$f(x) = 1 - x^{\alpha}, \text{ for } \alpha > 1.$$

It is helpful to define new constants Γ_1 and Γ_2 by

$$\Gamma_1 := \frac{r_1}{\mu + \langle x, P \rangle} \text{ and } \Gamma_2 := \frac{r_2}{\mu + \langle f, P \rangle},$$

so that (3.1) becomes

$$\varepsilon P''(x) - (\gamma - \Gamma_1 x - \Gamma_2 f(x)) P(x) = 0. \tag{4.2}$$

We are taking r_1, μ and r_2 as given, but Γ_1 and Γ_2 are not; equation (4.2) is essentially a linear problem, but the nonlinearity will arise when we have to fix the values of Γ_1 and Γ_2.

From the results obtained in §3 in the case of concave trade-offs, we anticipate that the solution of (4.2) is characterized by a bump centered at some point inside $[0, 1]$. Therefore we start by defining the co-ordinate X given by

$$x = x_0 + \varepsilon^{1/4} X;$$

we are expecting the solution to be concentrated in a layer of width $\mathcal{O}(\varepsilon^{1/4})$, but we do not yet know where, so x_0 is unknown and is to be found in the course of the following analysis.

Let us expand formally

$$P = P_0(X) + \varepsilon^{1/4} P_1(X) + \varepsilon^{1/2} P_2(X) + \mathcal{O}(\varepsilon^{3/4}), \tag{4.3}$$

$$\Gamma_1 = \Gamma_{10} + \varepsilon^{1/4} \Gamma_{11} + \varepsilon^{1/2} \Gamma_{12} + \mathcal{O}(\varepsilon^{3/4}), \tag{4.4}$$

$$\Gamma_2 = \Gamma_{20} + \varepsilon^{1/4} \Gamma_{21} + \varepsilon^{1/2} \Gamma_{22} + \mathcal{O}(\varepsilon^{3/4}), \tag{4.5}$$

as we are expecting to obtain the elements of (4.2) in powers of $\varepsilon^{1/4}$. Suppose also that $f(x)$ has a well-behaved Taylor series expansion about x_0, so we write

$$f = f_0 + \varepsilon^{1/4} X f_0' + \varepsilon^{1/2} X^2 f_0''/2 + \mathcal{O}(\varepsilon^{3/4}),$$

where $f_0 = f(x_0), f_0' = f'(x_0)$ and $f_0'' = f''(x_0)$.

An essential point is to recognise that the double derivative term in (4.2) is now formally $\mathcal{O}(\varepsilon^{1/2})$ (when the multiplying ε is taken into account) and we can then start to fix various constants appropriately. It follows that the expression $\gamma - \Gamma_1 x - \Gamma_2 f(x)$ must vanish at $\mathcal{O}(1)$ and $\mathcal{O}(\varepsilon^{1/4})$ and for this we require

$$\mathcal{O}(1): \quad \gamma - \Gamma_{10} x_0 - \Gamma_{20} f_0 = 0, \tag{4.6a}$$

$$\mathcal{O}(\varepsilon^{1/4}): \quad \Gamma_{10} + \Gamma_{20} f_0' = 0 \quad \text{and} \quad \Gamma_{11} x_0 + \Gamma_{21} f_0 = 0. \tag{4.6b}$$

Note that the second relation in (4.6b) is a higher-order effect which we shall not need again.

Taking the $\mathcal{O}(\varepsilon^{1/2})$ terms in equation (4.2) we obtain the expression

$$\frac{d^2 P_0}{dX^2} - \left[-\frac{1}{2} f_0'' X^2 \Gamma_{20} - (\Gamma_{11} + \Gamma_{21} f_0') X - (\Gamma_{22} f_0 + \Gamma_{12} x_0) \right] P_0 = 0.$$

Let us define

$$\lambda^2 := -\Gamma_{20} f_0''/2,$$

which is positive owing to the concavity of $f(x)$ and note that we may translate X by defining $X = \hat{X} + k$, where $k \in \mathbb{R}$ is a constant chosen so that the linear term in

the latter equation is removed. Details of the constant shift are not important for the leading-order analysis, but what is left is the determining equation

$$\frac{d^2 P_0}{d\hat{X}^2} - \left[\lambda^2 \hat{X}^2 - \hat{\gamma}\right] P_0 = 0, \tag{4.7}$$

where $\hat{\gamma}$ is some constant that depends on the coefficients in the expansions of Γ_1 and Γ_2 in (4.4) and (4.5). Now we would like solutions of this equation that decay away from x_0, that is as $\hat{X} \to \pm\infty$. This is perfectly feasible as the equation is just a scaled parabolic cylinder problem, the first solution of which is

$$P_0(\hat{X}) = c \exp(-\hat{\lambda}\hat{X}^2/2),$$

for an appropriate $\hat{\lambda}$ and any constant $c \in \mathbb{R}$. In determining the value of $\hat{\lambda}$, a further restriction must be imposed on the coefficients in (4.4) and (4.5) because $\lambda + \hat{\gamma}$ must hold for $\exp(-\hat{\lambda}\hat{X}^2/2) = 0$ to be a solution of (4.7).

Having identified the leading order eigenfunction, the stage is now set for us to tie down the actual location of the bump. Although we have not been concerned about the scaling of the eigenfunction yet, it is at this point that we observe that it must be of size $\mathcal{O}(\varepsilon^{-1/4})$ as it occupies a zone of width $\mathcal{O}(\varepsilon^{1/4})$ and we suspect that the overall integral of P is $\mathcal{O}(1)$. Suppose that

$$\int_0^1 P(\xi)d\xi = A;$$

then as P is concentrated around x_0, we find at leading order that

$$\int_0^1 \xi P(\xi)d\xi \approx A x_0 \text{ and } \int_0^1 f(\xi)P(\xi)d\xi \approx f_0 A.$$

Consequently we deduce that $\Gamma_{10} = r_1/(\mu + A\,x_0)$ and $\Gamma_{20} = r_2/(\mu + f_0 A)$. Equation (4.6b) now gives

$$\frac{r_1}{\mu + Ax_0} + \frac{r_2}{\mu + Af_0}f_0' = 0, \tag{4.8}$$

and (4.6a) tells us that

$$\gamma - \frac{x_0 r_1}{\mu + Ax_0} - \frac{f_0 r_2}{\mu + Af_0} = 0. \tag{4.9}$$

Thus we have found that for a given concave function the eigensolution is concentrated in a thin layer around $x = x_0$. This point is determined as the solution of (4.8-4.9) (for given r_1, r_2 and $f(x)$, the only unknowns are x_0 and A). The eigensolution itself is a Gaussian profile whose decay rate $\hat{\lambda}$ is determined as a function of the underlying trade-off $f(x)$.

5. Sigmoidal and Staircase Cases

Lemma 5.1. *If (F1) and (F2) hold with $\Theta = 2$, then any solution of (3.1) has at most four distinct local extrema in $[0, 1]$. Consequently, there are at most two local extrema of such a solution in $(0, 1)$.*

Proof. Suppose that a solution P of (3.1) has five, or more local extrema $0 = x_1 < x_2 < x_3 < x_4 < x_5 = 1$, so that $P'(x)$ vanishes at all of these points. This yields four zeros of P'' in $(0,1)$ and therefore r too has four zeros from which it follows that r' has three zeros in $(0,1)$. From (4.1) it follows that f'' has two zeros in $(0,1)$, a contradiction. Since the boundary conditions on P are Neumann, the boundary points provide two local extrema. $\square$

Lemma 5.2. *If (F1) and (F2) hold with $\Theta = 2$, then r has at least one zero in $(0,1)$ and at most three zeros in $[0,1]$ counted according to multiplicity. Moreover, r' has at most two zeros in $[0,1]$ counted according to multiplicity, so that r is either monotonic, has exactly one local extremum in $[0,1]$ or has two local extrema in $[0,1]$.*

Proof. Since $P'(0) = P'(1) = 0$, $P''(x_0)$ for some $x_0 \in (0,1)$, whence $r(x_0) = 0$. Now suppose that r has four zeros in $[0,1]$, $0 \le x_1 < x_2 < x_3 < x_4 \le 1$. This yields three zeros of r' in $(0,1)$ and consequently two zeros of r'' in $(0,1)$. Therefore it follows that f'' has two zeros in $(0,1)$, a contradiction.

The same contradiction is obtained if r has two non-transverse zeros, so that r has at most three zeros counted according to multiplicity and similar reasoning can be used to deduce that r' has at most two zeros, either two transverse zeros or one double zero. $\square$

Proposition 5.3. *If (F1) and (F2) hold with $\Theta = 2$ and the trade-off f is sigmoidal convex-concave, then any solution of (3.1) is either*
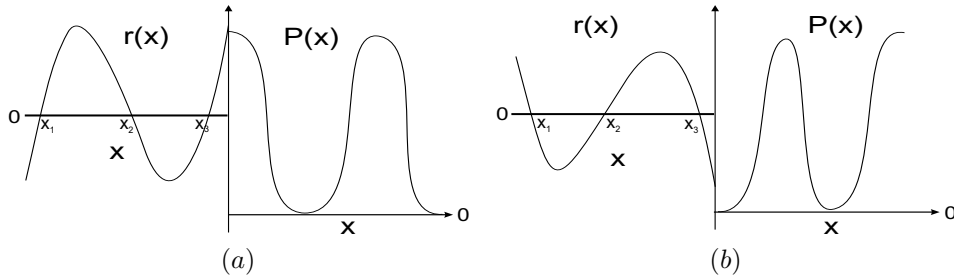
(i) *monotonic,*

(ii) *has a unique global minimum and two local maxima at $x = 0$ and $x = 1$,*

(iii) *has a unique global maximum and two local minima at $x = 0$ and $x = 1$ or*

(iv) *has one local maximum and one local minimum in $(0,1)$, with a local maximum at $x = 0$ and a local minimum at $x = 1$ (see Figure 3(a)).*

Proof. By Lemma 5.1, either there exists an $x_0 \in (0,1)$ such that $P'(x_0) = 0$, there exist $x_0, x_0' \in (0,1)$ such that $P'(x_0) = P'(x_0') = 0$, or P' has no zeros in $(0,1)$. In the latter case P is monotonic, so let us suppose that P' has one or two zeros in $(0,1)$. From Lemma 5.2 there are eight cases to consider, initially supposing that all the zeros of r are transverse:

1. r has one zero in $(0,1)$ and $r(0) = 0, r(1) \ne 0$,
2. r has one zero in $(0,1)$ and $r(0) \ne 0, r(1) = 0$,
3. r has one zero in $(0,1)$ and $r(0) \ne 0, r(1) \ne 0$,
4. r has one zero in $(0,1)$ and $r(0) = 0, r(1) = 0$,
5. r has two zeros in $(0,1)$ and $r(0) \ne 0, r(1) = 0$,
6. r has two zeros in $(0,1)$ and $r(0) = 0, r(1) \ne 0$,
7. r has two zeros in $(0,1)$ and $r(0) \ne 0, r(1) \ne 0$ or
8. r has three zeros in $(0,1)$.

In cases (1-4) the mean-value theorem ensures that P' cannot have a zero in $(0,1)$ and P is therefore monotonic on $(0,1)$. In cases (5-7), using the arguments from Proposition 4.2 and Proposition 4.3, it follows that if P is not monotone then

Figure 3. A possible non-monotonic shape of solutions of (3.1) when (a) f is sigmoidal (convex–concave) and (b) f is sigmoidal (concave–convex). In both (a) and (b), r is shown in left-hand figure (the thick black line denotes zero) and the corresponding solution in the right-hand figure.

P either has a unique global minimum and two local maxima at $x = 0$ and $x = 1$ or has a unique global maximum and two local minima at $x = 0$ and $x = 1$.

In case (8) it is possible for P' not to have a zero in $(0, 1)$ and P is therefore monotone. However, if P' has one zero then it must have another one. Since f is sigmoidal convex–concave, r is sigmoidal concave–convex and therefore has *sign pattern* $- + - +$ on $[0, 1]$ as shown in Figure 3(a) (see left-hand diagram). Therefore there are x_1, x_2 and x_3 such that $0 < x_1 < x_2 < x_3 < 1$ and $r(x) > 0$ on (x_1, x_2) and $r(x) < 0$ on (x_2, x_3), while $r(x_1) = r(x_2) = r(x_3) = 0$. Suppose that the two zeros of P' in $(0, 1)$ are x_0 and x_0'. The mean value theorem ensures that $x_0 \in (x_1, x_2)$ and $x_0' \in (x_2, x_3)$. Since $\text{sgn}(P''(x)) = \text{sgn}(r(x))$, $x = x_0$ is a local minimum and $x = x_0'$ is a local maximum. The boundary conditions ensure that $x = 0$ is a local maximum and $x = 1$ is a local minimum, as described in case (iv) in the statement of the theorem.

If r has a non-transverse zero of multiplicity three then this situation is analogous to case (3) above. The case whereby r has one transverse and one double zero is inadmissible as then P would have two consecutive non-degenerate local minima. This covers all possible non-transverse cases by Lemma 5.2. $\qquad\square$

Proposition 5.4. *If (F1) and (F2) hold with $\Theta = 2$ and f is sigmoidal concave-convex, then any solution of (3.1) is either*

(i) *monotonic,*

(ii) *has a unique global minimum and two local maxima at $x = 0$ and $x = 1$,*

(iii) *has a unique global maximum and two local minima at $x = 0$ and $x = 1$ or*

(iv) *has one local maximum and one local minimum in $(0, 1)$, with a local minimum at $x = 0$ and a local maximum at $x = 1$ (see Figure 3(b)).*

Proof. The proof is entirely analogous to Proposition 5.3. $\qquad\square$

From the preceding sequence of results the following can be deduced directly.

Proposition 5.5. *If (F1) and (F2) hold then the solution P of (3.1) can only have two local maxima in $(0, 1)$ if $\Theta \geq 3$ (see Figure 4).*

Figure 4. A possible non-monotonic shape of solutions of (3.1) when f has a staircase shape, r is shown in left-hand figure (the thick black line denotes zero) and the corresponding solution in the right-hand figure.
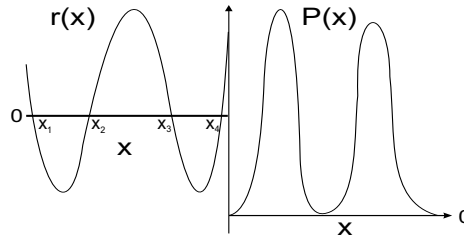
6. Concluding Remarks

Remark 1. The asymptotic analysis carried out in §4(a) gives a leading order approximation to the location of the modal phenotype for smooth, concave trade-offs, which are also ESSs in this context (for a definition of an ESS see Lawlor and Maynard-Smith, 1976). We can compare this with the location predicted by Gudelj *et al.* (2004) using the adaptive dynamics method and we now give the details of such a calculation.

Let us assume that f is a concave trade-off, so that from Gudelj *et al.* (2004), we know that the adaptive dynamics of (1.1) possesses a single evolutionary stable strategy (ESS) at some phenotypic location in $[0, 1]$. Let us consider the case where $i \in \{1, 2\}$, so that P_1 represents the resident and P_2 the mutant pathogen densities. The resulting four-dimensional system of ODEs has the mutant-free invariant set $\mathcal{M} := \{(P_1, P_2, H_1, H_2) \in \mathbb{R}^4 : P_2 = 0\}$, and when restricted to $\mathcal{M}$ this flow has the globally attractive steady-state $\mathbf{e} := (P_1(x_1), 0, H_1(x_1), H_2(x_1))$, where

$$H_1(x) = \frac{r_1}{\mu + xP_1(x)}, \quad H_2(x) = \frac{r_2}{\mu + f(x)P_1(x)}$$

and $P_1(x)$ is given by P in the equation

$$x \frac{r_1}{\mu + xP} + f(x) \frac{r_2}{\mu + f(x)P} - \gamma = 0. \tag{6.1}$$

Now, the introduction of a mutant corresponds to an examination of the flow of (1.1) when perturbed away from $\mathcal{M}$. However, $\mathcal{M}$ is unstable as an invariant manifold if $\mathbf{e}$ is unstable as an equilibrium, so we consider the right-most eigenvalue of the linearisation of (1.1) about $\mathbf{e}$. This is the quantity given by $\mathcal{F}(x_1, x_2)$, where

$$\mathcal{F}(x, y) := yH_1(x) + f(y)H_2(x) - \gamma;$$

$\mathcal{F}$ is often called the local fitness of the mutant.

From equation (3) of Lawlor and Maynard-Smith (1976), the phenotypic location of the ESS occurs where the rate of change of the fitness with respect to variations in the mutant phenotype is zero, namely at that value x which satisfies

$$\frac{\partial \mathcal{F}}{\partial y}(x, x) = 0.$$

Figure 5. Solutions of (3.1) for the *sigmoidal trade-off* $f(x) = \frac{1}{2}(1 - (2x - 1)^3)$. Here $\epsilon = 10^{-5}$, $r_1 = r_2 = 4$, $\mu = 1$ with $\gamma = 1$ (left) and $\gamma = 3\frac{1}{2}$ (right).
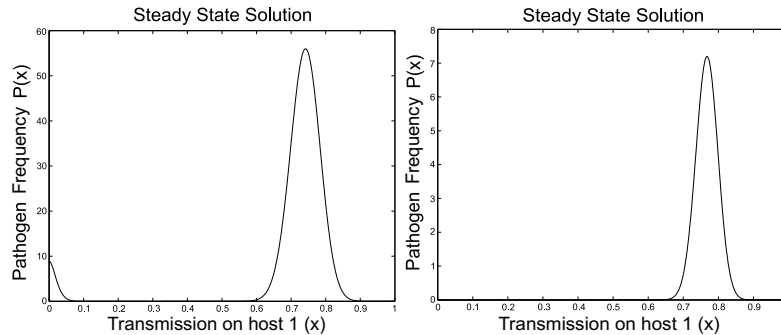
This is the equation

$$\frac{r_1}{\mu + xP} + f'(x)\frac{r_2}{\mu + f(x)P} = 0, \qquad (6.2)$$

where P is obtained from (6.1).

In §4(a) the location of the unique modal phenotype *was shown to be determined to leading order by (4.8) and (4.9) when the trade-off is concave, but these equations are identical to (6.2) and (6.1) respectively. Thus we have deduced that the location of the ESS obtained in Gudelj* et al. *which uses the fitness maximisation procedure of Lawlor and Maynard-Smith (1976), agrees to lowest order with our asymptotic findings, in the context of concave trade-off functions. Moreover, our classification also accords with the findings of Gudelj* et al. *even in the cases where f is not concave, but convex and even sigmoidal, although our results are more general because Gudelj* et al. *often rely on specific choices of trade-off function.*

Remark 2. Although the classification of phenotypic diversity given in this paper relies heavily on the second derivative of the trade-off, it only provides an upper bound on the number of modal phenotypes that exist for a given trade-off. Changes in other physical parameters could lead to a diminishing of phenotypic diversity. Consider, for example, Figure 5 (left), where for small γ the pathogen has a bimodal distribution with a specialist phenotype situated on the left boundary at $x = 0$. As γ increases the modal phenotype at this phenotype vanishes and the pathogen population becomes unimodal (see Figure 5 (right)) and infective to both hosts. In this case the trade-off is sigmoidal, but the same comment applies to other shapes too.

Remark 3. It would be a laudable, if intractable, goal to completely describe the set of modal phenotypes as a function of t for a given initial phenotypic distribution. To formalise a discussion in this direction, let us denote by $\sigma(t) \subset [0, 1]$ the set of modal phenotypes at time t, that is

$$\sigma(t) = \{x \in [0, 1] : P_x(x, t) = 0 \text{ and } P_{xx}(x, t) < 0\}.$$

If there is a $T > 0$ and a sequence $\epsilon_n \downarrow 0$ such that $\#\sigma(T + \epsilon_n) > \#\sigma(T - \epsilon_n)$ for all n (where a hash (#) denotes set cardinality) then *diversification* is said to

Figure 6. A schematic of possible solutions of $P_x(x,t) = 0$, plotting phenotype (x) against time (t). Gaussian-like bumps are centred on the curves of maxima (where $P_{xx}(x,t) < 0$) which can be created and destroyed in bifurcations where local maxima and minima of $P(\cdot,t)$ merge: a pitchfork bifurcation leads to evolutionary branching (left) and hysteresis leads to an evolutionary leap (right).
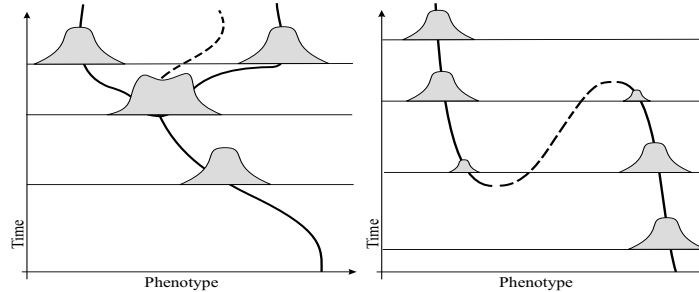
have taken place at $t = T$. On the other hand, if $\#\sigma(T + \epsilon_n) < \#\sigma(T - \epsilon_n)$ then *extinction* is said to have occurred.

Now, a function $s : [0, \infty) \to [0,1]$ such that $s(t) \in \sigma(t)$ for all $t > 0$ is called a selection function. If diversification takes place at $t = T$ and there exists a selection of $\sigma(\cdot)$ which is discontinuous, then an evolutionary *leap* is said to have taken place at $t = T$. On the other hand, if every selection function of $\sigma(\cdot)$ is continuous at $t = T$ then evolutionary *branching* is said to have taken place.

Different diversification and extinction events can be characterised according to bifurcations of the smooth bifurcation problem $P_x(x,t) = 0$, with t playing the role of parameter. It is clear that both evolutionary branching *and* leaping events can occur in smooth solutions of (HP) and are caused by well-known codimension-1 singularities. For instance, evolutionary branching occurs when $P_x(x,t) = 0$ has a pitchfork bifurcation, whereas an evolutionary leap occurs when the same equation has a saddle-node bifurcation point (see Figure 6 for pictorial representations of both cases). From this observation, it is apparent that evolutionary leaping ought to be the generic diversification mechanism and robust to perturbations in the initial state of the phenotypic distribution, whereas evolutionary branching should be a rare, non-generic event. A related discussion can be found in Stewart (2003) where the author reaches a similar conclusion but starting from different modelling assumptions. We also note that an evolutionary leap has some resonance with Gould's concept of punctuated equilibrium.

References

Bonhoeffer, S. & Nowak, M. A. 1994 Mutation and the evolution of virulence. *Proc. R. Soc. Lond.* B **258**, 133–140.

Boots, M. & Begon, M. 1993 Trade-offs with resistance to a granulosis virus in the Indian meal moth, examined by a laboratory evolution experiment. *Func. Ecol.* **7**, 528–534.

Bowers, R. G. & Hodgkinson, D. E. 2001 Community Dynamics, Trade-offs, Invasion Criteria and the Evolution of Host Resistence to Microparasites. *J. Theor. Biol.*, **212**, 315–331.

Calsina, A. & Cuadrado, S. 2004 Small mutation rate and evolutionary stable strategies in infinite dimensional adaptive dynamics *J. Math. Biol.*, **48**, 135–159.

Calsina, A. & Perelló, C. 1995 Equations for biological evolution. *Proc. R. Soc. Edin.*, **125**, 939–958.

Calsina, A., Perelló, C. & Saldaña, J. 1994 Non-local reaction-diffusion equations modelling predator-prey coevolution *Pub. Mat.*, **38**, 315–325.

Geritz, S. H. A. , Kisdi, E. , Meszena, G. & Metz, J. A. J. 1998 Evolutionary singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.*, **12**, 35–57.

Gudelj, I., van den Bosch, F. & Gilligan, C. A. 2004 Transmission rates and adaptive evolution of pathogens in sympatric heterogeneous plant populations. *Proc. R. Soc. Lond.* B **271**, 2187–2194. (DOI: 10.1098/rspb.2004.2837)

Henry, D. 1981 *Geometric Theory of Semilinear Parabolic Equations*, New York: Springer Verlag.

Kraaijeveld, A. R. & Godfray, H. C. J. 1997 Trade-off between parasitoid resistance and larval competitive ability in *Drosophila melanogaster. Nature*, **389**, 278–280.

Lawlor, L. & Maynard-Smith, J. 1976 The coevolution and stability of competing species. *Am. Nat.*, **110**, 79–99.

Levins, R. 1968 *Evolution in Changing Environments*, New Jersey: Princeton University Press.

Messenger, L. S., Molineux, I. J. & Bull J. J. 1999 Virulence evolution in a virus obeys a trade-off. *Proc. R. Soc. Lond.* B **266**, 397–404.

Saldaña, J., Elena, S. F. & Solé, R. V. 2003 Coinfection and superinfection in RNA virus populations: a selection-mutation model *Math. Biosci.*, **183**, 135–160.

Sasaki, A. & Godfray, H. C. J. 1999 A model for the coevolution of resistance and virulence in coupled host-parasitoid interactions. *Proc. R. Soc. Lond.* B **266**, 455–463.

Stewart, I. 2003 Self-organization in evolution: a mathematical perspective. *Phil. Trans. R. Soc. Lond.* A **361**, 1101–1123.

Tsmiring, L. S. & Levine, H. 1996 RNA Virus Evolution via a Fitness-Space Model. *Phys. Rev. Lett.*, **76**, 4440–4443.

Zeidler, E. 1986 *Nonlinear Functional Analysis and Its Applications: Part IV, Fixed Point Theorems*, Volume 1. New York: Springer-Verlag.